Imperial College London

Department of Mathematics
Imperial College London

On the Theory and Applications of Stochastic Gradient Descent in Continuous Time

Louis Sharrock

A thesis submitted to Imperial College London
for the degree of Doctor of Philosophy

Abstract

Stochastic optimisation problems are ubiquitous across machine learning, engineering, the natural sciences, economics, and operational research. One of the most popular and widely used methods for solving such problems is stochastic gradient descent. In this thesis, we study the theoretical properties and the applications of stochastic gradient descent in continuous time.

We begin by analysing the asymptotic properties of two-timescale stochastic gradient descent in continuous time, extending well known results in discrete time. The proposed algorithm, which arises naturally in the context of stochastic bilevel optimisation problems, consists of two coupled stochastic recursions which evolve on different timescales. Under weak and classical assumptions, we establish the almost sure convergence of this algorithm, and obtain an asymptotic convergence rate.

We next illustrate how the proposed algorithm can be applied to an important problem arising in continuous-time state-space models: joint online parameter estimation and optimal sensor placement. Under suitable conditions on the process consisting of the latent signal process, the filter, and the filter derivatives, we establish almost sure convergence of the online parameter estimates and optimal sensor placements generated by our algorithm to the stationary points of the asymptotic log-likelihood of the observations, and the asymptotic covariance of the state estimate, respectively. We also provide extensive numerical results illustrating the performance of our approach in the case that the hidden signal is governed by the two-dimensional stochastic advection-diffusion equation, a model arising in many meteorological and environmental monitoring applications.

In the final part of this thesis, we introduce a continuous-time stochastic gradient descent algorithm for recursive estimation of the parameters of a stochastic McKean-Vlasov equation equation, and the associated system of interacting particles. Such models arise in a variety of applications, including statistical physics, mathematical biology, and the social sciences. We prove that our estimator converges in $\mathbb{L}^1$ to the stationary points of the asymptotic log-likelihood of the McKean-Vlasov SDE in the joint limit as $t \to \infty$ and the number of particles $N \to \infty$, under suitable conditions which guarantee ergodicity and uniform-in-time propagation of chaos. We also establish, assuming also strong concavity for the asymptotic log-likelihood, an $\mathbb{L}^2$ convergence rate to the unique maximiser of this asymptotic log-likelihood function. Our theoretical results are demonstrated via a range of numerical examples, including a stochastic Kuramoto model and a stochastic opinion dynamics model.

Copyright Declaration

Declaration

I hereby certify that this thesis, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Louis Sharrock
May 2022

Acknowledgements

First and foremost, I would like to thank my supervisor, Nikolas Kantas. I have greatly benefited from his ideas, insight, and mentorship, without which the work in this thesis would not have been possible. His unwavering support, guidance, patience, and encouragement have been absolutely invaluable. I am proud to call Nikolas a friend.

I would also like to thank Panos Parpas and Greg Pavliotis, with whom I have greatly enjoyed collaborating over the last year, and whose ideas and feedback have been hugely valuable. I also owe thanks to Dan Crisan and Alistair Forbes for many interesting and insightful discussions.

I would also like to thank all of my friends and colleagues at Imperial College London, and particularly those in the Mathematics of Planet Earth Centre for Doctoral Training, for creating such an enjoyable and stimulating research environment. I am hugely grateful for their dedication, warmth, and generosity. I would also like to acknowledge the EPSRC and the National Physical Laboratory for funding this PhD.

On a more personal level, I am incredibly grateful to all of my friends, who have made the last three years infinitely more enjoyable. And above all, to my family, for their unconditional love, support, and encouragement. None of this would have been possible without them.

Publications

The material in the following papers appears in this thesis:

L. Sharrock and N. Kantas. Two-Timescale Stochastic Gradient Descent in Continuous Time with Applications to Joint Online Parameter Estimation and Optimal Sensor Placement. To appear in *Bernoulli*, 28(3), 2022.

L. Sharrock and N. Kantas. Joint Online Parameter Estimation and Optimal Sensor Placement for the Partially Observed Stochastic Advection-Diffusion Equation. *SIAM / ASA Journal on Uncertainty Quantification*, 10(1): 55-95, 2022.

L. Sharrock, N. Kantas, P. Parpas, and G.A. Pavliotis. Parameter Estimation for the McKean-Vlasov Stochastic Differential Equation. In submission to *Stochastic Processes and their Applications*, 2021. arXiv preprint: 2106.13751.

L. Sharrock. Two-Timescale Stochastic Approximation in Continuous Time: A Central Limit Theorem. In preparation for *Electronic Communications in Probability*, 2022.

The material in the following paper does not appear in this thesis:

C. Leadbeater*, L. Sharrock*, B. Coyle, and M. Benedetti. F-Divergences and Cost Function Locality in Generative Modelling with Quantum Circuits. *Entropy*, 23 (1281), 2021.

*Denotes equal first authors.

Table of Contents

List of Figures

List of Tables

1

Introduction

1.1 Background

Countless problems across machine learning, engineering, the natural sciences, economics, and operational research, involve the task of mathematical optimisation. That is, the task of obtaining $\alpha^* \in \Lambda \subseteq \mathbb{R}^d$ such that

$$\alpha^* = \arg\min_{\alpha \in \Lambda} f(\alpha) \tag{1.1}$$

where $f : \mathbb{R}^d \to \mathbb{R}$ is a suitably defined objective function. In many cases, this function cannot be computed directly; rather, its values are only available via noise corrupted observations. This is the topic of stochastic optimisation.

There are various methods available for solving problems of this kind (e.g., [185, 425]). One of the most popular and widely applicable of these is stochastic gradient descent, a stochastic optimisation method of the stochastic approximation type. Initialised at $\alpha_0 \in \mathbb{R}^d$, stochastic gradient descent methods generate a sequence of estimates $\{\alpha_n\}_{n\geq 0}$ according to the recursion

$$\alpha_{n+1} = \alpha_n - \gamma_n h(\alpha_n, \xi_n) \tag{1.2}$$

where $\{h(\alpha_n, \xi_n)\}_{n\geq 0}$ represents a sequence of noisy estimates of the gradients $\{\nabla f(\alpha_n)\}_{n\geq 0}$, $\{\xi_n\}_{n\geq 0}$ is a sequence of $\mathbb{R}^d$-valued random variables to be interpreted as noise, and $\{\gamma_n\}_{n\geq 0}$ is a sequence of positive real step-sizes.

Example 1. Let $\{\xi_n\}_{n\geq 0}$ be a sequence of independent identically distributed (i.i.d.)

random variables with common law μ. Suppose that the objective function to be optimised is given by $f(\alpha) = \mathbb{E}_\mu[F(\alpha, \xi)]$, with $\nabla f(\alpha) = \mathbb{E}_\mu[\nabla F(\alpha, \xi)]$. For any $n \geq 0$, suppose that it is possible to observe $\nabla F(\alpha_n, \xi_n)$, an unbiased estimate of $\nabla f(\alpha_n)$. Then the stochastic gradient descent algorithm for optimising $f(\alpha)$ is given by

$$\alpha_{n+1} = \alpha_n - \gamma_n \nabla F(\alpha_n, \xi_n). \tag{1.3}$$

Example 2. Let $\{\xi_n\}_{n\geq 0}$ be a sequence of i.i.d. random variables, each with distribution $\mathcal{U}\{1, N\}$. Suppose that the objective function to be optimised is given by $f(\alpha) = \sum_{i=1}^N f_i(\alpha)$, with $\nabla f(\alpha) = \sum_{i=1}^N \nabla f_i(\alpha)$. For any $n \geq 0$, suppose that it is only possible to observe $\nabla f_{\xi_n}(\alpha_n)$, which represents a single component of $\nabla f(\alpha_n)$. Then the stochastic gradient descent algorithm for estimating $f(\alpha)$ is given by

$$\alpha_{n+1} = \alpha_n - \gamma_n \nabla f_{\xi_n}(\alpha_n) \tag{1.4}$$

The theory of stochastic approximation was initiated in the early 1950s through the pioneering work of Robbins and Monro [396], who proposed a recursive algorithm to obtain the root(s) of an unknown function. This is often referred to as the Robbins-Monro (RM) algorithm. This approach was later applied to the setting of stochastic optimisation by solving for the zero of the gradient of an objective function. The gradient free setting was subsequently addressed by Kiefer and Wolfowitz [245], who proposed an alternative stochastic approximation algorithm - the Kiefer-Wolowitz (KW) algorithm - based on estimating the gradient using finite differences.

The key insight of Robbins and Monro was that if the sequence of step-sizes $\{\gamma_n\}_{n\geq 0}$ is chosen such that

$$\gamma_n \to 0 \quad , \quad \sum_{n=0}^{\infty} \gamma_n = \infty \quad , \quad \sum_{n=0}^{\infty} \gamma_n^2 < \infty, \tag{1.5}$$

then there is an averaging effect which eliminates the effect of the noise in the long run. Under some additional assumptions, it is then possible to show that $\alpha_n \to \alpha^*$ as $n \to \infty$ in $\mathbb{L}_2$ and thus in probability [245, 396]. Following this seminal result, there were significant advances in both the theory and applications of stochastic approximation. The original RM and KW algorithms applied to one-dimensional problems, but were subsequently extended by Blum to the multi-dimensional case [55, 56]. In addition, the conditions used to obtain convergence for both algorithms were weakened to obtain convergence in probability [471] and with probability one [55].

Many weaker conditions have since been obtained for the almost sure (a.s.) convergence of stochastic approximation algorithms, which apply in rather more general settings than the i.i.d. noise case considered in these early papers. These results have largely evolved out of two general approaches: a purely probabilistic approach, usually based on martingale theory (e.g., [46, 314, 402, 484]), and the 'ODE approach', originally due to Ljung [310] (see also [44, 265]). We will explore both of these approaches later in this thesis. A third, purely deterministic approach based on deterministic conditions on the noise sequence has also been proposed by Kulkarni and Horn [260] and Delyon [147].

Once the convergence of a stochastic approximation algorithm has been established, the natural next step is to obtain its convergence rate. General results on the asymptotic distribution of the stochastic approximation iterate were obtained by Fabian [175] who demonstrated that the iterate was asymptotically normal with rate of convergence $O(n^{-1/2})$. These results extended the earlier work of Chung [112] and Sacks [403]. For other relevant results, we refer also to [267, 269].

Other more recent advances in the theory and applications of stochastic approximation (and stochastic gradient descent) include asymptotic efficiency (e.g., [270, 380, 381, 401, 482]), finite-time convergence performance (e.g., [199] and references therein), the development of new algorithms, and modifications of existing ones. For a comprehensive account of such results, we refer to any one of a number of excellent monographs, including the books of Albert and Gardner [8], Wasan [463], Tsypkin [448], Nevel'son and Khasminskii [361], Kushner and Clark [272], Benveniste, Metivier and Priouret [44], Duflo [169], Solo and Kong [424], Benaim [36], Spall [425], Chen [107], and Kushner and Yin [265, 271], and Borkar [62]. For some more recent results, we point towards the work of Tadic [439, 440], Bottou [66], and Karimi et al. [240].

With some notable exceptions, the vast majority of the literature on stochastic approximation is formulated in discrete time. In this thesis, we focus instead on the continuous-time setting. In this framework, the RM algorithm takes the form

$$\alpha_t = \alpha_0 - \int_0^t \gamma_s \mathrm{d}h(\alpha_s, \xi_s), \tag{1.6}$$

where now $\{h(\alpha_s, \xi_s)\}_{s\geq 0}$ represents a continuous sequence of noisy estimates of the integrals $\{\int_0^s \nabla f(\alpha_u)\mathrm{d}u\}_{s\geq 0}$, $\{\xi_s\}_{s\geq 0}$ is an $\mathbb{R}^d$-valued continuous-time stochastic process which corresponds to the measurement noise, and $\gamma_s : \mathbb{R}_+ \to \mathbb{R}_+$ is a positive function known as the learning rate. This algorithm represents the natural continuous-time analogue of the discrete-time RM procedure, something which is clear upon rewriting the discrete-time recursion (1.2) in the form

$$\alpha_n = \alpha_0 - \sum_{i=0}^{n-1} \gamma_i h(\alpha_i, \xi_i) \tag{1.7}$$

The first rigorous treatment of stochastic approximation algorithms in continuous time was provided by Nevel'son and Khasminskii [361], who considered an algorithm of the form (1.6) in the case that

$$h(\alpha_s, \xi_s) = \int_0^s \nabla f(\alpha_u)\mathrm{d}u + \int_0^s \sigma_u(\alpha_u)\mathrm{d}w_u, \tag{1.8}$$

where $\sigma_u : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is a Borel measurable function, and $\{w_u\}_{u\geq 0}$ denotes an $\mathbb{R}^d$-valued Brownian motion. Under relatively weak assumptions, these authors established, using probabilistic arguments, a.s. convergence to the set $\{\alpha : \nabla f(\alpha) = 0\}$, a.s. and $\mathbb{L}^2$ convergence rates, as well as asymptotic normality. Under different assumptions, Sen and Athreya [411] later also established convergence of this algorithm, both a.s. and in $\mathbb{L}^2$.

Following these early results, Chen [103, 104, 105, 106, 108] made several significant contributions to the theory of continuous-time stochastic approximation. Using an approach which combined the probabilistic method and the ODE method, he obtained sufficient conditions for the a.s. convergence of continuous-time RM and KW procedures in the case that the measurement error is a process with dependent increments expressed as the output of a linear system driven by an Ito integral [103, 104], generalising the results in [361]. He later also demonstrated asymptotic normality in this case [105]. In [108], Chen and coworkers established the a.s. convergence of a modified continuous-time stochastic approximation procedure containing randomly varying truncations, under significantly weaker conditions than before. Finally, in [106], asymptotic efficiency was established for the time-averaged estimates of the form $\bar{\alpha}_t = \frac{1}{t}\int_0^t \alpha_s \mathrm{d}s$. Similar results, under slightly different conditions on the noise process, were also independently obtained by Yin and Gupta [483].

Several authors have also considered continuous-time stochastic approximation for the problem of recursive estimation, which will be a central theme of this thesis.[1] This problem can naturally be formulated as a stochastic optimisation problem, in which the objective function is given by the negative log-likelihood of the observed data. Aside from the book of Nevel'son and Khasminskii [361], one of the first results along these lines is due to Gerencser, Gyongy, and Michaletzky [196], who established the a.s. convergence of a recursive estimator for the parameters of a partially observed linear diffusion process. This analysis was later extended in [197]. A similar convergence result was obtained by DeWolf and Wiberg [157] via the ODE method. More recently, Sirignano and Spiolopolous [420, 422] established the a.s. convergence, $\mathbb{L}^2$ convergence rate, and asymptotic normality, of a continuous-time stochastic gradient descent method for recursively estimating the parameters of a fully observed diffusion process. The a.s. convergence result has since been extended to partially observed [430] and jump [50] diffusion processes.

While it is somewhat beyond the scope of this thesis, it would be remiss of us to conclude

[1] There is, unsurprisingly, also a substantial body of literature on recursive estimation in discrete time. We do not attempt to review this literature here, but refer to [312, 315] for some classical references.

this introduction without mentioning the work of Mel'nikov and Valkeila [341, 342, 343, 344, 456], who proposed a generalised stochastic approximation procedure which unified the discrete-time and continuous-time frameworks. In particular, these authors considered noisy observations of the form

$$h(\alpha_s, \xi_s) = \int_0^s \nabla f(\alpha_{u-}) \mathrm{d}a_u + \int_0^s \sigma_u(\alpha_{u-}) \mathrm{d}m_u \tag{1.9}$$

where $\{a_u\}_{u\geq 0}$ is a predictable, increasing càdlàg process, $\{m_u\}_{u\geq 0}$ is a locally square integrable martingale, and $\alpha_{u-} = \lim_{v\uparrow u} \alpha_v$. This defines the estimate $\{\alpha_t\}_{t\geq 0}$ as the strong solution of a stochastic integral equation with respect to a semi-martingale (see, e.g., [163, 174, 188, 331, 339, 340] for relevant results on the existence and uniqueness of such solutions). In the case that $a_u = u$ and $m_u = w_u$, one recovers the continuous-time stochastic gradient descent algorithm in (1.8). On the other hand, setting $a_u = \lfloor u \rfloor$, and defining a suitable discrete-time martingale, one can recover the original discrete-time RM procedure in (1.2) (e.g., [343]). In [343, 456], the asymptotic properties (a.s. convergence, asymptotic normality) of this procedure, as well as an averaged version of this procedure, were obtained under rather strong and technical conditions based on stochastic Lyapunov arguments. The same properties were later established by Lazrieva and coworkers, under somewhat weaker and more natural conditions, for an even more general version of this algorithm [283, 284, 285, 286, 287, 288, 447].

There are several reasons for considering stochastic approximation algorithms in continuous time. Firstly, continuous-time algorithms have not been studied nearly as widely as their discrete-time counterparts. Thus, inevitably, there are many results in discrete time which are yet to be established rigorously in continuous time. It is natural to ask whether such results can be extended to continuous-time, and whether any differences arise in this setting. Secondly, studying algorithms in continuous time can lead to new perspectives on discrete-time algorithms (see, e.g., [370]). Thirdly, one can often leverage powerful existing tools (e.g., the Itô calculus) to study the convergence properties of continuous-time algorithms. This can lead to cleaner proofs, as well as insights into the corresponding proofs in discrete time.

In addition to the mathematical interest, continuous-time stochastic approximation algorithms are the natural choice for solving continuous-time optimisation problems, which arise in many applications. For example, models in engineering, finance, and the natural sciences are often formulated in continuous time. There are often unknown parameters or functions in such models, which one may wish to recursively estimate from the continuous stream of data. This results in a continuous-time optimisation problem, since the objective function to be optimised (e.g., the negative log-likelihood), and its gradients, are defined in continuous time (i.e., in terms of the continuous-time process). Beyond recursive estimation, continuous-time optimisation problems are also common in stochastic optimal control (e.g., [378]) and reinforcement learning (e.g., [166, 481]). Regardless of the specific application, it is clear that the theoretically correct statistical learning equation in any of these cases must be defined in continuous time. Indeed, any statistical learning equation defined in discrete time can only be approximate, since it must rely on approximations for the objective function and its gradients derived from an (approximate) discretisation of the continuous-time model.

We should emphasise, at this point, that the continuous-time algorithms studied in this thesis are not directly applicable to discrete-time stochastic optimisation problems. Instead, as outlined above, the algorithms which we will analyse are designed for continuous-time optimisation problems. It is worth highlighting, however, that there is a growing body of literature which considers continuous-time stochastic gradient descent algorithms as an approximation to their discrete-time counterparts (e.g., [177, 298, 299, 330, 370, 470]). Indeed, continuous-time algorithms have long been viewed, at least formally, as good approximations to their discrete-time analogues in cases where the sampling is very frequent (e.g., [296, 483]). There is also a long tradition of deriving discrete-time stochastic optimisation algorithms from continuous-time dynamics (e.g., [111, 191, 192] for some classical references, and [347, 388, 475, 476] for some more recent contributions). While both of these directions are somewhat tangential to the line of work pursued in this thesis, we highlight them here with the expectation that some of the theoretical results obtained herein (and, in particular, in Chapter 2) may also be relevant in these contexts.

In practice, it is evident that any stochastic gradient scheme in continuous time must be discretised. Thus, when designing statistical learning algorithms for continuous-time optimisation problems, it is natural to ask why we prefer to use a discrete-time approximation of a continuous-time stochastic gradient descent algorithm (the 'continuous-time approach') over the traditional approach, which first discretises the continuous-time model, and then applies a classical discrete-time stochastic gradient descent algorithm (the 'discrete-time approach'). Providing a satisfactory answer to this question, namely, a detailed comparison of the relative advantages and disadvantages of these two approaches, is beyond the scope of this thesis. Nonetheless, let us provide some brief motivation for the continuous-time approach.

Firstly, this approach may overcome problems which arise when using the discrete-time approach, particular when the sampling rate increases (e.g., [355, 483]). These include ill conditioning [405], biased estimates [102, 420], or even divergence [355, 420]. One well known example, which highlights the challenges associated with first discretising the system dynamics, is the problem of estimating the value function associated with a continuous-time Markov (decision) process. This problem is central to continuous-time reinforcement learning (e.g., [20, 166, 246, 359, 481]), and also commonly arises in financial applications (e.g., [378, 420]). In this case, the continuous-time approach corresponds to applying continuous-time stochastic gradient descent to an objective function based on the Hamilton-Jacobi-Bellman equation [166, 420]. Meanwhile, the discrete-time approach corresponds to discretising the system dynamics, and then applying discrete-time stochastic gradient descent to a cost function based on the approximate discrete-time Bellman equation (see, e.g., [431]). One can show that the continuous-time approach is unbiased, while the discrete-time approach is biased. Moreover, as the time step size decreases (i.e., the sampling rate increases), the discrete-time approach can explode (see [420] for further details). Finally, numerical results indicate that the continuous-time approach can result in significantly faster convergence [166].

Another advantage of the continuous-time approach is that it allows one to apply any appropriate numerical discretisation scheme to the theoretically correct statistical learning equations. This can lead to entirely new discrete-time algorithms, with improved convergence properties (see, e.g., [282, 298, 299, 475, 476]). It can also be more computationally efficient, particularly when the dimensions of the model are significantly larger than the number of model parameters (see Chapter 4). This is common in large scale reinforcement learning problems [166]. Finally, it enables direct control of the numerical error of the resulting algorithm, which can result in more accurate and robust parameter updates. Indeed, there is no guarantee that discretising the model dynamics using a numerical scheme with certain numerical properties (e.g., higher order accuracy in time), and then applying traditional stochastic gradient descent, will result in a statistical learning algorithm which also has these properties. Conversely, the desired numerical properties will certainly hold if one applies the discretisation of choice directly to the continuous-time learning equation.

1.2 Contributions & Thesis Organisation

In this thesis, we make several contributions to the theory and applications of stochastic gradient descent in continuous time. The main contributions of each chapter are summarised below.

In **Chapter 2**, we analyse the asymptotic properties of two-timescale stochastic gradient descent in continuous time, extending well known results in discrete time. This algorithm, which arises naturally in the context of bilevel optimisation, consists of two coupled stochastic recursions which evolve on different timescales. Under relatively weak and classical assumptions, we establish the a.s. convergence of this algorithm in continuous time. Our analysis covers algorithms with both additive, state-dependent noise, and those with non-additive, state-dependent noise. Our proof of this result closely follows the classical ODE method, adapted appropriately to the continuous-time setting. We also obtain the asymptotic convergence rate of the proposed algorithm.

In **Chapter 3**, we illustrate how the continuous-time, two-timescale stochastic gradient descent algorithm analysed in Chapter 2 can be applied to an important problem arising in continuous time state-space models. The problem of interest is joint online parameter estimation and optimal sensor placement. Our approach represents a significant departure from the existing literature, in which these two problems have, until now, been studied separately. We first illustrate in detail how this problem can be formulated as a bilevel optimisation problem, with objective functions given by the asymptotic log-likelihood of the observations and the trace of the asymptotic filter covariance. Then, under suitable conditions on the process consisting of the latent signal process, the filter, and the filter derivatives, we establish a.s. convergence of the online parameter estimates and optimal sensor placements to the stationary points of these two objective functions.

In **Chapter 4**, we demonstrate how the methodology in Chapter 3 can be applied to the partially observed stochastic advection-diffusion partial differential equation, an equation which arises in many meteorological and environmental modelling applications. This represents a formal extension of the joint online parameter estimation and optimal sensor placement algorithm introduced in Chapter 3 to the case in which the hidden state is infinite-dimensional. We also provide extensive numerical results illustrating the performance of this method in different scenarios of practical interest.

In **Chapter 5**, we finally turn our attention away from two-timescale stochastic gradient descent in continuous time. We propose a continuous-time (single-timescale) stochastic gradient descent algorithm for online estimation of the parameters of a McKean-Vlasov stochastic differential equation (SDE), and the associated system of interacting particles. We prove that this estimator converges in $\mathbb{L}^1$ to the stationary points of the asymptotic log-likelihood of the McKean-Vlasov SDE in the joint limit as $t \to \infty$ and the number of particles $N \to \infty$, under suitable conditions which guarantee ergodicity and uniform-in-time propagation of chaos. We also establish, assuming also strong concavity for the asymptotic log-likelihood, an $\mathbb{L}^2$ convergence rate to the unique maximiser of this asymptotic log-likelihood function.

In **Chapter 6**, we provide some concluding remarks, summarising our main contributions and outlining some potential areas for future research.

2

Asymptotic Properties of Two-Timescale Stochastic Gradient Descent in Continuous Time

Summary. *In this chapter, we establish the almost sure convergence of two-timescale stochastic gradient descent algorithms in continuous time under general noise and stability conditions, extending well known results in discrete time. We analyse both algorithms with additive, state-dependent noise and those with non-additive, state-dependent noise. In the non-additive case, our analysis is carried out under the assumption that the noise is a continuous-time Markov process, controlled by the algorithm states. In the additive case, we also establish the weak convergence rate of the two-timescale stochastic gradient descent algorithm. The obtained results cover a broad class of highly non-linear two-timescale stochastic gradient descent algorithms in continuous time.*

2.1 Introduction

Many modern problems in engineering, the sciences, economics, and machine learning, involve the optimisation of two or more interdependent performance criteria. These include, among others, unsupervised learning [214], reinforcement learning [241, 250], meta-learning [389], game theory [399], and hyper-parameter optimisation [182]. In this chapter, we formulate such problems as unconstrained bilevel optimisation problems, in which the objective is to obtain $\alpha^* \in \Lambda_\alpha \subseteq \mathbb{R}^{d_1}$, $\beta^*(\alpha^*) \in \Lambda_\beta \subseteq \mathbb{R}^{d_2}$, such that

$$\alpha^* \in \underset{\alpha \in \Lambda_\alpha}{\arg\min}\, f\big(\alpha, \beta^*(\alpha)\big) \quad , \quad \beta^*(\alpha) \in \underset{\beta \in \Lambda_\beta}{\arg\min}\, g(\alpha, \beta) \tag{2.1}$$

where $f, g : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}$ are continuously differentiable functions, and Λ_α, Λ_β are closed subsets of $\mathbb{R}^{d_1}$, $\mathbb{R}^{d_2}$, respectively. We will assume, as in many applications, that we only have access to noisy estimates of f and g.

There are, unsurprisingly, several significant challenges associated with this optimisation problem. Firstly, in order to evaluate the upper-level objective function, $f(\cdot, \cdot)$, one must obtain the global minimiser of the lower-level objective function $g(\alpha, \cdot)$, for all $\alpha \in \Lambda_\alpha$. This may be very difficult, particularly if $g(\alpha, \cdot)$ is a complex function. In many practical applications of interest, one or both of the objective functions may be prohibitively costly to compute (e.g., they may depend on very high-dimensional data), which compounds this problem. Secondly, it may not be possible to compute the gradient of the function $\beta^*(\alpha)$. Thus, even if we could obtain $\beta^*(\alpha)$ and evaluate $f(\alpha, \beta^*(\alpha))$ for all $\alpha \in \Lambda_\alpha$, it would not be possible to solve the upper-level optimisation problem directly using gradient-based methods.

In practice, and with these considerations in mind, it is typical to consider a slightly weaker optimisation problem, in which the objective is to obtain α^*, β^* such that, simultaneously, α^* locally minimises $f(\cdot, \beta^*)$, and β^* locally minimises $g(\alpha^*, \cdot)$. That is, such that

$$\alpha^* = \underset{\alpha \in U_{\alpha^*}}{\arg\min}\, f(\alpha, \beta^*) \quad , \quad \beta^* = \underset{\beta \in U_{\beta^*}}{\arg\min}\, g(\alpha^*, \beta) \tag{2.2}$$

where $U_{\alpha^*} \subset \Lambda_\alpha$ and $U_{\beta^*} \subset \Lambda_\beta$ are local neighbourhoods of α^* and β^*, respectively. We will not assume any form of convexity, and thus we weaken this objective further, seeking values of α^* and β^* which satisfy the following local stationarity condition

$$\nabla_\alpha f(\alpha^*, \beta^*) = 0 \quad , \quad \nabla_\beta g(\alpha^*, \beta^*) = 0. \tag{2.3}$$

In this chapter, we analyse the use of gradient methods for this problem, under the assumption that we continuously observe noisy estimates of these gradients.

A natural candidate for a solution to this class of bilevel optimisation problems is two-timescale stochastic gradient descent. As outlined in the introduction, stochastic gradient descent is a sequential method for determining the minima or maxima of an objective function whose values are only available via noise-corrupted observations (e.g., [44, 62, 107, 271], and references therein).

Two-timescale stochastic gradient descent algorithms represent one of the most important and complex subclasses of stochastic gradient descent methods. These algorithms consist of two coupled recursions, which evolve on different timescales (e.g., [61, 62, 241, 437]). In particular, the step-sizes of the 'slow' recursion are considerably smaller than the step sizes of the 'fast' recursion. They can thus be considered as singularly perturbed SDEs.[1] in discrete time, this approach has found success in a wide variety of applications, including deep learning [214], reinforcement learning [29, 241, 249, 252, 432, 433], signal processing [49], power control in wireless networks [317], admission control in communication networks [48], optimisation [162, 462], and statistical inference [480], to name but a few. Consequently, the analysis of its asymptotic properties has been the subject of a large number of papers (e.g., [61, 62, 241, 251, 252, 354, 437, 442]).

Although these papers provide an excellent insight, they only explicitly consider two-timescale algorithms in discrete time. Indeed, to the best of our knowledge, there are no existing works which explicitly consider the a.s. convergence, or the convergence rate, of two-timescale stochastic gradient descent algorithms in continuous time, viz,

$$\mathrm{d}\alpha_t = -\gamma_t^1 \left[\nabla_\alpha f(\alpha_t,\beta_t)\mathrm{d}t + \mathrm{d}\xi_t^1\right], \tag{2.4a}$$

$$\mathrm{d}\beta_t = -\gamma_t^2 \left[\nabla_\beta g(\alpha_t,\beta_t)\mathrm{d}t + \mathrm{d}\xi_t^2\right], \tag{2.4b}$$

where $\{\gamma_t^i\}_{t\geq 0}$, $i=1,2$ are learning rates; and $\{\xi_t^i\}_{t\geq 0}$, $i=1,2$, are additive, possibly state-dependent noise processes.[2] Even upon restriction to the single timescale case, asymptotic results for continuous-time stochastic approximation are somewhat sparse, and generally apply only to algorithms with relatively simple dynamics (e.g., [103, 105, 106, 108, 283, 411, 456, 483]). There are, however, some notable recent exceptions. In particular, a.s. convergence of a continuous-time stochastic gradient descent algorithm for the parameters of a fully observed diffusion process was recently established in [420], and has since been extended to partially observed [430] and jump [50] diffusion processes. In the first case, the same authors have since also established an asymptotic $\mathbb{L}^p$ convergence rate and a central limit theorem [422].

2.1.1 Contributions

In this chapter, we establish the a.s. convergence of two-timescale stochastic gradient descent algorithms in continuous time, under general noise and stability conditions, namely,

$$\lim_{t\to\infty} \nabla_\alpha f(\alpha_t,\beta_t) = \lim_{t\to\infty} \nabla_\beta g(\alpha_t,\beta_t) = 0 \quad \text{a.s.} \tag{2.5}$$

[1] For more details on singularly perturbed SDEs, we refer to [375] and references therein.

[2] It is worth emphasising that the superscripts in $\{\gamma_t^i\}_{t\geq 0}$ are indices, rather than exponents. In what follows, whenever it is necessary to consider powers of the learning rates, we will use brackets to avoid any confusion. For example, we will write $(\gamma_t^1)^2$ and $(\gamma_t^2)^2$ to denote the square of the learning rates γ_t^1 and γ_t^2, respectively.

where $\{\alpha_t\}_{t\geq 0}$ and $\{\beta_t\}_{t\geq 0}$ are generated according to the recursions (2.4a) - (2.4b). The noise conditions, as well as the conditions on the objective functions, are some of the weakest under which a.s. convergence can still be obtained. We consider algorithms with additive, state-dependent noise, and, importantly, also those with non-additive, state-dependent noise. In the second case, our analysis is carried out under the assumption that the non-additive noise can be represented by an ergodic diffusion process, controlled by the algorithm states. To our knowledge, this is the first rigorous analysis of two-timescale stochastic approximation with Markovian dynamics in continuous time.

Our proof of these results closely follows the classical ODE method (e.g., [44, 59, 271, 310]), adapted appropriately to the continuous-time setting (e.g., [108, 272]). In the Markovian noise case, it also draws upon well known regularity results relating to the solution of the Poisson equation associated with the infinitesimal generator of the ergodic diffusion process (e.g., [371, 372]). The obtained results cover a broad class of non-linear, two-timescale stochastic gradient descent algorithms in continuous time. In particular, they can be applied to the two-timescale stochastic gradient descent algorithm which we develop for the problem of joint online parameter estimation and optimal sensor placement in a continuous-time state space model in Chapter 3. They also include, upon restriction to a single timescale, the continuous-time stochastic gradient descent algorithms recently studied in [420, 430].

In the non-additive noise case, we also establish, under some additional assumptions, the weak convergence rate of two-timescale stochastic gradient descent algorithms in continuous time. In particular, we obtain a central limit theorem of the form

$$\begin{pmatrix} (\gamma_t^1)^{-\frac{1}{2}}(\alpha_t - \alpha^*) \\ (\gamma_t^2)^{-\frac{1}{2}}(\beta_t - \beta^*) \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \begin{pmatrix} \Sigma_\alpha & 0 \\ 0 & \Sigma_\beta \end{pmatrix}\right). \tag{2.6}$$

This result, which indicates that asymptotically the two algorithm iterates 'decouple', represents a careful extension of a long standing and well known result in discrete time [354] to the continuous-time setting.

2.1.2 Chapter Organisation

The remainder of this chapter is organised as follows. In Section 2.2, we present our main results. In particular, in Section 2.2.1, we establish the a.s. convergence of continuous-time, two-timescale stochastic gradient descent algorithms with additive, state-dependent noise. In Section 2.2.2, we extend our analysis to continuous-time, two-timescale stochastic gradient descent algorithms with Markovian dynamics. In Section 2.3, we provide proofs of our main results. In Section 2.4, we provide a detailed discussion of several important extensions to our results, and establish the weak convergence rate of the two-timescale stochastic gradient descent algorithm with additive noise. Finally, in Section 2.5, we offer some concluding remarks.

2.2 Main Results

We will assume, throughout this section, that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, equipped with a filtration $(\mathcal{F}_t)_{t\geq 0}$ which satisfies the usual conditions.[3]

2.2.1 Two Timescale Stochastic Gradient Descent in Continuous Time

Let $f, g : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}$ be continuously differentiable functions. Suppose that, for any inputs $\{\alpha_t\}_{t\geq 0}$, $\{\beta_t\}_{t\geq 0}$, it is possible to obtain noisy estimates $\{h_t^1\}_{t\geq 0}$, $\{h_t^2\}_{t\geq 0}$ of $\nabla_\alpha f$ and $\nabla_\beta g$ as the output of the following SDEs

$$\mathrm{d}h_t^1 = \nabla_\alpha f(\alpha_t, \beta_t)\mathrm{d}t + \mathrm{d}\xi_t^1, \tag{2.7a}$$

$$\mathrm{d}h_t^2 = \nabla_\beta g(\alpha_t, \beta_t)\mathrm{d}t + \mathrm{d}\xi_t^2, \tag{2.7b}$$

where $\{\xi_t^1\}_{t\geq 0}$ and $\{\xi_t^2\}_{t\geq 0}$ are $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$ valued continuous semi-martingales on $(\Omega, \mathcal{F}, \mathbb{P})$, which are assumed to be measurable, random functions of $\{\alpha_s\}_{0\leq s\leq t}$ and $\{\beta_s\}_{0\leq s<t}$.[4] The functions f and g are to be regarded as the objective functions in the bilevel optimisation (2.2), while the semi-martingales $\{\xi_t^i\}_{t\geq 0}$, $i = 1, 2$, can be considered as additive noise.

On the basis of these noisy observations, it is natural to seek the stationary points of f and g via the following algorithm:

$$\mathrm{d}\alpha_t = -\gamma_t^1 \left[\nabla_\alpha f(\alpha_t, \beta_t)\mathrm{d}t + \mathrm{d}\xi_t^1\right], \tag{2.9a}$$

$$\mathrm{d}\beta_t = -\gamma_t^2 \left[\nabla_\beta g(\alpha_t, \beta_t)\mathrm{d}t + \mathrm{d}\xi_t^2\right], \tag{2.9b}$$

where $\{\gamma_t^i\}_{t\geq 0}$, $i = 1, 2$, are positive, non-increasing, deterministic functions known as the learning rates; and $\alpha_0 \in \mathbb{R}^{d_1}$, $\beta_0 \in \mathbb{R}^{d_2}$ are random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. We will assume directly the existence and uniqueness of strong solutions to (2.9a) - (2.9b). The interested reader can find some standard sufficient conditions in [343, Chapter II] or [383].

[3] See, for example, [19] for a definition of the usual conditions.

[4] In order to aid intuition, the reader may find it instructive to consider the formal time derivative of these measurement equations, viz

$$\dot{h}_t^1 = \nabla_\alpha f(\alpha_t, \beta_t) + \dot{\xi}_t^1, \tag{2.8a}$$

$$\dot{h}_t^2 = \nabla_\beta g(\alpha_t, \beta_t) + \dot{\xi}_t^2. \tag{2.8b}$$

This formulation, while lacking rigour, is useful in order to emphasise the connection with the standard form of noisy gradient measurements assumed in (two-timescale) stochastic approximation algorithms in discrete time (e.g., [62, 437]).

We will refer to this algorithm as two-timescale stochastic gradient descent in continuous time. This algorithm represents the continuous-time, gradient descent analogue of the two-timescale stochastic approximation algorithm originally introduced in [61]. For further details, see [62] and references therein. It can also be considered a two-timescale generalisation of the continuous-time stochastic approximation algorithms introduced in [167], and later studied in, for example, [103, 105, 361, 411, 483]. Finally, this algorithm can also be viewed as a two-timescale generalisation of the Robbins-Monro type semimartingale stochastic differential equations studied in [283, 284, 285, 286, 287, 288, 343, 456]. We should note, however, that in this final set of references, there is no requirement that the stochastic processes are continuous.

Before we proceed, it is worth noting that Algorithm (2.9a) - (2.9b) is not the only possible two-timescale stochastic gradient descent scheme that one can use to simultaneously optimise $f(\alpha,\beta)$ and $g(\alpha,\beta)$. This algorithm is certainly a natural choice if one only has access to noisy estimates of the partial derivatives $\nabla_\alpha f(\alpha,\beta)$ and $\nabla_\beta g(\alpha,\beta)$, and is interested in solving the bilevel optimisation problem in (2.2). It is less well suited, however, to the stronger version of the bilevel optimisation problem in (2.1), since it ignores the dependence of the true upper level objective $f(\alpha,\beta^*(\alpha))$ on α in its second argument. As such, if one has access to additional gradient information, then it may be preferable to use higher order updates to capture the dependence on $\beta^*(\alpha)$. We provide details of one such approach in Section 2.4.1 (see also [218] in discrete time).

We will analyse Algorithm (2.9a) - (2.9b) under the following set of assumptions. These are imposed in addition to any assumptions required for the existence and uniqueness of strong solutions to (2.9a) - (2.9b).[5] Broadly speaking, these assumptions represent the continuous-time analogues of standard assumptions used in the a.s. convergence analysis of two-timescale stochastic approximation algorithms in discrete time (see, e.g., [62, Chapter 6] or [437]).

Assumption 2.1.1. *The learning rates $\{\gamma_t^i\}_{t\geq 0}$, $i=1,2$, are positive, non-increasing functions which satisfy*

$$\lim_{t\to\infty}\gamma_t^1=\lim_{t\to\infty}\gamma_t^2=\lim_{t\to\infty}\frac{\gamma_t^1}{\gamma_t^2}=0, \tag{2.10a}$$

$$\int_0^\infty \gamma_t^1\mathrm{d}t=\int_0^\infty\gamma_t^2\mathrm{d}t=\infty. \tag{2.10b}$$

This assumption relates to the asymptotic properties of the learning rates $\{\gamma_t^i\}_{t\geq 0}$, $i=1,2$.

[5] It is worth noting that there is a non-empty intersection between the set of assumptions typically required for existence and uniqueness, and the set of assumptions required for a.s. convergence. Wherever an assumption required for a.s. convergence does coincide with an assumption already required for existence and uniqueness, we will highlight this in the corresponding discussion.

It is the continuous-time analogue of the standard step-size assumption used for the a.s. convergence analysis of two-timescale stochastic approximation algorithms in discrete time (e.g., [61, 62, 437]). This assumption implies that the process $\{\alpha_t\}_{t\geq 0}$ evolves on a slower time-scale than the process $\{\beta_t\}_{t\geq 0}$. Thus, intuitively speaking, the fast component, β_t, will see the slow component, α_t, as quasi-static, while the slow component will see the fast component as essentially equilibrated [61]. A standard choice of step sizes which satisfies this assumption is $\gamma_t^1 = \gamma_1^0(\delta_1 + t^{\eta_1})^{-1}$, $\gamma_t^2 = \gamma_2^0(\delta_2 + t^{\eta_2})^{-1}$ for $t \geq 0$, where $\gamma_1^0, \gamma_2^0 > 0$ and $\delta_1, \delta_2 > 0$ are positive constants, and $\eta_1, \eta_2 \in (0, 1]$ are constants such that $\eta_1 > \eta_2$.[6]

Assumption 2.1.2. *The functions $\nabla_\alpha f : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}^{d_1}$ and $\nabla_\beta g : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}^{d_2}$ are locally Lipschitz continuous. That is, for each $\alpha_1 \in \mathbb{R}^{d_1}$ and $\beta_1 \in \mathbb{R}^{d_2}$, there exist positive constants $L_\alpha, L_\beta > 0$ and $\delta_\alpha, \delta_\beta > 0$ such that, for all $\alpha_2 \in \mathbb{R}^{d_1}, \beta_2 \in \mathbb{R}^{d_2}$ with $||\alpha_2 - \alpha_1|| < \delta_\alpha$ and $||\beta_2 - \beta_1|| < \delta_\beta$,*

$$||\nabla_\alpha f(\alpha_1, \beta_1) - \nabla_\alpha f(\alpha_2, \beta_2)|| \leq L_\alpha \left[||\alpha_1 - \alpha_2|| + ||\beta_1 - \beta_2||\right], \tag{2.11a}$$

$$||\nabla_\beta g(\alpha_1, \beta_1) - \nabla_\alpha g(\alpha_2, \beta_2)|| \leq L_\beta \left[||\alpha_1 - \alpha_2|| + ||\beta_1 - \beta_2||\right]. \tag{2.11b}$$

This assumption relates to the smoothness of the objective functions $f(\cdot)$ and $g(\cdot)$, and is a standard assumption used in the convergence analysis of two-timescale stochastic approximation algorithms in discrete time [61, 241, 250], as well as single-timescale stochastic approximation algorithms in continuous time [106, 343, 411, 483], although slightly weaker assumptions may also be possible (see, e.g., [284]). This is also a standard condition required for the existence and uniqueness of strong solutions of (2.9a) - (2.9b).[7] This assumption implies, in particular, that the functions $\nabla_\alpha f(\cdot)$ and $\nabla_\beta g(\cdot)$ locally satisfy linear growth conditions.

Assumption 2.1.3. *For all $T \in [0, \infty)$, the noise processes $\{\xi_t^i\}_{t\geq 0}$, $i = 1, 2$, satisfy*

$$\lim_{s\to\infty} \sup_{t\in[s,s+T]} \left|\left| \int_s^t \gamma_v^i \mathrm{d}\xi_v^i \right|\right| = 0 \quad a.s. \tag{2.12}$$

This assumption relates to the asymptotic properties of the noise processes $\{\xi_t^i\}_{t\geq 0}$, $i = 1, 2$. It can be regarded as the continuous-time, two-timescale generalisation of the Kushner-Clark condition [272]. This assumption is significantly weaker than the

[6] The constants δ_1, δ_2 ensure regularity at $t = 0$. Another standard choice of step sizes which satisfies this assumption is $\gamma_t^1 = \gamma_1^0(\delta_1 + t)^{-\eta_1}$, $\gamma_t^2 = \gamma_2^0(\delta_2 + t)^{-\eta_2}$, with all constants defined as previously (e.g., [251].)

[7] To be precise, this condition, in addition to an analogous local Lipschitz condition on the components of the continuous semi-martingales $\{\xi_t^i\}_{t\geq 0}$, $i = 1, 2$, is sufficient for the existence and uniqueness of strong solutions to (2.9a) - (2.9b) up to some (possibly finite) explosion time (e.g., [383, Theorem 4.3]). If these conditions were replaced by rather stronger global Lipschitz conditions, one would instead have the existence and uniqueness of strong solutions to (2.9a) - (2.9b) for all $t \geq 0$ [383, Theorem 3.1]. .

noise conditions adopted in many of the existing results on a.s. convergence of continuous-time, single-timescale stochastic approximation algorithms. In particular, it includes the cases when $\{\xi_t^i\}_{t\geq 0}$, $i = 1, 2$, are continuous (local) martingales [411],[8] continuous finite variation processes with zero mean [483], or diffusion processes [106]. It also holds, under certain additional assumptions, for algorithms with Markovian dynamics [420, 430]. We discuss this case further in Section 2.2.2. The discrete-time analogue of this condition first appeared in [437], weakening the noise condition originally used in [61]. In fact, in the context of single-timescale stochastic approximation, the Kushner-Clark condition is the weakest condition under which it is possible to establish a.s. convergence (e.g., [271]). Furthermore, under certain stability conditions, the Kushner-Clark condition is both necessary and sufficient for the a.s. convergence of discrete-time, single-timescale stochastic approximation algorithms [460].

Assumption 2.1.4. *The iterates $\{\alpha_t\}_{t\geq 0}$, $\{\beta_t\}_{t\geq 0}$ are almost surely bounded:*

$$\sup_{t\geq 0} [||\alpha_t|| + ||\beta_t||] < \infty. \tag{2.13}$$

This assumption is necessary in order to prove a.s. convergence. This condition also ensures the existence of strong solutions of (2.9a) - (2.9b) for all times. In general, however, it is far from automatic, and not very straightforward to establish [62]. Indeed, sufficient conditions tend to be highly problem specific, or else somewhat restrictive (e.g., [283, 411, 456]). To circumvent this issue, a common approach is to include a truncation or projection device in the algorithm, which ensures that the iterates remain bounded with probability one, at the expense of an additional error term (e.g., [106, 108, 271, 430]). In addition, this may introduce spurious fixed points on the boundary of the domain (e.g., [430]). An alternative method, which avoids this shortcoming, is the 'continuous-time stochastic approximation procedure with randomly varying truncations', originally introduced in [108]. It is possible to partially extend this approach to the two-timescale setting, to establish a.s. boundedness of iterates on the fast-timescale $\{\beta_t\}_{t\geq 0}$. It is currently unclear, however, whether this approach can also be used to relax the assumption of boundedness for the slow-timescale.

Another common approach is to omit the boundedness assumption, and instead state asymptotic results which are local in nature (e.g., [44, 61, 437]). That is, which hold almost surely on the event

$$\Lambda = \{\sup_{t\geq 0} ||\alpha_t|| < \infty\} \cap \{\sup_{t\geq 0} ||\beta_t|| < \infty\}. \tag{2.14}$$

In the single-timescale setting, it is often then straightforward to establish the global coun-

[8] The case when the noise process is a local martingale is also considered by [283, 284, 285, 286, 343, 456]. In these works, however, there is no requirement that this local martingale is continuous.

terparts of these results, by combining them with existing methods for verifying stability (e.g., [44, 59, 265]). In contrast, the stability of two-timescale stochastic approximation algorithms has thus far not received much attention. Indeed, to the best our knowledge, the only existing result along these lines is [281].

Assumption 2.1.5. *For all $\alpha \in \mathbb{R}^{d_1}$, the ordinary differential equation*

$$\frac{\mathrm{d}\beta_t}{\mathrm{d}t} = -\nabla_\beta g(\alpha, \beta_t) \tag{2.15}$$

has a discrete, countable set of equilibria $\{\beta_i^*\}_{i\geq 1} = \{\beta_i^*(\alpha)\}_{i\geq 1}$, where $\beta_i^* : \mathbb{R}^{d_1} \to \mathbb{R}^{d_2}$, $i \geq 1$, are locally Lipschitz-continuous maps.

This is a stability condition relating to the fast recursion. It is somewhat weaker than the standard fast-timescale assumption used in the analysis of discrete-time, two-timescale stochastic approximation algorithms, which requires that this ordinary differential equation must have a unique global asymptotically stable equilibrium (e.g., [61, 252, 437]). We note, however, that a similar assumption has previously appeared in [241]. It may be possible to weaken this assumption further - that is, to remove the requirement for a discrete, countable set of equilibria - using the tools recently established in [439]. There, in the context of discrete-time, single-timescale stochastic gradient descent, a.s. single-limit point convergence is proved in the case of multiple or non-isolated equilibria, using tools from differential geometry (i.e., the Lojasiewicz gradient inequality). It remains an open problem to determine whether these results can be extended to the two-timescale setting.

In order to state our final assumption, we will require the following additional notation. Let $x \in \mathbb{R}^d$, and let $h : \mathbb{R}^d \to \mathbb{R}^d$. Consider an ordinary differential equation of the form $\dot{x}(t) = h(x(t))$. We say that a set $A \subset \mathbb{R}^d$ is invariant for this equation if any trajectory $x(t)$ satisfying $x(0) \in A$ satisfies $x(t) \in A$ for all $t \in \mathbb{R}$. In addition, we say that A is internally chain transitive for this equation if for any $x \in A$, and for any $\varepsilon > 0$, $T > 0$, there exists $n \in \mathbb{N}$, points $x_0, x_1, \ldots, x_n = x$ in A, and times $t_1, \ldots, t_n \geq T$, such that, for all $1 \leq i \leq n$, the trajectory of the equation initialised at x_{i-1} is in the ε-neighbourhood of x_i at time t_i. We can now state our final assumption.

Assumption 2.1.6. *For all $i \geq 1$, the only internally chain transitive invariant sets of the ordinary differential equation*

$$\frac{\mathrm{d}\alpha_t}{\mathrm{d}t} = -\nabla_\alpha f(\alpha_t, \beta_i^*(\alpha_t)) \tag{2.16}$$

are its equilibrium points.

This is a stability condition relating to the slow recursion. It can be regarded as a slightly weaker version of the standard slow-timescale assumption used in the analysis of two-timescale stochastic approximation algorithms, which stipulates that this ordinary differential equation must have a unique, globally asymptotically stable equilibrium (e.g., [61, 62, 252]). This assumption is required in order to rule out the possibility that (2.16) admits other internally chain transitive invariant sets aside from equilibria, such as cyclic orbit chains (see [36]). One can alternatively assume that this equation has a unique limit for each initial condition.

It is worth noting that, under additional assumptions on $\beta_i^*(\cdot)$, one can replace this with the weaker assumption that (2.16) has a discrete, countable set of isolated equilibria. Unfortunately, without additional assumptions on $\beta_i^*(\cdot)$, one cannot use this condition directly, since $f(\cdot,\beta_i^*(\cdot))$ is not, in general, a strict Lyapunov function for (2.16). We discuss this point in further detail in Section 2.4.1.

We conclude this commentary with the remark that our condition(s) on the objective function(s) are, broadly speaking, more general than those adopted in many of the existing results on the convergence of continuous-time, single-timescale stochastic approximation algorithms. In particular, we do not insist on the existence of a unique root for the gradient of the objective functions, as is the case in [106, 283, 343, 411, 456].

Our main result on the convergence of Algorithm (2.9a) - (2.9b) is contained in the following theorem.

Theorem 2.1. *Assume that Assumptions 2.1.1 - 2.1.6 hold. Then, almost surely,*

$$\lim_{t\to\infty}\nabla_\alpha f(\alpha_t,\beta_t)=\lim_{t\to\infty}\nabla_\beta g(\alpha_t,\beta_t)=0. \tag{2.17}$$

Proof. See Section 2.3.1. □

The proof of Theorem 2.1 follows the ODE method. This approach was first introduced in [310], and extensively developed by Kushner et al. (e.g., [44, 268, 271, 272]) and later Benaïm et al. [35, 36, 37]. It was first used to prove a.s. convergence of a two-timescale stochastic approximation algorithm in [61], which considered a discrete-time stochastic approximation algorithm with state-independent additive noise. It has since also been used to establish the convergence of more general discrete-time, two-timescale stochastic approximation algorithms [241, 252, 437].

In the context of continuous-time, single-timescale stochastic approximation, this method of proof has largely been neglected, with a small number of notable exceptions [103, 108, 272, 483]. While other approaches (e.g., [106, 283, 343, 411, 456]) may be more direct, they may also require slightly more restrictive assumptions. Moreover, it is unclear whether these approaches can straightforwardly be adapted to the two-timescale setting, or even to more complex single-timescale algorithms, such as those with Markovian dynamics (e.g., [420]). One other advantage of this method of proof is that it is straightforwardly adapted to other variations of Algorithm (2.9a) - (2.9b), as discussed prior to the statement of our assumptions. In Section 2.4.1, we show rigorously how to use this approach to establish an a.s. convergence result for one such algorithm.

We should emphasise, at this point, that Theorem 2.1 establishes a.s. convergence precisely to the stationary points of the objective functions f and g. In particular, the stated assumptions do not guarantee convergence to the set of local (or global) minima. On this point, two remarks are pertinent. Firstly, results of this type are standard in the recent literature on stochastic gradient descent in continuous time (e.g., [420, 430]), and the more classical literature on two-timescale stochastic approximation (e.g., [62]). Secondly, under additional assumptions, it should be possible to extend our analysis to guarantee that our algorithm converges a.s. to local minima of the two objective functions. Indeed, when a single timescale is considered, there are several existing 'avoidance of saddle' type results of this kind [67, 189, 346, 377]. While no explicit results of this type exist in the two-timescale framework, we outline details of the (minimal) assumptions which would be required to obtain such a result in Section 2.4.2, and discuss briefly how they can be used together with the results of this section.

2.2.2 Two Timescale Stochastic Gradient Descent in Continuous Time with Markovian Dynamics

Using the results obtained in Section 2.2.1, we now consider the situation in which the noisy estimates of $\nabla_\alpha f$ and $\nabla_\beta g$ are governed by some additional continuous-time dynamical process. In particular, we now analyse the convergence of the algorithm

$$\mathrm{d}\alpha_t = -\gamma_t^1\left[F(\alpha_t,\beta_t,\mathcal{X}_t)\mathrm{d}t + \mathrm{d}\zeta_t^1\right], \tag{2.18a}$$

$$\mathrm{d}\beta_t = -\gamma_t^2\left[G(\alpha_t,\beta_t,\mathcal{X}_t)\mathrm{d}t + \mathrm{d}\zeta_t^2\right], \tag{2.18b}$$

where $\{\gamma_t^i\}_{t\geq 0}$, $i=1,2$, are positive, decreasing functions; $F,G:\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}\times\mathbb{R}^{d_3}\rightarrow\mathbb{R}^{d_1},\mathbb{R}^{d_2}$ are Borel measurable functions; $\alpha_0\in\mathbb{R}^{d_1}$, $\beta_0\in\mathbb{R}^{d_2}$ are random variables on $(\Omega,\mathcal{F},\mathbb{P})$; and $\{\zeta_t^i\}_{t\geq 0}$, $i=1,2$, are $\mathbb{R}^{d_1}$, $\mathbb{R}^{d_2}$ valued continuous semi-martingales on the same probability space, which are measurable, random functions of $\{\alpha_s,\beta_s\}_{0\leq s<t}$. [9] In this algorithm, the functions $F(\cdot)$ and $G(\cdot)$ are to be regarded as noisy estimators of $\nabla_\alpha f(\cdot)$ and $\nabla_\beta g(\cdot)$; the precise relationship between these functions will be clarified below. The semi-martingales $\{\zeta_t^i\}_{t\geq 0}$, $i=1,2$, can once more be considered as additive noise; while the Markov process $\{\mathcal{X}_t\}_{t\geq 0}$ can be regarded as non-additive noise.

We will refer to Algorithm (2.18a) - (2.18b) as two-timescale stochastic gradient descent in continuous time with Markovian dynamics. This algorithm represents the continuous-

[9] Once more, we will assume directly the existence and uniqueness of strong solutions to (2.18a) - (2.18b).

time analogue of the discrete-time, two-timescale stochastic approximation algorithm with state-dependent non-additive noise analysed in [437, Section IV]. In fact, our presentation is slightly more general than in [437], as we also allow for the possibility of additive, state-dependent noise via the terms $\{\zeta_t^i\}_{t\geq 0}$, $i = 1, 2$. This increases the number of applications in which our algorithm can be applied, while not significantly complicating the analysis.

The a.s. convergence of discrete-time, two-timescale stochastic approximation algorithms with Markovian dynamics is also studied, under various assumptions, in [241, 249, 250, 419]. Conversely, there are no existing works which provide a rigorous analysis of two-timescale stochastic approximation algorithms with Markovian dynamics in continuous time. In fact, even in the single-timescale setting, such algorithms have only recently received attention [50, 420, 430]. In particular, [420] established the a.s. convergence of a continuous-time stochastic gradient descent algorithm for the parameters of a fully observed diffusion process. This analysis has since been extended to the case of a partially-observed diffusion process [430], and a fully observed jump-diffusion process [50].

We analyse this algorithm under the assumption that $\mathcal{X} = \{\mathcal{X}_t\}_{t\geq 0}$ is a diffusion process on $\mathbb{R}^{d_3}$, controlled by the algorithm states $\{\alpha_t\}_{t\geq 0}, \{\beta_t\}_{t\geq 0}$. In particular, we suppose that this process evolves according to

$$\mathrm{d}\mathcal{X}_t = \Phi(\alpha_t, \beta_t, \mathcal{X}_t)\mathrm{d}t + \Psi(\alpha_t, \beta_t, \mathcal{X}_t)\mathrm{d}b_t, \tag{2.19}$$

where, for all $\alpha \in \mathbb{R}^{d_1}$, $\beta \in \mathbb{R}^{d_2}$, $\Phi(\alpha, \beta, \cdot) : \mathbb{R}^{d_3} \to \mathbb{R}^{d_3}$ and $\Psi(\alpha, \beta, \cdot) : \mathbb{R}^{d_3} \to \mathbb{R}^{d_3 \times d_4}$ are Borel measurable functions; $\mathcal{X}_0$ is a random variable defined on $(\Omega, \mathcal{F}, \mathbb{P})$; and $\{b_t\}_{t\geq 0}$ is a $\mathbb{R}^{d_4}$ valued Wiener process on the same probability space. We should remark that, whenever $\alpha \in \mathbb{R}^{d_1}$, $\beta \in \mathbb{R}^{d_2}$ are fixed, we will denote the corresponding diffusion process by $\{\mathcal{X}_t(\alpha, \beta)\}_{t\geq 0}$, making explicit the dependence on these parameters.

Our motivation for this choice of dynamics is threefold: firstly, the existence, uniqueness, and asymptotic properties of this class of processes are very well studied (e.g., [229, 239, 371]). Secondly, this choice is sufficiently broad for many practical situations of interest. Finally, under the assumption that $\{\mathcal{X}_t(\alpha, \beta)\}_{t\geq 0}$ is ergodic for all $\alpha \in \mathbb{R}^{d_1}$, $\beta \in \mathbb{R}^{d_2}$, with unique invariant measure $\mu_{\alpha,\beta}(\cdot)$ (see Assumption 2.2.2a), one can obtain an explicit relation between the estimators $F(\cdot)$ and $G(\cdot)$ and the gradients of the objective functions $\nabla_\alpha f(\cdot)$ and $\nabla_\beta g(\cdot)$. In particular, in this case the gradients of the true objective functions are defined as ergodic averages of the noisy estimators:

$$\nabla_\alpha f(\alpha, \beta) = \int_{\mathbb{R}^{d_3}} F(\alpha, \beta, x)\mu_{\alpha,\beta}(\mathrm{d}x), \tag{2.20a}$$

$$\nabla_\beta g(\alpha, \beta) = \int_{\mathbb{R}^{d_3}} G(\alpha, \beta, x)\mu_{\alpha,\beta}(\mathrm{d}x). \tag{2.20b}$$

We remark that, in general, it is not possible to obtain the unique invariant measure $\mu_{\alpha,\beta}(\cdot)$ of the ergodic diffusion process $\mathcal{X}$ in closed form, let alone compute these integrals. Thus, in the Markovian framework we typically cannot compute the gradients $\nabla_\alpha f$ and $\nabla_\beta g$ exactly, even in the absence of the additive noise processes $\{\zeta_t^i\}_{t\geq 0}$, $i = 1, 2$.

We analyse this algorithm under the following set of assumptions. Similarly to before, these assumptions can be viewed both as the continuous-time analogues of standard assumptions used for the a.s. convergence analysis of two-timescale stochastic approximation algorithms with Markovian dynamics in discrete time (e.g., [437, Section IV]), and as the two-timescale generalisation of assumptions more recently introduced to analyse the convergence of single-timescale stochastic gradient descent algorithms with Markovian dynamics in continuous time [420, 430].

Assumption 2.2.1. *The learning rates $\{\gamma_t^i\}_{t\geq 0}$, $i = 1, 2$, satisfy Assumption 2.1.1. Furthermore,*

$$\int_0^\infty (\gamma_t^i)^2 \mathrm{d}t < \infty, \quad \int_0^\infty \left|\dot{\gamma}_t^i\right| \mathrm{d}t < \infty, \tag{2.21}$$

and there exist $r_i > 0$, $i = 1, 2$, such that $\lim_{t\to\infty} (\gamma_t^i)^2 t^{\frac{1}{2}+2r_i} = 0$.

This assumption corresponds to the asymptotic properties of the step sizes $\{\gamma_t^i\}_{t\geq 0}$, $i = 1, 2$. It can be regarded as the two-timescale generalisation of the standard step-size assumptions used for the a.s. convergence analysis of single-timescale stochastic gradient descent algorithms with Markovian dynamics in continuous time (e.g., [420, 430]). As previously, it implies that the process $\{\alpha_t\}_{t\geq 0}$ evolves on a slower time-scale than the process $\{\beta_t\}_{t\geq 0}$. A standard choice of step sizes which satisfies this assumption is $\gamma_t^1 = \gamma_1^0(\delta_1 + t^{\eta_1})^{-1}$, $\gamma_t^2 = \gamma_2^0(\delta_2 + t^{\eta_2})^{-1}$ for $t \geq 0$, where $\gamma_1^0, \gamma_2^0 > 0$ and $\delta_1, \delta_2 > 0$ are positive constants, and now $\eta_1, \eta_2 \in (\frac{1}{2}, 1]$ are constants such that $\eta_1 > \eta_2$. We remark, as in [420], that the condition relating to the derivatives, namely that $\int_0^\infty |\dot{\gamma}_t^i| \mathrm{d}t < \infty$, $i = 1, 2$, is satisfied automatically if the step sizes are monotonic functions of t.

Assumption 2.2.2a. *The process $\{\mathcal{X}_t(\alpha, \beta)\}_{t\geq 0}$ is ergodic for all $\alpha \in \mathbb{R}^{d_1}$, $\beta \in \mathbb{R}^{d_2}$, with unique invariant probability measure $\mu_{\alpha,\beta}$ on $(\mathbb{R}^{d_3}, \mathbb{B}_{d_3})$, where $\mathbb{B}_{d_3}$ denotes the Borel σ-algebra on $\mathbb{R}^{d_3}$.*

This assumption relates to the asymptotic properties of the non-additive, state-dependent noise process $\{\mathcal{X}_t(\alpha, \beta)\}_{t\geq 0}$. In the context of discrete-time stochastic approximation with Markovian dynamics, the requirement of ergodicity is relatively standard, in both single-timescale (e.g., [44, 268, 269]) and two-timescale (e.g., [249, 250, 437]) settings.[10] This

[10] It should be noted that it is possible to relax this assumption, and to establish a.s. convergence

assumption is also central to the existing results on the convergence of stochastic gradient descent with Markovian dynamics in continuous time [50, 420, 430].

Assumption 2.2.2b. *For any $q > 0$, $\alpha \in \mathbb{R}^{d_1}$, $\beta \in \mathbb{R}^{d_2}$, there exists constants $K_q, K_q^{\alpha}, K_q^{\beta} > 0$, such that*

$$\int_{\mathbb{R}^{d_3}} (1 + ||x||^q)\mu_{\alpha,\beta}(\mathrm{d}x) \leq K_q, \tag{2.22a}$$

$$\int_{\mathbb{R}^{d_3}} (1 + ||x||^q)|\nu^{(\alpha)}_{\alpha,\beta,i}(\mathrm{d}x)| \leq K_q^{\alpha}, \tag{2.22b}$$

$$\int_{\mathbb{R}^{d_3}} (1 + ||x||^q)|\nu^{(\beta)}_{\alpha,\beta,i}(\mathrm{d}x)| \leq K_q^{\beta}, \tag{2.22c}$$

where $|\nu^{(\alpha)}_{\alpha,\beta,i}(\mathrm{d}x)|, |\nu^{(\beta)}_{\alpha,\beta,i}(\mathrm{d}x)|$ denote the total variations of the finite signed measures $\nu^{(\alpha)}_{\alpha,\beta,i} = \partial_{\alpha_i}\mu_{\alpha,\beta}$, $i = 1, \dots, d_1$, and $\nu^{(\beta)}_{\alpha,\beta,i} = \partial_{\beta_i}\mu_{\alpha,\beta}$, $i = 1, \dots, d_2$.

This assumption relates to the regularity of the invariant measure and its derivatives. It can be regarded as a two-timescale extension of the regularity conditions used for the convergence analysis of the continuous-time, single-timescale stochastic gradient descent algorithm with Markovian dynamics in [430].[11] This condition ensures that the objective functions $f(\cdot)$ and $g(\cdot)$, and their first two derivatives, are uniformly bounded in both arguments.[12]

In order to state the remaining assumptions, we will require the following additional notation. We will say that a function $H : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^{d} \to \mathbb{R}$ satisfies the polynomial growth property (PGP) if there exist $q, K > 0$ such that, for all $\alpha \in \mathbb{R}^{d_1}, \beta \in \mathbb{R}^{d_2}$,

$$|H(\alpha, \beta, x)| \leq K(1 + ||x||^q). \tag{2.23}$$

We will write $\mathbb{H}^{i+\delta,j}(\mathbb{R}^d)$, $i, j \in \mathbb{N}$, $\delta \in (0, 1)$, to denote the space of all functions $H : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^d \to \mathbb{R}$ such that $H(\cdot, \cdot, x) \in C^j(\mathbb{R}^{d_1 \times d_2})$ and $H(\alpha, \beta, \cdot) \in C^i(\mathbb{R}^d)$; and such that $\nabla_x^{i'}\nabla_\alpha^{j'}H(\alpha, \beta, \cdot)$, $\nabla_x^{i'}\nabla_\beta^{j'}H(\alpha, \beta, \cdot)$ are Hölder continuous with exponent δ, uniformly in α and β, for $0 \leq i' \leq i$, $0 \leq j' \leq j$. We will also write $\mathbb{H}_c^{i+\delta,j}(\mathbb{R}^d)$ for the subspace consisting of all $H \in \mathbb{H}^{i+\delta,j}(\mathbb{R}^d)$ such that H is centered, in the sense that $\int_{\mathbb{R}^{d_3}} H(\alpha, \beta, x)\mu_{\alpha,\beta}(\mathrm{d}x) = 0$.

under the slightly weaker assumptions introduced in [348] (see also [44, 271]). These assumptions relate to the existence of solutions to a related Poisson equation, and automatically hold under the assumption of ergodicity.

[11] We refer to [44, Part II] for a detailed discussion of the corresponding conditions used in the convergence analysis of discrete-time stochastic approximation algorithms with Markovian dynamics. We remark only that, in this case, it is typical to require that the transition kernels of the Markov process satisfy certain regularity conditions, rather than the invariant measure (if this exists).

[12] In the analysis of discrete-time stochastic approximation algorithms with Markovian dynamics, it is not uncommon for boundedness to be assumed *a priori*. See, for example, [348] in the single-timescale case, and [249] in the two-timescale case.

Finally, we will write $\bar{\mathbb{H}}^{i+\delta,j}(\mathbb{R}^d)$ to denote the subspace consisting of $H \in \mathbb{H}^{i+\delta,j}(\mathbb{R}^d)$ such that H and all of its first and second derivatives with respect to α and β satisfy the PGP.

Assumption 2.2.2c. *There exist differentiable functions $f, g : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}$ such that $\nabla_\alpha f(\cdot)$ and $\nabla_\beta g(\cdot)$ are locally Lipschitz continuous, and unique Borel measurable functions $\tilde{F} : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^{d_3} \to \mathbb{R}^{d_1}$, $\tilde{G} : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^{d_3} \to \mathbb{R}^{d_2}$ such that, for all $\alpha \in \mathbb{R}^{d_1}$, $\beta \in \mathbb{R}^{d_2}$, $x \in \mathbb{R}^{d_3}$,*

$$\mathcal{A}_{\mathcal{X}} \tilde{F}(\alpha, \beta, x) = \nabla_\alpha f(\alpha, \beta) - F(\alpha, \beta, x), \tag{2.24a}$$

$$\mathcal{A}_{\mathcal{X}} \tilde{G}(\alpha, \beta, x) = \nabla_\beta g(\alpha, \beta) - G(\alpha, \beta, x), \tag{2.24b}$$

where $\mathcal{A}_{\mathcal{X}}$ is the infinitesimal generator of $\mathcal{X}$. In addition, the functions $\tilde{F}(\alpha, \beta, x)$ and $\tilde{G}(\alpha, \beta, x)$ are in $\bar{\mathbb{H}}^{1+\delta,2}(\mathbb{R}^{d_3})$, and their mixed first partial derivatives with respect to (α, x) and (β, x) have the PGP.

Assumption 2.2.2d. *The diffusion coefficient Ψ has the PGP componentwise. In particular, it grows no faster than polynomially with respect to the x variable.*

Assumption 2.2.2e. *For all $q > 0$, and for all $t \geq 0$, $\mathbb{E}[||\mathcal{X}_t||^q] < \infty$. Furthermore, there exists $K > 0$ such that for all t sufficiently large,*

$$\mathbb{E}\left[\sup_{s \leq t} ||\mathcal{X}_s(\alpha, \beta)||^q\right] \leq K\sqrt{t}\ , \quad \forall \alpha \in \mathbb{R}^{d_1}\ ,\ \forall \beta \in \mathbb{R}^{d_2}, \tag{2.25a}$$

$$\mathbb{E}\left[\sup_{s \leq t} ||\mathcal{X}_s||^q\right] \leq K\sqrt{t}. \tag{2.25b}$$

These three assumptions relate to the properties of the diffusion process $\{\mathcal{X}_t(\alpha, \beta)\}_{t \geq 0}$, and the definitions of the objective functions $f(\cdot)$ and $g(\cdot)$. In particular, the first condition establishes the relationship between the gradients of the true objective functions $\nabla_\alpha f(\cdot)$ and $\nabla_\beta g(\cdot)$, and the unbiased estimators $F(\cdot)$ and $G(\cdot)$. It also relates to the existence, uniqueness, and properties of solutions of the associated Poisson equations. The second condition pertains to the growth properties of the ergodic diffusion process, while the third condition provides bounds on its moments. Together, these conditions ensure that error terms which arise due to the noisy estimates of $\nabla_\alpha f(\cdot)$ and $\nabla_\beta g(\cdot)$, tend to zero sufficiently quickly as $t \to \infty$. They are therefore essential, whether or not they are required explicitly, to existing results on the a.s. convergence of continuous-time stochastic gradient descent with Markovian dynamics [50, 420, 430].[13]

The discrete-time analogues of these conditions, and variations thereof, also appear in almost all of the existing convergence results for stochastic approximation algorithms

[13] We remark that an analogue of these conditions only appears *explicitly* in [430]. Meanwhile, in [50, 420], it follows from the other stated assumptions and standard results on the solutions of the associated Poisson equation.

with Markovian dynamics in discrete time (e.g., [44, 271, 348, 439]), including those with two-timescales (e.g., [249, 250, 437]).[14,15] Our particular choice of assumptions can be considered as the two-timescale, continuous-time generalisation of the conditions appearing in [348, Section III] and [44, Part II]. It also closely resembles a continuous-time analogue of the assumptions used in [437, Section IV] for a discrete-time, two-timescale stochastic approximation algorithm with non-additive, state-dependent noise.

It remains only to provide our assumptions on the additive noise processes $\{\zeta_t^i\}_{t\geq 0}$, $i=1,2$. In order to state these assumptions, we will now require an explicit form for these semi-martingales. In particular, we will assume that they evolve according to

$$\mathrm{d}\zeta_t^i = \zeta_i^{(1)}(\alpha_t,\beta_t,\mathcal{X}_t)\mathrm{d}a_t^i + \zeta_i^{(2)}(\alpha_t,\beta_t,\mathcal{X}_t)\mathrm{d}z_t^i \tag{2.26}$$

where, for all $\alpha\in\mathbb{R}^{d_1}$, $\beta\in\mathbb{R}^{d_2}$, $\zeta_i^{(1)}(\alpha,\beta,\cdot):\mathbb{R}^{d_3}\to\mathbb{R}^{d_i}$, $\zeta_i^{(2)}(\alpha,\beta,\cdot):\mathbb{R}^{d_3}\to\mathbb{R}^{d_i\times d_5^i}$ are Borel measurable functions; $\{a_t^i\}_{t\geq 0}$ are predictable, increasing processes, and $\{z_t^i\}_{t\geq 0}$ are $\mathbb{R}^{d_5^i}$ valued Wiener processes defined on $(\Omega,\mathcal{F},\mathbb{P})$.

Assumption 2.2.3a. *For all $T>0$, the processes $\{\zeta_t^i\}_{t\geq 0}$, $i=1,2$ satisfy*

$$\lim_{s\to\infty}\sup_{t\in[s,s+T]}\left|\left|\int_s^t \gamma_v^i \mathrm{d}\zeta_v^i\right|\right| = 0\ , \quad a.s. \tag{2.27}$$

Assumption 2.2.3b. *The functions $\zeta_i^{(2)}$, $i=1,2$, have the PGP componentwise. In particular, they grow no faster than polynomially with respect to the x variable.*

Assumption 2.2.3c. *There exist constants $A_{z_1,z_2}, A_{z_i,b}>0$, $i=1,2$, such that, componentwise,*

$$c_t^{z_1,z_2} = \frac{\mathrm{d}[z_1,z_2]_t}{\mathrm{d}t} \leq A_{z_1,z_2}\ ,\ c_t^{z_i,b} = \frac{\mathrm{d}[z_i,b]_t}{\mathrm{d}t} \leq A_{z_i,b}. \tag{2.28}$$

where $[\cdot,\cdot]$ denotes the quadratic variation.

The first of these conditions is analogous to the noise condition which appeared in the analysis of the two-timescale stochastic gradient descent algorithm in Section 2.2.1. Once again, this can be regarded as a continuous-time version of the Kushner-Clark condition. The other two assumptions are unique to the continuous-time, two-timescale stochastic gradient descent algorithm with Markovian dynamics introduced in this chapter. We should note, however, that similar assumptions have previously appeared in the analysis of the single-timescale stochastic approximation schemes in [283, 284, 285, 286, 287, 288, 343, 456].

[14] in discrete time, these conditions are often stated in terms of the Markov transition kernel. They were first introduced in [348, Section III] (see also [44, Part II]), and later generalised in [271].

[15] Interestingly, the final two equations in Assumption 2.2.2e are peculiar to the continuous-time setting. in discrete time, only the first moment bound appears in the analysis of algorithms with Markovian dynamics (e.g., [44, 348, 439]), including the two-timescale case (e.g., [249, 437]).

The remaining assumptions required by Theorem 2.2 are identical to those required by Theorem 2.1.

Our main result on the convergence of Algorithm (2.18a) - (2.18b) is contained in the following theorem.

Theorem 2.2. *Assume that Assumptions 2.2.1 - 2.2.3c and 2.1.4 hold. In addition, assume that Assumptions 2.1.5 - 2.1.6 hold for the functions $f(\cdot)$ and $g(\cdot)$ defined in Assumption 2.2.2c. Then, almost surely,*

$$\lim_{t\to\infty} \nabla_\alpha f(\alpha_t,\beta_t) = \lim_{t\to\infty} \nabla_\beta g(\alpha_t,\beta_t) = 0. \tag{2.29}$$

Proof. See Section 2.3.2. □

Our proof of Theorem 2.2 is obtained by rewriting Algorithm (2.18a) - (2.18b) in the form of Algorithm (2.9a) - (2.9b), viz

$$\mathrm{d}\alpha_t = -\gamma_t^1\Big[\nabla_\alpha f(\alpha_t,\beta_t)\mathrm{d}t + \underbrace{\big(F(\alpha_t,\beta_t,\mathcal{X}_t) - \nabla_\alpha f(\alpha_t,\beta_t)\big)\mathrm{d}t + \mathrm{d}\zeta_t^1}_{=\mathrm{d}\xi_t^1}\Big], \tag{2.30a}$$

$$\mathrm{d}\beta_t = -\gamma_t^2\Big[\nabla_\beta g(\alpha_t,\beta_t)\mathrm{d}t + \underbrace{\big(G(\alpha_t,\beta_t,\mathcal{X}_t) - \nabla_\beta g(\alpha_t,\beta_t)\big)\mathrm{d}t + \mathrm{d}\zeta_t^2}_{=\mathrm{d}\xi_t^2}\Big], \tag{2.30b}$$

and proving that the conditions of Theorem 2.2 (Assumptions 2.2.1 - 2.2.3c) imply the conditions of Theorem 2.1 (Assumptions 2.1.1 - 2.1.3). Clearly, if this is the case, then Theorem 2.2 follows directly from Theorem 2.1. This statement holds trivially for all conditions except those relating to the noise processes. It thus remains to establish that, under the noise conditions in Theorem 2.2 (Assumptions 2.2.2a - 2.2.2e, 2.2.3a - 2.2.3c), the noise condition in Theorem 2.1 (Assumption 2.1.3) holds for the noise processes $\{\xi_t^i\}_{t\geq 0}$, $i = 1, 2$, as defined above. The central part of this proof is thus to control terms of the form,

$$\int_0^t \gamma_s^1 \left[F(\alpha_s,\beta_s,\mathcal{X}_s) - \nabla_\alpha f(\alpha_s,\beta_s)\right] \mathrm{d}s, \tag{2.31}$$

$$\int_0^t \gamma_s^2 \left[G(\alpha_s,\beta_s,\mathcal{X}_s) - \nabla_\beta g(\alpha_s,\beta_s)\right] \mathrm{d}s. \tag{2.32}$$

This is achieved by rewriting each such term using the solution of an appropriate Poisson equation, and applying regularity results. This approach - namely, the use of the Poisson equation - is standard in the a.s. convergence analysis of stochastic approximation algorithms with Markovian dynamics, both in discrete time, including the single-timescale case (e.g. [44, 90, 271, 315, 348]) and two-timescale case (e.g. [249, 250, 437]), and in continuous time (e.g. [50, 420, 430]).

This part of our proof most closely resembles the proofs of [420, Lemma 3.1] and [430, Lemma 1], adapted to the current, somewhat more general setting. In general, however, our proof follows an entirely different approach to those in [420, 430]. Indeed, the ODE method is central to our proof, while the proofs in these papers are based on more classical stochastic descent arguments. In particular, they represent a continuous-time, Markovian extension of the method introduced in [46], under the additional assumption that the objective function is bounded from below. This method, broadly speaking, demonstrates that whenever the magnitude of the gradient of the objective function is large, it remains so for a sufficiently long time interval, guaranteeing a decrease in the value of the objective function which is significant and dominates the noise effects. Under the additional assumption that the objective function is bounded from below, it must converge a.s. to some finite value, and its gradient must converge to zero [46].[16]

Crucially, these arguments do not rely on the assumption that the algorithm iterates remain bounded, which represents a significant advantage over the ODE method. It is thus of clear interest to extend this approach to the two-timescale setting. Thus far, however, our attempts to do so have been unsuccessful, due to the presence of the secondary process.[17] As such, this remains an interesting direction for future study.

We conclude this section with the remark that Theorem 2.2, and its proof, still hold upon restriction to a single-timescale (i.e., under the assumption that either α_t or β_t is held fixed). In this case, of course, we only require assumptions which pertain to that timescale. In this context, our theorem includes, as a particular case, the convergence result in [420]. Moreover, our proof provides an entirely different proof of that result.

2.3 Proof of Main Results

2.3.1 Proof of Theorem 2.1

In this Section, we provide a proof of Theorem 2.1. Our proof follows the approach in [62, Chapter 6], adapted appropriately to the continuous-time setting.

[16] If the boundedness assumption is removed, then either $f(\alpha_t) \to -\infty$, or else $f(\alpha_t)$ converges to a finite value and $\nabla_\alpha f(\alpha_t) \to 0$ [46].

[17] In the single-timescale case, one proves that when $\nabla f(\cdot)$ is 'large', the objective function $f(\cdot)$ decreases by at least $\delta > 0$, and that when $\nabla f(\cdot)$ is 'small', the objective function $f(\cdot)$ increases by no more than some smaller positive constant amount $0 < \delta_1 < \delta$. In the two-timescale case, the second of these steps is no longer possible, due to the presence of the secondary process.

2.3.1.1 Additional Notation

We will require the following additional notation. Firstly, in a slight abuse of notation, we will write (x_1, x_2) to denote the concatenation of $x_1 \in \mathbb{R}^{d_1}$ and $x_2 \in \mathbb{R}^{d_2}$. We will also write $\{q_t^i\}_{t\geq 0}$, $\{p_t^i\}_{t\geq 0}$, $i = 1, 2$, to denote the processes

$$q_t^i = \int_0^t \gamma_s^i \mathrm{d}s \tag{2.33a}$$

$$p_t^i = \left\{ s : \int_0^s \gamma_s^i \mathrm{d}v = t \right\} = (q_t^i)^{-1}. \tag{2.33b}$$

where in the second line we have used $(\cdot)^{-1}$ to denote the inverse function. We then define the time-scaled processes $\{\alpha_t^{\gamma_i}\}_{t\geq 0}$, $\{\beta_t^{\gamma_i}\}_{t\geq 0}$, $i = 1, 2$, by

$$\alpha_t^{\gamma_i} = \alpha_{p_t^i}, \tag{2.34a}$$

$$\beta_t^{\gamma_i} = \beta_{p_t^i}. \tag{2.34b}$$

2.3.1.2 The Fast Timescale

2.3.1.2.1 Additional Notation

We will write $\{\bar{\alpha}_t\}_{t\geq 0}$, $\{\bar{\beta}_t\}_{t\geq 0}$ to denote the solutions of the coupled ordinary differential equations

$$\dot{\bar{\alpha}}_t = 0, \tag{2.35a}$$

$$\dot{\bar{\beta}}_t = -\nabla_\beta g(\bar{\alpha}_t, \bar{\beta}_t). \tag{2.35b}$$

We can then define $\{\bar{\alpha}_t^{(s)}\}_{0\leq s\leq t}$, $\{\bar{\beta}_t^{(s)}\}_{0\leq s\leq t}$, as the unique solutions of equations (2.35a)-(2.35b) which 'start at s', and coincide with the time-scaled processes $\{\alpha_t^{\gamma_2}\}_{t\geq 0}$, $\{\beta_t^{\gamma_2}\}_{t\geq 0}$, at s. That is,

$$\dot{\bar{\alpha}}_t^{(s)} = 0 \ , \qquad \bar{\alpha}_s^{(s)} = \alpha_s^{\gamma_2}, \quad t \geq s, \tag{2.36a}$$

$$\dot{\bar{\beta}}_t^{(s)} = -\nabla_\beta g(\bar{\alpha}_t^{(s)}, \bar{\beta}_t^{(s)}) \ , \quad \bar{\beta}_s^{(s)} = \beta_s^{\gamma_2}, \quad t \geq s. \tag{2.36b}$$

We can similarly define $\{\bar{\alpha}_t^{[s]}\}_{0\leq t\leq s}$, $\{\bar{\beta}_t^{[s]}\}_{0\leq t\leq s}$, as the unique solutions of equations (2.35a)-(2.35b) which 'end at s', and coincide with the time-scaled processes $\{\alpha_t^{\gamma_2}\}_{t\geq 0}$, $\{\beta_t^{\gamma_2}\}_{t\geq 0}$, at s. That is,

$$\dot{\bar{\alpha}}_t^{[s]} = 0 \ , \qquad \bar{\alpha}_s^{[s]} = \alpha_s^{\gamma_2}, \quad t \leq s, \tag{2.37a}$$

$$\dot{\bar{\beta}}_t^{[s]} = -\nabla_\beta g(\bar{\alpha}_t^{[s]}, \bar{\beta}_t^{[s]}) \ , \ \ \bar{\beta}_s^{[s]} = \beta_s^{\gamma_2}, \quad t \leq s. \tag{2.37b}$$

2.3.1.2.2 Proof of Convergence

We first establish, using the processes defined above, that the process $(\alpha_t^{\gamma_2}, \beta_t^{\gamma_2})$ is an asymptotic pseudo-trajectory (APT) of the flow induced by the coupled ODEs (2.35a) - (2.35b). Broadly speaking, this means that $(\alpha_t^{\gamma_2}, \beta_t^{\gamma_2})$ tracks the flow induced by these coupled ODEs with arbitrary accuracy over windows of arbitrary length as time goes to infinity. This provides a notion of "asymptotic closeness" between the paths generated by Algorithm (2.9a) - (2.9b), and the flow of the coupled ODEs. The motivation for this comparison is that, provided the trajectories generated by Algorithm (2.9a) - (2.9b) are "good enough" approximations to the solutions of the coupled ODEs, one can expect that the two sets of equations will enjoy similar convergence properties. For further details, we refer to [35, 36].

Lemma 2.1.1. *Assume that Assumptions 2.1.1-2.1.4 hold. Then, for all $T > 0$,*

$$\lim_{s\to\infty} \sup_{t\in[s,s+T]} \left\| \begin{pmatrix} \alpha_t^{\gamma_2} \\ \beta_t^{\gamma_2} \end{pmatrix} - \begin{pmatrix} \bar{\alpha}_t^{(s)} \\ \bar{\beta}_t^{(s)} \end{pmatrix} \right\| = 0 \ , \quad a.s. \tag{2.38a}$$

$$\lim_{s\to\infty} \sup_{t\in[s-T,s]} \left\| \begin{pmatrix} \alpha_t^{\gamma_2} \\ \beta_t^{\gamma_2} \end{pmatrix} - \begin{pmatrix} \bar{\alpha}_t^{[s]} \\ \bar{\beta}_t^{[s]} \end{pmatrix} \right\| = 0 \ , \quad a.s. \tag{2.38b}$$

Proof. We will prove only the first part of this Lemma, as the method for proving the second part is entirely analogous. We will begin by considering $\{\alpha_t\}_{0\leq s\leq t}$. By definition, we have

$$\alpha_t = \alpha_s - \int_s^t \gamma_u^1 \nabla_\alpha f(\alpha_u, \beta_u)\mathrm{d}u - \int_s^t \gamma_u^1 \mathrm{d}\xi_u^1 \tag{2.39}$$

$$= \alpha_s - \int_s^t \frac{\gamma_u^1}{\gamma_u^2}\gamma_u^2 \nabla_\alpha f(\alpha_u, \beta_u)\mathrm{d}u - \int_s^t \gamma_u^1 \mathrm{d}\xi_u^1 \tag{2.40}$$

It follows immediately from the definition of $\{\alpha_t^{\gamma_2}\}_{0\leq s\leq t}$ that

$$\alpha_t^{\gamma_2} = \alpha_s^{\gamma_2} - \int_{p_s^2}^{p_t^2} \frac{\gamma_u^1}{\gamma_u^2}\gamma_u^2 \nabla_\alpha f(\alpha_u, \beta_u)\mathrm{d}u - \int_{p_s^2}^{p_t^2} \gamma_u^1 \mathrm{d}\xi_u^1 \tag{2.41}$$

$$= \alpha_s^{\gamma_2} - \int_s^t \frac{\gamma_{p_u^2}^1}{\gamma_{p_u^2}^2} \nabla_\alpha f(\alpha_u^{\gamma_2}, \beta_u^{\gamma_2})\mathrm{d}u - \int_{p_s^2}^{p_t^2} \gamma_u^1 \mathrm{d}\xi_u^1. \tag{2.42}$$

We also have, making use of the ODE for $\{\bar{\alpha}_t^{(s)}\}_{0\leq s\leq t}$, that

$$\bar{\alpha}_t^{(s)} = \alpha_s^{\gamma_2}. \tag{2.43}$$

It follows straightforwardly from equations (2.42), (2.43) that

$$||\alpha_t^{\gamma_2} - \bar{\alpha}_t^{(s)}|| = \left|\left| - \int_s^t \frac{\gamma^1_{p_u^2}}{\gamma^2_{p_u^2}} \nabla_\alpha f(\alpha_u^{\gamma_2}, \beta_u^{\gamma_2}) \mathrm{d}u - \int_{p_s^2}^{p_t^2} \gamma_u^1 \mathrm{d}\xi_u^1 \right|\right| \tag{2.44}$$

$$\leq \underbrace{\left|\left| \int_s^t \frac{\gamma^1_{p_u^2}}{\gamma^2_{p_u^2}} \nabla_\alpha f(\alpha_u^{\gamma_2}, \beta_u^{\gamma_2}) \mathrm{d}u \right|\right|}_{\Omega_{1,\alpha}(s,t)} + \underbrace{\left|\left| \int_{p_s^2}^{p_t^2} \gamma_u^1 \mathrm{d}\xi_u^1 \right|\right|}_{\Omega_{2,\alpha}(s,t)} \tag{2.45}$$

For the first term, by Assumptions 2.1.2 and 2.1.4, which together imply the boundedness of $||\nabla_\alpha f(\cdot,\cdot)||$, we have that for all $T > 0$,

$$\sup_{t\in[s,s+T]} \Omega_{1,\alpha}(s,t) = \sup_{t\in[s,s+T]} \left|\left| \int_s^t \frac{\gamma^1_{p_u^2}}{\gamma^2_{p_u^2}} \nabla_\alpha f(\alpha_u^{\gamma_2}, \beta_u^{\gamma_2}) \mathrm{d}u \right|\right| \tag{2.46}$$

$$\leq \sup_{t\in[s,s+T]} ||\nabla_\alpha f(\alpha_t^{\gamma_2}, \beta_t^{\gamma_2})|| \int_s^{s+T} \frac{\gamma^1_{p_u^2}}{\gamma^2_{p_u^2}} \mathrm{d}u \tag{2.47}$$

$$\leq K \int_s^{s+T} \frac{\gamma^1_{p_u^2}}{\gamma^2_{p_u^2}} \mathrm{d}u \tag{2.48}$$

$$\leq KT \frac{\gamma^1_{p_s^2}}{\gamma^2_{p_s^2}}. \tag{2.49}$$

It follows immediately, using also Assumption 2.1.1, that, for all $T > 0$,

$$\lim_{s\to\infty} \sup_{t\in[s,s+T]} \Omega_{1,\alpha}(s,t) = 0 \ , \quad \text{a.s.} \tag{2.50}$$

For the second term, using the definition of $\{p_t^2\}_{t\geq 0}$, we have that, for sufficiently large s,

$$\sup_{t\in[s,s+T]} \Omega_{2,\alpha}(s,t) = \sup_{t\in[s,s+T]} \left|\left| \int_{p_s^2}^{p_t^2} \gamma_u^1 \mathrm{d}\xi_u^1 \right|\right| \tag{2.51}$$

$$\leq \sup_{t\in[s,s+T]} \left|\left| \int_s^{p_t^2} \gamma_u^1 \mathrm{d}\xi_u^1 \right|\right| = \sup_{t\in[s,s+\tau]} \left|\left| \int_s^t \gamma_u^1 \mathrm{d}\xi_u^1 \right|\right| \tag{2.52}$$

where, in the second line, we have used the fact that $s \leq p_s^2$ for sufficiently large s, and in the final line, we have defined $\tau = \tau_T = p_T^2$. It then follows directly from the first part of Assumption 2.1.3 that, for all $T > 0$,

$$\lim_{s\to\infty} \sup_{t\in[s,s+T]} \Omega_{2,\alpha}(s,t) = 0 \ , \quad \text{a.s.} \tag{2.53}$$

We will now consider $\{\beta_t\}_{0\leq s\leq t}$. By definition, we have that

$$\beta_t = \beta_s - \int_s^t \gamma_u^2 \nabla_\beta g(\alpha_u, \beta_u)\mathrm{d}u - \int_s^t \mathrm{d}\xi_u^2 \tag{2.54}$$

It follows immediately from the definition of $\{\beta_t^{\gamma_2}\}_{0\leq s\leq t}$ that

$$\beta_t^{\gamma_2} = \beta_s^{\gamma_2} - \int_{p_s^2}^{p_t^2} \gamma_u^2 \nabla_\beta g(\alpha_u, \beta_u)\mathrm{d}u - \int_{p_s^2}^{p_t^2} \gamma_u^2 \mathrm{d}\xi_u^2 \tag{2.55}$$

$$= \beta_s^{\gamma_2} - \int_s^t \nabla_\beta g(\alpha_u^{\gamma_2}, \beta_u^{\gamma_2})\mathrm{d}u - \int_{p_s^2}^{p_t^2} \gamma_u^2 \mathrm{d}\xi_u^2. \tag{2.56}$$

We also have, now making use of the ODE for $\{\bar{\beta}_t^{(s)}\}_{0\leq s\leq t}$, that

$$\bar{\beta}_t^{(s)} = \beta_s^{\gamma_2} - \int_s^t \nabla_\beta g(\bar{\alpha}_u^{(s)}, \bar{\beta}_u^{(s)})\mathrm{d}u \tag{2.57}$$

It follows straightforwardly from equations (2.56), (2.57) that

$$||\beta_t^{\gamma_2} - \bar{\beta}_t^{(s)}|| = \left|\left| - \int_s^t \left[\nabla_\beta g(\alpha_u^{\gamma_2}, \beta_u^{\gamma_2}) - \nabla_\beta g(\bar{\alpha}_u^{(s)}, \bar{\beta}_u^{(s)})\right]\mathrm{d}u - \int_{p_s^2}^{p_t^2} \gamma_u^2 \mathrm{d}\xi_u^2 \right|\right| \tag{2.58}$$

$$\leq \underbrace{\left|\left| \int_{p_s^2}^{p_t^2} \gamma_u^2 \mathrm{d}\xi_u^2 \right|\right|}_{\Omega_{1,\beta}(s,t)} + \underbrace{\left|\left| \int_s^t \left[\nabla_\beta g(\alpha_u^{\gamma_2}, \beta_u^{\gamma_2}) - \nabla_\beta g(\bar{\alpha}_u^{(s)}, \bar{\beta}_u^{(s)})\right]\mathrm{d}u \right|\right|}_{\Omega_{2,\beta}(s,t)} \tag{2.59}$$

For the first term, using the second part of Assumption 2.1.3, and arguing as in equations (2.51)-(2.52), we have that, for all $T > 0$,

$$\lim_{s\to\infty} \sup_{t\in[s,s+T]} \Omega_{1,\beta}(s,t) = 0 \ , \quad \text{a.s.} \tag{2.60}$$

For the second term, using elementary properties of the Euclidean norm, and Assumption

2.1.2 (i.e., Lipschitz continuity of $\nabla_\alpha g(\cdot,\cdot)$), we have that, for all $T>0$,

$$\Omega_{2,\beta}(s,t) = \left|\left|\int_s^t \left[\nabla_\beta g(\alpha_u^{\gamma_2},\beta_u^{\gamma_2}) - \nabla_\beta g(\bar{\alpha}_u^{(s)},\bar{\beta}_u^{(s)})\right] \mathrm{d}u\right|\right| \tag{2.61}$$

$$\leq \int_s^t \left|\left|\nabla_\beta g(\alpha_u^{\gamma_2},\beta_u^{\gamma_2}) - \nabla_\beta g(\bar{\alpha}_u^{(s)},\bar{\beta}_u^{(s)})\right|\right| \mathrm{d}u \tag{2.62}$$

$$\leq \int_s^t L_\beta \left|\left|\begin{pmatrix}\alpha_u^{\gamma_2} - \bar{\alpha}_u^{(s)}\\ \beta_u^{\gamma_2} - \bar{\beta}_u^{(s)}\end{pmatrix}\right|\right| \mathrm{d}u \tag{2.63}$$

It remains to observe that, combining inequalities (2.45) and (2.59), and using Grömwall's Inequality, we have

$$\left|\left|\begin{pmatrix}\alpha_t^{\gamma_2}\\ \beta_t^{\gamma_2}\end{pmatrix} - \begin{pmatrix}\bar{\alpha}_t^{(s)}\\ \bar{\beta}_t^{(s)}\end{pmatrix}\right|\right| \leq ||\alpha_t^{\gamma_2} - \bar{\alpha}_t^{(s)}|| + ||\beta_t^{\gamma_2} - \bar{\beta}_t^{(s)}|| \tag{2.64}$$

$$\leq \underbrace{\Omega_{1,\alpha}(s,t) + \Omega_{2,\alpha}(s,t) + \Omega_{1,\beta}(s,t)}_{\Omega(s,t)} + \Omega_{2,\beta}(s,t) \tag{2.65}$$

$$= \Omega(s,t) + \int_s^t L_\beta \left|\left|\begin{pmatrix}\alpha_u^{\gamma_2} - \bar{\alpha}_u^{(s)}\\ \beta_u^{\gamma_2} - \bar{\beta}_u^{(s)}\end{pmatrix}\right|\right| \mathrm{d}u \tag{2.66}$$

$$\leq \Omega(s,t)\exp\left[\int_s^t L_\beta \mathrm{d}u\right] \tag{2.67}$$

$$= \Omega(s,t)\exp\left[L_\beta(t-s)\right], \tag{2.68}$$

where, from (2.50), (2.53) and (2.60), we have that, for all $T>0$,

$$\lim_{s\to\infty}\sup_{t\in[s,s+T]} \Omega(s,t) = 0\ , \quad \text{a.s.} \tag{2.69}$$

It follows immediately from (2.68) and (2.69) that, for all $T>0$,

$$\lim_{s\to\infty}\sup_{t\in[s,s+T]} \left|\left|\begin{pmatrix}\alpha_t^{\gamma_2}\\ \beta_t^{\gamma_2}\end{pmatrix} - \begin{pmatrix}\bar{\alpha}_t^{(s)}\\ \bar{\beta}_t^{(s)}\end{pmatrix}\right|\right| \leq \lim_{s\to\infty}\sup_{t\in[s,s+T]} \left[\Omega(s,t)\exp\left[L_\beta(t-s)\right]\right] \tag{2.70}$$

$$\leq \exp\left[L_\beta T\right] \lim_{s\to\infty}\sup_{t\in[s,s+T]} \Omega(s,t) \tag{2.71}$$

$$= 0\ , \quad \text{a.s.} \tag{2.72}$$

□

Lemma 2.1.2. *Assume that Assumptions 2.1.1-2.1.5 hold. Then, almost surely, for some*

$i \geq 1$,

$$(\alpha_t, \beta_t) \overset{t\to\infty}{\longrightarrow} \{(\alpha, \beta_i^*(\alpha)) \ : \ \alpha \in \mathbb{R}^{d_1}\}. \tag{2.73}$$

Proof. We begin with the observation that, by Lemma 2.1.1, $(\alpha_t^{\gamma_2}, \beta_t^{\gamma_2})$ are asymptotic pseudo-trajectories of (2.35a) - (2.35b). Moreover, by Assumption 2.1.4, they are pre-compact. We can thus apply Theorem 5.7 in Benaïm [36] to conclude that $(\alpha_t^{\gamma_2}, \beta_t^{\gamma_2})$ converges to an internally chain transitive set for (2.35a) - (2.35b).

We next observe that the function $g : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}$ is a strict Lyapunov function for (2.35a) - (2.35b) in the sense of Benaïm [36, Chapter 6.2]. In particular, $g(\bar{\alpha}_t, \bar{\beta}_t)$ is strictly decreasing in t, unless $(\bar{\alpha}_t, \bar{\beta}_t)$ is an equilibrium point of (2.35a) - (2.35b). This follows straightforwardly from

$$\dot{g}(\bar{\alpha}_t, \bar{\beta}_t) = -||\nabla_\beta g(\bar{\alpha}_t, \bar{\beta}_t)||^2 \leq 0. \tag{2.74}$$

with equality if and only if $\nabla_\beta g(\bar{\alpha}_t, \bar{\beta}_t) = 0$. By Assumption 2.1.5, the set of critical values of g is given by $E_g = \cup_{i=1}\{(\alpha, \beta_i^*(\alpha) : \alpha \in \mathbb{R}^d\}$. Since $\beta_i^*(\cdot)$ are discrete and countable, this set has Lebesgue measure zero, and hence empty topological interior. Thus, by Proposition 6.4 in Benaïm [36], every internally chain transitive set for (2.35a) - (2.35b) is contained in E_g. Moreover, by Assumption 2.1.5, the internally chain transitive sets of E_g are precisely the sets $\{(\alpha, \beta_i^*(\alpha)) : \alpha \in \mathbb{R}^{d_1}\}$.

It follows from our two observations that, for some $i \geq 1$, we have that $(\alpha_t^{\gamma_2}, \alpha_t^{\gamma_2}) := (\alpha_{p_t^2}, \beta_{p_t^2}) \to \{(\alpha, \beta_i^*(\alpha)) : \alpha \in \mathbb{R}^{d_1}\}$ as $t \to \infty$. Finally, noting that $t \geq p_t^2$ for sufficiently large t, the result holds. □

2.3.1.3 The Slow Timescale

2.3.1.3.1 Additional Notation

For $i = 1, 2, \ldots$, we will write $\{\underline{\alpha}_t^i\}_{t\geq 0}$ to denote the solutions of the ordinary differential equations

$$\dot{\underline{\alpha}}_t^i = -\nabla_\alpha f(\underline{\alpha}_t^i, \beta_i^*(\underline{\alpha}_t^i)) \tag{2.75}$$

where $\beta_i^*(\cdot) : \mathbb{R}^{d_1} \to \mathbb{R}^{d_2}$, $i = 1, \ldots$, are defined in Assumption 2.1.5.

We can then define $\{\underline{\alpha}_t^{i,(s)}\}_{0\leq s\leq t}$, $i = 1, 2, \ldots$, as the unique solutions of (2.75) which 'start at s', and coincide with the time-scaled process $\{\alpha_t^{\gamma_1}\}_{t\geq 0}$ at s. That is,

$$\dot{\underline{\alpha}}_t^{i,(s)} = -\nabla_\alpha f(\underline{\alpha}_t^{i,(s)}, \beta_i^*(\underline{\alpha}_t^{i,(s)})) \ , \quad \underline{\alpha}_s^{i,(s)} = \alpha_s^{\gamma_1}. \tag{2.76}$$

We can also define $\{\underline{\alpha}_t^{i,[s]}\}_{0\leq s\leq t}$, $i = 1, 2, \ldots$, as the unique solutions of (2.75) which 'end at s', and coincide with the time-scaled process $\{\alpha_t^{\gamma_1}\}_{t\geq 0}$ at s. That is,

$$\dot{\underline{\alpha}}_t^{i,[s]} = -\nabla_\alpha f(\underline{\alpha}_t^{i,[s]}, \beta_i^*(\underline{\alpha}_i^{[s]}(t))) \ , \quad \underline{\alpha}_s^{i,[s]} = \alpha_s^{\gamma_1}. \tag{2.77}$$

2.3.1.3.2 Proof of Convergence

We now demonstrate, using these processes, that for some $i \geq 1$, the time-scaled process $\alpha_t^{\gamma_1}$ is an asymptotic pseudo-trajectory of the flow induced by the ODE for $\underline{\alpha}_t^i$.

Lemma 2.1.3. *Assume that Assumptions 2.1.1-2.1.5 hold. Then, for any $T > 0$, and the $i \geq 1$ given in Lemma 2.1.2,*

$$\lim_{s\to\infty} \sup_{t\in[s,s+T]} ||\alpha_t^{\gamma_1} - \underline{\alpha}_t^{i,(s)}|| = 0 \ , \quad a.s. \tag{2.78a}$$

$$\lim_{s\to\infty} \sup_{t\in[s-T,s]} ||\alpha_t^{\gamma_1} - \underline{\alpha}_t^{i,[s]}|| = 0 \ , \quad a.s. \tag{2.78b}$$

Proof. This proof is similar in style to the proof of Lemma 2.1.1. Once more, we will prove only the first part of this Lemma, as the method for proving the second part is entirely analogous. By definition of the process $\{\alpha_t\}_{0\leq s\leq t}$, we have

$$\alpha_t = \alpha_s - \int_s^t \gamma_u^1 \nabla_\alpha f(\alpha_u, \beta_u)\mathrm{d}u - \int_s^t \gamma_u^1 \mathrm{d}\xi_u^1 \tag{2.79}$$

It follows immediately from the definition of $\{\alpha_t^{\gamma_2}\}_{0\leq s\leq t}$ that

$$\alpha_t^{\gamma_1} = \alpha_s^{\gamma_1} - \int_{p_s^1}^{p_t^1} \gamma_u^1 \nabla_\alpha f(\alpha_u, \beta_u)\mathrm{d}u - \int_{p_s^1}^{p_t^1} \gamma_u^1 \mathrm{d}\xi_u^1 \tag{2.80}$$

$$= \alpha_s^{\gamma_1} - \int_s^t \nabla_\alpha f(\alpha_u^1, \beta_u^1)\mathrm{d}u - \int_{p_s^1}^{p_t^1} \gamma_u^1 \mathrm{d}\xi_u^1. \tag{2.81}$$

We also have, making use of the ODE for $\{\underline{\alpha}_t^{i,(s)}\}_{0\leq s\leq t}$, that

$$\underline{\alpha}_i^{(s)}(t) = \underline{\alpha}_i^{(s)}(s) - \int_s^t \nabla_\alpha f(\underline{\alpha}_u^{i,(s)}, \beta_i^*(\underline{\alpha}_u^{i,(s)}))\mathrm{d}s. \tag{2.82}$$

It follows straightforwardly from equations (2.81), (2.82) that

$$\left|\left|\alpha_t^{\gamma_1} - \underline{\alpha}_t^{i,(s)}\right|\right| = \left|\left| -\int_s^t \left[\nabla_\alpha f(\alpha_u^{\gamma_1}, \beta_u^{\gamma_1}) - \nabla_\alpha f(\underline{\alpha}_u^{i,(s)}, \beta_i^*(\underline{\alpha}_u^{i,(s)}))\right] \mathrm{d}u - \int_{p_s^1}^{p_t^1} \gamma_u^1 \mathrm{d}\xi_u^1 \right|\right| \tag{2.83}$$

$$\leq \underbrace{\left|\left|\int_{p_s^1}^{p_t^1} \gamma_u^1 \mathrm{d}\xi_u^1\right|\right|}_{\Pi_{1,\alpha}(s,t)} + \underbrace{\left|\left|\int_s^t \left[\nabla_\alpha f(\alpha_u^{\gamma_1}, \beta_u^{\gamma_1}) - \nabla_\alpha f(\underline{\alpha}_u^{i,(s)}, \beta_i^*(\underline{\alpha}_u^{i,(s)}))\right] \mathrm{d}u\right|\right|}_{\Pi_{2,\alpha}(s,t)} \tag{2.84}$$

For the first term, using the first part of Assumption 2.1.3, and arguing as in equations (2.51)-(2.52), we have that, a.s., for all $T > 0$,

$$\lim_{s\to\infty} \sup_{t\in[s,s+T]} \Pi_{1,\alpha}(s,t) = 0. \tag{2.85}$$

For the second term, using the triangle inequality, Assumptions 2.1.2 and 2.1.4, which together imply the boundedness of $||\nabla_\alpha f(\cdot,\cdot)||$, and Assumption 2.1.5, which guarantees the Lipschitz continuity of $\beta_i^*(\cdot)$, we have that, a.s., for all $T > 0$,

$$\Pi_{2,\alpha}(s,t) \leq \int_s^t \left|\left|\nabla_\alpha f(\alpha_u^{\gamma_1}, \beta_u^{\gamma_1}) - \nabla_\alpha f(\underline{\alpha}_u^{i,(s)}, \beta_i^*(\underline{\alpha}_u^{i,(s)}))\right|\right| \mathrm{d}u \tag{2.86}$$

$$\begin{aligned}
&\leq \int_s^t \left|\left|\nabla_\alpha f(\alpha_u^{\gamma_1}, \beta_u^{\gamma_1}) - \nabla_\alpha f(\alpha_u^{\gamma_1}, \beta_i^*(\alpha_u^{\gamma_1}))\right|\right| \mathrm{d}u \\
&\quad + \int_s^t \left|\left|\nabla_\alpha f(\alpha_u^{\gamma_1}, \beta_i^*(\alpha_u^{\gamma_1})) - \nabla_\alpha f(\underline{\alpha}_u^{i,(s)}, \beta_i^*(\underline{\alpha}_u^{i,(s)})\right|\right| \mathrm{d}u
\end{aligned} \tag{2.87}$$

$$\begin{aligned}
&\leq \int_s^t L_\alpha \left[\left|\left|\alpha_u^{\gamma_1} - \alpha_u^{\gamma_1}\right|\right| + \left|\left|\beta_u^{\gamma_1} - \beta_i^*(\alpha_u^{\gamma_1})\right|\right|\right] \mathrm{d}u \\
&\quad + \int_s^t L_\alpha \left[\left|\left|\alpha_u^{\gamma_1} - \underline{\alpha}_u^{i,(s)}\right|\right| + \left|\left|\beta_i^*(\alpha_u^{\gamma_1}) - \beta_i^*(\underline{\alpha}_u^{i,(s)})\right|\right|\right] \mathrm{d}u
\end{aligned} \tag{2.88}$$

$$\leq \underbrace{\int_s^t L_\alpha \left|\left|\beta_u^{\gamma_1} - \beta_i^*(\alpha_u^{\gamma_1})\right|\right| \mathrm{d}u}_{\Pi_{2,\alpha}^{(1)}(s,t)} + \underbrace{\int_s^t L_\alpha (1 + L_{\beta_i^*}) \left|\left|\alpha_u^{\gamma_1} - \underline{\alpha}_u^{i,(s)}\right|\right| \mathrm{d}u}_{\Pi_{2,\alpha}^{(2)}(s,t)}. \tag{2.89}$$

For the first term, we have that, a.s., for all $T > 0$,

$$\sup_{t\in[s,s+T]} \Pi_{2,\alpha}^{(1)}(s,t) = \sup_{t\in[s,s+T]} \int_s^t L_\alpha \left|\left|\beta_u^{\gamma_1} - \beta_i^*(\alpha_u^{\gamma_1})\right|\right| \mathrm{d}u \tag{2.90}$$

$$\leq L_\alpha T \sup_{t\geq s} \left|\left|\beta_t^{\gamma_1} - \beta_i^*(\alpha_t^{\gamma_1})\right|\right|. \tag{2.91}$$

It then follows, using Lemma 2.1.2, that, a.s., for all $T > 0$,

$$\lim_{s\to\infty} \sup_{t\in[s,s+T]} \Pi^{(1)}_{2,\alpha}(s,t) \leq L_\alpha T \limsup_{s\to\infty} \left|\left|\beta^{\gamma_1}_s - \beta^*_i(\alpha^{\gamma_1}_s)\right|\right| \tag{2.92}$$

$$= L_\alpha T \lim_{s\to\infty} \left|\left|\beta^{\gamma_1}_s - \beta^*_i(\alpha^{\gamma_1}_s)\right|\right| \tag{2.93}$$

$$= 0. \tag{2.94}$$

It remains to observe, combining inequalities (2.84) and (2.89), and making use of Grömwall's Inequality, that

$$\left|\left|\alpha^{\gamma_1}_t - \underline{\alpha}^{i,(s)}_t\right|\right| \leq \underbrace{\Pi_{1,\alpha}(s,t) + \Pi^{(1)}_{2,\alpha}(s,t)}_{\Pi(s,t)} + \Pi^{(2)}_{2,\alpha}(s,t) \tag{2.95}$$

$$= \Pi(s,t) + \int_s^t L_\alpha(1 + L_{\beta^*_i})\left|\left|\alpha^{\gamma_1}_u - \underline{\alpha}^{i,(s)}_u\right|\right| \mathrm{d}u \tag{2.96}$$

$$\leq \Pi(s,t)\exp\left[\int_s^t L_\alpha(1+L_{\beta^*_i})\mathrm{d}u\right] \tag{2.97}$$

$$= \Pi(s,t)\exp\left[L_\alpha(1+L_{\beta^*_i})(t-s)\right] \tag{2.98}$$

where, from (2.85) and (2.94), we have that, a.s., for all $T > 0$,

$$\lim_{s\to\infty} \sup_{t\in[s,s+T]} \Pi(s,t) = 0. \tag{2.99}$$

It follows immediately from (2.98) and (2.99) that, a.s., for all $T > 0$,

$$\lim_{s\to\infty} \sup_{t\in[s,s+T]} \left|\left|\alpha^{\gamma_1}_t - \underline{\alpha}^{i,(s)}_t\right|\right| \leq \lim_{s\to\infty} \sup_{t\in[s,s+T]} \left[\Pi(s,t)\exp\left[L_\alpha(1+L_{\beta^*_i})(t-s)\right]\right] \tag{2.100}$$

$$\leq \exp\left[L_\alpha(1+L_{\beta^*_i})\right] \lim_{s\to\infty} \sup_{t\in[s,s+T]} \Pi(s,t) \tag{2.101}$$

$$= 0. \tag{2.102}$$

□

Lemma 2.1.4. *Assume that Assumptions 2.1.1-2.1.6 hold. Then, almost surely*

$$\alpha_t \overset{t\to\infty}{\longrightarrow} \{\alpha \in \mathbb{R}^{d_1} : \nabla_\alpha f(\alpha, \beta^*_i(\alpha)) = 0\}. \tag{2.103}$$

Proof. The proof follows a similar trajectory to the proof of Lemma 2.1.2, now with some simplifications. By Lemma 2.1.3, $\alpha^{\gamma_1}_t$ is an asymptotic pseudo-trajectory for (2.75). Moreover, it is pre-compact by Assumption 2.1.4. Thus, applying Theorem 5.7 in Benaïm [36], it follows that $\alpha^{\gamma_1}_t := \alpha_{p^1_t}$ converges to an internally chain transitive set of (2.75). The same is thus also true for α_t, noting as before that $t \geq p_t^1$ for sufficiently large t. Finally, by Assumption 2.1.6, the only internally chain transitive sets of (2.75) are its (possibly non-isolated) equilibrium points. The result follows immediately. $\square$

2.3.1.4 Proof of Theorem 2.1

Theorem 2.1. *Assume that Assumptions 2.1.1-2.1.6 hold. Then, almost surely,*

$$\lim_{t\to\infty} \nabla_\alpha f(\alpha_t, \beta_t) = \lim_{t\to\infty} \nabla_\beta g(\alpha_t, \beta_t) = 0. \tag{2.104}$$

Proof. The result is an immediate consequence of Lemmas 2.1.2 and 2.1.4. By Lemma 2.1.2, the process $(\alpha_t, \beta_t) \to \{(\alpha, \beta_i^*(\alpha)) : \alpha \in \mathbb{R}^{d_1}\}$ a.s., for some $i \geq 1$. By Lemma 2.1.4, the process $\alpha_t \to \{\alpha \in \mathbb{R}^{d_1} : \nabla_\alpha f(\alpha, \beta_i^*(\alpha)) = 0\}$ a.s.. Together, these lemmas imply that, for some $i \geq 1$,

$$(\alpha_t, \beta_t) \to \{(\alpha, \beta_i^*(\alpha)) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} : \nabla_\alpha f(\alpha, \beta_i^*(\alpha)) = 0\} \quad \text{a.s.} \tag{2.105}$$

It follows, in particular, that $\nabla_\alpha f(\alpha_t, \beta_t) \to 0$ and $\nabla_\beta g(\alpha_t, \beta_t) \to 0$ with probability one. $\square$

2.3.2 Proof of Theorem 2.2

In this Section, we provide a proof of Theorem 2.2. Our proof combines the methods in [420, Lemma 3.1] and [430, Lemma 1], adapted appropriately to the two-timescale setting, with the results of Theorem 2.1.

Lemma 2.2.1. *For $0 \leq s \leq t$, define*

$$\Gamma_\alpha(s,t) = \int_s^t \gamma_u^1 \left[F(\alpha_u, \beta_u, \mathcal{X}_u) - \nabla_\alpha f(\alpha_u, \beta_u)\right] \mathrm{d}u \tag{2.106a}$$

$$\Gamma_\beta(s,t) = \int_s^t \gamma_u^2 \left[G(\alpha_u, \beta_u, \mathcal{X}_u) - \nabla_\beta g(\alpha_u, \beta_u)\right] \mathrm{d}u. \tag{2.106b}$$

Assume that Assumptions 2.2.1-2.2.3a hold. Then, for all $T \in [0, \infty)$, with probability one,

$$\lim_{s\to\infty} \sup_{t\in[s,s+T]} ||\Gamma_\alpha(s,t)|| = \lim_{s\to\infty} \sup_{t\in[s,s+T]} ||\Gamma_\beta(s,t)|| = 0. \tag{2.107}$$

Proof. We will prove only the first part of the Lemma, as the method for proving the second part is entirely analogous. By Assumption 2.2.2c, there exists a differentiable function $f : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}$, and a unique Borel-measurable function $\tilde{F} : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^{d_3}$ such that $\nabla_\alpha f(\cdot)$ is Lipschitz continuous, and moreover, such that the Poisson equation

$$\mathcal{A}_{\mathcal{X}}\tilde{F}(\alpha,\beta,x) = F(\alpha,\beta,x) - \nabla_\alpha f(\alpha,\beta) \tag{2.108}$$

has a unique, twice-differentiable solution which grows at most polynomially in x. In particular, there exist $K', q' > 0$ such that

$$\sum_{i=0}^{2} ||\partial_\alpha^i \tilde{F}(\alpha,\beta,x)|| + ||\partial_x\partial_\alpha \tilde{F}(\alpha,\beta,x)|| \leq K(1+||x||^{q'}), \tag{2.109a}$$
$$\sum_{i=0}^{2} ||\partial_\beta^i \tilde{F}(\alpha,\beta,x)|| + ||\partial_x\partial_\beta \tilde{F}(\alpha,\beta,x)|| \leq K(1+||x||^{q'}). \tag{2.109b}$$

Now consider the vector-valued function $\hat{F}(\alpha,\beta,x,t) = \gamma_t^1\tilde{F}(\alpha,\beta,x)$, with $\tilde{F}$ as defined in (2.108). Applying Itô's Lemma to each component of $\hat{F}$, we obtain, for $i = 1,\dots,d_1$,

$$\begin{aligned}
&\hat{F}_i(\alpha_t,\beta_t,\mathcal{X}_t,t) - \hat{F}_i(\alpha_s,\beta_s,\mathcal{X}_s,s) \qquad (2.110)\\
&= \int_s^t \partial_\tau \hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\mathrm{d}\tau + \int_s^t \mathcal{A}_{\mathcal{X}}\hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\mathrm{d}\tau \\
&- \int_s^t \gamma_\tau^1 F(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau)\cdot\nabla_\alpha\hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\mathrm{d}\tau \\
&- \int_s^t \gamma_\tau^2 G(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau)\cdot\nabla_\beta\hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\mathrm{d}\tau \\
&+ \frac{1}{2}\int_s^t (\gamma_\tau^1)^2\nabla_\alpha\nabla_\alpha\hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau) : \mathrm{d}\,[\zeta_1,\zeta_1]_\tau \\
&+ \frac{1}{2}\int_s^t (\gamma_\tau^2)^2\nabla_\beta\nabla_\beta\hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau) : \mathrm{d}\,[\zeta_2,\zeta_2]_\tau \\
&- \int_s^t \gamma_\tau^1\nabla_\alpha\hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\cdot\mathrm{d}\zeta_\tau^1 \\
&- \int_s^t \gamma_\tau^2\nabla_\beta\hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\cdot\mathrm{d}\zeta_\tau^2 \\
&+ \int_s^t \nabla_x\hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\cdot\Psi(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau)\mathrm{d}b_\tau
\end{aligned}$$

$$
\begin{aligned}
&+\int_s^t \gamma_\tau^1\gamma_\tau^2 \nabla_\alpha\nabla_\beta \hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau) : \mathrm{d}\left[\zeta_1,\zeta_2\right]_\tau \\
&-\int_s^t \gamma_\tau^1 \nabla_\alpha\nabla_x \hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau) : \Psi(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau)\mathrm{d}\left[\zeta_1,b\right]_\tau \\
&-\int_s^t \gamma_\tau^2 \nabla_\beta\nabla_x \hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau) : \Psi(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau)\mathrm{d}\left[\zeta_2,b\right]_\tau
\end{aligned}
$$

For the sake of brevity, we will proceed under the assumption that the continuous semi-martingales $\{\zeta_t^i\}_{t\geq 0}$, $i=1,2$ are, in fact, diffusion processes. We should emphasise, however, that this assumption does not change the subsequent analysis in any meaningful way, and can be easily relaxed. In particular, we will thus assume that

$$
\mathrm{d}\zeta_t^i = \zeta_i^{(1)}(\alpha_t,\beta_t,\mathcal{X}_t)\mathrm{d}t + \zeta_i^{(2)}(\alpha_t,\beta_t,\mathcal{X}_t)\mathrm{d}z_t^i \tag{2.111}
$$

where $\zeta_i^{(1)}(\alpha,\beta,\cdot):\mathbb{R}^{d_3}\rightarrow\mathbb{R}^{d_i}$ and $\zeta_i^{(2)}(\alpha,\beta,\cdot):\mathbb{R}^{d_3}\rightarrow\mathbb{R}^{d_i\times d_5^i}$ are Borel measurable functions; and $\{z_t^i\}_{t\geq 0}$ are $\mathbb{R}^{d_5^i}$ valued Wiener processes. In this case, recalling the definition of the functions c_{z_1,z_2}, $c_{z_1,b}$ and $c_{z_2,b}$ in Assumption 2.2.3c, the previous equation becomes

$$
\begin{aligned}
&\hat{F}_i(\alpha_t,\beta_t,\mathcal{X}_t,t) - \hat{F}_i(\alpha_s,\beta_s,\mathcal{X}_s,s) \qquad (2.112)\\
&=\int_s^t \partial_\tau \hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\mathrm{d}\tau + \int_s^t \mathcal{A}_{\mathcal{X}}\hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\mathrm{d}\tau \\
&+\int_s^t \mathcal{A}_\alpha\hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\mathrm{d}\tau + \int_s^t \mathcal{A}_\beta\hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\mathrm{d}\tau \\
&-\int_s^t \gamma_\tau^1\nabla_\alpha\hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\cdot\zeta_1^{(2)}(\tau)\mathrm{d}z_\tau^1 \\
&-\int_s^t \gamma_\tau^2\nabla_\beta\hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\cdot\zeta_2^{(2)}(\tau)\mathrm{d}z_\tau^2 \\
&+\int_s^t \nabla_x\hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\cdot\Psi(\tau)\mathrm{d}b_\tau \\
&-\int_s^t \gamma_\tau^1\mathrm{Tr}\left[\nabla_\alpha\nabla_x\hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\Psi(\tau)\zeta_1^{(2)}(\tau)c_\tau^{z_1,b}\right]\mathrm{d}\tau \\
&-\int_s^t \gamma_\tau^2\mathrm{Tr}\left[\nabla_\beta\nabla_x\hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\Psi(\tau)\zeta_2^{(2)}(\tau)c_\tau^{z_2,b}\right]\mathrm{d}\tau
\end{aligned}
$$

$$+\int_s^t \gamma_\tau^1\gamma_\tau^2 \mathrm{Tr}\left[\nabla_\alpha\nabla_\beta \hat{F}_i(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\zeta_1^{(2)}(\tau)\zeta_2^{(2)}(\tau)c_\tau^{z_1,z_2}\right]\mathrm{d}\tau,$$

where, $\mathcal{A}_\alpha$ and $\mathcal{A}_\beta$ are the infinitesimal generators of the processes $\{\alpha_t\}_{t\geq 0}$ and $\{\beta_t\}_{t\geq 0}$; and $\nabla_\alpha\nabla_\beta u_k(\alpha,\beta,x,\tau)_{ij} = \partial_{\alpha_i}\partial_{\beta_j}u_k(\alpha,\beta,x,\tau)$, with $\nabla_\alpha\nabla_x$ and $\nabla_\beta\nabla_x$ defined similarly. For the sake of simplicity, we have temporarily suppressed the dependence of the functions $\zeta_i^{(1)},\zeta_i^{(2)}$, $i=1,2$, and Ψ on $\{\alpha_t\}_{t\geq 0}, \{\beta_t\}_{t\geq 0}$ and $\{\mathcal{X}_t\}_{t\geq 0}$. It follows straightforwardly that

$$\begin{aligned}
\Gamma_\alpha(s,t) &= \int_s^t \gamma_\tau^1\left[F(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau) - \nabla_\alpha f(\alpha_\tau,\beta_\tau)\right]\mathrm{d}\tau && (2.113)\\
&= \int_s^t \gamma_\tau^1 \mathcal{A}_\mathcal{X}\tilde{F}(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau)\mathrm{d}\tau && (2.114)\\
&= \int_s^t \mathcal{A}_\mathcal{X}\hat{F}(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau)\mathrm{d}\tau && (2.115)\\
&= \gamma_t^1\tilde{F}(\alpha_t,\beta_t,\mathcal{X}_t) - \gamma_s^1\tilde{F}(\alpha_s,\beta_s,\mathcal{X}_s) && (2.116)\\
&- \int_s^t \dot{\gamma}_\tau^1\partial_\tau\tilde{F}(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau)\mathrm{d}\tau\\
&- \int_s^t \gamma_\tau^1\mathcal{A}_\alpha\tilde{F}(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau)\mathrm{d}\tau - \int_s^t \gamma_\tau^1\mathcal{A}_\beta\tilde{F}(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau)\mathrm{d}\tau\\
&+ \int_s^t (\gamma_\tau^1)^2\nabla_\alpha\tilde{F}(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\cdot\zeta_1^{(2)}(\tau)\mathrm{d}z_\tau^1\\
&+ \int_s^t \gamma_\tau^1\gamma_\tau^2\nabla_\beta\tilde{F}(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau,\tau)\cdot\zeta_2^{(2)}(\tau)\mathrm{d}z_\tau^2\\
&- \int_s^t \gamma_\tau^1\nabla_x\tilde{F}(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau)\cdot\Psi(\tau)\mathrm{d}b_\tau\\
&+ \int_s^t (\gamma_\tau^1)^2\mathrm{Tr}\left[\nabla_\alpha\nabla_x\tilde{F}(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau)\Psi(\tau)\zeta_1^{(2)}(\tau)c_\tau^{z_1,b}\right]\mathrm{d}\tau\\
&+ \int_s^t \gamma_\tau^1\gamma_\tau^2\mathrm{Tr}\left[\nabla_\beta\nabla_x\tilde{F}(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau)\Psi(\tau)\zeta_2^{(2)}(\tau)c_\tau^{z_2,b}\right]\mathrm{d}\tau\\
&- \int_s^t (\gamma_\tau^1)^2\gamma_\tau^2\mathrm{Tr}\left[\nabla_\alpha\nabla_\beta\tilde{F}(\alpha_\tau,\beta_\tau,\mathcal{X}_\tau)\zeta_1^{(2)}(\tau)\zeta_2^{(2)}(\tau)c_\tau^{z_1,z_2}\right]\mathrm{d}\tau
\end{aligned}$$

We will now bound each of these terms in turn. We first define

$$J_t^{(1)} = \gamma_t^1 \sup_{\tau\in[0,t]} ||\tilde{F}(\alpha_\tau, \beta_\tau, \mathcal{X}_\tau)||. \tag{2.117}$$

By Assumption 2.2.2c and Assumption 2.2.2e, there exists $q > 0$, and $K, K' > 0$ such that for all t sufficiently large, we have

$$\mathbb{E}\big[\big(J_t^{(1)}\big)^2\big] = \mathbb{E}\big[(\gamma_t^1)^2 \sup_{\tau\in[0,t]} ||\tilde{F}(\alpha_\tau, \beta_\tau, \mathcal{X}_\tau)||^2\big] \tag{2.118}$$

$$\leq K(\gamma_t^1)^2\big[1 + \mathbb{E}\sup_{\tau\in[0,t]} ||\mathcal{X}_\tau||^q\big] \tag{2.119}$$

$$\leq K(\gamma_t^1)^2\big[1 + K'\sqrt{t}\big] \tag{2.120}$$

$$\leq K''(\gamma_t^1)^2\sqrt{t}. \tag{2.121}$$

By Assumption 2.2.1, there exists $r_1 > 0$ such that $\lim_{t\to\infty}(\gamma_t^1)^2 t^{\frac{1}{2}+2r_1} = 0$. In particular, there exists $T > 0$ such that for all $t \geq T$,

$$(\gamma_t^1)^2 t^{\frac{1}{2}+2r_1} \leq 1. \tag{2.122}$$

Now suppose that, for any $0 < \delta < r_1$, we define the event $A_t^\delta = \{J_t^{(1)} \cdot t^{r_1-\delta} \geq 1\}$. Then, by Markov's inequality, equation (2.121), and equation (2.122), we have that, for all $t \geq T$,

$$\mathbb{P}(A_t^\delta) \leq \mathbb{E}\left[(J_t^{(1)})^2\right] t^{2(r_1-\delta)} \leq K''(\gamma_t^1)^2 t^{\frac{1}{2}+2r_1-2\delta} \leq K'' t^{-2\delta} \tag{2.123}$$

It follows that $\sum_{n=1}^\infty \mathbb{P}(A_{2^n}^\delta) < \infty$. By the Borel-Cantelli Lemma, this observation implies that only finitely many events $A_{2^n}^\delta$ can occur. Therefore, there exists a random index $n_0(\omega)$ such that

$$J_{2^n}^{(1)} \cdot 2^{n(r_1-\delta)} \leq 1 \tag{2.124}$$

for all $n \geq n_0$. Equivalently, there exists a finite positive random variable $d(\omega)$ and a deterministic $0 < n_1 < \infty$ such that for all $n \geq n_1$,

$$J_{2^n}^{(1)} \cdot 2^{n(r-\delta)} \leq d(\omega). \tag{2.125}$$

Thus, for $t \in [2^n, 2^{n+1}]$, and $n \geq n_1$, we have, for some constant $0 < K < \infty$,

$$J_t^{(1)} = \gamma_t^1 \sup_{\tau\in[0,t]} ||\tilde{F}(\alpha_\tau, \beta_\tau, \mathcal{X}_\tau)|| \tag{2.126}$$

$$\leq K\gamma_{2^{n+1}}^1 \sup_{\tau\in[0,2^{n+1}]} ||\tilde{F}(\alpha_\tau, \beta_\tau, \mathcal{X}_\tau)|| \tag{2.127}$$

$$= KJ_{2^{n+1}}^{(1)} \tag{2.128}$$

$$\leq K\frac{d(\omega)}{2^{(n+1)(r_1-\delta)}} \tag{2.129}$$

$$\leq K\frac{d(\omega)}{t^{r_1-\delta}}. \tag{2.130}$$

It follows that, for all $t \geq 2^{n_0}$, with probability one,

$$J_t^{(1)} \leq K\frac{d(\omega)}{t^{r_1-\delta}} \to 0 \text{ as } t \to \infty. \tag{2.131}$$

We next define

$$\begin{aligned} J_t^{(2)} = \int_0^t \Bigg|\Bigg| & \dot{\gamma}_\tau^1 \partial_\tau \tilde{F}(\alpha_\tau, \beta_\tau, \mathcal{X}_\tau) \\ & + \gamma_\tau^1 \mathcal{A}_\alpha \tilde{F}(\alpha_\tau, \beta_\tau, \mathcal{X}_\tau) + \gamma_\tau^1 \mathcal{A}_\beta \tilde{F}(\alpha_\tau, \beta_\tau, \mathcal{X}_\tau) \\ & - (\gamma_\tau^1)^2 \mathrm{Tr}\left[\nabla_\alpha \nabla_x \tilde{F}(\alpha_\tau, \beta_\tau, \mathcal{X}_\tau)\Psi(\tau)\zeta_1^{(2)}(\tau)c_\tau^{z_1,b}\right] \\ & - \gamma_\tau^1\gamma_\tau^2 \mathrm{Tr}\left[\nabla_\beta \nabla_x \tilde{F}(\alpha_\tau, \beta_\tau, \mathcal{X}_\tau)\Psi(\tau)\zeta_2^{(2)}(\tau)c_\tau^{z_2,b}\right] \\ & + (\gamma_\tau^1)^2\gamma_\tau^2 \mathrm{Tr}\left[\nabla_\alpha \nabla_\beta \tilde{F}(\alpha_\tau, \beta_\tau, \mathcal{X}_\tau)\zeta_1^{(2)}(\tau)\zeta_2^{(2)}(\tau)c_\tau^{z_1,z_2}\right] \Bigg|\Bigg| \mathrm{d}\tau. \end{aligned} \tag{2.132}$$

By Assumptions 2.2.1, 2.2.2d, 2.2.2c, 2.2.2e, 2.2.3b and 2.2.3c, there exists $q > 0$, and constants $K, K', K'' > 0$ such that

$$\begin{aligned} \sup_{t\geq 0} \mathbb{E}[J_t^{(2)}] \leq K \int_0^\infty & \left(\dot{\gamma}_\tau^1 + (\gamma_\tau^1)^2 + \gamma_\tau^1\gamma_\tau^2 + (\gamma_\tau^1)^2\right. \\ & \left. + \gamma_\tau^1\gamma_\tau^2 + (\gamma_\tau^1)^2\gamma_\tau^2\right)(1 + \mathbb{E}||\mathcal{X}(\tau)||^q)\mathrm{d}\tau \end{aligned} \tag{2.133}$$

$$\begin{aligned} \leq KK' \int_0^\infty & \left(\dot{\gamma}_\tau^1 + (\gamma_\tau^1)^2 + (\gamma_\tau^2)^2 + (\gamma_\tau^1)^2\right. \\ & \left. + \gamma_\tau^1\gamma_\tau^2 + (\gamma_\tau^1)^2\gamma_\tau^2\right)\mathrm{d}\tau \end{aligned} \tag{2.134}$$

$$\leq KK'K'' < \infty. \tag{2.135}$$

In particular, the first inequality follows from Assumptions 2.2.2d, 2.2.2c, 2.2.3b and 2.2.3c, using additionally the fact that $\mathcal{A}_\alpha$ contains at least a factor of γ_t^1, and $\mathcal{A}_\beta$ contains at least a factor of γ_t^2. The second inequality follows from the first part of Assumption 2.2.2e. The final inequality follows from Assumption 2.2.1. It follow that there exists a finite random variable, say $\bar{J}_\infty^{(2)}$, such that, with probability one,

$$\lim_{t\to\infty} J_t^{(2)} = \bar{J}_\infty^{(2)}. \tag{2.136}$$

Finally, we define

$$J_t^{(3)} = \int_0^t (\gamma_\tau^1)^2 \nabla_\alpha \tilde{F}(\alpha_\tau, \beta_\tau, \mathcal{X}_\tau, \tau) \cdot \zeta_1^{(2)}(\tau) \mathrm{d}z_\tau^1 \tag{2.137}$$
$$+ \gamma_\tau^1 \gamma_\tau^2 \nabla_\beta \tilde{F}(\alpha_\tau, \beta_\tau, \mathcal{X}_\tau, \tau) \cdot \zeta_2^{(2)}(\tau) \mathrm{d}z_\tau^2$$
$$- \gamma_\tau^1 \nabla_x \tilde{F}(\alpha_\tau, \beta_\tau, \mathcal{X}_\tau) \cdot \Psi(\tau) \mathrm{d}b_\tau$$

By the Itô Isometry, and Assumptions 2.2.1, 2.2.2d, 2.2.2c, 2.2.2e, 2.2.3b and 2.2.3c, similar calculations to those for $J_t^{(2)}$ show there exists $q > 0$, and constants $K, K', K'' > 0$ such that

$$\sup_{t \geq 0} \mathbb{E}[||J_t^{(3)}||^2] \leq K \int_0^\infty \big((\gamma_\tau^1)^4 + (\gamma_\tau^1)^2 (\gamma_\tau^2)^2 + (\gamma_\tau^1)^2 + 2(\gamma_\tau^1)^3 \gamma_\tau^2 \tag{2.138}$$
$$+ 2(\gamma_\tau^1)^2 \gamma_\tau^2 + 2(\gamma_\tau^1)^3\big) \left(1 + \mathbb{E}||\mathcal{X}(\tau)||^q\right) \mathrm{d}\tau$$
$$\leq KK' \int_0^\infty \big((\gamma_\tau^1)^4 + (\gamma_\tau^1)^2 (\gamma_\tau^2)^2 + (\gamma_\tau^1)^2 \tag{2.139}$$
$$+ 2(\gamma_\tau^1)^3 \gamma_\tau^2 + 2(\gamma_\tau^1)^2 \gamma_\tau^2 + 2(\gamma_\tau^1)^3\big) \,\mathrm{d}\tau$$
$$\leq KK'K'' < \infty. \tag{2.140}$$

Thus, by Doob's martingale convergence theorem, there exists a square integrable random variable, say $\bar{J}_\infty^{(3)}$, such that, with probability one and in L^2,

$$\lim_{t \to \infty} J_t^{(3)} = \bar{J}_\infty^{(3)}. \tag{2.141}$$

It remains only to observe that

$$||\Gamma_\alpha(s,t)|| \leq J_t^{(1)} + J_s^{(1)} + J_t^{(2)} - J_s^{(2)} + ||J_t^{(3)} - J_s^{(3)}||. \tag{2.142}$$

Together with (2.131), (2.136) and (2.141), this expression implies that for all $T \in [0, \infty)$, with probability one,

$$\lim_{s \to \infty} \sup_{t \in [s, s+T]} ||\Gamma_\alpha(s, s+T)|| = 0. \tag{2.143}$$

□

Theorem 2.2. *Assume that Assumptions 2.2.1 - 2.2.3c and 2.1.4 hold. In addition, assume that Assumptions 2.1.5 - 2.1.6 hold for the functions $f(\cdot)$ and $g(\cdot)$ defined in*

Assumption 2.2.2c. Then, almost surely,

$$\lim_{t\to\infty} \nabla_\alpha f(\alpha_t,\beta_t) = \lim_{t\to\infty} \nabla_\beta g(\alpha_t,\beta_t) = 0. \tag{2.144}$$

Proof. We begin with the observation that Algorithm (2.18a) - (2.18b) can be written in the form of Algorithm (2.9a) - (2.9b), viz

$$\mathrm{d}\alpha_t = -\gamma_t^1\Big[\nabla_\alpha f(\alpha_t,\beta_t)\mathrm{d}t + \underbrace{\big(F(\alpha_t,\beta_t,\mathcal{X}_t) - \nabla_\alpha f(\alpha_t,\beta_t)\big)\mathrm{d}t + \mathrm{d}\zeta_t^1}_{=\mathrm{d}\xi_t^1}\Big], \tag{2.145a}$$

$$\mathrm{d}\beta_t = -\gamma_t^2\Big[\nabla_\beta g(\alpha_t,\beta_t)\mathrm{d}t + \underbrace{\big(G(\alpha_t,\beta_t,\mathcal{X}_t) - \nabla_\beta g(\alpha_t,\beta_t)\big)\mathrm{d}t + \mathrm{d}\zeta_t^2}_{=\mathrm{d}\xi_t^2}\Big]. \tag{2.145b}$$

It is thus sufficient to prove that the alternative conditions in Theorem 2.2 (Assumptions 2.2.1 - 2.2.3c) imply the original conditions of Theorem 2.1 (Assumptions 2.1.1 - 2.1.3). Indeed, if this is the case, then Theorem 2.2 follows directly from Theorem 2.1. This statement holds trivially for all conditions except those relating to the noise processes. It thus remains to establish that, under the noise conditions in Theorem 2.2 (Assumptions 2.2.2a - 2.2.2e, 2.2.3a - 2.2.3c), the noise condition in Theorem 2.1 (Assumption 2.1.3) holds for the noise processes $\{\xi_t^i\}_{t\geq 0}$, $i=1,2$, as defined above. That is, for all $T>0$, and $i=1,2$,

$$\lim_{s\to\infty}\sup_{t\in[s,s+T]}\left|\left|\int_s^t \gamma_v^i \mathrm{d}\xi_v^i\right|\right| = 0 \tag{2.146}$$

But this is an immediate consequence of Assumption 2.2.3a and Lemma 2.2.1. The result follows immediately. □

2.4 Extensions

In this section, we provide details of several possible extensions to Theorem 2.1. In particular, we discuss how to obtain an a.s. convergence result for an alternative version of Algorithm (2.9a) - (2.9b) which makes use of higher order gradient information, as well as the additional assumptions required in order to establish convergence of Algorithm (2.9a) - (2.9b) to the set of local and global minima of the two objective functions (in a sense to be made precise below). We also establish an asymptotic convergence rate.

2.4.1 Higher Order Updates

In Theorem 2.1, we analysed a two-timescale stochastic gradient descent algorithm designed to solve a weak formulation of the original bilevel optimisation problem in (2.1), as stated in (2.2). This, we now recall for convenience, refers to the task of obtaining (α^*, β^*) which jointly satisfy

$$\alpha^* = \underset{\alpha \in U_{\alpha^*}}{\arg\min}\, f(\alpha, \beta^*) \quad , \quad \beta^* = \underset{\beta \in U_{\beta^*}}{\arg\min}\, g(\alpha^*, \beta) \tag{2.147}$$

where $U_{\alpha^*} \subset \mathbb{R}^{d_1}$ and $U_{\beta^*} \subset \mathbb{R}^{d_2}$ are local neighbourhoods of α^* and β^*, respectively. That is, equivalently, values (α^*, β^*) such that α^* locally minimises $f(\alpha, \beta^*)$ with respect to α, and β^* which locally minimises $g(\alpha^*, \beta)$ with respect to β. To tackle this problem using gradient methods, it is natural to consider an algorithm which only utilises noisy estimates of the partial derivatives $\nabla_\alpha f(\alpha, \beta)$ and $\nabla_\beta g(\alpha, \beta)$. This is precisely the form of Algorithm (2.9a) - (2.9b).

Suppose, instead, that we would like to solve a local version of the original bilevel optimisation problem (2.1) more directly. In particular, suppose that we wish to obtain $(\alpha^*, \beta^*) = (\alpha^*, \beta^*(\alpha^*))$ which satisfy

$$\alpha^* = \underset{\alpha \in U_{\alpha^*}}{\arg\min}\, f\big(\alpha, \beta^*(\alpha)\big) \quad \text{s.t.} \quad \beta^*(\alpha) = \underset{\beta \in U_{\beta*(\alpha)}}{\arg\min}\, g(\alpha, \beta) \tag{2.148}$$

where, similarly to above, $U_{\alpha^*} \subset \mathbb{R}^{d_1}$ and $U_{\beta^*(\alpha)} \subset \mathbb{R}^{d_2}$ are local neighbourhoods of α^* and $\beta^*(\alpha)$, respectively. The crucial difference between (2.147) and (2.148) is that, in the latter, we insist that α^* minimises $f(\alpha, \beta^*(\alpha))$ with respect to α. In particular, the second argument in the upper level optimisation function now depends explicitly on α.

To tackle this problem using gradient methods, we can still use the partial derivative $\nabla_\beta g(\alpha, \beta)$ to minimise the lower-level objective function. If possible, however, we should now use the total derivative $\nabla f(\alpha, \beta^*(\alpha))$ to minimise the upper-level objective function. This will, of course, require additional assumptions on the two-objective functions (to be specified below). To make progress in this direction, first note that, via the chain rule, we have

$$\nabla f(\alpha, \beta^*(\alpha)) = \nabla_\alpha f(\alpha, \beta^*(\alpha)) + [\nabla_\alpha \beta^*(\alpha)]^T \nabla_\beta f(\alpha, \beta^*(\alpha)). \tag{2.149}$$

Moreover, owing to the first order optimality condition for $\beta^*(\alpha)$, under appropriate additional assumptions on g, it holds that (see, e.g., [200])

$$\nabla_\alpha \beta^*(\alpha) = -\nabla^2_{\alpha\beta} g(\alpha, \beta^*(\alpha)) \left[\nabla^2_{\beta\beta} g(\alpha, \beta^*(\alpha))\right]^{-1}. \tag{2.150}$$

In practice, $\beta^*(\alpha)$ is not available in closed form. Thus, one typically approximates $\nabla f(\alpha, \beta^*(\alpha))$ by replacing $\beta^*(\alpha)$ with $\beta \in \mathbb{R}^{d_2}$. This yields, instead of (2.149), the 'surrogate' gradient

$$\bar{\nabla} f(\alpha, \beta) = \nabla_\alpha f(\alpha, \beta) - \nabla^2_{\alpha\beta} g(\alpha, \beta) \left[\nabla^2_{\beta\beta} g(\alpha, \beta)\right]^{-1} \nabla_\beta f(\alpha, \beta). \tag{2.151}$$

Using this surrogate gradient, we can now obtain an alternative version of Algorithm (2.9a) - (2.9b). In particular, suppose that we continuously observe noisy gradients of $\bar{\nabla} f(\alpha,\beta)$ and $\nabla_\beta g(\alpha,\beta)$, as in (2.7a) - (2.7b). Then it is natural to consider the following continuous-time two-timescale stochastic gradient descent algorithm:

$$\mathrm{d}\alpha_t = -\gamma_t^1 \left[\bar{\nabla} f(\alpha_t,\beta_t)\mathrm{d}t + \mathrm{d}\xi_t^1\right], \tag{2.152a}$$

$$\mathrm{d}\beta_t = -\gamma_t^2 \left[\nabla_\beta g(\alpha_t,\beta_t)\mathrm{d}t + \mathrm{d}\xi_t^2\right], \tag{2.152b}$$

where $f : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}$ is a continuously differentiable function, $g : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}$ is now a *twice* continuously differentiable function, $\bar{\nabla} f : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}^{d_1}$ is defined in (2.151), and all other terms are as defined in Section 2.2.1. This represents the continuous-time version of the two-timescale stochastic gradient descent algorithm in discrete time recently introduced in [218].

We can analyse this algorithm using similar assumptions to those used to establish the convergence of Algorithm (2.9a) - (2.9b) in Theorem 2.1. Let us briefly highlight the required modifications. We will first require the following stronger version of Assumption 2.1.2.

Assumption 2.1.2.i'. *The outer function $f : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}$ has the following properties*

- *The function $\nabla_\alpha f : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}^{d_1}$ is locally Lipschitz continuous.*
- *The function $\nabla_\beta f : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}^{d_2}$ is locally Lipschitz continuous.*

Assumption 2.1.2.ii'. *The inner function $g : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}$ has the following properties.*

- *The function $\nabla_\beta g : \mathbb{R}^{d_2} \times \mathbb{R}^{d_2} \to \mathbb{R}^{d_2}$ is locally Lipschitz continuous.*
- *For all $\alpha \in \mathbb{R}^{d_1}$, the function $g(\alpha,\cdot) : \mathbb{R}^{d_1} \to \mathbb{R}$ is strongly convex.*
- *The function $\nabla^2_{\alpha\beta} g : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}^{d_1 \times d_2}$ is bounded.*

These assumptions, which also appear in [200, 218], ensure that the surrogate $\bar{\nabla} f(\alpha,\beta)$ is well-defined and locally Lipschitz continuous. In particular, it is necessary to assume strong convexity for $g(\alpha,\cdot)$ to ensure that the Hessian of this function, whose inverse appears in the definition of $\bar{\nabla} f(\alpha,\beta)$ in (2.151), is bounded away from zero. An immediate consequence of this assumption is that, for all $\alpha \in \mathbb{R}^{d_1}$, $g(\alpha,\cdot)$ has a single global minimiser. This, in turn, implies that, for all $\alpha \in \mathbb{R}^{d_1}$, the equation $\dot{\beta}_t = -\nabla_\beta g(\alpha,\beta_t)$ has a globally asymptotically stable equilibrium $\beta^*(\alpha)$, thus doing away with with the need for Assumption 2.1.5.

Finally, we will now replace Assumption 2.1.6 with the following condition.

Assumption 2.1.6'. *The set* $f(E_f, \beta_i^*(E_f))$ *contains no open sets of* $\mathbb{R}^{d_1}$ *other than the empty set (i.e., has empty interior), where*

$$E_f = \left\{\alpha \in \mathbb{R}^{d_1} : \nabla f(\alpha, \beta^*(\alpha)) = 0\right\}. \tag{2.153}$$

Interestingly, this assumption is actually slightly weaker than Assumption 2.1.6. This condition was first introduced in [36], in the context of single-timescale stochastic approximation, and later also appeared in [436] in a slightly different form: namely, that the set $f(E_f, \beta^*(E_f)) \cap f(E_f^c, \beta^*(E_f^c))$ has Lebesgue measure zero. Both versions have since also appeared in the two-timescale setting [241, 437].

Broadly speaking, this condition ensures that the function $f(\cdot, \beta^*(\cdot))$ admits a certain topological property: namely, that each closed continuous path starting and ending in E_f^c has a subpath contained in E_f^c along which $f(\cdot, \beta^*(\cdot))$ does not increase. This property prevents the noise processes from forcing the slow process to drift from one connected component of E_f to another. In turn, this ensures that the slow process converges to a connected component of E_f. This assumption is satisfied under several more easily verifiable conditions. In particular, it holds if E_f or $f(E_f)$ are countable (e.g., [36]). By the Morse-Sard Theorem [217], it also holds if the function $f(\cdot, \beta^*(\cdot))$ is d_1-times differentiable, a situation which is somewhat common in two-timescale stochastic approximation algorithms (e.g., [250]).

Our main result on the convergence of Algorithm (2.152a) - (2.152b) is contained in the following theorem.

Theorem 2.1'. *Assume that Conditions 2.1.1, 2.1.2.i' - 2.1.2.ii', 2.1.3, 2.1.4, and 2.1.6' hold. Then, almost surely,*

$$\lim_{t\to\infty} \nabla f(\alpha_t, \beta^*(\alpha_t)) = \lim_{t\to\infty} \nabla_\beta g(\alpha_t, \beta_t) = 0. \tag{2.154}$$

The second limit implies, in particular, that $\lim_{t\to\infty} ||\beta_t - \beta^*(\alpha_t)|| = 0$, *where* $\beta^*(\alpha) = \arg\min_{\beta \in \mathbb{R}^{d_2}} g(\alpha, \beta)$.

Proof. Under the stated assumptions, the proof of Theorem 2.1' is essentially identical to the proof of Theorem 2.1. Let us briefly highlight the main changes. Lemma 2.1.1 goes through unchanged, replacing $\nabla_\alpha f(\alpha, \beta)$ with $\bar{\nabla} f(\alpha, \beta)$ where required. Lemma 2.1.2, which now states that $(\alpha_t, \beta_t) \to \{(\alpha, \beta^*(\alpha)) : \alpha \in \mathbb{R}^{d_1}\}$, is now much more straightforward, since the only internally chain transitive set for (the analogue of) (2.35a) - (2.35b) is now the globally asymptotically stable equilibrium point $\beta^*(\alpha)$.

Lemma 2.1.3 is essentially unchanged, again replacing $\nabla_\alpha f(\alpha, \beta)$ with $\bar{\nabla} f(\alpha, \beta)$, and also replacing $\beta_i^*(\alpha)$ with $\beta^*(\alpha)$. Finally, Lemma 2.1.4 is proved along the same lines as the original proof of Lemma 2.1.2. In particular, this proof begins by noting that, by (the modified version of) Lemma 2.1.3, $\alpha_t^{\gamma_1}$ is an asymptotic pseudo-trajectory for the ODE

$$\dot{\underline{\alpha}}_t = -\bar{\nabla} f(\underline{\alpha}_t, \beta^*(\underline{\alpha}_t)), \tag{2.155}$$

or, using the fact that $\nabla f(\alpha, \beta^*(\alpha))$ and $\bar{\nabla} f(\alpha, \beta)$ coincide when $\beta = \beta^*(\alpha)$, the ODE

$$\dot{\underline{\alpha}}_t = -\nabla f(\underline{\alpha}_t, \beta^*(\underline{\alpha}_t)). \tag{2.156}$$

By Assumption 2.1.4, this trajectory is also pre-compact. Thus, by Theorem 5.7 in Benaim [36], $\alpha_t^{\gamma_1}$ converges to an internally chain transitive set for (2.156). We next observe that $f : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}$ is a strict Lyapunov function for (2.156). Indeed, this follows immediately from

$$\begin{aligned} \dot{f}(\underline{\alpha}_t, \beta^*(\underline{\alpha}_t)) &= \dot{\underline{\alpha}}_t \nabla_\alpha f(\underline{\alpha}_t, \beta^*(\underline{\alpha}_t)) + \dot{\beta}^*(\underline{\alpha}_t) \nabla_\beta f(\underline{\alpha}_t, \beta^*(\underline{\alpha}_t)) \\ &= \dot{\underline{\alpha}}_t \left[\nabla_\alpha f(\underline{\alpha}_t, \beta^*(\underline{\alpha}_t)) + \left[\nabla_\alpha \beta^*(\underline{\alpha}_t)\right]^T \nabla_\beta f(\underline{\alpha}_t, \beta^*(\underline{\alpha}_t)) \right] \\ &= -||\nabla f(\underline{\alpha}_t, \beta^*(\underline{\alpha}_t))||^2 \leq 0. \end{aligned} \tag{2.157}$$

In addition, by Assumption 2.1.6', the set of critical values of f, namely E_f, has Lebesgue measure zero. We can thus apply Proposition 6.4 in Benaïm [36] to conclude that every internally chain transitive set for (2.156) is contained in E_f. Lemma 2.1.4 now follows straightforwardly.

Finally, combining the results of the modified versions of Lemma 2.1.2 and Lemma 2.1.4, one obtains the result of Theorem 2.1'. □

2.4.2 Convergence to Local or Global Minima

In Theorem 2.1, we established the convergence of Algorithm (2.9a) - (2.9b) to the stationary points of the two objective functions. In this section, we outline the additional assumptions required in order to establish a.s. convergence to the set of local or global minima of these functions, in the sense of (2.147). That is, (α^*, β^*) such that jointly α^* (locally) minimises $f(\alpha, \beta^*)$ with respect to α, and β^* which (locally) minimises $g(\alpha^*, \beta)$ with respect to β.

2.4.2.1 Local Minima

In order to guarantee convergence to the set of local minima, we will require the following conditions in addition to Assumptions 2.1.1 - 2.1.6.

Assumption 2.1.3 (L). *The quadratic variations of the noise processes* $\{\xi_t^i\}_{t\geq 0}$, $i = 1, 2$, *are uniformly positive definite.*

Assumption 2.1.5 (L). *For all* $\alpha \in \mathbb{R}^{d_1}$, *the function* $g(\alpha, \cdot) : \mathbb{R}^{d_2} \to \mathbb{R}$ *is twice continuously differentiable. Moreover, this function is strict saddle.*[18]

Assumption 2.1.6 (L). *For all* $\beta \in \mathbb{R}^{d_2}$, *the function* $f(\cdot, \beta) : \mathbb{R}^{d_1} \to \mathbb{R}$ *is twice continuously differentiable. Moreover, this function is strict saddle.*[19]

The analogue of these assumptions (for a single objective function) appear in both classical [67, 377] and more recent [189, 346] results on the 'avoidance of saddles' in the discrete-time stochastic approximation literature. See also [479] for a related result in continuous time. Broadly speaking, the first of these assumptions is required in order to ensure that the additive noise processes are 'sufficiently exciting', that is, that they have sufficiently large components in all directions. Meanwhile, the final two assumptions rule out degenerate cases in which the Hessian does not contain sufficient information to characterise the nature of a critical point.

Using these assumptions, one can establish (in single-timescale, discrete-time stochastic gradient descent) that unstable equilibria (i.e., saddle points) are avoided with probability one via the central manifold theorem (e.g., [416]). We leave the rigorous extension of these results to the continuous-time, two-timescale framework to future work.

2.4.2.2 Global Minima

It is somewhat more straightforward to establish convergence to the global minima of the two objective functions. In this case, we can simply replace Assumption 2.1.5 and Assumption 2.1.6 with the following conditions.

Assumption 2.1.5 (G). *For all* $\alpha \in \mathbb{R}^{d_1}$, *the ordinary differential equation*

$$\frac{\mathrm{d}\beta_t}{\mathrm{d}t} = -\nabla_\beta g(\alpha, \beta_t) \tag{2.158}$$

has a globally asymptotically stable equilibrium $\beta^*(\alpha)$, *where* $\beta^* : \mathbb{R}^{d_1} \to \mathbb{R}^{d_2}$ *is a Lipschitz-continuous map.*

Assumption 2.1.6 (G). *The ordinary differential equation*

$$\frac{\mathrm{d}\alpha_t}{\mathrm{d}t} = -\nabla_\alpha f(\alpha_t, \beta^*(\alpha_t)) \tag{2.159}$$

[18] A twice differentiable function $h : \mathbb{R}^d \to \mathbb{R}^d$ is said to be strict saddle if all of its local minima satisfy $\nabla_x^2 h(x) \succ 0$, and all of its other stationary points satisfy $\lambda_{\min}(\nabla_x^2 h(x)) < 0$ (i.e., the minimum eigenvalue of the Hessian evaluated at the critical points is negative).

[19] Strictly speaking one would only require that this holds for $\beta \in \mathbb{R}^{d_2}$ such that $\beta = \beta_i^*(\alpha)$ for some $i \geq 1$, $\alpha \in \mathbb{R}^{d_1}$.

has a globally asymptotically stable equilibrium α^*.

These are rather classical assumptions used to establish a.s. convergence of discrete-time two-timescale stochastic approximation algorithms to a unique equilibrium point $(\alpha^*, \beta^*(\alpha^*))$ (e.g., [61, 62, 252, 437]). In the context of two-timescale stochastic gradient descent, it is common to instead use the assumptions that the functions $g(\alpha, \cdot) : \mathbb{R}^{d_2} \to \mathbb{R}^{d_2}$ and $f(\alpha, \beta^*(\alpha)) : \mathbb{R}^{d_1} \to \mathbb{R}^{d_1}$ are (strongly) convex, with some additional assumptions on the mixed partial derivatives of g to ensure that $\beta^*(\alpha)$ is Lipschitz continuous (e.g., [160, 200, 218]).

Under these assumptions, the proof of Theorem 2.1 only requires minor modifications in order to establish convergence to the global minima, that is, $(\alpha_t, \beta_t) \to (\alpha^*, \beta^*(\alpha^*))$ as $t \to \infty$. In particular, Assumption 2.1.5 (G) implies that Lemma 2.1.2 can now conclude $(\alpha_t, \beta_t) \to \{(\alpha, \beta^*(\alpha)) : \alpha \in \mathbb{R}^{d_1}\}$. Meanwhile, Assumption 2.1.6 (G) implies that Lemma 2.1.4 yields $\alpha_t \to \alpha^*$. The remainder of the proof is unchanged.

2.4.3 Convergence Rates

In this section, we make some progress towards obtaining the convergence rate of two-timescale stochastic gradient descent in continuous time. Once more, we restrict our attention to the algorithm analysed in Theorem 2.1 (i.e., the additive noise case). In order to do so, we will require the following assumptions. These will be required either in place of, or in addition to, our previous assumptions.

Assumption 2.1.1'. *The learning rates $\{\gamma_i(t)\}_{t\geq 0}$, $i = 1, 2$, are of the form*

$$\gamma_t^1 = \gamma_0^1(\delta_1 + t)^{-\eta_1} \quad , \quad \gamma_t^2 = \gamma_0^2(\delta_2 + t)^{-\eta_2} \tag{2.160}$$

where $\gamma_0^1, \gamma_0^2 > 0$, $\delta_1, \delta_2 > 0$ are positive constants, and $\eta_1, \eta_2 \in (\frac{1}{2}, 1]$ are positive constants such that $\eta_2 < \eta_1$. In the case that $\eta_1 = 1$, we have $\gamma_0^1 \Lambda_H > \frac{1}{2}$, where Λ_H is the constant defined in Assumption 2.1.7'b.

This is a standard choice of learning rates for two-timescale stochastic approximation algorithms. In particular, these learning rates satisfy the conditions required to establish a.s. convergence of the algorithm iterates (see Assumptions 2.1.1 and 2.2.1). While one can use a slightly more general form for the learning rates (e.g., [422, Proposition 2.13] in the single-timescale case), this choice significantly simplifies our subsequent analysis, and corresponds to the choice most commonly used in practice.

Assumption 2.1.3'a. *The additive noise processes $\{\xi_t^i\}_{t\geq 0}$, $i = 1, 2$, are of the form*

$$\mathrm{d}\xi_t^i = \xi_i^{(1)}(\alpha_t, \beta_t)\mathrm{d}t + \xi_i^{(2)}(\alpha_t, \beta_t)\mathrm{d}w_t^i \tag{2.161}$$

where $\xi_i^1 : \mathbb{R}^{d_1}\times\mathbb{R}^{d_2} \to \mathbb{R}^{d_i}$, $\xi_i^{(2)} : \mathbb{R}^{d_1}\times\mathbb{R}^{d_2} \to \mathbb{R}^{d_i\times d_i}$, *and* $\{w_t^i\}_{t\geq 0}$ *are* $\mathbb{R}^{d_i}$ *valued standard Brownian motions, whose components may coincide.*

Assumption 2.1.3'b. *The functions* $\xi_i^{(1)} : \mathbb{R}^{d_1}\times\mathbb{R}^{d_2} \to \mathbb{R}^{d_i}$ *satisfy*

$$\xi_i^{(1)}(\alpha_t,\beta_t) = o\left((\gamma_t^1)^{\frac{1}{2}}\right). \tag{2.162}$$

Assumption 2.1.3'c. *The functions* $\xi_i^{(2)} : \mathbb{R}^{d_1}\times\mathbb{R}^{d_2} \to \mathbb{R}^{d_i}$ *have the following property. Let* $\Gamma_t^{11}\in\mathbb{R}^{d_1\times d_1}$, $\Gamma_t^{22}\in\mathbb{R}^{d_2\times d_2}$, $\Gamma_t^{12}\in\mathbb{R}^{d_1\times d_2}$ *be the matrices such that*

$$\left[\xi_1^{(2)}\mathrm{d}w_1\right]_t = \Gamma_t^{11}\mathrm{d}t \tag{2.163a}$$

$$\left[\xi_2^{(2)}\mathrm{d}w_2\right]_t = \Gamma_t^{22}\mathrm{d}t \tag{2.163b}$$

$$\left[\xi_1^{(2)}\mathrm{d}w_1,\xi_2^{(2)}\mathrm{d}w_2\right]_t = \Gamma_t^{12}\mathrm{d}t \tag{2.163c}$$

where $[\cdot]_t$ *and* $[\cdot,\cdot]_t$ *denote the quadratic and cross variation, respectively.*[20] *Then there exist* $\Gamma_{11}\in\mathbb{R}^{d_1\times d_1}$, $\Gamma_{22}\in\mathbb{R}^{d_2\times d_2}$, $\Gamma_{12}\in\mathbb{R}^{d_1\times d_2}$ *such that*

$$\lim_{t\to\infty}\Gamma_t^{11} = \Gamma_{11} \tag{2.165a}$$

$$\lim_{t\to\infty}\Gamma_t^{22} = \Gamma_{22} \tag{2.165b}$$

$$\lim_{t\to\infty}\Gamma_t^{12} = \Gamma_{12} \tag{2.165c}$$

These assumptions relate to the properties of the additive noise processes $\{\xi_t^i\}_{t\geq 0}$, $i=1,2$. They represent a continuous-time analogue of the conditions used to establish the weak convergence rate of two-timescale stochastic gradient in discrete time in [354]. They are also sufficient for the noise conditions used to establish a.s. convergence in Theorem 2.1 and Theorem 2.2. While more general assumptions are possible, these assumptions are sufficiently broad to cover most cases of practical interest.

Assumption 2.1.7'a. *There exists a neighbourhood* $\mathcal{U}_{\alpha^*,\beta^*}$ *of* (α^*,β^*) *such that, for all*

[20] In particular, using the notation $[A]_{ij}$ to denote the ij^{th} element of the matrix A, and writing $[z^i]_k$ to denote the k^{th} element of the d_i-dimensional Brownian motion z^i, $i=1,2$, we have

$$[\Gamma_t^{11}]_{ij} = \sum_{k=1}^{d_1}\left[\xi_1^{(2)}(\alpha_t,\beta_t)\right]_{ik}\left[\xi_1^{(2)}(\alpha_t,\beta_t)\right]_{jk} \tag{2.164a}$$

$$[\Gamma_t^{22}]_{ij} = \sum_{k=1}^{d_1}\left[\xi_2^{(2)}(\alpha_t,\beta_t)\right]_{ik}\left[\xi_2^{(2)}(\alpha_t,\beta_t)\right]_{jk} \tag{2.164b}$$

$$[\Gamma_t^{12}]_{ij} = \sum_{k=1}^{d_1}\left[\xi_1^{(2)}(\alpha_t,\beta_t)\right]_{ik}\left[\xi_2^{(2)}(\alpha_t,\beta_t)\right]_{jk}\mathbb{1}_{\{[z^1]_k=[z^2]_k\}}. \tag{2.164c}$$

We remark that, in the case that z^1 and z^2 are independent, so that $[z^1]_k\neq[z^2]_k$ for any k, it follows from (2.164c) that $\Gamma_t^{12}\equiv 0$.

$(\alpha,\beta)\in\mathcal{U}_{\alpha^*,\beta^*}$,

$$-\begin{pmatrix}\nabla_\alpha f(\alpha,\beta)\\ \nabla_\beta g(\alpha,\beta)\end{pmatrix}=\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}\begin{pmatrix}\alpha-\alpha^*\\ \beta-\beta^*\end{pmatrix}+O\left(||\begin{matrix}\alpha-\alpha^*\\ \beta-\beta^*\end{matrix}||^2\right) \tag{2.166}$$

where $\alpha^*\in\mathbb{R}^{d_1}$, $\beta^*\in\mathbb{R}^{d_2}$ are the values such that $\lim_{t\to\infty}\alpha_t=\alpha^*$ and $\lim_{t\to\infty}\beta_t=\beta^*$.

Assumption 2.1.7'b. *The matrices A_{22} and $H=A_{11}-A_{12}A_{22}^{-1}A_{21}$ are stable (or Hurwitz). That is,*

$$\Lambda_H=-\lambda_{\max}(H)>0 \quad , \quad \Lambda_{A_{22}}=-\lambda_{\max}(A_{22})>0 \tag{2.167}$$

where the notation $\lambda_{\max}(M)$ denotes the maximum real part of the eigenvalues of the matrix M.

These assumptions relate to the nature of the equilibrium point (α^*,β^*), and are essential in establishing the (asymptotic) convergence rate of two-timescale stochastic approximation algorithms. Assumption 2.1.7'b was first introduced by Konda and Tsitsiklis to establish the asymptotic convergence rate of linear two-timescale stochastic approximation [251] and has since also be used to obtain non-asymptotic convergence rates for this algorithm (e.g., [134, 158, 159, 206, 236]).[21] Meanwhile, Assumption 2.1.7'a was used by Mokkadem and Pelletier to obtain asymptotic convergence rates in the non-linear case [354].

We are now ready to state our main result on the weak convergence rate of Algorithm (2.9a) - (2.9b).

Theorem 2.3. *Assume that Assumptions 2.1.1', 2.1.2, 2.1.3'a - 2.1.3'c, 2.1.4, 2.1.5, 2.1.6, and 2.1.7'a - 2.1.7'b are satisfied. Then*

$$\begin{pmatrix}(\gamma_t^1)^{-\frac{1}{2}}(\alpha_t-\alpha^*)\\ (\gamma_t^2)^{-\frac{1}{2}}(\beta_t-\beta^*)\end{pmatrix}\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\begin{pmatrix}\Sigma_\alpha & 0\\ 0 & \Sigma_\beta\end{pmatrix}\right). \tag{2.168}$$

where, defining $\Gamma_\alpha=\Gamma_{11}+A_{12}A_{22}^{-1}\Gamma_{22}[A_{22}^{-1}]^TA_{12}^T-\Gamma_{12}[A_{22}^{-1}]^TA_{12}^T-A_{12}A_{22}^{-1}\Gamma_{21}$,

$$\Sigma_\alpha=\int_0^\infty\exp\left[\left(H+\frac{\mathbb{1}_{\eta_1=1}}{2\gamma_0^0}I\right)t\right]\Gamma_\alpha\exp\left[\left(H+\frac{\mathbb{1}_{\eta_1=1}}{2\gamma_0^0}I\right)t\right]\mathrm{d}t \tag{2.169a}$$

$$\Sigma_\beta=\int_0^\infty\exp\left[A_{22}t\right]\Gamma_{22}\exp\left[A_{22}t\right]\mathrm{d}t. \tag{2.169b}$$

[21] We should remark that several of these papers actually use a slightly weaker assumption; in particular, in [134, 158, 159], it is assumed that these matrices are negative definite.

Proof. In what follows, we sketch the proof of this result. Our proof relies on several key lemmas, whose own proofs are deferred to Appendix 2.A.

Without loss of generality, we will assume throughout this proof that $\alpha^* = \beta^* = 0$. We begin with the observation that Assumptions 2.1.1', 2.1.2, 2.1.3'a - 2.1.3'c, 2.1.4, 2.1.5, 2.1.6 are sufficient for Assumptions 2.1.1 - 2.1.6. Thus, we can apply Theorem 2.1, which implies in particular that $\alpha_t \to \alpha^*$ and $\beta_t \to \beta^*$ a.s. as $t \to \infty$. This observation, together with Assumption 2.1.7'a, means that for sufficiently large t we can write

$$\mathrm{d}\alpha_t = \gamma_t^1 \left[A_{11}\alpha_t \mathrm{d}t + A_{12}\beta_t \mathrm{d}t + \varepsilon_t^1 \mathrm{d}t + \mathrm{d}\xi_t^1\right], \tag{2.170a}$$

$$\mathrm{d}\beta_t = \gamma_t^2 \left[A_{21}\alpha_t \mathrm{d}t + A_{22}\beta_t \mathrm{d}t + \varepsilon_t^2 \mathrm{d}t + \mathrm{d}\xi_t^2\right]. \tag{2.170b}$$

where $\varepsilon_t^i = O(||\alpha_t||^2 + ||\beta_t||^2)$, $i = 1, 2$. Rearranging (2.170b), we obtain

$$\beta_t \mathrm{d}t = A_{22}^{-1} \left[(\gamma_t^2)^{-1}\mathrm{d}\beta_t - A_{21}\alpha_t \mathrm{d}t - \varepsilon_t^2 \mathrm{d}t - \mathrm{d}\xi_t^2\right]. \tag{2.171}$$

Substituting this expression into (2.170a), it follows that

$$\begin{aligned}\mathrm{d}\alpha_t &= \gamma_t^1 \left[A_{11}\alpha_t \mathrm{d}t + A_{12}A_{22}^{-1}\left[(\gamma_t^2)^{-1}\mathrm{d}\beta_t - A_{21}\alpha_t \mathrm{d}t - \varepsilon_t^2 \mathrm{d}t - \mathrm{d}\xi_t^2\right] + \varepsilon_t^1 \mathrm{d}t + \mathrm{d}\xi_t^1\right] \\ &= \gamma_t^1 \underbrace{\left(A_{11} - A_{12}A_{22}^{-1}A_{21}\right)}_{H} \alpha_t \mathrm{d}t + \frac{\gamma_t^1}{\gamma_t^2} A_{12}A_{22}^{-1}\mathrm{d}\beta_t \\ &\quad + \gamma_t^1 \left[\mathrm{d}\xi_t^1 - A_{12}A_{22}^{-1}\mathrm{d}\xi_t^2\right] + \gamma_t^1 \left[\varepsilon_t^1 - A_{12}A_{22}^{-1}\varepsilon_t^2\right]\mathrm{d}t\end{aligned} \tag{2.172}$$

Let us now define the matrices

$$\Phi_{s,t}^{(1)} = \exp\left[H \int_s^t \gamma_u^1 \mathrm{d}u\right] \quad , \quad \Phi_{s,t}^{(2)} = \exp\left[A_{22} \int_s^t \gamma_u^2 \mathrm{d}u\right] \tag{2.173}$$

and the real numbers

$$\Psi_{s,t}^{(1)} = \exp\left[-\mu_1 \int_s^t \gamma_u^1 \mathrm{d}u\right] \quad , \quad \Psi_{s,t}^{(2)} = \exp\left[-\mu_2 \int_s^t \gamma_u^2 \mathrm{d}u\right] \tag{2.174}$$

We remark that, by Assumption 2.1.7'b, there exists $K > 0$ such that, for any $\mu_1 \in (0, \Lambda_H)$ and $\mu_2 \in (0, \Lambda_{A_{22}})$, it holds that (see, e.g., [106, page 121])

$$||\Phi_{s,t}^{(1)}|| \leq K\Psi_{s,t}^{(1)} \tag{2.175a}$$

$$||\Phi_{s,t}^{(2)}|| \leq K\Psi_{s,t}^{(2)} \tag{2.175b}$$

We can now write the solution of (2.172) as

$$\alpha_t = \underbrace{\alpha_0\Phi^{(1)}_{0,t}}_{I^\alpha_t} + \underbrace{\int_0^t \Phi^{(1)}_{s,t}\gamma^1_s\left[\xi^{(2)}_1(\alpha_s,\beta_s)\mathrm{d}w^1_s - A_{12}A_{22}^{-1}\xi^{(2)}_2(\alpha_s,\beta_s)\mathrm{d}w^2_s\right]}_{L^\alpha_t} + \underbrace{\int_0^t \Phi^{(1)}_{s,t}\frac{\gamma^1_s}{\gamma^2_s}A_{12}A_{22}^{-1}\mathrm{d}\beta_s}_{R^\alpha_t}$$
$$\underbrace{+\int_0^t \Phi^{(1)}_{s,t}\gamma^1_s\left[\xi^{(1)}_1(\alpha_s,\beta_s) - A_{12}A_{22}^{-1}\xi^{(1)}_2(\alpha_s,\beta_s)\right]\mathrm{d}s + \int_0^t \Phi^{(1)}_{s,t}\gamma^1_s\left[\varepsilon^1_s - A_{12}A_{22}^{-1}\varepsilon^2_s\right]\mathrm{d}s}_{\Delta^\alpha_t} \tag{2.176}$$

$$:= I^\alpha_t + L^\alpha_t + R^\alpha_t + \Delta^\alpha_t \tag{2.177}$$

and, thus, the solution of (2.170b) as

$$\beta_t = \underbrace{\beta_0\Phi^{(2)}_{0,t}}_{I^\beta_t} + \underbrace{\int_0^t \Phi^{(2)}_{s,t}\gamma^2_s\xi^{(2)}_2(\alpha_s,\beta_s)\mathrm{d}w^2_s}_{L^\beta_t} \quad \underbrace{+\int_0^t \Phi^{(2)}_{s,t}\gamma^2_s A_{21}\left[L^\alpha_s + R^\alpha_s\right]\mathrm{d}s}_{R^\beta_t} \tag{2.178}$$

$$\underbrace{+\int_0^t \Phi^{(2)}_{s,t}\gamma^2_s A_{21}\Delta^\alpha_s\mathrm{d}s + \int_0^t \Phi^{(2)}_{s,t}\gamma^2_s\xi^{(1)}_2(\alpha_s,\beta_s)\mathrm{d}s + \int_0^t \Phi^{(2)}_{s,t}\gamma^2_s\varepsilon^2_s\mathrm{d}s}_{\Delta^\beta_t} \tag{2.179}$$

$$:= I^\beta_t + L^\beta_t + R^\beta_t + \Delta^\beta_t. \tag{2.180}$$

By Lemmas 2.1, 2.3, 2.4 and 2.5 (see Appendix 2.A), we have that

$$||I^\alpha_t|| = o\left(\sqrt{\gamma^1_t}\right) \tag{2.181}$$

$$||L^\alpha_t|| = O\left(\sqrt{\gamma^1_t \log \int_0^t \gamma^1_s\mathrm{d}s}\right) \tag{2.182}$$

$$||R^\alpha_t|| = o\left(\sqrt{\gamma^1_t}\right) \tag{2.183}$$

$$||\Delta^\alpha_t|| = o\left(\sqrt{\gamma^1_t}\right) \tag{2.184}$$

and, in addition, that

$$||I^\beta_t|| = o\left(\sqrt{\gamma^2_t}\right) \tag{2.185}$$

$$||L_t^{\beta}|| = O\left(\sqrt{\gamma_t^2 \log \int_0^t \gamma_s^2 \mathrm{d}s}\right) \tag{2.186}$$

$$||R_t^{\beta}|| = O\left(\sqrt{\gamma_t^1 \log \int_0^t \gamma_s^1 \mathrm{d}s}\right) \tag{2.187}$$

$$||\Delta_t^{\beta}|| = o\left(\sqrt{\gamma_t^1}\right) \tag{2.188}$$

Thus, the asymptotic convergence rates of $\{\alpha_t\}_{t\geq 0}$ and $\{\beta_t\}_{t\geq 0}$ are determined by $\{L_t^{\alpha}\}_{t\geq 0}$ and $\{L_t^{\beta}\}_{t\geq 0}$. It remains only to note that, by Lemma 2.2 (see Appendix 2.A), we have

$$\begin{pmatrix} \left(\gamma_t^1\right)^{-\frac{1}{2}} L_t^{\alpha} \\ \left(\gamma_t^2\right)^{-\frac{1}{2}} L_t^{\beta} \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \begin{pmatrix} \Sigma_{\alpha} & 0 \\ 0 & \Sigma_{\beta} \end{pmatrix}\right), \tag{2.189}$$

which completes the proof. □

Our proof of Theorem 2.3 represents a careful adaptation of the approach in [354] to the continuous-time setting. Indeed, in discrete time, convergence rates of two-timescale stochastic approximation algorithms are the subject of several classical papers (e.g., [251, 354]), and have also received renewed attention in recent years (e.g., [133, 160, 477]). In the presence of Markovian dynamics (see Section 2.2.2), the analysis required to establish asymptotic normality for the two-timescale stochastic gradient descent algorithm is rather more involved (see [160, 236] for some relevant results in discrete time). We leave this extension for future work, noting that the recent results in [422] will likely prove useful in this direction.

2.5 Conclusions

In this chapter, we have analysed the asymptotic properties of two-timescale stochastic gradient in continuous time, under general noise and stability conditions. Our analysis covers algorithms with both additive, state-dependent noise, and also those with non-additive, state-dependent noise. In the second case, our results were obtained under the rather weak assumption that the non-additive noise process can be represented by an ergodic diffusion process controlled by the algorithm states.

We conclude this chapter with some remarks regarding possible directions for future work. There are a number of additional extensions to the results presented in this chapter which may be of theoretical or practical interest. These include relaxing the assumption that the algorithm iterates are continuous, and thus considering a more general algorithm in which $\{\xi_t^i\}_{t\geq 0}$, $i = 1, 2$, are arbitrary semi-martingales. This can be regarded, in some sense, as as a two-timescale extension of the Robbins-Monro type semimartingale SDEs studied in [283, 284, 285, 286, 287, 288, 343, 456]. Obtaining asymptotic results under this somewhat more general framework is of considerable interest, as such results would apply to two-timescale stochastic gradient descent schemes in both discrete-time and continuous-time. Another natural extension is to transfer other existing results from discrete-time to continuous-time. These include establishing asymptotic normality in the Markovian case, a finite-time analysis (e.g., [133, 134, 158, 159, 160, 161, 206, 218, 236]), and obtaining concentration bounds [60].

Appendices

2.A Proof of Lemmas for Theorem 2.3

.

In this Appendix, we state and prove the Lemmas required for the proof of Theorem 2.3.

2.A.0.1 Main Lemmas

Lemma 2.1. *The processes* $\{I_t^\alpha\}_{t\geq 0}$, $\{I_t^\beta\}_{t\geq 0}$ *satisfy*

$$||I_t^\alpha|| = o(\left(\gamma_t^1\right)^{\frac{1}{2}}) \tag{2.190}$$

$$||I_t^\beta|| = o(\left(\gamma_t^2\right)^{\frac{1}{2}}). \tag{2.191}$$

Proof. This result is an immediate consequence of Lemma 2.10. In particular, we have that

$$||I_t^\alpha|| \leq ||\alpha_0||||\Phi_{0,t}^{(1)}|| \leq K||\Psi_{0,t}^{(1)}|| = o((\gamma_t^1)^{\frac{1}{2}}) \tag{2.192}$$

$$||I_t^\beta|| \leq ||\beta_0||||\Phi_{0,t}^{(2)}|| \leq K||\Psi_{0,t}^{(2)}|| = o((\gamma_t^2)^{\frac{1}{2}}) \tag{2.193}$$

where in the second inequality we have used the stability of the matrices H and A_{22}, c.f. (2.175a) - (2.175b), and in the final inequality we have used Lemma 2.10. □

Lemma 2.2. *The processes* $\{L_t^\alpha\}_{t\geq 0}$, $\{L_t^\beta\}_{t\geq 0}$ *satisfy*

$$\begin{pmatrix} (\gamma_t^1)^{-\frac{1}{2}}L_t^\alpha \\ (\gamma_t^2)^{-\frac{1}{2}}L_t^\beta \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \begin{pmatrix} \Sigma_\alpha & 0 \\ 0 & \Sigma_\beta \end{pmatrix}\right). \tag{2.194}$$

Proof. Let us define $\Sigma = \{\Sigma_t\}_{t\geq 0}$ according to

$$\Sigma_t = \begin{pmatrix} (\gamma_t^1)^{-\frac{1}{2}}L_t^\alpha \\ (\gamma_t^2)^{-\frac{1}{2}}L_t^\beta \end{pmatrix} \tag{2.195}$$

This process is a (continuous) martingale, with quadratic variation given by

$$[\Sigma]_t = \begin{pmatrix} \Sigma_t^{1,1} & \Sigma_t^{1,2} \\ (\Sigma_t^{1,2})^T & \Sigma_t^{2,2} \end{pmatrix} \tag{2.196}$$

where, due to Assumptions 2.1.3'a - 2.1.3'c,

$$\Sigma_t^{1,1} = (\gamma_t^1)^{-1} \int_0^t (\gamma_s^1)^2 \Phi_{s,t}^{(1)} \left[\Gamma_s^{11} + A_{12}A_{22}^{-1}\Gamma_s^{22}[A_{22}^{-1}]^T A_{12}^T \right.$$
$$\left. -\Gamma_s^{12}[A_{22}^{-1}]^T A_{12}^T - A_{12}A_{22}^{-1}(\Gamma_s^{12})^T\right] [\Phi_{s,t}^{(1)}]^T \mathrm{d}s \tag{2.197}$$

$$\Sigma_t^{1,2} = \sqrt{(\gamma_t^1)^{-1}(\gamma_t^2)^{-1}} \int_0^t \gamma_s^1\gamma_s^2 \Phi_{s,t}^{(1)} \left[\Gamma_s^{22} - A_{12}A_{22}^{-1}\Gamma_s^{22}\right] [\Phi_{s,t}^{(2)}]^T \mathrm{d}s \tag{2.198}$$

$$\Sigma_t^{2,2} = (\gamma_t^2)^{-1} \int_0^t (\gamma_s^2)^2 \Phi_{s,t}^{(2)} \Gamma_s^{22} [\Phi_{s,t}^{(2)}]^T \mathrm{d}s \tag{2.199}$$

We will begin by showing that

$$\Sigma_t^{2,2} \xrightarrow{\mathbb{P}} \Sigma_\beta = \int_0^\infty \exp\left[A_{22}t\right] \Gamma_{22} \exp\left[A_{22}t\right] \mathrm{d}t \tag{2.200}$$

We will do so by using the decomposition

$$\mathbb{E}\left[||\Sigma_t^{2,2} - \Sigma_\beta||\right] \leq \mathbb{E}\left[||\Sigma_t^{2,2} - \bar{\Sigma}_t^{2,2}||\right] + \mathbb{E}\left[||\bar{\Sigma}_t^{2,2} - \Sigma_\beta||\right], \tag{2.201}$$

where we have introduced

$$\bar{\Sigma}_t^{2,2} = (\gamma_t^2)^{-1} \int_0^t (\gamma_s^2)^2 \Phi_{s,t}^{(2)} \Gamma_{22} [\Phi_{s,t}^{(2)}]^T \mathrm{d}s \tag{2.202}$$

In particular, if we can prove that both terms on the RHS of (2.201) converge to zero as $t \to \infty$, then it follows that $\mathbb{E}\left[||\Sigma_t^{2,2} - \Sigma_\beta||\right] \to 0$. In turn, this implies that $\Sigma_t^{2,2} \xrightarrow{\mathbb{P}} \Sigma_\beta$.

We will begin with the first term in (2.201). Observe that

$$||\Sigma_t^{2,2} - \bar{\Sigma}_t^{2,2}|| \leq (\gamma_t^2)^{-1} \int_0^t (\gamma_s^2)^2 ||\Phi_{s,t}^{(2)}||^2 ||\Gamma_s^{22} - \Gamma_{22}|| \mathrm{d}s \tag{2.203}$$

$$\leq (\gamma_t^2)^{-1} \int_0^t (\gamma_s^2)^2 [\Psi_{s,t}^{(2)}]^2 ||\Gamma_s^{22} - \Gamma_{22}|| \mathrm{d}s \tag{2.204}$$

$$\leq (\gamma_t^2)^{-1} \int_0^t \gamma_s^2 \Psi_{s,t}^{(2)} \left[\gamma_s^2 ||\Gamma_s^{22} - \Gamma_{22}||\right] \mathrm{d}s \tag{2.205}$$

$$\leq K(\gamma_t^2)^{-1}\gamma_t^2 ||\Gamma_s^{22} - \Gamma_{22}|| = K||\Gamma_s^{22} - \Gamma_{22}|| \tag{2.206}$$

where in the second line we have used the stability of A_{22} to pass from $\Phi^{(2)}_{s,t}$ to $\Psi^{(2)}_{s,t}$ (e.g., [106, pg 121], and in the final line we have applied Lemma 2.9 with $w_s = \gamma_s^2||\Gamma_s^{22} - \Gamma_{22}||$. Taking expectations, and using Assumption 2.1.3'c, we obtain

$$\mathbb{E}\left[||\Sigma_t^{2,2} - \bar{\Sigma}_t^{2,2}||\right] \leq K\mathbb{E}\left[||\Gamma_s^{22} - \Gamma_{22}||\right] \to 0. \tag{2.207}$$

We now turn our attention to the second term in (2.201). We are required to show that

$$\mathbb{E}\left[\left|\left|(\gamma_t^2)^{-1}\int_0^t(\gamma_s^2)^2\Phi^{(2)}_{s,t}\Gamma_{22}[\Phi^{(2)}_{s,t}]^T\mathrm{d}s - \int_0^\infty \exp[A_{22}s]\Gamma_{22}\exp[A_{22}s]\mathrm{d}s\right|\right|\right] \to 0. \tag{2.208}$$

To do so, we begin by writing the inside of this expectation as

$$(\gamma_t^2)^{-1}\Phi^{(2)}_{0,t}\left[\int_0^t(\gamma_s^2)^2\Phi^{(2)}_{s,0}\Gamma_{22}[\Phi^{(2)}_{s,0}]^T\mathrm{d}s - \gamma_t^2\Phi^{(2)}_{t,0}\int_0^\infty \exp[A_{22}s]\Gamma_{22}\exp[A_{22}s]\mathrm{d}s[\Phi^{(2)}_{t,0}]^T\right][\Phi^{(2)}_{0,t}]^T \tag{2.209}$$

$$:= (\gamma_t^2)^{-1}\Phi^{(2)}_{0,t}\left[\int_0^t(\gamma_s^2)^2\Phi^{(2)}_{s,0}\Gamma_{22}[\Phi^{(2)}_{s,0}]^T\mathrm{d}s - P_t\right][\Phi^{(2)}_{0,t}]^T := B_t \tag{2.210}$$

where we have defined

$$P_t := \gamma_t^2\Phi^{(2)}_{t,0}\underbrace{\int_0^\infty \exp[A_{22}s]\Gamma_{22}\exp[A_{22}s]\mathrm{d}s}_{\Sigma_\beta}[\Phi^{(2)}_{t,0}]^T := \gamma_t^2\Phi^{(2)}_{t,0}\Sigma_\beta[\Phi^{(2)}_{t,0}]^T \tag{2.211}$$

Next, we observe that

$$\frac{\mathrm{d}}{\mathrm{d}t}P_t = \dot{\gamma}_t^2\Phi^{(2)}_{t,0}\Sigma_\beta[\Phi^{(2)}_{t,0}]^T - (\gamma_t^2)^2\Phi^{(2)}_{t,0}A_{22}\Sigma_\beta[\Phi^{(2)}_{t,0}]^T - (\gamma_t^2)^2\Phi^{(2)}_{t,0}\Sigma_\beta A_{22}^T[\Phi^{(2)}_{t,0}]^T \tag{2.212}$$

$$= \dot{\gamma}_t^2\Phi^{(2)}_{t,0}\Sigma_\beta[\Phi^{(2)}_{t,0}]^T - (\gamma_t^2)^2\Phi^{(2)}_{t,0}\left[A_{22}\Sigma_\beta + \Sigma_\beta A_{22}^T\right][\Phi^{(2)}_{t,0}]^T \tag{2.213}$$

$$= \dot{\gamma}_t^2\Phi^{(2)}_{t,0}\Sigma_\beta[\Phi^{(2)}_{t,0}]^T + (\gamma_t^2)^2\Phi^{(2)}_{t,0}\Gamma_{22}[\Phi^{(2)}_{t,0}]^T \tag{2.214}$$

where in the final line we have used the well known fact that Σ_β is the solution of the Lyapunov equation (e.g., [354, page 5])

$$A_{22}\Sigma_\beta + \Sigma_\beta A_{22}^T = -\Gamma_{22}. \tag{2.215}$$

We thus have

$$P_t - P_0 = \int_0^t \dot{P}_s\mathrm{d}s = \int_0^t \dot{\gamma}_s^2\Phi^{(2)}_{s,0}\Sigma_\beta[\Phi^{(2)}_{s,0}]^T + (\gamma_s^2)^2\Phi^{(2)}_{s,0}\Gamma_{22}[\Phi^{(2)}_{s,0}]^T\mathrm{d}s \tag{2.216}$$

$$= \int_0^t \gamma_s^2 \Phi_{s,0}^{(2)} \left[\frac{\dot{\gamma}_s^2}{\gamma_s^2} \Sigma_\beta + \gamma_s^2 \Gamma_{22} \right] [\Phi_{s,0}^{(2)}]^T \mathrm{d}s \tag{2.217}$$

Substituting back into (2.210), we have

$$\gamma_t^2 \Phi_{t,0}^{(2)} B_t [\Phi_{t,0}^{(2)}]^T = \int_0^t \gamma_s^2 \Phi_{s,0}^{(2)} \left[\gamma_s^2 \Gamma_{22} - \frac{\dot{\gamma}_s^2}{\gamma_s^2} \Sigma_\beta - \gamma_s^2 \Gamma_{22} \right] [\Phi_{s,0}^{(2)}]^T \mathrm{d}s + P_0 \tag{2.218}$$

$$= \int_0^t \gamma_s^2 \Phi_{s,0}^{(2)} \left[- \frac{\dot{\gamma}_s^2}{\gamma_s^2} \Sigma_\beta \right] [\Phi_{s,0}^{(2)}]^T \mathrm{d}s + \gamma_0^2 \Sigma_\beta \tag{2.219}$$

It follows that

$$||B_t|| \leq K(\gamma_t^2)^{-1} \int_0^t \gamma_s^2 \left[\frac{\dot{\gamma}_s^2}{\gamma_s^2} \right] ||\Sigma_\beta|| ||\Phi_{s,t}^{(2)}||^2 \mathrm{d}s + \frac{\gamma_0^2}{\gamma_t^2} ||\Sigma_\beta|| ||\Phi_{0,t}^{(2)}||^2 \tag{2.220}$$

$$\leq K(\gamma_t^2)^{-1} \int_0^t \gamma_s^2 \Phi_{s,t}^{(2)} \left[\frac{\dot{\gamma}_s^2}{\gamma_s^2} \right] \mathrm{d}s + K \left[(\gamma_t^2)^{-\frac{1}{2}} \Psi_{0,t}^{(2)} \right]^2 \tag{2.221}$$

For the first term, we begin with the observation that

$$\frac{\dot{\gamma}_t^2}{\gamma_t^2} = o(\gamma_t^2) = \mathcal{O}((\gamma_t^2)^{1+\delta}) \tag{2.222}$$

for some $\delta > 0$. It follows from Lemma 2.9 that, for sufficiently large t,

$$\int_0^t \gamma_s^2 \Phi_{s,t}^{(2)} \left[\frac{\dot{\gamma}_s^2}{\gamma_s^2} \right] \mathrm{d}s \leq K(\gamma_t^2)^{1+\delta}. \tag{2.223}$$

Thus, the first term converges to zero as $t \to \infty$. Meanwhile, the second term converges to zero by Lemma 2.11. This completes the proof that $\Sigma_t^{2,2} \xrightarrow{\mathbb{P}} \Sigma_\beta$ as $t \to \infty$.

Using an essentially identical argument, but now appealing to Lemma 2.8, 2.10, and the fact that Σ_α is the solution of the Lyapunov equation (e.g., [354, page 5])

$$\left(H + \frac{\mathbb{1}_{\eta_1 = 1}}{2\gamma_0^0} I \right) \Sigma_\alpha + \Sigma_\alpha \left(H + \frac{\mathbb{1}_{\eta_1 = 1}}{2\gamma_0^0} I \right)^T = -\Gamma_\alpha \tag{2.224}$$

we can also show that $\Sigma_t^{1,1} \xrightarrow{\mathbb{P}} \Sigma_\alpha$ as $t \to \infty$. It remains to deal with $\Sigma_t^{1,2}$. For this term, observe that, for sufficiently large t,

$$||\Sigma_t^{1,2}|| \leq K\sqrt{(\gamma_t^1)^{-1}(\gamma_t^2)^{-1}} \int_0^t \gamma_s^1\gamma_s^2\Psi_{s,t}^{(1)}\Psi_{s,t}^{(2)}\mathrm{d}s \tag{2.225}$$

$$\leq K\sqrt{(\gamma_t^1)^{-1}(\gamma_t^2)^{-1}} \int_0^t \gamma_s^2\Psi_{s,t}^{(2)}\gamma_s^1\mathrm{d}s \tag{2.226}$$

$$\leq K\sqrt{(\gamma_t^1)^{-1}(\gamma_t^2)^{-1}\gamma_t^1} = K\sqrt{\gamma_t^1(\gamma_t^2)^{-1}} \to 0 \tag{2.227}$$

where in the first line we have used the stability of H and A_{22}, in the third line we have used Lemma 2.8, and in the final line we have used Assumption 2.1.1'. It follows that

$$[\Sigma]_t \xrightarrow{\mathbb{P}} \begin{pmatrix} \Sigma_\alpha & 0 \\ 0 & \Sigma_\beta \end{pmatrix} \tag{2.228}$$

The convergence in distribution to a normal random variable with zero mean and this covariance now follows from standard results (e.g., [276, Section 1.2.2]). □

Lemma 2.3. *The processes $\{L_t^\alpha\}_{t\geq 0}$, $\{L_t^\beta\}_{t\geq 0}$ satisfy, for sufficiently large t,*

$$||L_t^\alpha|| \leq K\sqrt{\gamma_t^1 \log \int_0^t \gamma_s^1\mathrm{d}s} \tag{2.229}$$

$$||L_t^\beta|| \leq K\sqrt{\gamma_t^2 \log \int_0^t \gamma_s^2\mathrm{d}s} \tag{2.230}$$

Proof. We will prove that the desired bound holds for L_t^β; an almost identical argument can be used for L_t^α. Let $-\mu_2$ be an eigenvalue of A_{22}^T, and let v_2 be an eigenvector associated with $-\mu_2$. Let $\{N_t\}_{t\geq 0}$ be the martingale defined according to

$$N_t = \int_0^t \Psi_{s,0}^{(2)}\gamma_s^2 v_2^T\xi_2^{(2)}(\alpha_s,\beta_s)\mathrm{d}w_s^2. \tag{2.231}$$

where we recall, c.f. (2.174), that

$$\Psi_{s,0}^{(2)} = \exp\left[-\mu_2\int_s^0 \gamma_u^2\mathrm{d}u\right] = \exp\left[\mu_2\int_0^s \gamma_u^2\mathrm{d}u\right]. \tag{2.232}$$

The (predictable) quadratic variation of this martingale is given by

$$[N]_t = \int_0^t (\gamma_s^2)^2 \Psi_{s,0}^{(2)} v_2^T \Gamma_s^{22} [\Psi_{s,0}^{(2)}]^T v_2 \mathrm{d}s \tag{2.233}$$

Arguing as in the proof of Lemma 2.2, c.f. (2.199) - (2.223), we can show that

$$\lim_{t\to\infty} (\gamma_t^2)^{-1} \Psi_{0,t}^{(2)} [N]_t [\Psi_{0,t}^{(2)}]^T = v_2^T \Sigma_\beta v_2 \tag{2.234}$$

That is, re-ordering (note that Ψ is a scalar, and thus re-ordering is permitted here)

$$\lim_{t\to\infty} (\gamma_t^2)^{-1} [\Psi_{0,t}^{(2)}]^2 [N]_t = v_2^T \Sigma_\beta v_2 \tag{2.235}$$

Now, according to Lemma 2.11, we have that

$$\lim_{t\to\infty} (\gamma_t^2)^{-1} [\Psi_{0,t}^{(2)}]^2 = 0. \tag{2.236}$$

Thus, it follows that $\lim_{t\to\infty}[N]_t = \infty$ a.s. Now, applying the law of the iterated logarithm for stochastic integrals (e.g, [456, 461]), we have that

$$\limsup_{t\to\infty} \frac{|N_t|}{\sqrt{2[N]_t \ln\ln[N]_t}} \leq 1 \tag{2.237}$$

In view of (2.234), it follows that, for sufficiently large t,

$$\begin{aligned} |N_t| &\leq K\sqrt{[N]_t \ln\ln[N]_t} && (2.238)\\ &\leq K\sqrt{\gamma_t^2 [\Psi_{t,0}^{(2)}]^2 v_2^T \Sigma_\beta v_2 \ln\ln\left(\gamma_t^2 [\Psi_{t,0}^{(2)}]^2 v_2^T \Sigma_\beta v_2\right)} && (2.239)\\ &\leq K\Psi_{t,0}^{(2)} \sqrt{\gamma_t^2 \ln \int_0^t \gamma_s^2 \mathrm{d}s} && (2.240) \end{aligned}$$

It remains to note, using the fact that v_2 is an eigenvector of A_{22}^T, that

$$\begin{aligned} v_2^T L_t^\beta &:= \int_0^t v_2^T \Phi_{s,t}^{(2)} \gamma_s^2 \xi_2^{(2)}(\alpha_s, \beta_s) \mathrm{d}w_s^2 && (2.241)\\ &= \int_0^t \Psi_{s,t}^{(2)} \gamma_s^2 v_2^T \xi_2^{(2)}(\alpha_s, \beta_s) \mathrm{d}w_s^2 && (2.242)\\ &= \Psi_{0,t}^{(2)} \int_0^t \Psi_{s,0}^{(2)} \gamma_s^2 v_2^T \xi_2^{(2)}(\alpha_s, \beta_s) \mathrm{d}w_s^2 := \Psi_{0,t}^{(2)} N_t && (2.243) \end{aligned}$$

It follows, combining (2.240) and (2.243), that for any eigenvector w_2 of A_{22}^T, and for sufficiently large t

$$|v_2^T L_t^\beta| \leq K\Psi_{0,t}^{(2)}\Psi_{t,0}^{(2)}\sqrt{\gamma_t^2 \ln \int_0^t \gamma_s^2 \mathrm{d}s} = K\sqrt{\gamma_t^2 \ln \int_0^t \gamma_s^2 \mathrm{d}s} \tag{2.244}$$

and thus, in particular, that

$$||L_t^\beta|| \leq K\sqrt{\gamma_t^2 \ln \int_0^t \gamma_s^2 \mathrm{d}s}. \tag{2.245}$$

□

Lemma 2.4. *There exists $\delta > 0$ such that, for sufficiently large t, the processes $\{R_t^\alpha\}_{t\geq 0}$, $\{R_t^\beta\}_{t\geq 0}$ satisfy*

$$||R_t^\alpha|| \leq K(\gamma_t^1)^{\frac{1}{2}+\delta} \tag{2.246}$$

$$||R_t^\beta|| \leq K\sqrt{\gamma_t^1 \int_0^t \log \gamma_s^1 \mathrm{d}s} \tag{2.247}$$

Lemma 2.5. *There exist $\delta > 0$ such that, for sufficiently large t, the processes $\{\Delta_t^\alpha\}_{t\geq 0}$, $\{\Delta_t^\beta\}_{t\geq 0}$ satisfy*

$$||\Delta_t^\alpha|| + ||\Delta_t^\beta|| \leq K(\gamma_t^1)^{\frac{1}{2}+\delta} \tag{2.248}$$

Proof. Suppose that $\{w_t\}_{t\geq 0}$ is a sequence of positive numbers satisfying Conditions (A_1) and (A_2) (see Definition 2.1). Suppose also that $||\beta_t|| = O(w_t)$. Since $\lim_{t\to\infty} \Delta_t^\beta = 0$ (see the proof of Lemma 2.7), we know that $||\Delta_t^\beta|| = O(1)$. We can thus apply Lemma 2.7 with $w_t^1 := w_t$, $w_t^2 := 1$, to obtain

$$||\Delta_t^\alpha|| + ||\Delta_t^\beta|| \leq K\left[\frac{(\gamma_t^1)^2}{(\gamma_t^2)^2}(w_t)^2 + \frac{\gamma_t^1}{\gamma_t^2} + (\gamma_t^1)^{\frac{1}{2}+\delta}\right] \tag{2.249}$$

We in fact claim that, for all $k \in \mathbb{N}$, it holds that

$$||\Delta_t^\alpha|| + ||\Delta_t^\beta|| \leq K\left[\frac{(\gamma_t^1)^2}{(\gamma_t^2)^2}(w_t)^2 + \left[\frac{\gamma_t^1}{\gamma_t^2}\right]^k + (\gamma_t^1)^{\frac{1}{2}+\delta}\right] \tag{2.250}$$

We will prove this inductively. Clearly, the base case is true. Let us assume that the hypothesis holds for some $k \in \mathbb{N}$. Since w_t satisfies Conditions (A_1) and (A_2), so too does $\frac{(\gamma_t^1)^2}{(\gamma_t^2)^2}(w_t)^2 + \frac{\gamma_1^k(t)}{\gamma_2^k(t)} + (\gamma_t^1)^{\frac{1}{2}}$. It follows from Lemma 2.7, with $w_t^1 = w_t$, w_t^2 equal to the sequence just defined, that

$$||\Delta_t^{\alpha}|| + ||\Delta_t^{\beta}|| \leq K \left[\frac{(\gamma_t^1)^2}{(\gamma_t^2)^2}(w_t)^2 + \left[\frac{\gamma_t^1}{\gamma_t^2}\right]^{k+1} + (\gamma_t^1)^{\frac{1}{2}+\delta} \right] \tag{2.251}$$

Thus, by induction, the bound (2.250) holds for all $k \in \mathbb{N}$. By Assumption 2.1.1', there exists $k \in \mathbb{N}$ such that, for sufficiently large t, $\frac{(\gamma_t^1)^k}{(\gamma_t^2)^k} \leq K(\gamma_t^1)^{\frac{1}{2}+\delta}$. Thus, it follows from (2.250) that, for any sequence of positive numbers $\{w_t\}_{t\geq 0}$ satisfying Conditions (A_1) and (A_2) and such that $||\beta_t|| = O(w_t)$, we have

$$||\Delta_t^{\alpha}|| + ||\Delta_t^{\beta}|| \leq K \left[\frac{(\gamma_t^1)^2}{(\gamma_t^2)^2}(w_t)^2 + (\gamma_t^1)^{\frac{1}{2}+\delta} \right] \tag{2.252}$$

To complete the proof, it remains for us to find an appropriate w. We claim that, for all $k \in \mathbb{N}_0$, the function defined according to

$$w_t = \sqrt{\gamma_t^2 \log \int_0^t \gamma_s^2 \mathrm{d}s} + \left[\frac{\gamma_t^1}{\gamma_t^2}\right]^k \tag{2.253}$$

is one such function. Certainly, this function satisfies Conditions (A_1) and (A_2). It remains to show that $||\beta_t|| = O(w_t)$. We will once more prove this by induction. By Assumption 2.1.4, we have that β_t is bounded a.s., and so this bound holds for $k = 0$. For the inductive step, using Lemma 2.3, Lemma 2.6, and the bound in (2.252), we have that for sufficiently large t,

$$\begin{aligned}
||\beta_t|| &\leq K \left[||L_t^{\beta}|| + ||R_t^{\beta}|| + ||\Delta_t^{\beta}|| \right] && (2.254)\\
&\leq K \left[\sqrt{\gamma_t^2 \log \int_0^t \gamma_s^2 \mathrm{d}s} + \frac{\gamma_t^1}{\gamma_t^2} \left[\sqrt{\gamma_t^2 \log \int_0^t \gamma_s^2 \mathrm{d}s} + \left[\frac{\gamma_t^1}{\gamma_t^2}\right]^k \right] \right. && (2.255)\\
&\quad \left. + \sqrt{\gamma_t^1 \log \int_0^t \gamma_s^1 \mathrm{d}s} + \frac{(\gamma_t^1)^2}{(\gamma_t^2)^2} \left[\sqrt{\gamma_t^2 \log \int_0^t \gamma_s^2 \mathrm{d}s} + \left[\frac{\gamma_t^1}{\gamma_t^2}\right]^k \right]^2 + \gamma_1^{\frac{1}{2}+\delta} \right] && (2.256)\\
&\leq K \left[\sqrt{\gamma_t^2 \log \int_0^t \gamma_s^2 \mathrm{d}s} + \left[\frac{\gamma_t^1}{\gamma_t^2}\right]^{k+1} \right] && (2.257)
\end{aligned}$$

where in going from the first line to the second line we have made repeated use of Assumption 2.1.1'. This completes the inductive proof of (2.253). In fact, Assumption 2.1.1' ensures that, for sufficiently large k, there exists K such that, for sufficiently large t,

$[\frac{\gamma_t^1}{\gamma_t^2}]^k \leq K[\gamma_t^2 \int_0^t \gamma_s^2 \mathrm{d}s]^{\frac{1}{2}}$. Thus, we in fact have that

$$||\beta_t|| \leq K\sqrt{\gamma_t^2 \log \int_0^t \gamma_s^2 \mathrm{d}s}. \tag{2.258}$$

and we can improve upon the choice of w_t in (2.253) by choosing

$$w_t = \sqrt{\gamma_t^2 \log \int_0^t \gamma_s^2 \mathrm{d}s} \tag{2.259}$$

To prove Lemma 2.4, it remains to apply Lemma 2.6 with this choice of w_t. In particular, this yields that for all $s \in (\frac{1}{2}, \gamma_0^1 \Lambda_H)$,

$$||R_t^\alpha|| \leq K\left[\frac{\gamma_t^1}{\gamma_t^2}\sqrt{\gamma_t^2 \log \int_0^t \gamma_s^2 \mathrm{d}s} + t^{-s}\right] \leq K\left[t^{-(\eta_0^1 - \frac{\eta_0^2}{2})}\sqrt{\log t} + t^{-s}\right]. \tag{2.260}$$

By Assumption 2.1.1', we have that $\eta_0^1 - \frac{\eta_0^2}{2} > \frac{\eta_0^1}{2}$ and $s > \frac{\eta_0^1}{2}$. It follows that, for some $s > \frac{\eta_0^1}{2}$, for sufficiently large t it holds that $||R_t^\alpha|| \leq Kt^{-s}$. Equivalently, for some $\delta > 0$, for sufficiently large t we have that

$$||R_t^\alpha|| \leq K(\gamma_t^1)^{\frac{1}{2}+\delta}. \tag{2.261}$$

In addition, once more using Lemma 2.6, we have that

$$||R_t^\beta|| \leq K\left[\frac{\gamma_t^1}{\gamma_t^2}\sqrt{\gamma_t^2 \int_0^t \gamma_s^2 \mathrm{d}s} + \sqrt{\gamma_t^1 \int_0^t \gamma_s^1 \mathrm{d}s}\right] \tag{2.262}$$

$$= K\left[\frac{(\gamma_t^1)^{\frac{1}{2}}}{(\gamma_t^2)^{\frac{1}{2}}}\sqrt{\gamma_t^1 \int_0^t \gamma_s^2 \mathrm{d}s} + \sqrt{\gamma_t^1 \int_0^t \gamma_s^1 \mathrm{d}s}\right] \tag{2.263}$$

$$\leq K\sqrt{\gamma_t^1 \int_0^t \gamma_s^1 \mathrm{d}s}. \tag{2.264}$$

To prove Lemma 2.5, we will apply (2.252) with $w_t = \sqrt{\gamma_t^2 \int_0^t \gamma_s^2 \mathrm{d}s}$. This gives that, for sufficiently large t, there exists $K > 0$ and (small) $\delta > 0$ such that

$$||\Delta_t^\alpha|| + ||\Delta_t^\beta|| \leq K\left[\frac{(\gamma_t^1)^2}{(\gamma_t^2)^2}\left[\gamma_t^2 \int_0^t \gamma_s^2 \mathrm{d}s\right] + (\gamma_t^1)^{\frac{1}{2}+\delta}\right] \leq K(\gamma_t^1)^{\frac{1}{2}+\delta} \tag{2.265}$$

where the second inequality follows from Assumption 2.1.1'. □

2.A.0.2 Intermediate Lemmas

Definition 2.1. *Let $\{w_t\}_{t\geq 0}$ be a positive, bounded, non-increasing, continuous sequence of real numbers. We say that w_t satisfies Condition* (A_1) *if:*

(i) In the case $\eta_1 = 1$, there exists $\omega \geq 0$ and a non-decreasing slowly varying function $\mathcal{L}$ such that

$$w_t = (\delta_0 + t)^{-\omega}\mathcal{L}(t). \tag{2.266}$$

(ii) In the case $\eta_1 < 1$, for all $s \leq t$,

$$\frac{w_s}{w_t} \leq \exp\left[o(1)\int_s^t \gamma_u^1 \mathrm{d}u\right] \tag{2.267}$$

where $o(1) \to 0$ as $s \to \infty$.

Definition 2.2. *Let $\{w_t\}_{t\geq 0}$ be a positive, bounded, non-increasing, continuous sequence of real numbers. We say that w_t satisfies Condition* (A_2) *if, for all $s \leq t$,*

$$\frac{w_s}{w_t} \leq \exp\left[o(1)\int_s^t \gamma_u^2 \mathrm{d}u\right] \tag{2.268}$$

where $o(1) \to 0$ as $s \to \infty$.

Lemma 2.6. *Assume there exists a function w_t satisfying Conditions (A_1) and (A_2) such that $||\beta_t|| = O(w_t)$ for sufficiently large t. Then, for sufficiently large t, there exists $K > 0$ such that, for all $s \in (\frac{1}{2}, \gamma_0^1\Lambda_H)$,*

$$||R_t^\alpha|| \leq K\left[\frac{\gamma_t^1}{\gamma_t^2}w_t + t^{-s}\right] \tag{2.269}$$

$$||R_t^\beta|| \leq K\left[\frac{\gamma_t^1}{\gamma_t^2}w_t + \sqrt{\gamma_t^1 \log\int_0^t \gamma_s^1 \mathrm{d}s}\right] \tag{2.270}$$

Proof. We will begin by considering R_t^α. Applying Itô's formula to the function $h(s,\beta_s) = \Phi_{s,t}^{(1)}\frac{\gamma_s^1}{\gamma_s^2}A_{12}A_{22}^{-1}\beta_s$, we obtain

$$\Phi_{t,t}^{(1)}\frac{\gamma_t^1}{\gamma_t^2}A_{12}A_{22}^{-1}\beta_t - \Phi_{0,t}^{(1)}\frac{\gamma_0^1}{\gamma_0^2}A_{12}A_{22}^{-1}\beta_0 \tag{2.271}$$

$$= \underbrace{\int_0^t \Phi^{(1)}_{s,t}\frac{\gamma_s^1}{\gamma_s^2}A_{12}A_{22}^{-1}\mathrm{d}\beta_s}_{R_t^\alpha} + \int_0^t \dot{\Phi}^{(1)}_{s,t}\frac{\gamma_s^1}{\gamma_s^2}A_{12}A_{22}^{-1}\beta_s\mathrm{d}s + \int_0^t \Phi^{(1)}_{s,t}\frac{\dot{\gamma}_s^1}{\gamma_s^2}A_{12}A_{22}^{-1}\beta_s\mathrm{d}s$$

from which it follows that

$$R_t^\alpha = \frac{\gamma_t^1}{\gamma_t^2}A_{12}A_{22}^{-1}\beta_t - \Phi^{(1)}_{0,t}\frac{\gamma_0^1}{\gamma_0^2}A_{12}A_{22}^{-1}\beta_0 \tag{2.272}$$

$$- \int_0^t \Phi^{(1)}_{s,t}\frac{(\gamma_s^1)^2}{\gamma_s^2}HA_{12}A_{22}^{-1}\beta_s\mathrm{d}s - \int_0^t \Phi^{(1)}_{s,t}\frac{\dot{\gamma}_s^1}{\gamma_s^2}A_{12}A_{22}^{-1}\beta_s\mathrm{d}s. \tag{2.273}$$

We thus have

$$||R_t^\alpha|| \leq K\frac{\gamma_t^1}{\gamma_t^2}||\beta_t|| + K||\Phi^{(1)}_{0,t}||\;||\beta_0|| + K\int_0^t ||\Phi^{(1)}_{s,t}||\left[\frac{(\gamma_s^1)^2}{\gamma_s^2} + \frac{\dot{\gamma}_s^1}{\gamma_s^2}\right]||\beta_s||\mathrm{d}s \tag{2.274}$$

$$\leq K\frac{\gamma_t^1}{\gamma_t^2}w_t + K||\Phi^{(1)}_{0,t}|| + K\int_0^t \Psi^{(1)}_{s,t}\gamma_s^1\left[\frac{\gamma_s^1}{\gamma_s^2}\right]w_s\mathrm{d}s. \tag{2.275}$$

Now, observe that, for all $\mu \in (0, \Lambda_H)$, it holds that

$$||\Phi^{(1)}_{0,t}|| \leq K\Phi^{(1)}_{0,t} := K\exp\left[-\mu\int_0^t \gamma_u^1\mathrm{d}u\right] \leq \exp\left[-\mu\int_0^t \gamma_0^1(\delta_1+s)^{-1}\mathrm{d}s\right] \leq Kt^{-\mu\gamma_0^1}. \tag{2.276}$$

Moreover, if w_t satisfies Condition (A$_1$), then so too does $\tilde{w}_t = \frac{\gamma_t^1}{\gamma_t^2}w_t$. We can thus apply Lemma 2.8 to $\tilde{w}_t$ to conclude that, for all $\mu \in (0, \Lambda_H)$,

$$\int_0^t \Psi^{(1)}_{s,t}\gamma_s^1\left[\frac{\gamma_s^1}{\gamma_s^2}\right]w_s\mathrm{d}s \leq K\left[\frac{\gamma_t^1}{\gamma_t^2}w_t + t^{-\mu\gamma_0^1}\mathbb{1}_{\eta_1=1}\right]. \tag{2.277}$$

Combining (2.276) and (2.277), we thus have that

$$||R_t^\alpha|| \leq K\left[\frac{\gamma_t^1}{\gamma_t^2}w_t + t^{-\mu\gamma_0^1}\right] \tag{2.278}$$

for all $\mu \in (0, \Lambda_H)$. Thus, in particular, this bound holds for all $\mu \in (\frac{1}{2\gamma_0^1}, \Lambda_H)$ or, equivalently, for all $\mu\gamma_0^1 \in (\frac{1}{2}, \gamma_0^1\Lambda_H)$. It follows that, for all $s \in (\frac{1}{2}, \gamma_0^1\Lambda_H)$, for sufficiently large t we have that

$$||R_t^\alpha|| \leq K\left[\frac{\gamma_t^1}{\gamma_t^2}w_t + t^{-s}\right]. \tag{2.279}$$

We now turn our attention to R_t^β. For this term, substituting the existing bounds for

$||L_t^\alpha||$ and $||R_t^\alpha||$, c.f. (2.229) and (2.269), we obtain

$$||R_t^\beta|| \leq K \int_0^t ||\Phi_{s,t}^{(2)}||\gamma_s^2 \left[||L_s^\alpha|| + ||R_s^\alpha||\right] \mathrm{d}s \tag{2.280}$$

$$\leq K \int_0^t \Psi_{s,t}^{(2)}\gamma_s^2 \left[\left[\gamma_s^1 \log \int_0^s \gamma_u^1 \mathrm{d}u\right]^{\frac{1}{2}} + \left[\frac{\gamma_s^1}{\gamma_s^2}w_s + t^{-s}\right]\right] \mathrm{d}s \tag{2.281}$$

$$\leq K \int_0^t \Psi_{s,t}^{(2)}\gamma_s^2 \left[\left[\gamma_s^1 \log \int_0^s \gamma_u^1 \mathrm{d}u\right]^{\frac{1}{2}} + \left[\frac{\gamma_s^1}{\gamma_s^2}w_s\right]\right] \mathrm{d}s \tag{2.282}$$

where in going from (2.281) to (2.282), we have used the fact that the final term is smaller than the first term. It remains to note that both terms in the square brackets in (2.282) satisfy Condition (A_2), and thus we can apply Lemma 2.9 to conclude that, for sufficiently large t,

$$||R_t^\beta|| \leq K \left[\frac{\gamma_t^1}{\gamma_t^2}w_t + \sqrt{\gamma_t^1 \log \int_0^t \gamma_s^1 \mathrm{d}s}\right]. \tag{2.283}$$

□

Lemma 2.7. *Assume that there exist functions w_t^i satisfying Conditions (A_1) and (A_2) such that $||\beta_t|| \leq K w_t^1$ and $||\Delta_t^\beta|| \leq K w_t^2$ for sufficiently large t. Then there exist $K, \delta > 0$ such that, for sufficiently large t,*

$$||\Delta_t^\alpha|| \leq K \left[\frac{(\gamma_t^1)^2}{(\gamma_t^2)^2}(w_t^1)^2 + \frac{\gamma_t^1}{\gamma_t^2}w_t^2 + (\gamma_t^1)^{\frac{1}{2}+\delta}\right] \tag{2.284}$$

$$||\Delta_t^\beta|| \leq K \left[\frac{(\gamma_t^1)^2}{(\gamma_t^2)^2}(w_t^1)^2 + \frac{\gamma_t^1}{\gamma_t^2}w_t^2 + (\gamma_t^1)^{\frac{1}{2}+\delta}\right]. \tag{2.285}$$

Proof. From the definition, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta_t^\alpha - \gamma_t^1 H \Delta_t^\alpha = \gamma_t^1\left[\xi_1^{(1)}(\alpha_t,\beta_t) - A_{12}A_{22}^{-1}\xi_2^{(1)}(\alpha_t,\beta_t)\right] + \gamma_t^1\left[\varepsilon_t^1 - A_{12}A_{22}^{-1}\varepsilon_t^2\right] \tag{2.286}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta_t^\beta - \gamma_t^2 A_{22}\Delta_t^\beta = \gamma_t^2\left[A_{21}\Delta_t^\alpha + \xi_2^{(1)}(\alpha_t,\beta_t) + \varepsilon_t^2\right] \tag{2.287}$$

Let $\mu_1 \in (\frac{\mathbb{1}_{\eta_1=1}}{2\gamma_0^1}, \Lambda_H)$ and $\mu_2 \in (0, \Lambda_{A_{22}})$. It follows, taking norms, using the triangle inequality, and arguing in a similar fashion to [354, page 12], that

$$\frac{\mathrm{d}}{\mathrm{d}t}||\Delta_t^\alpha|| \leq -\mu_1\gamma_t^1||\Delta_t^\alpha|| + K_1\gamma_t^1\left[||\xi_1^{(1)}(\alpha_t,\beta_t)|| + ||\xi_2^{(1)}(\alpha_t,\beta_t)|| + ||\varepsilon_t^1|| + ||\varepsilon_t^2||\right] \tag{2.288}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}||\Delta_t^\beta|| \leq -\mu_2\gamma_t^2||\Delta_t^\beta|| + K_2\gamma_t^2\left[||\Delta_t^\alpha|| + ||\xi_2^{(1)}(\alpha_t,\beta_t)|| + ||\varepsilon_t^2||\right] \tag{2.289}$$

Now, using Assumption 2.1.7'a and Assumption 2.1.3'b, and the decomposition (2.177) - (2.180), it follows, allowing the values of the constants K_1, K_2 to vary from line to line, that

$$\frac{\mathrm{d}}{\mathrm{d}t}||\Delta_t^\alpha|| \leq -\mu_1\gamma_t^1||\Delta_t^\alpha|| + K_1\gamma_t^1\left[||L_t^\alpha||^2 + ||L_t^\beta||^2 + ||R_t^\alpha||^2 + ||R_t^\beta||^2 + ||\Delta_t^\alpha||^2 + ||\Delta_t^\beta||^2\right] \tag{2.290}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}||\Delta_t^\beta|| \leq -\mu_2\gamma_t^2||\Delta_t^\beta|| + K_2\gamma_t^2\Big[||L_t^\alpha||^2 + ||L_t^\beta||^2 + ||R_t^\alpha||^2 + ||R_t^\beta||^2 + ||\Delta_t^\alpha|| + ||\Delta_t^\alpha||^2 + ||\Delta_t^\beta||^2\Big] \tag{2.291}$$

By Theorem 2.1, $\lim_{t\to\infty}\alpha_t = \lim_{t\to\infty}\beta_t = 0$ a.s. In addition, Lemma 2.6 implies that $\lim_{t\to\infty}R_t^\alpha = \lim_{t\to\infty}R_t^\beta = 0$ a.s. and Lemma 2.3 implies that $\lim_{t\to\infty}L_t^\alpha = \lim_{t\to\infty}L_t^\beta = 0$ a.s. It follows immediately that $\lim_{t\to\infty}\Delta_t^\alpha = \lim_{t\to\infty}\Delta_t^\beta = 0$. Let $\tilde{\mu}_1 \in (\frac{\mathbb{1}_{\eta_1=1}}{2\gamma_0^1}, \mu_1)$ and $\mu_2 \in (0, \mu_2)$. Then, from the previous expressions, it follows that there exists $\tilde{K}_2 > 0$ such that for sufficiently large t, we have (see also [354, page 13])

$$\frac{\mathrm{d}}{\mathrm{d}t}||\Delta_t^\alpha|| \leq -\tilde{\mu}_1\gamma_t^1||\Delta_t^\alpha|| + K_1\gamma_t^1\left[||L_t^\alpha||^2 + ||L_t^\beta||^2 + ||R_t^\alpha||^2 + ||R_t^\beta||^2 + ||\Delta_t^\beta||^2\right] \tag{2.292}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}||\Delta_t^\beta|| \leq -\tilde{\mu}_2\gamma_t^2||\Delta_t^\beta|| + \tilde{K}_2\tilde{\mu}_2\gamma_t^2\left[||L_t^\alpha||^2 + ||L_t^\beta||^2 + ||R_t^\alpha||^2 + ||R_t^\beta||^2 + ||\Delta_t^\alpha||\right] \tag{2.293}$$

Now, from (2.294), we have that

$$||\Delta_t^\beta|| \leq -\frac{1}{\tilde{\mu}_2\gamma_t^2}\frac{\mathrm{d}}{\mathrm{d}t}||\Delta_t^\beta|| + \tilde{K}_2\left[||L_t^\alpha||^2 + ||L_t^\beta||^2 + ||R_t^\alpha||^2 + ||R_t^\beta||^2 + ||\Delta_t^\alpha||\right] \tag{2.294}$$

Let $\tilde{K}_1 \in \left(0, \frac{1}{\tilde{K}_2}(\tilde{\mu}_1 - \frac{\mathbb{1}_{\eta_1=1}}{2\gamma_0^1})\right)$. Then substituting (2.294) into (2.292), and once more making use of the fact that $\lim_{t\to\infty}\Delta_t^\beta = 0$, it follows that, for sufficiently large t, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}||\Delta_t^\alpha|| \leq -\tilde{\mu}_1\gamma_t^1||\Delta_t^\alpha|| + K_1\gamma_t^1\left[||L_t^\alpha||^2 + ||L_t^\beta||^2 + ||R_t^\alpha||^2 + ||R_t^\beta||^2\right] + \tilde{K}_1\gamma_t^1||\Delta_t^\beta|| \tag{2.295}$$

$$\begin{aligned}\leq & -\tilde{\mu}_1\gamma_t^1||\Delta_t^\alpha|| + K_1\gamma_t^1\left[||L_t^\alpha||^2 + ||L_t^\beta||^2 + ||R_t^\alpha||^2 + ||R_t^\beta||^2\right] \\ & - \frac{\tilde{K}_1\gamma_t^1}{\tilde{\mu}_2\gamma_t^2}\frac{\mathrm{d}}{\mathrm{d}t}||\Delta_t^\beta|| + \tilde{K}_1\tilde{K}_2\gamma_t^1\left[||L_t^\alpha||^2 + ||L_t^\beta||^2 + ||R_t^\alpha||^2 + ||R_t^\beta||^2 + ||\Delta_t^\alpha||\right]\end{aligned} \tag{2.296}$$

Let $\bar{\mu}_1 \in (\frac{\mathbb{1}_{\eta_1=1}}{2\gamma_0^1}, \tilde{\mu}_1 - \tilde{K}_1\tilde{K}_2)$, and $\bar{K}_1 = \max\{K_1, \tilde{K}_1\tilde{K}_2\}$. Then, for sufficiently large t, it follows from (2.296) that

$$\frac{\mathrm{d}}{\mathrm{d}t}||\Delta_t^\alpha|| \leq -\bar{\mu}_1\gamma_t^1||\Delta_t^\alpha|| + \bar{K}_1\gamma_t^1\left[||L_t^\alpha||^2 + ||L_t^\beta||^2 + ||R_t^\alpha||^2 + ||R_t^\beta||^2\right] - \frac{\tilde{K}_1\gamma_t^1}{\tilde{\mu}_2\gamma_t^2}\frac{\mathrm{d}}{\mathrm{d}t}||\Delta_t^\beta|| \tag{2.297}$$

Let $\bar{\Psi}^{(1)}_{s,t} = \exp\left[-\bar{\mu}_1 \int_s^t \gamma_u^1 \mathrm{d}u\right]$. Then it follows from the previous expression that

$$||\Delta_t^\alpha|| \leq ||\Delta_0^\alpha||\bar{\Psi}^{(1)}_{0,t} + \bar{K}_1 \int_0^t \bar{\Psi}^{(1)}_{1,t}\gamma_s^1 \left[||L_s^\alpha||^2 + ||L_s^\beta||^2 + ||R_s^\alpha||^2 + ||R_s^\beta||^2\right] \mathrm{d}s$$

$$+ \frac{\tilde{K}_1}{\tilde{\mu}_2} \int_0^t \bar{\Psi}^{(1)}_{1,t} \frac{\gamma_s^1}{\gamma_s^2} \frac{\mathrm{d}}{\mathrm{d}s} ||\Delta_s^\beta|| \mathrm{d}s \tag{2.298}$$

$$:= ||\Delta_t^{\alpha,1}|| + ||\Delta_t^{\alpha,2}|| + ||\Delta_t^{\alpha,3}|| \tag{2.299}$$

Let us consider each of these terms in turn. Using Assumption 2.1.1', and the fact that $||\Delta_0^\alpha|| = O(1)$, Lemma 2.10 implies that, for sufficiently large t,

$$||\Delta_t^{\alpha,1}|| \leq K\bar{\Psi}^{(1)}_{0,t} \leq K(\gamma_t^1)^{\frac{1}{2}+\delta} \tag{2.300}$$

for some $\delta > 0$. Using the existing bounds for $||L_t^\alpha||$, $||L_t^\beta||$, $|R_t^\alpha||$, and $||R_t^\beta||$ in Lemma 2.3 and Lemma 2.6, for the second term we have that

$$||\Delta_t^{\alpha,2}|| \leq K \int_0^t \bar{\Psi}^{(1)}_{s,t}\gamma_t^1 \left[\gamma_t^1 \log \int_0^t \gamma_s^1 \mathrm{d}s + \gamma_t^2 \log \int_0^t \gamma_s^2 \mathrm{d}s + \frac{(\gamma_t^1)^2}{(\gamma_t^2)^2}(w_t^1)^2 + t^{-2s}\right] \tag{2.301}$$

$$\leq K \int_0^t \bar{\Psi}^{(1)}_{s,t}\gamma_t^1 \left[\gamma_t^2 \log \int_0^t \gamma_s^2 \mathrm{d}s + \frac{(\gamma_t^1)^2}{(\gamma_t^2)^2}(w_t^1)^2\right] \tag{2.302}$$

Now, since $\{w_t^1\}_{t\geq 0}$ satisfies Condition A_1, so too does $\frac{(\gamma_t^1)^2}{(\gamma_t^2)^2}(w_t^1)^2$. It follows from Lemma 2.8 that, for sufficiently large t, there exists $K > 0$ such that

$$||\Delta_t^{\alpha,2}|| \leq K \left[\gamma_t^2 \log \int_0^t \gamma_s^2 \mathrm{d}s + \frac{(\gamma_t^1)^2}{(\gamma_t^2)^2}(w_t^1)^2 + (\gamma_t^1)^{\frac{1}{2}+\delta} + t^{-\mu\gamma_0^1}\mathbb{1}_{\{\eta_1=1\}}\right]. \tag{2.303}$$

for all $\mu \in (0, \bar{\mu}_1)$. By definition, we have that $\bar{\mu}_1 \in (\frac{\mathbb{1}_{\eta_1=1}}{2\gamma_0^1}, \tilde{\mu}_1 - \tilde{K}_1\tilde{K}_2)$. Thus, in particular, there exists $\mu \in (0, \bar{\mu}_1)$ such that $\mu > \frac{1}{2\gamma_0^1}$, or equivalently, $\mu\gamma_0^1 > \frac{1}{2} = \frac{1}{2} + \delta$, for some $\delta > 0$. Noting also that $\gamma_t^2 \log \int_0^t \gamma_s^2 \mathrm{d}s \leq (\gamma_t^1)^{\frac{1}{2}+\delta}$ for sufficiently large t, it follows that

$$||\Delta_t^{\alpha,2}|| \leq K \left[\frac{(\gamma_t^1)^2}{(\gamma_t^2)^2}(w_t^1)^2 + (\gamma_t^1)^{\frac{1}{2}+\delta}\right]. \tag{2.304}$$

We now turn our attention to $||\Delta_t^{\alpha,3}||$. For this term, integrating by parts yields

$$||\Delta_t^{\alpha,3}|| \leq \left[\bar{\Psi}^{(1)}_{s,t}\frac{\gamma_s^1}{\gamma_s^2}||\Delta_s^\beta||\right]_{s=0}^{s=t} + \int_0^t \left|\left|\frac{\mathrm{d}}{\mathrm{d}s}\left[\bar{\Psi}^{(1)}_{s,t}\frac{\gamma_s^1}{\gamma_s^2}\right]\right|\right| ||\Delta_s^\beta|| \mathrm{d}s \tag{2.305}$$

$$\leq \frac{\gamma_t^1}{\gamma_t^2}||\Delta_t^\beta|| + \bar{\Psi}_{0,t}\frac{\gamma_0^1}{\gamma_0^2}||\Delta_0^1|| + \int_0^t \left|\left|\frac{\mathrm{d}}{\mathrm{d}s}\left[\bar{\Psi}_{s,t}^{(1)}\frac{\gamma_s^1}{\gamma_s^2}\right]\right|\right| ||\Delta_s^\beta||\mathrm{d}s \tag{2.306}$$

$$\leq K\left[\frac{\gamma_t^1}{\gamma_t^2}w_t^2 + (\gamma_t^1)^{\frac{1}{2}+\delta}\right] + \int_0^t \left|\left|\frac{\mathrm{d}}{\mathrm{d}s}\left[\bar{\Psi}_{s,t}^{(1)}\frac{\gamma_s^1}{\gamma_s^2}\right]\right|\right| ||\Delta_s^\beta||\mathrm{d}s \tag{2.307}$$

where in the final line we have used the assumed bound for $||\Delta_t^\beta||$ in the first term, and a similar argument to that used in (2.300) for the second term. For the third term, observe

$$\left|\left|\frac{\mathrm{d}}{\mathrm{d}s}\left[\bar{\Psi}_{s,t}^{(1)}\frac{\gamma_s^1}{\gamma_s^2}\right]\right|\right| \leq K\left[\bar{\Psi}_{s,t}^{(1)}\gamma_s^1\frac{\gamma_s^1}{\gamma_s^2} + \bar{\Psi}_{s,t}^{(1)}\left[\dot{\frac{\gamma_s^1}{\gamma_s^2}}\right]\right] \leq K\left[\bar{\Psi}_{s,t}^{(1)}\gamma_s^1\frac{\gamma_s^1}{\gamma_s^2}\right] \tag{2.308}$$

where the final inequality follows from Assumption 2.1.1' after some straightforward calculations. It follows, again using the assumed bound for $||\Delta_t^\beta||$, and also now Lemma 2.8, that

$$\int_0^t \left|\left|\frac{\mathrm{d}}{\mathrm{d}s}\left[\bar{\Psi}_{s,t}^{(1)}\frac{\gamma_s^1}{\gamma_s^2}\right]\right|\right| ||\Delta_s^\beta||\mathrm{d}s \leq K\int_0^t \bar{\Psi}_{s,t}^{(1)}\gamma_s^1\frac{\gamma_s^1}{\gamma_s^2}w_s^2\mathrm{d}s \tag{2.309}$$

$$\leq K\left[\frac{\gamma_t^1}{\gamma_t^2}w_t^2 + t^{-\mu\gamma_0^1}\mathbb{1}_{\{\eta_1=1\}}\right] \tag{2.310}$$

$$\leq K\left[\frac{\gamma_t^1}{\gamma_t^2}w_t^2 + (\gamma_t^1)^{\frac{1}{2}+\delta}\right] \tag{2.311}$$

where, as previously, the second line holds for all $\mu \in (0, \bar{\mu}_1)$, and the final line follows by arguing as we did above to obtain (2.304). Combining this with (2.307), we have that

$$||\Delta_t^{\alpha,3}|| \leq K\left[\frac{\gamma_t^1}{\gamma_t^2}w_t^2 + (\gamma_t^1)^{\frac{1}{2}+\delta}\right] \tag{2.312}$$

Finally, putting everything together, we have that

$$||\Delta_t^\alpha|| \leq K\left[\frac{(\gamma_t^1)^2}{(\gamma_t^2)^2}(w_t^1)^2 + \frac{\gamma_t^1}{\gamma_t^2}w_t^2 + (\gamma_t^1)^{\frac{1}{2}+\delta}\right] \tag{2.313}$$

It remains to handle $||\Delta_t^\beta||$. Let $\tilde{\Phi}_{s,t}^{(2)} = \exp[-\tilde{\mu}_2\int_s^t \gamma_u^2\mathrm{d}u]$. Then, returning to (2.294), classical computations yield

$$||\Delta_t^\beta|| \leq \tilde{\Phi}_{0,t}^{(2)}||\Delta_0^\beta|| + \tilde{K}_2\int_0^t \tilde{\Phi}_{s,t}^{(2)}\gamma_s^2\left[||L_s^\alpha||^2 + ||L_s^\beta||^2 + ||R_s^\alpha||^2 + ||R_s^\beta||^2 + ||\Delta_s^\alpha||\right]\mathrm{d}s$$

Arguing similarly to above, this time using Lemma 2.9, Lemma 2.11, and the bound just obtained for $||\Delta_t^\alpha||$, it follows straightforwardly that, for sufficiently large t, there exists

$K > 0$ such that

$$||\Delta_t^\beta|| \leq K\left[\frac{(\gamma_t^1)^2}{(\gamma_t^2)^2}(w_t^1)^2 + \frac{\gamma_t^1}{\gamma_t^2}w_t^2 + (\gamma_t^1)^{\frac{1}{2}+\delta}\right]. \tag{2.314}$$

□

2.A.0.3 Technical Lemmas

Lemma 2.8. *Let $\{w_t\}_{t\geq 0}$ be a sequence of numbers satisfying Condition (A_1). Let $\{v_t\}_{t\geq 0}$ be $\mathbb{R}^{d_2}$-valued random sequence such that $||v_t|| = O(w_t)$. Then, for sufficiently large t, we have*

(i) For all $\mu \in (0, \mu_1)$,

$$\int_0^t \Psi_{s,t}^{(1)}\gamma_s^1 w_s \mathrm{d}s \leq K\left(w_t + t^{-\mu\gamma_0^1}\mathbb{1}_{\{\eta_1=1\}}\right). \tag{2.315}$$

(ii) For all $\mu \in (0, \Lambda_H)$,

$$\left|\left|\int_0^t \Phi_{s,t}^{(1)}\gamma_s^1 v_s \mathrm{d}s\right|\right| \leq K\left(w_t + t^{-\mu\gamma_0^1}\mathbb{1}_{\{\eta_1=1\}}\right). \tag{2.316}$$

Proof. We will begin by proving (i). Let us consider the case $\eta_1 = 1$. That is, $\gamma_t^1 = \gamma_0^1(\delta_0 + t)^{-1}$. In this case, we have

$$\Psi_{s,t}^{(1)} = \exp\left[-\mu_1\int_s^t \gamma_u^1 \mathrm{d}u\right] = \exp\left[-\mu_1\gamma_0^1\ln\left(\frac{\delta_0+t}{\delta_0+s}\right)\right] = \left(\frac{\delta_0+t}{\delta_0+s}\right)^{-\mu_1\gamma_0^1}. \tag{2.317}$$

We thus have, allowing the value of the constant K to increase from line to line, that

$$\int_0^t \Psi_{s,t}^{(1)}\gamma_s^1 w_s \mathrm{d}s = \gamma_0^1(\delta_0+t)^{-\mu_1\gamma_0^1}\int_0^t(\delta_0+s)^{\mu_1\gamma_0^1-1}w_s \mathrm{d}s \tag{2.318}$$

$$= \gamma_0^1(\delta_0+t)^{-\mu_1\gamma_0^1}\int_0^t(\delta_0+s)^{\mu_1\gamma_0^1-1-\omega}\mathcal{L}(s)\mathrm{d}s \tag{2.319}$$

$$\leq K(\delta_0+t)^{-\mu_1\gamma_0^1}\mathcal{L}(t)\left[(\delta_0+t)^{\mu_1\gamma_0^1-\omega} + \log(\delta_0+t)\right] \tag{2.320}$$

$$\leq K\left[(\delta_0+t)^{-\omega}\mathcal{L}(t) + (\delta_0+t)^{-\mu_1\gamma_0^1}\mathcal{L}(t)\log(\delta_0+t)\right] \tag{2.321}$$

$$\leq K\left[t^{-\omega}\mathcal{L}(t) + t^{-\mu_1\gamma_0^1}\mathcal{L}(t)\log(t)\right] = K\left[w_t + t^{-\mu_1\gamma_0^1}\mathcal{L}(t)\log(t)\right] \tag{2.322}$$

where in (2.320) we have made use of the Karamata's integral theorem for slowly varying functions (e.g., [51]), and in (2.322) we have used some elementary inequalities and the definition of $(w_t)_{t\geq 0}$. Since $\mathcal{L}(t)$ is a slowly varying function, it follows that, for all $\mu \in (0,\mu_1)$, and for sufficiently large t, we have (see also [354, Lemma 9])

$$\int_0^t \Psi_{s,t}^{(1)}\gamma_s^1 w_s \mathrm{d}s \leq K\left[w_t + t^{-\mu\gamma_0^1}\right]. \tag{2.323}$$

We now turn our attention to the case $\eta_1 < 1$. For sufficiently large t_0, we have, using Condition (A_1), that

$$\begin{aligned}
\frac{1}{w_t}\int_0^t \Psi_{s,t}^{(1)}\gamma_s^i w_s \mathrm{d}s &= \int_0^{t_0} \Psi_{s,t}^{(1)}\gamma_s^1 \frac{w_s}{w_t}\mathrm{d}s + \int_{t_0}^t \Psi_{s,t}^{(1)}\gamma_s^1 \frac{w_s}{w_t}\mathrm{d}s &(2.324)\\
&\leq K + \int_{t_0}^t \exp\left[-\mu_1\int_s^t \gamma_u^1 \mathrm{d}u\right]\gamma_s^1 \exp\left[\frac{\mu_1}{2}\int_s^t \gamma_u^1\mathrm{d}u\right]\mathrm{d}s &(2.325)\\
&\leq K + \int_{t_0}^t \exp\left[-\frac{\mu_1}{2}\int_s^t \gamma_u^1\mathrm{d}u\right]\gamma_s^1 \mathrm{d}s \leq K. &(2.326)
\end{aligned}$$

It remains to prove (ii). In fact, this now follows straightforwardly as a consequence of (i). In particular, using the stability of the matrix A, we have that

$$\left|\left|\int_0^t \Phi_{s,t}^{(1)}\gamma_s^1 v_s \mathrm{d}s\right|\right| \leq \int_0^t ||\Phi_{s,t}^{(1)}||\gamma_s^1||v_s||\mathrm{d}s \leq \int_0^t \Psi_{s,t}^{(1)}\gamma_s^1 w_s\mathrm{d}s \tag{2.327}$$

where the final inequality holds for any $\mu_1 \in (0,\Lambda_H)$. The result now follows straightforwardly from (i). □

Lemma 2.9. *Let $\{w_t\}_{t\geq 0}$ be a sequence of numbers satisfying Condition (A_2). Let $\{v_t\}_{t\geq 0}$ be $\mathbb{R}^{d_2}$-valued random sequence such that $||v_t|| = O(w_t)$. Then, for sufficiently large t,*

$$\left|\int_0^t \Psi_{s,t}^{(2)}\gamma_s^2 w_s \mathrm{d}s\right| \leq K w_t. \tag{2.328}$$

and

$$\left|\left|\int_0^t \Phi_{s,t}^{(2)}\gamma_s^2 w_s \mathrm{d}s\right|\right| \leq K w_t. \tag{2.329}$$

Proof. The proof follows the proof of Lemma 2.8 in the case $\eta_1 < 1$. □

Lemma 2.10. *Suppose γ_t^1 satisfies Condition (A_1). Then, for sufficiently large t,*

$$\Psi_{0,t}^{(1)}(\gamma_t^1)^{-\frac{1}{2}} \to 0. \tag{2.330}$$

Proof. The proof proceeds in much the same way as the proof of Lemma 2.8. First consider the case $\eta_1 = 1$. In this case, it is straightforward to compute

$$\Psi_{0,t}^{(1)} = \exp\left[-\mu_1 \int_0^t \gamma_u^1 \mathrm{d}u\right] = \exp\left[-\mu_1\gamma_0^1 \ln\left(\frac{\delta_0 + t}{\delta_0}\right)\right] \leq K\left(\delta_0 + t\right)^{-\mu_1\gamma_0^1}. \tag{2.331}$$

It follows, on account of Assumption 2.1.1', that

$$\Psi_{0,t}^{(1)}(\gamma_t^1)^{-\frac{1}{2}} \leq K\left(\delta_0 + t\right)^{\frac{1}{2}-\mu_1\gamma_0^1} \to 0. \tag{2.332}$$

Now consider the case $\eta_1 < 1$. In this case, we have

$$(\gamma_t^1)^{-\frac{1}{2}} \leq (\gamma_t^1)^{-1} \leq \exp\left[o(1)\int_0^t \gamma_u^1 \mathrm{d}u\right]. \tag{2.333}$$

It follows straightforwardly that, for sufficiently large t,

$$\Psi_{0,t}^{(1)}(\gamma_t^1)^{-\frac{1}{2}} \leq \exp\left[-\mu_1\int_0^t \gamma_u^1 \mathrm{d}u\right] \exp\left[\frac{\mu_1}{2}\int_0^t \gamma_u^1 \mathrm{d}u\right] = \exp\left[-\frac{\mu_1}{2}\int_0^t \gamma_u^1 \mathrm{d}u\right] \to 0. \tag{2.334}$$

□

Lemma 2.11. *Suppose γ_t^2 satisfies Condition (A_2). Then, for sufficiently large t,*

$$\Psi_{0,t}^{(2)}(\gamma_t^2)^{-\frac{1}{2}} \to 0. \tag{2.335}$$

Proof. The proof follows the proof of Lemma 2.10 in the case $\eta_1 < 1$. □

3

An Application of Two-Timescale Stochastic Gradient Descent in Continuous Time

Summary. *In this chapter, we analyse the problem of joint online parameter estimation and optimal sensor placement for a partially observed, finite-dimensional, non-linear diffusion process. We demonstrate in detail how this problem can be formulated as a bilevel optimisation problem, and solved using a continuous-time, two-timescale, stochastic gradient descent algorithm. Under suitable conditions on the latent signal, the filter, and the filter derivatives, we establish almost sure convergence of the online parameter estimates and optimal sensor placements to the stationary points of the asymptotic log-likelihood and asymptotic filter covariance, respectively. In addition, we provide a simple numerical example, illustrating the application of the proposed methodology to a partially observed Beneš equation.*

3.1 Introduction

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with filtration $\{\mathcal{F}_t\}_{t\geq 0}$, which satisfies the usual conditions. In this chapter, we consider a problem arising in the following family of partially observed finite dimensional diffusion processes[1] under the probability measure $\mathbb{P}_\theta$,

$$\mathrm{d}x_t = A(\theta, x_t)\mathrm{d}t + B(\theta, x_t)\mathrm{d}v_t \ , \quad x(0) = x_0, \tag{3.1}$$

$$\mathrm{d}y_t = C(\theta, \boldsymbol{o}, x_t)\mathrm{d}t + \mathrm{d}w_t \ , \qquad y(0) = 0, \tag{3.2}$$

where $\{x_t\}_{t\geq 0}$ denotes a hidden $\mathbb{R}^{n_x}$-valued signal process, $\{y_t\}_{t\geq 0}$ denotes a $\mathbb{R}^{n_y}$-valued observation process, and $\{v_t\}_{t\geq 0}$, $\{w_t\}_{t\geq 0}$ are independent $\mathbb{R}^{n_x}$-, $\mathbb{R}^{n_y}$-valued Wiener processes, with incremental covariances $\mathcal{Q}(\theta) \in \mathbb{R}^{n_x\times n_x}$ and $\mathcal{R}(\boldsymbol{o}) \in \mathbb{R}^{n_y\times n_y}$, which correspond to signal noise and measurement noise, respectively. Meanwhile, $\theta \in \Theta \subseteq \mathbb{R}^{n_\theta}$ is an n_θ-dimensional parameter, and $\boldsymbol{o} = \{\boldsymbol{o}_i\}_{i=1}^{n_y} \in \Omega^{n_y} \subseteq \mathbb{R}^{n_y n_{\boldsymbol{o}}}$ is a set of n_y sensor locations, with $\boldsymbol{o}_i \in \Omega \subseteq \mathbb{R}^{n_{\boldsymbol{o}}}$ for $i = 1, \ldots, n_y$. We assume that, for all $(\theta, \boldsymbol{o}) \in \Theta \times \Omega^{n_y}$, the initial conditions $x_0 \sim p_0(\theta, \boldsymbol{o})$ are independent of $\{w_t\}_{t\geq 0}$ and $\{v_t\}_{t\geq 0}$. We also suppose that $A(\theta, \cdot)$, $B(\theta, \cdot)$, and $C(\theta, \boldsymbol{o}, \cdot)$ are measurable functions which ensure the existence and uniqueness of strong solutions to these equations for all $t \geq 0$ (e.g., [19]).

The central problem underlying this partially observed stochastic dynamical system is that of optimal state estimation, or filtering. This consists in determining the conditional probability distribution of the latent signal process (i.e., the filter), given the history of observations, under the assumption that any model parameters are known, and the locations of the measurement sensors are fixed (e.g., [19]). In practical applications, however, it is often the case that the parameters of this model are unknown, and must be inferred from the data. Indeed, inferring the model parameters is often the primary problem of interest (e.g., [261, 353]).

It is often also the case that the locations of the measurement sensors are not fixed, and thus it may be possible to improve upon the optimal state estimate by determining an 'optimal sensor placement'. Alternatively, one may have access to a large number of measurement sensors, but it may only be possible to utilise a small subset of these at any given time instant (e.g., due to communication constraints). In this case, it is of interest to obtain an 'optimal sensor selection' or 'optimal sensor schedule'.

The first of these two scenarios is particularly relevant to applications in engineering and the applied sciences, including meteorology, environmental monitoring, and fluid dynamics. In such applications, the process of interest, even if defined continuously over space

[1] We will sometimes also refer to such processes as finite-dimensional state space models, or lumped parameter systems.

and in time, can only be measured at a finite number of spatial locations. Moreover, the spatial density of observations is generally very low, due either to prohibitive expense (i.e., the sensors are expensive, or expensive to place), or geographical inaccessibility (i.e., the sensors cannot be placed in particular locations). Furthermore, measurements at certain points in the domain may yield more information about the system than measurements at other points, due to correlations in the signal. Thus, to a greater or lesser extent, the accuracy of the estimate of the signal is dependent on location of the measurement sensors.

In this chapter, we address, for the first time, the problems of parameter estimation and optimal sensor placement (for the purpose of optimal state estimation) together. This represents a significant departure from the existing literature, in which these two problems have, until now, been studied separately. Before we provide further details on our approach, let us briefly review the existing literature on these two important problems.

3.1.1 Literature Review

3.1.1.1 Parameter Estimation

The problem of parameter estimation for partially observed stochastic processes in continuous time has been somewhat well studied, particularly in the offline setting (e.g., [14, 132, 148, 275]). This being said, it remains the case that the majority of literature on this subject has been formulated for discrete-time systems (e.g., [88, 238] for an overview).[2] Among the different methods that have been considered for this problem, those based on direct maximisation of the likelihood are arguably the most ubiquitous (e.g., [259, 275, 430]), although other approaches based on the maximum likelihood principle such as expectation-maximisation (EM) have also been considered (e.g., [83, 148]).

In the offline setting, maximum likelihood (ML) methods seek the value of θ that maximises the log-likelihood of the observations, or incomplete data log-likelihood, after some fixed time-period T. That is,

$$\hat{\theta}_T = \operatorname*{arg\,max}_{\theta \in \Theta} \mathcal{L}_T(\theta, \boldsymbol{o}). \tag{3.3}$$

The asymptotic properties of this estimator, including asymptotic consistency, asymptotic efficiency, and asymptotic normality, have been the subject of several papers (see, e.g., the survey paper [275] and references therein). The results in these papers only hold, however, under somewhat restrictive assumptions on the model dynamics: namely, linearity [14, 15, 16, 17, 22, 196, 451, 452], small noise [231, 253, 274], or both [277, 278].[3]

[2] We also note that there is a significant body of work on parameter estimation for fully observed stochastic processes in continuous time (e.g., [52, 63, 273, 296]).

[3] We should also mention [83, 132, 148, 293], which consider the offline parameter estimation problem for general non-linear, partially observed, continuous-time diffusion processes, but do not provide any asymptotic results.

We are primarily concerned with online parameter estimation methods, which recursively estimate the unknown model parameters based on the continuous stream of observations. In comparison to classical methods, which process the observed data in a batch fashion, online methods perform inference in real time, can track changes in parameters over time, are more computationally efficient, and have significantly smaller storage requirements.

Various methods have been proposed for online parameter estimation in partially observed state space models, although these methods are typically formulated in the discrete-time setting. These include an online variant of the EM algorithm (e.g., [10, 86, 87, 119, 172, 180, 289] in discrete time and [96, 97, 148, 173, 487] in continuous time), state augmentation algorithms (e.g., [179, 305, 428] in discrete time), and recursive minimum prediction error schemes (e.g., [115, 180, 181, 295] in discrete time).[4] In this chapter, we will focus exclusively on recursive maximum likelihood (RML) methods, which use stochastic gradient descent to recursively seek the value of θ which maximises an asymptotic log-likelihood function (e.g., [196, 197, 430])

$$\tilde{\mathcal{L}}(\theta, \boldsymbol{o}) = \lim_{t\to\infty} \frac{1}{t}\mathcal{L}_t(\theta, \boldsymbol{o}). \tag{3.4}$$

The asymptotic properties of this method for partially observed, discrete-time systems (e.g., [116, 165, 258, 294, 295, 382, 409, 438, 441]), and for fully-observed, continuous-time systems (e.g., [52, 63, 273, 296]), have been studied extensively. In comparison, the partially observed, continuous-time case has received relatively little attention. The use of a continuous-time RML method for online parameter estimation in a partially-observed linear diffusion process was first proposed in [196], and later extended in [197]. In this paper, a recursive maximum likelihood estimator for the parameters of a partially-observed, linear diffusion process was derived using the Itô-Venzel formula (e.g., [457]), and an a.s. convergence result for this estimator was provided without proof. In particular, it was established that the estimator $\hat{\theta}_t \to \theta^*$ a.s. on the event $\Omega = \{\hat{\theta}_t < \infty \text{ for all } t \geq t_0\}$, for some fixed, non-random initial time t_0. This analysis was later extended in [197], which established the a.s. convergence of a modified version of the estimator in [196], which included an additional resetting mechanism.[5]

The use of a continuous-time RML method for online parameter estimation was more recently revisited in [430]. In this paper, the authors derived a RML estimator for the parameters of a general, non-linear partially observed diffusion process, and established the a.s. convergence of this estimator under appropriate conditions on the process consisting of the latent state, the filter, and the filter derivative. This paper extended the results in

[4]For a more comprehensive overview of online parameter estimation methods in *discrete-time* partially observed state space models, the reader is referred to [88, 238, 244].

[5]We remark that the evolution equations for the estimators considered in [196, 197] include an additional second order term compared to the estimator analysed later in this chapter. This arises when the Itô-Venzel formula is applied to the score function.

[420] to the partially-observed setting. We should remark that the use of a continuous-time RML method for non-linear partially observed diffusion processes was also considered in [311, 358]. In these papers, however, in addition to the model parameters, the hidden state was estimated via maximum likelihood, rather than the usual filtering paradigm.

3.1.1.2 Optimal Sensor Placement

The problem of optimal sensor placement (and optimal sensor selection) for optimal state estimation in partially observed finite-dimensional diffusion processes has been studied by a very large number of authors, and in a wide variety of contexts.[6] Arguably the first mathematically rigorous treatment of this problem for linear systems was provided by Athans [11], who formulated it as an application of optimal control on the Ricatti equation governing the covariance of the optimal filter (see also [211, 256, 279, 337, 349, 350]).

In this framework, the sensor placement $\boldsymbol{o}_T$ is treated as a control variable, and the optimal sensor placement $\hat{\boldsymbol{o}}_T$ is obtained as the minima of a suitable objective function, that is,

$$\hat{\boldsymbol{o}}_T = \arg\min_{\boldsymbol{o}\in\mathcal{O}} \mathcal{J}_T(\theta, \boldsymbol{o}). \tag{3.5}$$

This objective function is typically defined as the trace of the filter covariance at some finite time (e.g., [114]), or the integral of the trace of the filter covariance over some finite time interval (e.g., [109]), which are designed to minimise the uncertainty in the state estimate. One can also consider optimal sensor placement with respect to asymptotic versions of these functions (e.g [3, 9, 386, 485]) in which case the optimal sensor placements are obtained, possibly recursively, as the minima of an asymptotic version of the objective function

$$\tilde{\mathcal{J}}(\theta, \boldsymbol{o}) = \frac{1}{t}\lim_{t\to\infty} \mathcal{J}_t(\theta, \boldsymbol{o}). \tag{3.6}$$

In the case of optimal sensor placement, the design variable $\boldsymbol{o}$ is continuous, denoting the location of the measurement sensors. One can thus optimise the objective function directly, and at relatively low computational cost, by using gradient based methods (e.g., [2, 3, 9, 114, 151, 154, 155]).

3.1.2 Contributions

In this chapter, we present a principled method for performing joint online parameter estimation and optimal sensor placement in a partially observed, possibly non-linear diffusion process. We show how to formulate this as a bilevel optimisation problem, in which the

[6] While we do not consider it here, the problem of optimal sensor placement for the purpose of parameter estimation, rather than optimal state estimation, has also been studied extensively (see, e.g., [373, 453] and references therein).

objective is to obtain estimates $\hat{\theta} \in \Theta$ and $\hat{\boldsymbol{o}}(\hat{\theta}) \in \Pi^{n_y}$ which simultaneously maximise the log-likelihood of the observations and minimise an appropriately chosen sensor placement objective function. That is,

$$\hat{\theta} = \underset{\theta \in \Theta}{\arg\max}\, \tilde{\mathcal{L}}\big(\theta, \hat{\boldsymbol{o}}(\theta)\big) \quad , \quad \hat{\boldsymbol{o}}(\theta) = \underset{\boldsymbol{o} \in \Pi^{n_y}}{\arg\min}\, \tilde{\mathcal{J}}\big(\theta, \boldsymbol{o}\big). \tag{3.7}$$

On the basis of the theoretical results established in Chapter 2, we propose a solution to this bilevel optimisation problem in the form of a two-timescale, stochastic gradient descent algorithm in continuous time. Moreover, under reasonable conditions on the process consisting of the latent signal process, the filter, and the filter derivatives, we establish a.s. convergence of the online parameter estimates and recursive optimal sensor placements generated by this algorithm to the stationary points of the asymptotic log-likelihood and the asymptotic filter covariance, respectively. The effectiveness of this algorithm is demonstrated via a one-dimensional, partially observed stochastic differential equation (SDE) of Beneš class.

3.1.3 Chapter Organisation

The remainder of this chapter is organised as follows. In Section 3.2, we demonstrate rigorously how to formulate the problem of joint online parameter estimation and optimal sensor placement as an unconstrained bilevel optimisation problem; and propose a solution to this problem in the form of a continuous-time, two-timescale stochastic gradient descent algorithm. In Section 3.3, we prove the a.s. convergence of this algorithm. In Section 3.4, we provide a numerical example illustrating the performance of the proposed algorithm. Finally, in Section 3.5, we offer some concluding remarks.

3.2 Main Results

3.2.1 Parameter Estimation

We first review the problem of parameter estimation. We will suppose that the model generates the observation process $\{y_t\}_{t\geq 0}$ according to a true, but unknown, static parameter θ^*. The objective is then to obtain an estimator $\{\theta_t\}_{t\geq 0}$ of θ^* which is both $\mathcal{F}_t^Y$-measurable and recursively computable. That is, an estimator which can be computed online using the continuous stream of observations, without revisiting the past. In this subsection, we will assume that the sensor locations $\boldsymbol{o} \in \Omega^{n_y}$ are fixed. We will, however, make explicit the dependence of functions on $\boldsymbol{o}$, where appropriate.

One such estimator can be obtained as a modification of the classical offline maximum likelihood estimator (e.g., [358, 430]). We thus recall the expression for the log-likelihood of the observations, or incomplete data log-likelihood, for a partially observed diffusion process (e.g., [22, 184, 196, 430, 452]), namely

$$\mathcal{L}_t(\theta,\boldsymbol{o}) = \int_0^t R^{-1}(\boldsymbol{o})\hat{C}_s(\theta,\boldsymbol{o})\cdot \mathrm{d}y_s - \frac{1}{2}\int_0^t ||\mathcal{R}^{-\frac{1}{2}}(\boldsymbol{o})\hat{C}_s(\theta,\boldsymbol{o})||^2 \mathrm{d}s, \tag{3.8}$$

where $\hat{C}_s(\theta,\boldsymbol{o})$ denotes the conditional expectation of $C(\theta,\boldsymbol{o},x_s)$, given the observation sigma-algebra $\mathcal{F}_s^Y$, viz

$$\hat{C}_s(\theta,\boldsymbol{o}) = \mathbb{E}_{\theta,\boldsymbol{o}}\left[C(\theta,\boldsymbol{o},x_s)|\mathcal{F}_s^Y\right]. \tag{3.9}$$

3.2.1.1 Offline Parameter Estimation

In the offline setting, one seeks to obtain the value of θ that maximises the incomplete data log-likelihood after some fixed time-period, say T. In particular, the maximum likelihood estimator (MLE) is defined as

$$\hat{\theta}_T = \arg\max_{\theta\in\Theta} \mathcal{L}_T(\theta,\boldsymbol{o}). \tag{3.10}$$

In practice, various methods can be used to solve this optimisation problem. Perhaps the most popular of these is gradient ascent, which, initialised at $\theta_0 \in \Theta$, generates a sequence of parameter estimates $\{\theta_k\}_{k\geq 1}$ via the recursion

$$\theta_{k+1} = \theta_k + \gamma_k \underbrace{\int_0^T R^{-1}(\boldsymbol{o})[\mathrm{d}y_s - \hat{C}_s(\theta,\boldsymbol{o})\mathrm{d}s]^T \nabla_\theta[\hat{C}_s(\theta,\boldsymbol{o})]}_{\nabla_\theta \mathcal{L}_T(\theta,\boldsymbol{o})}\Bigg|_{\theta=\theta_k}, \tag{3.11}$$

where $\{\gamma_k\}_{k\geq 1}$ is a non-negative, non-increasing sequence of step-sizes.[7] Clearly, at each iteration of this algorithm, the derivative of the log-likelihood function must be recomputed using the current values of the parameters.

[7] Another popular approach, which we will not consider here, is the Expectation-Maximisation (EM) algorithm (e.g., [83, 148]).

3.2.1.2 Online Parameter Estimation

In the online setting, a standard approach is to recursively seek the value of θ which maximises the asymptotic log-likelihood, viz

$$\tilde{\mathcal{L}}(\theta,\boldsymbol{o}) = \lim_{t\rightarrow\infty}\frac{1}{t}\mathcal{L}_t(\theta,\boldsymbol{o}) = \lim_{t\rightarrow\infty}\bigg[\frac{1}{t}\int_0^t R^{-1}(\boldsymbol{o})\hat{C}_s(\theta,\boldsymbol{o})\cdot \mathrm{d}y_s \tag{3.12}$$

$$-\frac{1}{2t}\int_0^t R^{-1}(\boldsymbol{o})||\hat{C}_s(\theta,\boldsymbol{o})||^2\mathrm{d}s\bigg]. \tag{3.13}$$

Typically, neither the asymptotic log-likelihood, nor its gradient, are available in analytic form. It is, however, possible to compute noisy estimates of these quantities at any finite time, using the integrand of the log-likelihood and the integrand of its gradient, respectively. This optimisation problem can thus be tackled using continuous-time stochastic gradient ascent, whereby the parameters follow a noisy ascent direction given by the integrand of the gradient of the log-likelihood, evaluated with the current parameter estimate. In particular, initialised at $\theta_0\in\Theta$, the parameter estimates $\{\theta_t\}_{t\geq 0}$ are generated according to the SDE [430]

$$\mathrm{d}\theta_t = \begin{cases} \gamma_t\big[\hat{C}_t^{\theta}(\theta_t,\boldsymbol{o})\big]^T\mathcal{R}^{-1}(\boldsymbol{o})\big[\mathrm{d}y_t - \hat{C}_t(\theta_t,\boldsymbol{o})\mathrm{d}t\big] & , \quad \theta_t\in\Theta, \\ 0 & , \quad \theta_t\notin\Theta, \end{cases} \tag{3.14}$$

where $\gamma:\mathbb{R}_+\rightarrow\mathbb{R}_+$ is a non-negative, non-increasing, continuous function (i.e., the learning rate), and where we have written $\hat{C}_t^{\theta}(\theta,\boldsymbol{o}) = \nabla_\theta\hat{C}_t(\theta,\boldsymbol{o})$ to denote the gradient of $\hat{C}_t(\theta,\boldsymbol{o})$ with respect to the parameter vector.[8] Following [430], this algorithm includes a projection device which ensures that the parameter estimates $\{\theta_t\}_{t\geq0}$ remain in $\Theta\subset\mathbb{R}^{n_\theta}$ with probability one. This is common for algorithms of this type (e.g., [107, 313]).[9] In the literature on statistical inference and system identification, this algorithm is commonly referred to as recursive maximum likelihood (RML).

3.2.2 Optimal Sensor Placement

We now turn our attention to the problem of optimal sensor placement. We will suppose that the observation process $\{y_t\}_{t\geq0}$ is generated using a finite set of n_y sensors. Our objective is to obtain an estimator of the set of n_y sensor locations $\boldsymbol{o}^* = \{\boldsymbol{o}_i\}_{i=1}^{n_y}$ which are optimal with respect to some pre-determined criteria, possibly subject to constraints.

[8] We use the convention that the gradient operator adds a covariant dimension to the tensor field upon which it acts. Thus, for example, since $\hat{C}_t(\theta,\boldsymbol{o}) = \mathbb{E}_{\theta,\boldsymbol{o}}[C(\theta,\boldsymbol{o},x_t)|\mathcal{F}_t^Y]$ takes values in $\mathbb{R}^{n_y}$, its gradient $\hat{C}_t^{\theta}(\theta,\boldsymbol{o}) = \nabla_\theta\hat{C}_t(\theta,\boldsymbol{o})$, takes values in $\mathbb{R}^{n_y\times n_\theta}$.

[9] Other, more complex, projection devices are of course possible, but will not be considered here.

Once more, we require our estimator to be $\mathcal{F}_t^Y$-measurable and recursively computable. In this subsection, we will assume that the parameter $\theta \in \Theta$ is fixed. We will, however, make explicit the dependence of functions on θ, where appropriate.

A standard approach to this problem is to define a suitable objective function, say $\mathcal{J}_t(\theta, \cdot) : \Omega^{n_y} \to \mathbb{R}$, and then to define the optimal estimator as

$$\hat{\boldsymbol{o}}_t = \underset{\boldsymbol{o} \in \Omega^{n_y}}{\arg\min}\, \mathcal{J}_t(\theta, \boldsymbol{o}). \tag{3.15}$$

We focus on the objective of optimal state estimation. In this case, following [80, 109, 211, 254], we consider the following objective function

$$\mathcal{J}_t(\theta, \boldsymbol{o}) = \int_0^t \underbrace{\mathrm{Tr}\left[M_s \hat{\Sigma}_s(\theta, \boldsymbol{o})\right]}_{\hat{j}_s(\theta, \boldsymbol{o})} \mathrm{d}s := \int_0^t \hat{j}_s(\theta, \boldsymbol{o}) \mathrm{d}s, \tag{3.16}$$

where $M_s : \mathbb{R}^{d_x \times d_x} \to \mathbb{R}^{d_x \times d_x}$ is a matrix which allows one to weight significant parts of the state estimate, and

$$\begin{aligned} \hat{\Sigma}_s(\theta, \boldsymbol{o}) &= \mathrm{Cov}_{\theta, \boldsymbol{o}}\left[x_s | \mathcal{F}_s^Y\right] &(3.17)\\ &= \mathbb{E}_{\theta, \boldsymbol{o}}\left[x_s x_s^T | \mathcal{F}_s^Y\right] - \mathbb{E}_{\theta, \boldsymbol{o}}\left[x_s | \mathcal{F}_s^Y\right] \mathbb{E}_{\theta, \boldsymbol{o}}\left[x_s | \mathcal{F}_s^Y\right]^T &(3.18) \end{aligned}$$

denotes the conditional covariance of the latent state x_s, given the history of observations $\mathcal{F}_s^Y$. Broadly speaking, the use of this objective corresponds to seeking the sensor placement which minimises the uncertainty in the estimate of the latent state. Other choices for the objective function are, of course, possible. These include, among many others, the trace of the conditional covariance at some finite, terminal time (e.g., [41, 43, 84, 109, 128, 368, 472]), and variants thereof (e.g., [262, 360]), and the trace of the steady-state conditional covariance (e.g., [3, 9, 386, 485, 488]).

3.2.2.1 Offline Optimal Sensor Placement

In the offline setting, the objective is to obtain the optimal sensor placement with respect to state estimation over some fixed-time period, say $[0, T]$. One thus seeks to minimise the value of the objective function at time T. In this instance, the optimal estimator is straightforwardly defined as

$$\hat{\boldsymbol{o}}_T = \underset{\boldsymbol{o} \in \Omega^{n_y}}{\arg\min}\, \mathcal{J}_T(\theta, \boldsymbol{o}). \tag{3.19}$$

This optimisation problem can be tackled via a simple gradient descent scheme, which, initialised at $\boldsymbol{o}_0 \in \Omega^{n_y}$, generates a sequence of sensor placement estimates $\{\boldsymbol{o}_k\}_{k \geq 1}$ via the recursion

$$\boldsymbol{o}_{k+1} = \boldsymbol{o}_k - \gamma_k \underbrace{\int_0^T \nabla_{\boldsymbol{o}} \mathrm{Tr}\left[M_t \hat{\Sigma}_t(\theta, \boldsymbol{o})\right] \mathrm{d}t \bigg|_{\boldsymbol{o}=\boldsymbol{o}_k}}_{\nabla_{\boldsymbol{o}} \mathcal{J}_T(\theta,\boldsymbol{o})}, \tag{3.20}$$

where $\{\gamma_k\}_{k\geq 1}$ is a non-negative, non-increasing sequence of step-sizes.[10] The use of gradient descent for the optimal sensor placement problem is very well established (e.g., [2, 3, 9, 114]). Most recently, a gradient descent scheme for the objective function considered in this paper was proposed in [80], and implemented in a numerical example involving the stochastic convection-diffusion equation, for both stationary and moving sensor networks.

3.2.2.2 Online Optimal Sensor Placement

In the online setting, the objective is to recursively estimate the optimal sensor locations $\hat{\boldsymbol{o}}$ in real time using the continuous stream of observations. In this case, in the spirit of the previous section, one approach is to recursively seek the value of $\boldsymbol{o}$ which minimises the asymptotic objective function (e.g., [488]), namely

$$\tilde{\mathcal{J}}(\theta, \boldsymbol{o}) = \lim_{t\to\infty} \frac{1}{t} \mathcal{J}_t(\theta, \boldsymbol{o}) = \lim_{t\to\infty} \left[\frac{1}{t} \int_0^t \hat{j}_s(\theta, \boldsymbol{o}) \mathrm{d}s\right]. \tag{3.21}$$

Typically, neither the asymptotic objective function, nor its gradient, are available in analytic form.[11] It is, however, possible to compute noisy estimates of these quantities at any finite time, using the integrand of the objective function and its gradient, respectively. Similar to online parameter estimation, this optimisation problem can thus also be tackled using continuous-time stochastic gradient descent, whereby the sensor locations follow a noisy descent direction given by the integrand of the gradient of the objective function, evaluated with the current estimates of the sensor placements. In particular, initialised at $\boldsymbol{o}_0 \in \Omega^{n_y}$, the sensor locations $\{\boldsymbol{o}_t\}_{t\geq 0}$ are generated according to the ordinary differential

[10] Another popular approach, somewhat different in spirit to the approach considered here, first discretises the spatial domain, and then solves the combinatorial optimisation problem of determining the optimal subset of locations at which to place the available sensors (at each time point). This approach, as mentioned in the introduction, is commonly referred to as 'optimal sensor selection' or 'optimal sensor scheduling' (e.g., [30, 178, 291, 362, 363, 473]).

[11] A notable exception to this is the linear Gaussian case, in which case the asymptotic objective function is the solution of the so-called algebraic Ricatti equation, which is independent of the observation process (e.g., [237]). This independence no longer holds, however, when online parameter estimation and optimal sensor placement are coupled (see Section 4.3.3). In this case, the (asymptotic) objective function depends on the parameter estimates via equation (3.21), and the parameter estimates depend on the observations via equation (3.14). Thus, implicitly, the sensor placements estimates do now depend on the observations. We explore this in more detail in Chapter 4.

equation

$$\mathrm{d}\boldsymbol{o}_t = \begin{cases} -\gamma_t \big[\hat{j}_t^{\boldsymbol{o}}(\theta, \boldsymbol{o}_t)\big]^T \mathrm{d}t & , \quad \boldsymbol{o}_t \in \Omega^{n_y} \\ 0 & , \quad \boldsymbol{o}_t \notin \Omega^{n_y} \end{cases} \tag{3.22}$$

where $\{\gamma_t\}_{t\geq 0}$ is a non-negative, non-increasing continuous sequence of real step-sizes, and where $\hat{j}_t^{\boldsymbol{o}}(\theta, \boldsymbol{o}) = \nabla_{\boldsymbol{o}} \hat{j}_t(\theta, \boldsymbol{o}) = \nabla_{\boldsymbol{o}} \mathrm{Tr}[M_t \hat{\Sigma}_t(\theta, \boldsymbol{o})]$ is used to denote the gradient of $\hat{j}_t(\theta, \boldsymbol{o})$ with respect to the sensor locations. Similar to the online parameter estimation algorithm, this recursion includes a projection device to ensure that the sensor placements $\{\boldsymbol{o}_t\}_{t\geq 0}$ remain in $\Omega^{n_y} \subset \mathbb{R}^{n_y n_o}$ with probability one.

3.2.3 The Filter and Its Gradients

In order to implement either of these algorithms, it is necessary to compute the conditional expectations $\hat{C}_t(\theta, \boldsymbol{o})$ and $\hat{j}_t(\theta, \boldsymbol{o})$, as well as their gradients, $\hat{C}_t^{\theta}(\theta, \boldsymbol{o})$ and $\hat{j}_t^{\boldsymbol{o}}(\theta, \boldsymbol{o})$. In principle, this requires one to obtain solutions of the Kushner-Stratonovich equation for arbitrary integrable $\varphi : \mathbb{R}^{n_x} \to \mathbb{R}$, viz (e.g., [19, 266])

$$\mathrm{d}\hat{\varphi}_t = (\hat{\mathcal{A}_x \varphi})_t + \big((\hat{C\varphi})_t - \hat{C}_t \hat{\varphi}_t\big) \cdot \big(\mathrm{d}y_t - \hat{C}_t \mathrm{d}t\big), \tag{3.23}$$

where $\hat{\varphi}_t = \mathbb{E}[\varphi(x_t)|\mathcal{F}_t^Y]$ denotes the conditional expectation of $\varphi(x_t)$ given the history of observations $\mathcal{F}_t^Y$, and $\mathcal{A}_x$ denotes the infinitesimal generator of the latent signal process. In general, exact solutions to the Kushner-Stratonovich equation are very rarely available [297, 332, 365, 366]. In order to make any progress, we must therefore introduce the following additional assumption.

Assumption 3.2.1. *The Kushner-Stratonovich equation admits a finite dimensional recursive solution, or a finite-dimensional recursive approximation.*

There are a small but important class of filters for which finite-dimensional recursive solutions do exist, namely, the Kalman-Bucy filter [176, 237], the Beneš filter [39, 40] and extensions thereof [98, 99, 136, 137, 173, 208, 364]. In addition, there are a much larger class of processes for which finite-dimensional recursive approximations are available, and thus, crucially, for which the proposed algorithm can still be applied. Standard approximation schemes include, among others, the extended Kalman-Bucy filter [144], the unscented Kalman-Bucy filter [406], projection filters [69], assumed-density filters [70, 230] the ensemble Kalman-Bucy filter (EnKBF) [143], and other particle filters (e.g., [142], [19, Chapter 9], and references therein).

This assumption implies, in particular, that there exists a finite-dimensional, $\mathcal{F}_t^Y$-adapted process $M(\theta, \boldsymbol{o}) = \{M_t(\theta, \boldsymbol{o})\}_{t\geq 0}$, taking values in $\mathbb{R}^p$, and functions $\psi_C(\theta, \boldsymbol{o}, \cdot) : \mathbb{R}^p \to \mathbb{R}^{n_y}$,

$\psi_j(\theta, \boldsymbol{o}, \cdot) : \mathbb{R}^p \to \mathbb{R}^{n_x}$ such that, in the case of an exact solution,

$$\hat{C}_t(\theta, \boldsymbol{o}) = \psi_C(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o})), \tag{3.24a}$$

$$\hat{j}_t(\theta, \boldsymbol{o}) = \psi_j(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o})), \tag{3.24b}$$

or, in the case of an approximate solution, such that these equations hold only approximately. The process $M(\theta, \boldsymbol{o})$ is typically referred to as the finite-dimensional (approximate) filter representation, or more simply, the filter. We provide an illustrative example of one such finite-dimensional filter representation after stating our remaining assumptions.

We are also required to compute the gradients $\hat{C}_t^{\theta}(\theta, \boldsymbol{o})$ and $\hat{j}_t^{\boldsymbol{o}}(\theta, \boldsymbol{o})$ in order to implement our algorithm. We must therefore also introduce the following additional assumption.

Assumption 3.2.2. *The finite-dimensional filter representation is continuously differentiable with respect to θ and $\boldsymbol{o}$.*

Following this assumption, we can define $M^{\theta}(\theta, \boldsymbol{o}) = \{M_t^{\theta}(\theta, \boldsymbol{o})\}_{t\geq 0} = \{\nabla_{\theta} M_t(\theta, \boldsymbol{o})\}_{t\geq 0}$ and $M^{\boldsymbol{o}}(\theta, \boldsymbol{o}) = \{M_t^{\boldsymbol{o}}(\theta, \boldsymbol{o})\}_{t\geq 0} = \{\nabla_{\boldsymbol{o}} M_t(\theta, \boldsymbol{o})\}_{t\geq 0}$ as the $\mathbb{R}^{p\times n_{\theta}}$ and $\mathbb{R}^{p\times n_y n_{\boldsymbol{o}}}$ valued processes consisting of the gradients of the finite dimensional filter representation with respect to θ and $\boldsymbol{o}$, respectively. We will refer to these processes as the (finite-dimensional) tangent filters.

It follows, upon formal differentiation of equations (3.24a) and (3.24b), that, either exactly or approximately, we have

$$\hat{C}_t^{\theta}(\theta, \boldsymbol{o}) = \psi_C^{\theta}(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o}), M_t^{\theta}(\theta, \boldsymbol{o})) \tag{3.25a}$$

$$= \nabla_{\theta}\psi_C(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o})) + \nabla_M \psi_C(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o})) M_t^{\theta}(\theta, \boldsymbol{o}), \tag{3.25b}$$

and

$$\hat{j}_t^{\boldsymbol{o}}(\theta, \boldsymbol{o}) = \psi_j^{\boldsymbol{o}}(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o}), M_t^{\boldsymbol{o}}(\theta, \boldsymbol{o})) \tag{3.26a}$$

$$= \nabla_{\boldsymbol{o}}\psi_j(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o})) + \nabla_M \psi_j(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o})) M_t^{\boldsymbol{o}}(\theta, \boldsymbol{o}). \tag{3.26b}$$

We are now ready to introduce our final assumption on the filter. This assumption will allow us to rewrite the joint online parameter estimation and optimal sensor placement algorithm in the form of Algorithm (2.18a) - (2.18b), and thus to apply Theorem 2.2.

Assumption 3.2.3. *The finite-dimensional filter representation satisfies a stochastic differential equation of the form*

$$\mathrm{d}M_t(\theta, \boldsymbol{o}) = S(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o}))\mathrm{d}t + T(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o}))\mathrm{d}y_t \tag{3.27}$$

$$+ U(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o}))\mathrm{d}a_t,$$

where $a = \{a_t\}_{t\geq 0}$ is a $\mathbb{R}^q$ valued Wiener process independent of $\mathcal{F}_t^{X,Y}$, and the functions S, T, and U map $\mathbb{R}^{n_\theta} \times \mathbb{R}^{n_y n_{\boldsymbol{o}}} \times \mathbb{R}^p$ to $\mathbb{R}^p$, $\mathbb{R}^{p\times n_y n_{\boldsymbol{o}}}$, and $\mathbb{R}^{p\times q}$, respectively.

This assumption can be shown to hold for a broad class of filters. In particular, the inclusion of the independent noise process means that this SDE holds for a large number of approximate filters, including many of those mentioned after Assumption 3.2.1. It follows from this assumption, upon differentiation of (3.27), that the finite-dimensional tangent filters satisfy the SDEs

$$\begin{aligned}\mathrm{d}M_t^\theta(\theta, \boldsymbol{o}) &= S'_\theta(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o}), M_t^\theta(\theta, \boldsymbol{o}))\mathrm{d}t \\ &\quad + T'_\theta(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o}), M_t^\theta(\theta, \boldsymbol{o}))\mathrm{d}y_t \\ &\quad + U'_\theta(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o}), M_t^\theta(\theta, \boldsymbol{o}))\mathrm{d}a_t,\end{aligned} \tag{3.28a}$$

$$\begin{aligned}\mathrm{d}M_t^{\boldsymbol{o}}(\theta, \boldsymbol{o}) &= S'_{\boldsymbol{o}}(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o}), M_t^{\boldsymbol{o}}(\theta, \boldsymbol{o}))\mathrm{d}t \\ &\quad + T'_{\boldsymbol{o}}(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o}), M_t^{\boldsymbol{o}}(\theta, \boldsymbol{o}))\mathrm{d}y_t \\ &\quad + U'_{\boldsymbol{o}}(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o}), M_t^{\boldsymbol{o}}(\theta, \boldsymbol{o}))\mathrm{d}a_t.\end{aligned} \tag{3.28b}$$

where, for example, the tensor field S'_θ is obtained explicitly according to

$$\begin{aligned}S'_\theta(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o}), M_t^\theta(\theta, \boldsymbol{o})) &= \nabla_\theta S(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o})) \\ &\quad + \nabla_M S(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o}))M_t^\theta(\theta, \boldsymbol{o}).\end{aligned} \tag{3.29}$$

with analogous expressions for the tensor fields T'_θ, U'_θ, $S'_{\boldsymbol{o}}$, $T'_{\boldsymbol{o}}$ and $U'_{\boldsymbol{o}}$.

We can now summarise the evolution equations for the latent signal, the finite-dimensional filter, and the finite-dimensional tangent filters, into a single SDE. In particular, let us define $\mathcal{X}(\theta, \boldsymbol{o}) = \{\mathcal{X}_t(\theta, \boldsymbol{o})\}_{t\geq 0}$ as the $\mathbb{R}^N$ valued diffusion process consisting of the concatenation of the latent signal, the (vectorised) finite-dimensional filter, and the (vectorised) finite-dimensional tangent filters, with $N = n_x + p + pn_\theta + pn_y n_{\boldsymbol{o}}$. That is, in a slight abuse of notation,

$$\mathcal{X}_t(\theta, \boldsymbol{o}) = \left(x_t, \mathrm{vec}(M_t(\theta, \boldsymbol{o})), \mathrm{vec}(M_t^\theta(\theta, \boldsymbol{o})), \mathrm{vec}(M_t^{\boldsymbol{o}}(\theta, \boldsymbol{o}))\right)^T. \tag{3.30}$$

It then follows straightforwardly, stacking the equation for the signal process (3.1), the filter (3.27), and tangent filters (3.28a) - (3.28b), and substituting the equation for the observation process (3.2), that

$$\mathrm{d}\mathcal{X}_t(\theta, \boldsymbol{o}) = \Phi(\theta, \boldsymbol{o}, \mathcal{X}_t(\theta, \boldsymbol{o}))\mathrm{d}t + \Psi(\theta, \boldsymbol{o}, \mathcal{X}_t(\theta, \boldsymbol{o}))\mathrm{d}b_t, \tag{3.31}$$

where the functions Φ and Ψ take values in $\mathbb{R}^N$ and $\mathbb{R}^{N\times(n_x+n_y+q)}$, respectively, and where

$b = \{b_t\}_{t\geq 0}$ is the $\mathbb{R}^{n_x+n_y+q}$ valued Wiener process obtained by concatenating the signal noise process $v = \{v_t\}_{t\geq 0}$, the observation noise process $w = \{v_t\}_{t\geq 0}$, and the independent noise process arising in the equations for the finite-dimensional filter representation $a = \{a_t\}_{t\geq 0}$.

Example. *To help to illustrate the notation introduced in this section, let us consider a simple one-dimensional linear Gaussian model with a single unknown parameter, and a single sensor location, viz*

$$\mathrm{d}x_t = -\theta x_t \mathrm{d}t + \mathrm{d}v_t \ , \quad x(0) = x_0, \tag{3.32}$$

$$\mathrm{d}y_t = x_t \mathrm{d}t + \mathrm{d}w_t \quad , \quad y(0) = 0, \tag{3.33}$$

where $v = \{w_t\}_{t\geq 0}$ and $w = \{v_t\}_{t\geq 0}$ are one-dimensional Brownian motions with incremental variances $Q(\theta) = 1$ and $R(o) = (o - o_0)^2$, and $x_0 \sim \mathcal{N}(0, \frac{1}{2\theta})$. Clearly, this is an example of a partially observed diffusion process of the form (3.1) - (3.2)*, with $\theta \in \mathbb{R}$, $o \in \mathbb{R}$, and operators $A(\theta, x) = -\theta x$, $B(\theta, x) = 1$, and $C(\theta, \boldsymbol{o}, x) = x$. We can also identify, using* (3.9) *and* (3.16)*, the conditional expectations*

$$\hat{C}_t(\theta, \boldsymbol{o}) = \mathbb{E}_{\theta,\boldsymbol{o}}[C(\theta, \boldsymbol{o}, x_t)|\mathcal{F}_t^Y] = \mathbb{E}_{\theta,\boldsymbol{o}}[x_t|\mathcal{F}_t^Y] \tag{3.34a}$$

$$\hat{j}_t(\theta, \boldsymbol{o}) = \mathrm{Tr}[\mathrm{Var}_{\theta,\boldsymbol{o}}[x_t|\mathcal{F}_t^Y]] = \mathrm{Var}_{\theta,\boldsymbol{o}}\left[x_t|\mathcal{F}_t^Y\right]. \tag{3.34b}$$

Let us consider each of the assumptions introduced in this section in turn, starting with Assumption 3.2.1. For the linear Gaussian model, the optimal filter has a Gaussian distribution with mean $\hat{x}_t(\theta, o) = \mathbb{E}_{\theta,o}[x_t|\mathcal{F}_t^Y]$ and variance $\hat{\Sigma}_t(\theta, o) = \mathrm{Var}_{\theta,o}[x_t|\mathcal{F}_t^Y]$, both of which can be computed recursively. The precise form of these equations is presented below in (3.37)*. This is known as the Kalman-Bucy filter [237]. We thus have a $p = 2$ dimensional filter representation $M_t(\theta, \boldsymbol{o}) = (\hat{x}_t(\theta, \boldsymbol{o}), \hat{\Sigma}_t(\theta, \boldsymbol{o}))^T$. It follows straightforwardly that, in the case,*

$$\psi_C(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o})) := \hat{C}_t(\theta, \boldsymbol{o}) = \hat{x}_t(\theta, \boldsymbol{o}) \tag{3.35a}$$

$$\psi_j(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o})) := \hat{j}_t(\theta, \boldsymbol{o}) = \hat{\Sigma}_t(\theta, \boldsymbol{o}). \tag{3.35b}$$

We next consider Assumption 3.2.2. In the current example, it is clear that the two-dimensional filter $M_t(\theta, \boldsymbol{o})$ is continuously differentiable with respect to both θ and o. Indeed, this follows directly from the differentiability of $A(\theta, x)$, $B(\theta, x)$, $C(\theta, \boldsymbol{o}, x)$, $Q(\theta)$ and $R(\boldsymbol{o})$ with respect to these variables. We can thus define the finite-dimensional tangent filters $M_t^{\theta}(\theta, \boldsymbol{o}) = (\hat{x}_t^{\theta}(\theta, \boldsymbol{o}), \hat{\Sigma}_t^{\theta}(\theta, \boldsymbol{o}))^T$ and $M_t^{\boldsymbol{o}}(\theta, \boldsymbol{o}) = (\hat{x}_t^{\boldsymbol{o}}(\theta, \boldsymbol{o}), \hat{\Sigma}_t^{\boldsymbol{o}}(\theta, \boldsymbol{o}))^T$, and compute

$$\psi_C^{\theta}(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o}), M_t^{\theta}(\theta, \boldsymbol{o})) := \hat{C}_t^{\theta}(\theta, \boldsymbol{o}) = \hat{x}_t^{\theta}(\theta, \boldsymbol{o}) \tag{3.36a}$$

$$\psi_j^{\boldsymbol{o}}(\theta, \boldsymbol{o}, M_t(\theta, \boldsymbol{o}), M_t^{\boldsymbol{o}}(\theta, \boldsymbol{o})) := \hat{j}_t^{\boldsymbol{o}}(\theta, \boldsymbol{o}) = \hat{\Sigma}_t^{\boldsymbol{o}}(\theta, \boldsymbol{o}) \tag{3.36b}$$

Finally, we consider Assumption 3.2.3. The Kalman-Bucy filter evolves according to the following SDE

$$\underbrace{\begin{pmatrix} \mathrm{d}\hat{x}_t(\theta,\boldsymbol{o}) \\ \mathrm{d}\hat{\Sigma}_t(\theta,\boldsymbol{o}) \end{pmatrix}}_{\mathrm{d}M_t(\theta,\boldsymbol{o})} = \underbrace{\begin{pmatrix} -\theta\hat{x}_t(\theta,\boldsymbol{o}) - (\boldsymbol{o}-\boldsymbol{o}_0)^{-2}\hat{x}_t(\theta,\boldsymbol{o})\hat{\Sigma}_t(\theta,\boldsymbol{o}) \\ 1 - 2\theta\hat{\Sigma}_t(\theta,\boldsymbol{o}) - (\boldsymbol{o}-\boldsymbol{o}_0)^{-2}\hat{\Sigma}_t^2(\theta,\boldsymbol{o}) \end{pmatrix}}_{S(\theta,\boldsymbol{o},M_t(\theta,\boldsymbol{o}))}\mathrm{d}t + \underbrace{\begin{pmatrix} (\boldsymbol{o}-\boldsymbol{o}_0)^{-2}\hat{\Sigma}_t(\theta,\boldsymbol{o}) \\ 0 \end{pmatrix}}_{T(\theta,\boldsymbol{o},M_t(\theta,\boldsymbol{o}))}\mathrm{d}y_t, \tag{3.37}$$

Thus, the filter does indeed evolve according to an SDE of the form (3.27)*, with the final term identically equal to zero. Taking formal derivatives of this SDE, we can obtain the SDEs for the tangent filters, namely*

$$\underbrace{\begin{pmatrix} \mathrm{d}\hat{x}_t^{\theta}(\theta,\boldsymbol{o}) \\ \\ \mathrm{d}\hat{\Sigma}_t^{\theta}(\theta,\boldsymbol{o}) \end{pmatrix}}_{\mathrm{d}M_t^{\theta}(\theta,\boldsymbol{o})} = \underbrace{\begin{pmatrix} -\hat{x}_t(\theta,\boldsymbol{o}) - \theta\hat{x}_t^{\theta}(\theta,\boldsymbol{o}) \\ -(\boldsymbol{o}-\boldsymbol{o}_0)^{-2}\hat{x}_t^{\theta}(\theta,\boldsymbol{o})\hat{\Sigma}_t(\theta,\boldsymbol{o}) \\ -(\boldsymbol{o}-\boldsymbol{o}_0)^{-2}\hat{x}_t(\theta,\boldsymbol{o})\hat{\Sigma}_t^{\theta}(\theta,\boldsymbol{o}) \\ -2\hat{\Sigma}_t(\theta,\boldsymbol{o}) - 2\theta\hat{\Sigma}_t^{\theta}(\theta,\boldsymbol{o}) \\ -2(\boldsymbol{o}-\boldsymbol{o}_0)^{-2}\hat{\Sigma}_t(\theta,\boldsymbol{o})\hat{\Sigma}_t^{\theta}(\theta,\boldsymbol{o}) \end{pmatrix}}_{S'_{\theta}(\theta,\boldsymbol{o},M_t(\theta,\boldsymbol{o}))}\mathrm{d}t + \underbrace{\begin{pmatrix} (\boldsymbol{o}-\boldsymbol{o}_0)^{-2}\hat{\Sigma}_t^{\theta}(\theta,\boldsymbol{o}) \\ \\ 0 \end{pmatrix}}_{T'_{\theta}(\theta,\boldsymbol{o},M_t(\theta,\boldsymbol{o}))}\mathrm{d}y_t. \tag{3.38}$$

and, similarly,

$$\underbrace{\begin{pmatrix} \mathrm{d}\hat{x}_t^{\boldsymbol{o}}(\theta,\boldsymbol{o}) \\ \\ \mathrm{d}\hat{\Sigma}_t^{\boldsymbol{o}}(\theta,\boldsymbol{o}) \end{pmatrix}}_{\mathrm{d}M_t^{\boldsymbol{o}}(\theta,\boldsymbol{o})} = \underbrace{\begin{pmatrix} -\theta\hat{x}_t^{\boldsymbol{o}}(\theta,\boldsymbol{o}) \\ +2(\boldsymbol{o}-\boldsymbol{o}_0)^{-3}\hat{x}_t(\theta,\boldsymbol{o})\hat{\Sigma}_t(\theta,\boldsymbol{o}) \\ -(\boldsymbol{o}-\boldsymbol{o}_0)^{-2}\hat{x}_t^{\boldsymbol{o}}(\theta,\boldsymbol{o})\hat{\Sigma}_t(\theta,\boldsymbol{o}) \\ -(\boldsymbol{o}-\boldsymbol{o}_0)^{-2}\hat{x}_t(\theta,\boldsymbol{o})\hat{\Sigma}_t^{\boldsymbol{o}}(\theta,\boldsymbol{o}) \\ -2\theta\hat{\Sigma}_t^{\boldsymbol{o}}(\theta,\boldsymbol{o}) + 2(\boldsymbol{o}-\boldsymbol{o}_0)^{-3}\hat{\Sigma}_t^2(\theta,\boldsymbol{o}) \\ -2(\boldsymbol{o}-\boldsymbol{o}_0)^{-2}\hat{\Sigma}_t(\theta,\boldsymbol{o})\hat{\Sigma}_t^{\boldsymbol{o}}(\theta,\boldsymbol{o}) \end{pmatrix}}_{S'_{\boldsymbol{o}}(\theta,\boldsymbol{o},M_t(\theta,\boldsymbol{o}))}\mathrm{d}t \tag{3.39}$$

$$+\underbrace{\begin{pmatrix} -2(\boldsymbol{o}-\boldsymbol{o}_0)^{-3}\hat{\Sigma}_t(\theta,\boldsymbol{o}) \\ +(\boldsymbol{o}-\boldsymbol{o}_0)^{-2}\hat{\Sigma}_t^{\boldsymbol{o}}(\theta,\boldsymbol{o}) \\ \\ 0 \\ \\ \end{pmatrix}}_{T'_{\boldsymbol{o}}(\theta,\boldsymbol{o},M_t(\theta,\boldsymbol{o}))}\mathrm{d}y_t.$$

Finally, we can concatenate the (one-dimensional) signal, the (two-dimensional) filter, and the two (two-dimensional) tangent filters into a single diffusion process, namely,

$$\mathcal{X}_t(\theta,\boldsymbol{o}) = \big(x_t, \underbrace{\hat{x}_t(\theta,\boldsymbol{o}),\hat{\Sigma}_t(\theta,\boldsymbol{o})}_{\mathrm{vec}(M_t(\theta,\boldsymbol{o}))}, \underbrace{\hat{x}_t^{\theta}(\theta,\boldsymbol{o}),\hat{\Sigma}_t^{\theta}(\theta,\boldsymbol{o})}_{\mathrm{vec}(M_t^{\theta}(\theta,\boldsymbol{o}))}, \underbrace{\hat{x}_t^{\boldsymbol{o}}(\theta,\boldsymbol{o}),\hat{\Sigma}_t^{\boldsymbol{o}}(\theta,\boldsymbol{o})}_{\mathrm{vec}(M_t^{\boldsymbol{o}}(\theta,\boldsymbol{o}))}\big)^T. \tag{3.40}$$

This process evolves according to an SDE of the form (3.31)*, which we obtain by stacking the signal equation* (3.32)*, the Kalman-Bucy filtering equations* (3.37)*, and the tangent Kalman-Bucy filtering equations* (3.38) - (3.39)*, before substituting the observation equation* (3.33)*. For brevity, the explicit form of this equation is omitted.*

3.2.4 Joint Parameter Estimation and Optimal Sensor Placement

We can finally now turn our attention to the problem of simultaneous online parameter estimation and online optimal sensor placement. As outlined in the introduction, we cast this as a bilevel optimisation problem, in which the objective is to obtain $\hat{\theta}\in\Theta$, $\hat{\boldsymbol{o}}(\hat{\theta})\in\Omega^{n_y}$ such that

$$\hat{\theta}\in\underset{\theta\in\Theta}{\arg\min}\Big[-\tilde{\mathcal{L}}\big(\theta,\hat{\boldsymbol{o}}(\theta)\big)\Big] \quad , \quad \hat{\boldsymbol{o}}(\theta)\in\underset{\boldsymbol{o}\in\Omega^{n_y}}{\arg\min}\,\tilde{\mathcal{J}}\big(\theta,\boldsymbol{o}\big). \tag{3.41}$$

We should remark that, depending on our primary objective, we may instead specify $\tilde{\mathcal{J}}$ as the upper-level objective function, and $-\tilde{\mathcal{L}}$ as the lower-level objective function. Indeed, the subsequent methodology is generic to either case. As in Chapter 2, we will consider a weaker version of this problem, in which we simply seek to obtain joint stationary points of $\tilde{\mathcal{L}}$ and $\tilde{\mathcal{J}}$.

There are two possible approaches to this task. The first is to alternate between online parameter estimation and optimal sensor placement, periodically updating the locations of the measurement sensors on the basis of the current parameter estimates. The second is to jointly perform online parameter estimation and optimal sensor placement, simultaneously and recursively updating the parameter estimates and the locations of the measurement sensors. We strongly advocate the second approach, which is not only more numerically convenient, but can be implemented in a truly online fashion. Moreover, in the case of mobile sensors, this approach can provide real-time motion guidance.

3.2.4.1 The 'Ideal' Algorithm

To solve this bilevel optimisation problem, we propose a continuous-time, stochastic gradient descent algorithm, which combines the schemes in Sections 3.2.1 and 3.2.2, c.f., (3.14) and (3.22). In particular, suppose some initialisation at $\theta_0 \in \Theta$, $\boldsymbol{o}_0 \in \Omega^{n_y}$. Then, simultaneously, we generate parameter estimates $\{\theta_t\}_{t\geq 0}$ and optimal sensor locations $\{\boldsymbol{o}_t\}_{t\geq 0}$ according to

$$\mathrm{d}\theta_t = \begin{cases} \gamma_t^1 \big[\hat{C}_t^{\theta}(\theta_t, \boldsymbol{o}_t)\big]^T R^{-1}(\boldsymbol{o}_t)\big[\mathrm{d}y_t - \hat{C}_t(\theta_t, \boldsymbol{o}_t)\mathrm{d}t\big] & , \ \theta_t \in \Theta, \\ 0 & , \ \theta_t \notin \Theta, \end{cases} \tag{3.42a}$$

$$\mathrm{d}\boldsymbol{o}_t = \begin{cases} -\gamma_t^2 \big[\hat{j}_t^{\boldsymbol{o}}(\theta_t, \boldsymbol{o}_t)\big]^T \mathrm{d}t & , \ \boldsymbol{o}_t \in \Omega^{n_y}, \\ 0 & , \ \boldsymbol{o}_t \notin \Omega^{n_y}. \end{cases} \tag{3.42b}$$

3.2.4.2 The Implementable Algorithm

In general, it is not possible to implement Algorithm (3.42a) - (3.42b) in its current form, since it depends on the possibly intractable conditional expectations $\hat{C}, \hat{C}^{\theta}$, and $\hat{j}^{\boldsymbol{o}}$. We can, however, obtain an implementable version of this algorithm by replacing these quantities by their (possibly approximate) finite-dimensional filter representations ψ_C, ψ_C^{θ}, and $\psi_j^{\boldsymbol{o}}$.

For the purpose of our theoretical analysis, it will also be useful to rewrite Algorithm (3.42a) - (3.42b) in the form of Algorithm (2.18a) - (2.18b), the generic two-timescale algorithm analysed in Section 2.2.2. It is worth noting that this requires us to rewrite $\mathrm{d}y_t$ using the observation equation (3.2). After taking these steps, we finally arrive at

$$\mathrm{d}\theta_t = \begin{cases} -\gamma_t^1 \big[F(\theta_t, \boldsymbol{o}_t, \mathcal{X}_t)\mathrm{d}t + \mathrm{d}\zeta_t^1\big] & , \ \theta_t \in \Theta, \\ 0 & , \ \theta_t \notin \Theta, \end{cases} \tag{3.43a}$$

$$\mathrm{d}\boldsymbol{o}_t = \begin{cases} -\gamma_t^2 \big[G(\theta_t, \boldsymbol{o}_t, \mathcal{X}_t)\mathrm{d}t\big] & , \ \boldsymbol{o}_t \in \Omega^{n_y}, \\ 0 & , \ \boldsymbol{o}_t \notin \Omega^{n_y}, \end{cases} \tag{3.43b}$$

where F and G are the $\mathbb{R}^{n_\theta}$- and $\mathbb{R}^{n_y n_{\boldsymbol{o}}}$-valued functions defined according to

$$F(\theta_t, \boldsymbol{o}_t, \mathcal{X}_t) = -\big[\psi_C^{\theta}(\theta_t, \boldsymbol{o}_t, M_t, M_t^{\theta})\big]^T R^{-1}(\boldsymbol{o}_t) \big[C(\theta^*, \boldsymbol{o}_t, x_t) - \psi_C(\theta_t, \boldsymbol{o}_t, M_t)\big], \tag{3.44}$$

$$G(\theta_t, \boldsymbol{o}_t, \mathcal{X}_t) = \psi_j^{\boldsymbol{o}}(\theta_t, \boldsymbol{o}_t, M_t, M_t^{\boldsymbol{o}})^T, \tag{3.45}$$

where ζ_1 is the $\mathbb{R}^{n_\theta}$-valued semi-martingale which evolves according to the SDE

$$\mathrm{d}\zeta_t^1(\theta_t,\boldsymbol{o}_t) = \underbrace{\left[\psi_C^\theta(\theta_t,\boldsymbol{o}_t,M_t,M_t^\theta)\right]^T\mathcal{R}^{-1}(\boldsymbol{o}_t)}_{\zeta_1^{(2)}(\theta_t,\boldsymbol{o}_t,\mathcal{X}_t)}\mathrm{d}w_t, \tag{3.46}$$

and where $\mathcal{X} = \{\mathcal{X}_t\}_{t\geq 0} = \{\mathcal{X}_t(\theta_t,\boldsymbol{o}_t)\}_{t\geq 0}$ is the $\mathbb{R}^N$-valued diffusion process defined in (3.31), consisting of latent state, the filter, and the tangent filters, now integrated along the path of the algorithm iterates. We emphasise that this algorithm can be implemented for both exact (e.g. Kalman-Bucy, Beneš) and approximate (e.g., ensemble Kalman-Bucy, unscented Kalman-Bucy, projection) filters.

Example. *Let us return to the one-dimensional linear Gaussian example considered in the previous section. We can now provide the specific joint online parameter estimation and optimal sensor placement algorithm for this model. In particular, substituting our previous expressions for* $C(\theta,\boldsymbol{o},x)$*,* $R(\boldsymbol{o})$*,* $\psi_C(\theta,\boldsymbol{o},M)$*,* $\psi_C^\theta(\theta,\boldsymbol{o},M,M^\theta)$*, and* $\psi_j^{\boldsymbol{o}}(\theta,\boldsymbol{o},M,M^{\boldsymbol{o}})$*, c.f.* (3.33)*,* (3.35a)*,* (3.36a) *and* (3.36b)*, into the equations for* F*,* G*, and* ζ_1*, c.f.* (3.44) *,* (3.45) *and* (3.46)*, we obtain the update equations*

$$\mathrm{d}\theta_t = -\gamma_t^1\Big[-\hat{x}_t^\theta(\theta_t,\boldsymbol{o}_t)(\boldsymbol{o}_t-\boldsymbol{o}_0)^{-2}(x_t-\hat{x}_t(\theta_t,\boldsymbol{o}_t))\mathrm{d}t \tag{3.47a}$$
$$+\hat{x}_t^\theta(\theta_t,\boldsymbol{o}_t)(\boldsymbol{o}_t-\boldsymbol{o}_0)^{-2}\mathrm{d}w_t)\Big]$$

$$\mathrm{d}\boldsymbol{o}_t = -\gamma_t^2\left[\hat{\Sigma}_t^{\boldsymbol{o}}(\theta_t,\boldsymbol{o}_t)\right]\mathrm{d}t. \tag{3.47b}$$

where the filter mean $\hat{x}_t(\theta_t,\boldsymbol{o}_t)$*, and filter derivatives* $\hat{x}_t^\theta(\theta_t,\boldsymbol{o}_t)$*,* $\hat{\Sigma}_t^{\boldsymbol{o}}(\theta_t,\boldsymbol{o}_t)$*, evolve according to the Kalman-Bucy filter equation* (3.37)*, and the tangent Kalman-Bucy filter equations* (3.38) - (3.39)*, now evaluated along the path of the algorithm iterates.*

3.2.4.3 Main Result

We will analyse Algorithm (3.43a) - (3.43b) under most of the assumptions introduced in Section 2.2.2 for the general two-timescale gradient descent algorithm with Markovian dynamics,[12] in addition to the assumptions introduced in Section 3.2.3 for the filter and filter derivatives. In order to state our main result, we must first define the (possibly approximate) representations of the asymptotic log-likelihood and the asymptotic sensor placement objective, c.f. (3.13) and (3.21), in terms of the (possibly approximate) finite

[12] In particular, we now no longer require two of the conditions relating to the additive, state-dependent noise processes $\{\zeta_t^i\}_{t\geq 0}$, namely Conditions 2.2.3a and 2.2.3c, as these can be shown to follow directly from Condition 2.2.3b.

dimensional filter. In particular, we will write

$$\tilde{\mathcal{L}}^{(\text{filter})}(\theta,\boldsymbol{o}) = \lim_{t\to\infty}\frac{1}{t}\left[\int_0^t R^{-1}(\boldsymbol{o})\psi_C(\theta,\boldsymbol{o},M_s(\theta,\boldsymbol{o}))\cdot \mathrm{d}y_s \right. \left. -\frac{1}{2}\int_0^t ||R^{-\frac{1}{2}}(\boldsymbol{o})\psi_C(\theta,\boldsymbol{o},M_s(\theta,\boldsymbol{o}))||^2 \mathrm{d}s\right] \tag{3.48}$$

$$\tilde{\mathcal{J}}^{(\text{filter})}(\theta,\boldsymbol{o}) = \lim_{t\to\infty}\frac{1}{t}\left[\int_0^t \psi_j(\theta,\boldsymbol{o},M_s(\theta,\boldsymbol{o}))\mathrm{d}s\right]. \tag{3.49}$$

We are now ready to state our main result on the convergence of Algorithm (3.43a) - (3.43b).

Proposition 3.1. *Assume that Conditions 2.2.1, 2.2.2a - 2.2.2e, 2.2.3b, 2.1.4 - 2.1.6, and 3.2.1 - 3.2.3 hold.*[13] *Then, with probability one,*

$$\lim_{t\to\infty}\nabla_\theta\tilde{\mathcal{L}}^{(\text{filter})}(\theta_t,\boldsymbol{o}_t) = \lim_{t\to\infty}\nabla_{\boldsymbol{o}}\tilde{\mathcal{J}}^{(\text{filter})}(\theta_t,\boldsymbol{o}_t) = 0, \tag{3.50}$$

or

$$\lim_{t\to\infty}(\theta_t,\boldsymbol{o}_t)\in\{(\theta,\boldsymbol{o}):\theta\in\partial\Theta\cup\boldsymbol{o}\in\partial\Omega^{n_y}\}. \tag{3.51}$$

Proof. See Section 3.3. □

Proposition 3.1 is obtained as a corollary of Theorem 2.2. In particular, Algorithm (3.43a) - (3.43b) is a special case of Algorithm (2.18a) - (2.18b), in which the additive noise for the slow process is defined by equation (3.46), and the additive noise for the fast process is identically equal to zero. Aside from notational differences, the modifications in the statement of this theorem, when compared to Theorem 2.2, are due solely to the inclusion of the projection which ensures that the algorithm iterates remain in the open sets $\Theta\in\mathbb{R}^{n_\theta}$, $\Omega^{n_y}\in\mathbb{R}^{n_y n_{\boldsymbol{o}}}$ with probability one.

Proposition 3.1 extends Theorem 1 in [430], in which a.s. convergence of the online parameter estimate was established under slightly weaker conditions. In particular, the a.s. convergence results in [430] does not depend on a.s. boundedness of the algorithm iterates. The method of proof, however, is entirely different (see discussion in Section 2.2.2). We remark, as in the previous section, that our theorem (and its proof) still

[13] In particular, we assume that these conditions hold for $\mathcal{X}$, the $\mathbb{R}^N$- valued diffusion process defined in (3.31), F and G, the functions defined in (3.44) and (3.45), and ζ_1, the semi-martingale defined in (3.46), replacing the algorithm iterates $\alpha\leftrightarrow\theta$, $\beta\leftrightarrow\boldsymbol{o}$ and the functions $f\leftrightarrow\tilde{\mathcal{L}}^{(\text{filter})}$, $g\leftrightarrow\tilde{\mathcal{J}}^{(\text{filter})}$ where necessary, with $\tilde{\mathcal{L}}^{(\text{filter})}$ and $\tilde{\mathcal{J}}^{(\text{filter})}$ the functions defined in (3.48) and (3.49).

holds upon restriction to a single-timescale; that is, under the assumption that only the parameters are estimated, while the sensor locations are fixed, or vice versa. In this case, of course, we only require assumptions which relate to the quantity of interest. Thus, upon restriction to a single-timescale (i.e., assuming that the sensors are fixed), our theorem reduces to the result in [430], while our proof provides an entirely different proof for that result.

3.2.4.4 Extensions for Approximate Filters

Proposition 3.1 guarantees that the online parameter estimates and the optimal sensor placements generated by Algorithm (3.43a) - (3.43b) converge to the stationary points of the finite-dimensional filter representations of the asymptotic log-likelihood and the sensor placement objective function, namely, $\tilde{\mathcal{L}}^{(\text{filter})}(\theta,\boldsymbol{o})$ and $\tilde{\mathcal{J}}^{(\text{filter})}(\theta,\boldsymbol{o})$. In the case that one can obtain exact solutions to the Kushner Stratonovich equation (e.g., using the Kalman-Bucy filter for a linear Gaussian model), these representations will be exact, and thus this proposition implies convergence to the stationary points of the 'true' objective functions $\tilde{\mathcal{L}}(\theta,\boldsymbol{o})$ and $\tilde{\mathcal{J}}(\theta,\boldsymbol{o})$.

On the other hand, if it is only possible to obtain approximate solutions to the Kushner-Stratonovich equation (e.g., using a continuous-time particle filter for a non-linear model), Proposition 3.1 still guarantees convergence, but now to the stationary points of an 'approximate' asymptotic log-likelihood and an 'approximate' asymptotic sensor placement objective function, namely, the representations of these functions in terms of the approximate finite-dimensional filter. In this case, it is clear that the asymptotic properties of the online parameter estimates and optimal sensor placements with respect to the 'true' objective functions will be determined by the properties of the approximate filter. In particular, in order to obtain convergence (e.g., in $\mathbb{L}^p$) to the stationary points of the true objective functions, one now requires bounds, preferably uniform in time, on terms such as

$$\mathbb{E}\left[||\psi_C(\theta,\boldsymbol{o},M_t)-\hat{C}_t(\theta,\boldsymbol{o})||^n\right]\ ,\ \mathbb{E}\left[||\psi_j(\theta,\boldsymbol{o},M_t)-\hat{j}_t(\theta,\boldsymbol{o})||^n\right] \tag{3.52}$$

We discuss this point in greater depth in Appendix 3.A, and sketch the details of how one can obtain an $\mathbb{L}^p$ convergence result of this type for the Ensemble Kalman-Bucy Filter (EnKBF) (e.g., [141, 143]).

3.2.4.5 Other Extensions

It is worth emphasising that Proposition 3.1 only establishes the convergence of the parameter estimates and the sensor placements generated by Algorithm (3.43a) - (3.43b) to stationary points of $\tilde{\mathcal{L}}^{(\text{filter})}(\theta, \boldsymbol{o})$ and $\tilde{\mathcal{J}}^{(\text{filter})}(\theta, \boldsymbol{o})$. In particular, it does not guarantee convergence of the parameter estimates to a global maximum of the asymptotic log-likelihood function, or of the sensor placements to a global minimum of the asymptotic sensor placement objective function. It is possible to extend Proposition 3.1 in this direction, establishing a.s. convergence of the algorithm iterates to global optima of these two objective functions (see Section 2.4). This implies, under the assumption that the model is well-specified, that the parameter estimates converge to the true parameter value, and the sensor placements converge to an optimal sensor placement, respectively.[14,15]

However, such an extension comes at the expense of rather strong conditions which are often not satisfied in practice (e.g., global convexity of the objective functions). Thus, in practice, we can generally only hope that our two-timescale stochastic gradient descent descent algorithm will converge to local optima of the two objective functions. This being said, there are various ways in which one can improve the chance of converging to 'good' local optima, or even the global optima. These include multiple random restarts, interacting particles (e.g., [64]), annealing noise (e.g., [111]), momentum (e.g., [306]), amongst others. While many of these techniques have only rigorously been analysed in the single-timescale case, it is reasonable to expect that they could also improve the performance of our two-timescale algorithm.

3.2.4.6 Sufficient Conditions

Let us make some brief remarks on the assumptions required for Proposition 3.1 (see also [430]). The majority of these assumptions are fairly classical, namely, those on the learning rate (Condition 2.2.1), the additive noise process ζ_1 (Condition 2.2.3b), the stability of the algorithm iterates (Condition 2.1.4), and the stationary points of the asymptotic objective functions (Conditions 2.1.5 - 2.1.6). Meanwhile, our assumptions on the filter (Conditions 3.2.1 - 3.2.3) are relatively weak, and are satisfied by many exact and approximate filters. In fact, using the results recently established in [50], it may be possible to relax Condition 3.2.3 further, and allow the evolution equation for the filter to include a jump process. This would further extend the applicability of this result, allowing for a broader class of continuous-time particle filters (e.g., [142]).

It remains to consider the assumptions relating to the diffusion process $\mathcal{X}$ consisting of the latent signal, the filter, and the tangent filters (Conditions 2.2.2a - 2.2.2e). These include

[14] To be precise, in order to identify the global maximum of the asymptotic log-likelihood with the true parameter value, we also require some weak identifiability assumptions (e.g., [296]).

[15] In the case that the model is not well-specified, we can no longer identify the global maximum of the asymptotic log-likelihood with the true parameter value, since we no longer have a notion of a 'true parameter'. However, one can still show that the global maximum of the asymptotic log-likelihood occurs when the 'distance' between the true observation operator and the filter estimate of the observation operator is minimised (e.g., [333] for some relevant results in the fully observed case). Thus, in this case, the online parameter estimates would still converge to a parameter which is in some sense 'optimal'.

ergodicity (Condition 2.2.2a), uniformly bounded moments (Condition 2.2.2b, Condition 2.2.2e), polynomial growth for the diffusion term in the associated SDE (Condition 2.2.2c), and existence and regularity of solutions of the Poisson equations associated with the generator of this process and the asymptotic objective functions (Condition 2.2.2d). One can find sufficient conditions for one of these assumptions (Condition 2.2.2c) in [413, Appendix F]. In particular, there we show that this assumption can be replaced by the slightly weaker assumptions that (i) certain functions appearing in the definition of Algorithm (3.43a) - (3.43b) have the polynomial growth property and (ii) for certain functions satisfying the polynomial growth property, the Poisson equation admits a unique solution which also has this property.

In general, while our conditions are necessary in order to establish a.s. convergence, they are somewhat strong, and in general must be verified on a case by case basis. This being said, in the linear Gaussian case, one can obtain sufficient conditions which are straightforward to verify [412]. In particular, the required conditions coincide with standard conditions required for stability of the Kalman-Bucy filter, and are thus arguably the weakest under which an asymptotic result of this type can be established.

More broadly, the problem of obtaining more easily verifiable sufficient conditions remains open. Indeed, the diffusion process is generally highly degenerate, and thus standard sufficient conditions for non degenerate elliptic diffusion processes (see [413, Appendix D]), do not apply (e.g., [371, 372]). In the case that an exact, finite-dimensional solution to the Kushner-Stratonovich equation exists, ergodicity of the optimal filter follows directly from ergodicity of the latent signal process and the non-degeneracy of the observation process (e.g., [75, 263]), but ergodicity of the tangent filter(s) must still established. Meanwhile, in the case that only an approximate, finite-dimensional solution to the Kushner-Stratonovich equation exists, there is no guarantee that the approximate filter is ergodic, let alone the tangent filter(s).

3.2.4.7 Discussion

In practice, it is evident that our continuous-time, two-timescale stochastic gradient scheme must be discretised. It is thus natural to ask why we prefer to use a discrete-time approximation of this continuous-time algorithm over the traditional approach, which first discretises the continuous-time model, and then applies a classical discrete-time, two-timescale stochastic gradient descent algorithm (see, e.g., [62]). This question was discussed in some detail in Chapter 1, where we provided motivation for this approach. Let us briefly make some additional remarks here.

Using the first approach, the parameter update equations are defined in terms of stochastic estimates of the gradient of the true, continuous-time log-likelihood function. That is, the log-likelihood function defined in terms of the continuous-time model and the corresponding continuous-time filter. On the other hand, using the second approach, the parameter update equations are defined in terms of stochastic estimates of the gradient of an approximate, discrete-time log-likelihood function. That is, a log-likelihood function defined in terms of a discretisation of the continuous-time model, and the corresponding discrete-time filter.

The approximate, discrete-time log-likelihood will always differ from the true, continuous-time log-likelihood.[16] It follows, in particular, that stochastic estimates of the gradient of the approximate, discrete-time log-likelihood are biased estimates of the gradient of the true, continuous-time log-likelihood. Thus, at least in principle, the parameter update equations obtained from discretising the continuous-time parameter update equations are defined in terms of unbiased gradient estimates, while the parameter update equations obtained by first discretising the model are necessarily defined in terms of biased gradient estimates.[17]

From a mathematical perspective, it thus seems reasonable to prefer the first of these approaches over the second. Indeed, while it is possible to analyse the asymptotic behaviour of stochastic gradient algorithms with biased gradient estimates, one can only hope to obtain rather weaker convergence results in this case (see, e.g., [440]). Moreover, while there are several approaches for obtaining unbiased estimates of the gradient of the log-likelihood for partially observed diffusion processes (e.g., [26, 210]), incorporating these into our theoretical analysis would undoubtedly incur a significant technical overhead.

From a practical perspective, the picture is somewhat less clear. Indeed, while in theory the discretisation of the continuous-time learning equations can still be defined in terms of the continuous-time log-likelihood function, in practice this quantity will also need to be approximated. This will certainly require a discretisation of the continuous-time filter, and possibly also a discretisation of the model dynamics. Thus, numerically, there may be little difference between the two approaches, particularly if one restricts attention to simple (e.g., first-order) discretisation schemes for the model, the filter, and the learning equations. For example, in the linear Gaussian case, one can show that a first-order discretisation of the continuous-time Kalman-Bucy filter coincides with the discrete-time Kalman filter, applied to a first-order discretisation of the continuous-time model, up to first order terms (e.g., [423]).

This being said, it is worth emphasising that the first approach allows one to apply any

[16] The error in this approximation will depend, amongst other factors, on the model dynamics, the discretisation scheme, the discrete-time filter, and the method used to approximate the log-likelihood function.

[17] We note that, while we have framed this discussion in terms of the recursive parameter estimates and the log-likelihood function, analogous statements hold for the recursive optimal sensor placements and the sensor placement objective function.

numerical discretisation scheme directly to the continuous-time equations. This has the potential to result in algorithms with improved convergence properties (e.g., [298, 299]). There may also be other practical reasons to prefer the first approach. In particular, if the dimensions of the model parameters (and the filter) are significantly smaller than the dimensions of the model, then discretising the low-dimensional learning equations (and the low-dimensional filter) while avoiding the need to discretise the high-dimensional model dynamics could result in significant computational benefits. We leave a more detailed analysis of these issues to future work.

3.3 Proof of Main Results

In this Section, we provide a proof of Proposition 3.1. For convenience, we now recall the statement of this result.

Proposition 3.1. *Assume that Conditions 2.2.1, 2.2.2a - 2.2.2e, 2.2.3b, 2.1.4 - 2.1.6, and 3.2.1 - 3.2.3 hold.*[18] *Then, with probability one,*

$$\lim_{t\to\infty} \nabla_\theta \tilde{\mathcal{L}}^{(\text{filter})}(\theta_t, \boldsymbol{o}_t) = \lim_{t\to\infty} \nabla_{\boldsymbol{o}} \tilde{\mathcal{J}}^{(\text{filter})}(\theta_t, \boldsymbol{o}_t) = 0, \tag{3.53}$$

or

$$\lim_{t\to\infty} (\theta_t, \boldsymbol{o}_t) \in \{(\theta, \boldsymbol{o}) : \theta \in \partial\Theta \cup \boldsymbol{o} \in \partial\Omega^{n_y}\}. \tag{3.54}$$

Proof. We begin by defining the first exit times from Θ, Ω^{n_y}, respectively, as

$$\tau_\theta = \inf\{t \geq 0 : \theta_t \notin \Theta\}, \tag{3.55}$$

$$\tau_{\boldsymbol{o}} = \inf\{t \geq 0 : \boldsymbol{o}_t \notin \Omega^{n_y}\}. \tag{3.56}$$

First suppose that $\tau_\theta < \infty$. Since the paths of $\{\theta_t\}_{t\geq 0}$ are continuous, it follows that $\theta_{\tau_\theta} \in \partial\Theta$. Furthermore, since $\mathrm{d}\theta_t = 0$ on $\partial\Theta$, we in fact have $\theta_t \in \partial\Theta$ for all $t \geq \tau_\theta$. In particular, it follows that

$$\lim_{t\to\infty} (\theta_t, \boldsymbol{o}_t) \in \{(\theta, \boldsymbol{o}) : \theta \in \partial\Theta \cup \boldsymbol{o} \in \partial\Omega^{n_y}\}. \tag{3.57}$$

Using an identical argument, the same conclusion holds under the assumption that $\tau_{\boldsymbol{o}} < \infty$.

It remains to consider the case when $\tau_\theta = \tau_{\boldsymbol{o}} = \infty$. That is, equivalently, when $\theta_t \in \Theta$ and $\boldsymbol{o}_t \in \Omega^m$ for all $t \geq 0$. In this instance, it is straightforward to see that Algorithm (3.43a)

[18] In particular, we assume that these conditions hold for $\mathcal{X}$, the $\mathbb{R}^N$- valued diffusion process defined in (3.31), F and G, the functions defined in (3.44) and (3.45), and ζ_1, the semi-martingale defined in (3.46), replacing the algorithm iterates $\alpha \leftrightarrow \theta$, $\beta \leftrightarrow \boldsymbol{o}$ and the functions $f \leftrightarrow \tilde{\mathcal{L}}^{(\text{filter})}$, $g \leftrightarrow \tilde{\mathcal{J}}^{(\text{filter})}$ where necessary, with $\tilde{\mathcal{L}}^{(\text{filter})}$ and $\tilde{\mathcal{J}}^{(\text{filter})}$ the functions defined in (3.48) and (3.49).

- (3.43b) is a special case of Algorithm (2.18a) - (2.18b), in which the noise sequences $\{\zeta_t^i\}_{t\geq 0}$, $i = 1, 2$, are defined according to

$$\mathrm{d}\zeta_t^1 = -\zeta_1^{(2)}(\theta_t, \boldsymbol{o}_t, \mathcal{X}_t)\mathrm{d}w_t, \tag{3.58}$$

$$\mathrm{d}\zeta_t^2 = 0, \tag{3.59}$$

and where the function $\zeta_1^{(2)} : \mathbb{R}^{n_\theta} \times \mathbb{R}^{n_y n_{\boldsymbol{o}}} \times \mathbb{R}^N \to \mathbb{R}^{n_\theta \times n_y}$ is defined in equation (3.46). It is thus sufficient to prove that the single condition relating to these noise sequences in Proposition 3.1 (Condition 2.2.3b) is sufficient for the additional conditions in Proposition 2.2 (Conditions 2.2.3a, 2.2.3c). Indeed, in this case, it follows immediately from Proposition 2.2 that

$$\lim_{t\to\infty} \nabla_\theta \tilde{\mathcal{L}}(\theta_t, \boldsymbol{o}_t) = \lim_{t\to\infty} \nabla_{\boldsymbol{o}} \tilde{\mathcal{J}}(\theta_t, \boldsymbol{o}_t) = 0. \tag{3.60}$$

We begin by considering Condition 2.2.3a. We wish to prove that for all $T > 0$, the noise sequences $\{\zeta_t^i\}_{t\geq 0}$, $i = 1, 2$, a.s. satisfy

$$\lim_{s\to\infty} \sup_{t\in[s,s+T]} \left|\left| \int_s^t \gamma_v^i \mathrm{d}\zeta_v^i \right|\right| = 0 \tag{3.61}$$

This condition holds trivially for $\{\zeta_t^2\}_{t\geq 0}$. We thus turn our attention to $\{\zeta_t^1\}_{t\geq 0}$. Using the Itô Isometry, and Conditions 2.2.1, 2.2.2e, and 2.2.3b, there exist constants $q > 0$ and constants $K, K', K'' > 0$ such that

$$\sup_{t\geq 0} \mathbb{E}\left[\left(\int_0^\infty \gamma_t^1 \mathrm{d}\zeta_t^1\right)^2\right] = \mathbb{E}\left[\left(\int_0^t \gamma_t^1 \zeta_1^{(2)}(\theta_t, \boldsymbol{o}_t, \mathcal{X}_t)\mathrm{d}w_t\right)^2\right] \tag{3.62}$$

$$\leq K\mathbb{E}\left[\int_0^t (\gamma_t^1)^2 (1 + \mathbb{E}||\mathcal{X}_t||^q)\mathrm{d}t\right] \tag{3.63}$$

$$\leq KK' \int_0^t (\gamma_t^1)^2 \mathrm{d}t \tag{3.64}$$

$$\leq KK'K'' < \infty. \tag{3.65}$$

Thus, by Doob's martingale convergence theorem, there exists a square integrable random variable, say M_∞, such that, both a.s. and in L^2, $\lim_{t\to\infty} \int_0^t \gamma_t^1 \mathrm{d}\zeta_t^1 = M_\infty$. The required result follows.

It remains to consider Condition 2.2.3c. We wish to prove that there exist constants

$A_{z_1,z_2}, A_{z_i,b} > 0$, $i = 1, 2$, such that, componentwise,

$$c_t^{z_1,z_2} = \frac{\mathrm{d}[z_1, z_2]_t}{\mathrm{d}t} \leq A_{z_1,z_2} \ , \ c_t^{z_i,b} = \frac{\mathrm{d}[z_i, b]_t}{\mathrm{d}t} \leq A_{z_i,b}. \tag{3.66}$$

where, in the general case, $\{z_t^i\}_{t\geq 0}$ are the $\mathbb{R}^{d_5^i}$-valued Wiener processes appearing in the definition of the noise processes $\{\zeta_t^i\}_{t\geq 0}$, c.f. (2.26), and $\{b_t\}_{t\geq 0}$ is the $\mathbb{R}^{d_4}$-valued Wiener process appearing in the definition of the ergodic diffusion process $\{\mathcal{X}_t\}_{t\geq 0}$, c.f. (2.19).

In the case of Algorithm (3.43a) - (3.43b), we identify $z_t^1 = w_t$, $z_t^2 = 0$ from equations (3.58) - (3.59), and $b_t = (v_t, w_t, a_t)^T$ from equation (3.31). Thus, using elementary properties of the quadratic variation, we have, componentwise,

$$\frac{\mathrm{d}[z_1, z_2]_t}{\mathrm{d}t} = 0 \ , \ \frac{\mathrm{d}[z_1, b]_t}{\mathrm{d}t} = 1 \text{ or } 0 \ , \ \frac{\mathrm{d}[z_2, b]_t}{\mathrm{d}t} = 0. \tag{3.67}$$

In particular, Condition 2.2.3c is satisfied. The result now follows immediately from our previous remarks. □

3.4 Numerical Example

To illustrate the results of Section 3.2, we now provide an example of joint online parameter estimation and optimal sensor placement. In particular, we study the numerical performance of the proposed two-timescale stochastic gradient descent algorithm, and verify numerically the convergence of the parameter estimates and the sensor placements. We also provide explicit derivations of the parameter and sensor update equations.

3.4.1 One-Dimensional Benes Filter

We will consider a one-dimensional, partially observed diffusion process defined by

$$\mathrm{d}x_t = \mu\sigma \tanh\left[\frac{\mu}{\sigma}x_t\right]\mathrm{d}t + \mathrm{d}w_t \ , \quad x(0) = 0, \tag{3.68}$$

$$\mathrm{d}y_t = cx_t\mathrm{d}t + \mathrm{d}v_t \ , \qquad y(0) = 0, \tag{3.69}$$

where $w = \{w_t\}_{t\geq 0}$ and $v = \{v_t\}_{t\geq 0}$ are independent, one-dimensional Brownian motions with incremental variances $q(\theta) = \sigma^2$ and $r(\boldsymbol{o}) = \tau^2 + (\boldsymbol{o} - \boldsymbol{o}_*)^2$, respectively, for some fixed positive constant $\tau \in \mathbb{R}_+$. We assume that the initial condition $x_0 \in \mathbb{R}$, that the parameters $\mu, c \in \mathbb{R}$ and $\sigma \in \mathbb{R}_+$, respectively, and that the sensor location $\boldsymbol{o} \in \mathbb{R}$. We thus have a three-dimensional parameter vector $\theta = (\mu, c, \sigma) \in \mathbb{R}^2 \times \mathbb{R}_+$, and a single, one-dimensional sensor location $\boldsymbol{o} \in \mathbb{R}$.

This system has an analytic, finite-dimensional solution, known as the Beneš filter [39].

Namely, the conditional law of the latent signal process $x = \{x_t\}_{t\geq 0}$ given the history of observations $\mathcal{F}_t^Y = \sigma(y_s : 0 \leq s \leq t)$ is a weighted mixture of two normal distributions [19, Chapter 6], which takes the form

$$\begin{aligned}\pi_t = w_t^+(\theta,\boldsymbol{o})\mathcal{N}\left(\frac{A_t^+(\theta,\boldsymbol{o})}{2B_t(\theta,\boldsymbol{o})},\frac{1}{2B_t(\theta,\boldsymbol{o})}\right) \\ + w_t^-(\theta,\boldsymbol{o})\mathcal{N}\left(\frac{A_t^-(\theta,\boldsymbol{o})}{2B_t(\theta,\boldsymbol{o})},\frac{1}{2B_t(\theta,\boldsymbol{o})}\right),\end{aligned} \tag{3.70}$$

where

$$w_t^{\pm}(\theta,\boldsymbol{o}) = \frac{\exp\left(\frac{A_t^{\pm}(\theta,\boldsymbol{o})^2}{4B_t(\theta,\boldsymbol{o})}\right)}{\exp\left(\frac{A_t^{+}(\theta,\boldsymbol{o})^2}{4B_t(\theta,\boldsymbol{o})}\right)+\exp\left(\frac{A_t^{-}(\theta,\boldsymbol{o})^2}{4B_t(\theta,\boldsymbol{o})}\right)} \tag{3.71a}$$

$$A_t^{\pm}(\theta,\boldsymbol{o}) = \pm\frac{\mu}{\sigma} + cr^{-1}(\boldsymbol{o})\int_0^t \frac{\sinh(c\sigma r^{-\frac{1}{2}}(\boldsymbol{o})s)}{\sinh(c\sigma r^{-\frac{1}{2}}(\boldsymbol{o})t)}\mathrm{d}y_s \tag{3.71b}$$

$$B_t(\theta,\boldsymbol{o}) = \frac{cr^{-\frac{1}{2}}(\boldsymbol{o})}{2\sigma}\coth(c\sigma r^{-\frac{1}{2}}(\boldsymbol{o})t). \tag{3.71c}$$

It follows, in particular, that the optimal filter has a (non-unique) two-dimensional representation, which we will write as $M_t(\theta,\boldsymbol{o}) = (m_t(\theta,\boldsymbol{o}),P_t(\theta,\boldsymbol{o}))^T$. In this case, we choose to define

$$m_t(\theta,\boldsymbol{o}) = c\frac{A_t^{\pm}(\theta,\boldsymbol{o})\mp\frac{\mu}{\sigma}}{2B_t(\theta,\boldsymbol{o})} = c\sigma r^{-\frac{1}{2}}(\boldsymbol{o})\frac{\int_0^t\sinh\left(cr^{-\frac{1}{2}}(\boldsymbol{o})\sigma s\right)\mathrm{d}y_t}{\cosh\left(cr^{-\frac{1}{2}}(\boldsymbol{o})\sigma t\right)} \tag{3.72a}$$

$$P_t(\theta,\boldsymbol{o}) = \frac{1}{2B_t(\theta,\boldsymbol{o})} = \frac{\sigma r^{\frac{1}{2}}(\boldsymbol{o})}{c}\tanh\left(cr^{-\frac{1}{2}}(\boldsymbol{o})\sigma t\right). \tag{3.72b}$$

This choice implies that the finite-dimensional filter evolves according to an SDE of the required form, namely (e.g., [407])

$$\mathrm{d}M_t(\theta,\boldsymbol{o}) = \begin{pmatrix}-c^2r^{-1}(\boldsymbol{o})P_t(\theta,\boldsymbol{o})m_t(\theta,\boldsymbol{o})\\ \sigma^2-c^2r^{-1}(\boldsymbol{o})P_t^2(\theta,\boldsymbol{o})\end{pmatrix}\mathrm{d}t+\begin{pmatrix}cr^{-1}(\boldsymbol{o})P_t(\theta,\boldsymbol{o})\\ 0\end{pmatrix}\mathrm{d}y_t. \tag{3.73}$$

The equations for the tangent filters, namely $M_t^{\mu}(\theta,\boldsymbol{o})$, $M_t^{\sigma}(\theta,\boldsymbol{o})$, $M_t^{c}(\theta,\boldsymbol{o})$ and $M_t^{\boldsymbol{o}}(\theta,\boldsymbol{o})$, can then be obtained by (formal) differentiation of this equation with respect to the relevant variable. Illustratively, for the first parameter, we have

$$\mathrm{d}M_t^{\mu}(\theta,\boldsymbol{o}) = \begin{pmatrix}-c^2r^{-1}(\boldsymbol{o})P_t^{\mu}(\theta,\boldsymbol{o})m_t(\theta,\boldsymbol{o})\\ -c^2r^{-1}(\boldsymbol{o})P_t(\theta,\boldsymbol{o})m_t^{\mu}(\theta,\boldsymbol{o})\\ \\ -2c^2r^{-1}(\boldsymbol{o})P_t(\theta,\boldsymbol{o})P_t^{\mu}(\theta,\boldsymbol{o})\end{pmatrix}\mathrm{d}t+\begin{pmatrix}cr^{-1}(\boldsymbol{o})P_t^{\mu}(\theta,\boldsymbol{o})\\ \\ \\ 0\end{pmatrix}\mathrm{d}y_t. \tag{3.74}$$

We should remark that $m_t(\theta,\boldsymbol{o})$ and $P_t(\theta,\boldsymbol{o})$ do not correspond directly to the mean $\hat{x}_t(\theta,\boldsymbol{o})$ and variance $\hat{\Sigma}_t(\theta,\boldsymbol{o})$ of the optimal filter. However, these quantities can be computed as [407]

$$\hat{x}_t(\theta,\boldsymbol{o}) = m_t(\theta,\boldsymbol{o}) + \frac{\mu}{\sigma}P_t(\theta,\boldsymbol{o})\tanh\left(\frac{\mu}{\sigma}m_t(\theta,\boldsymbol{o})\right), \tag{3.75a}$$

$$\hat{\Sigma}_t(\theta,\boldsymbol{o}) = P_t(\theta,\boldsymbol{o}) + \frac{\mu^2}{\sigma^2}\left(1-\tanh^2(\frac{\mu}{\sigma}m_t(\theta,\boldsymbol{o}))\right)P_t^2(\theta,\boldsymbol{o}). \tag{3.75b}$$

We can then compute the conditional expectations

$$\hat{C}_t(\theta,\boldsymbol{o}) = \psi_C(\theta,\boldsymbol{o},M_t(\theta,\boldsymbol{o})) = c\hat{x}_t(\theta,\boldsymbol{o}), \tag{3.76a}$$

$$\hat{j}_t(\theta,\boldsymbol{o}) = \psi_j(\theta,\boldsymbol{o},M_t(\theta,\boldsymbol{o})) = \mathrm{Tr}\big[\hat{\Sigma}_t(\theta,\boldsymbol{o})\big]. \tag{3.76b}$$

It is now straightforward to obtain the explicit form of the two-timescale, joint online parameter estimation and optimal sensor placement algorithm for this system. In particular, we have

$$\mathrm{d}\mu_t = -\gamma_t^{1,\mu}c\hat{x}_t^{\mu}(\theta_t,\boldsymbol{o}_t)\left[c\hat{x}_t(\theta_t,\boldsymbol{o}_t)\mathrm{d}t - \mathrm{d}y_t\right] \tag{3.77a}$$

$$\mathrm{d}\sigma_t = -\gamma_t^{1,\sigma}c\hat{x}_t^{\sigma}(\theta_t,\boldsymbol{o}_t)\left[c\hat{x}_t(\theta_t,\boldsymbol{o}_t)\mathrm{d}t - \mathrm{d}y_t\right] \tag{3.77b}$$

$$\mathrm{d}c_t = -\gamma_t^{1,c}\left[\hat{x}_t(\theta_t,\boldsymbol{o}_t) + c\hat{x}_t^{c}(\theta_t,\boldsymbol{o}_t)\right]\left[c\hat{x}_t(\theta_t,\boldsymbol{o}_t)\mathrm{d}t - \mathrm{d}y_t\right] \tag{3.77c}$$

$$\mathrm{d}\boldsymbol{o}_t = -\gamma_t^{2,\boldsymbol{o}}\mathrm{Tr}\big[\hat{\Sigma}_t^{\boldsymbol{o}}(\theta_t,\boldsymbol{o}_t)\big]. \tag{3.77d}$$

where $\hat{x}^{\mu}$, $\hat{x}^{\sigma}$, $\hat{x}^{c}$ and $\hat{\Sigma}^{\boldsymbol{o}}$ are the filter derivatives of the posterior mean and the posterior variance, respectively. These quantities are obtained by differentiating (3.75a) - (3.75b) with respect to the relevant variable, and substituting the filter and the relevant tangent filter where appropriate.

Illustratively, we can compute the first of these quantities, by formal differentiation of equation (3.75a), as

$$\begin{aligned}\hat{x}_t^{\mu}(\theta_t,\boldsymbol{o}_t) &= m_t^{\mu}(\theta_t,\boldsymbol{o}_t) \\ &\quad+ \frac{1}{\sigma_t}P_t(\theta_t,\boldsymbol{o}_t)\tanh\left(\frac{\mu_t}{\sigma_t}m_t(\theta_t,\boldsymbol{o}_t)\right) \\ &\quad+ \frac{\mu_t}{\sigma_t}P_t^{\mu}(\theta_t,\boldsymbol{o}_t)\tanh\left(\frac{\mu_t}{\sigma_t}m_t(\theta_t,\boldsymbol{o}_t)\right) \\ &\quad+ \frac{\mu_t}{\sigma_t^2}P_t(\theta_t,\boldsymbol{o}_t)m_t(\theta_t,\boldsymbol{o}_t)\,\mathrm{sech}^2\left(\frac{\mu_t}{\sigma_t}m_t(\theta_t,\boldsymbol{o}_t)\right) \\ &\quad+ \frac{\mu_t^2}{\sigma_t^2}P_t(\theta_t,\boldsymbol{o}_t)m_t^{\mu}(\theta_t,\boldsymbol{o}_t)\,\mathrm{sech}^2\left(\frac{\mu_t}{\sigma_t}m_t(\theta_t,\boldsymbol{o}_t)\right),\end{aligned} \tag{3.78}$$

and where, for example, the filter derivative with respect to the first parameter evolves according to

$$\mathrm{d}m_t^{\mu}(\theta_t, \boldsymbol{o}_t) = - c^2 R^{-1}(\boldsymbol{o}_t) P_t^{\mu}(\theta_t, \boldsymbol{o}_t) m_t(\theta_t, \boldsymbol{o}_t)\mathrm{d}t \tag{3.79a}$$
$$- c^2 R^{-1}(\boldsymbol{o}_t) P_t(\theta_t, \boldsymbol{o}_t) m_t^{\mu}(\theta_t, \boldsymbol{o}_t)\mathrm{d}t$$
$$+ c R^{-1}(\boldsymbol{o}_t) P_t^{\mu}(\theta_t, \boldsymbol{o}_t)\mathrm{d}y_t$$
$$\mathrm{d}P_t^{\mu}(\theta_t, \boldsymbol{o}_t) = - 2c^2 R^{-1}(\boldsymbol{o}_t) P_t(\theta_t, \boldsymbol{o}_t) P_t^{\mu}(\theta_t, \boldsymbol{o}_t)\mathrm{d}t. \tag{3.79b}$$

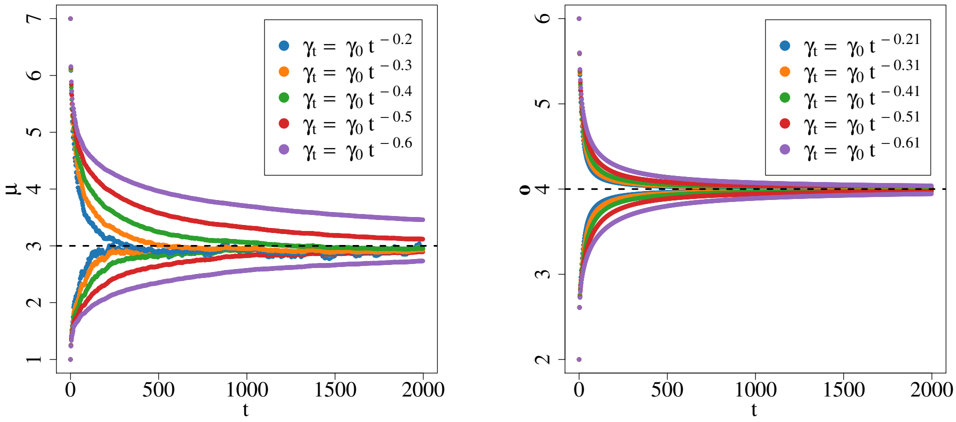
The performance of the two-timescale stochastic gradient descent algorithm is illustrated in Figure 3.1. In this simulation, we assume that the parameters $\sigma^2 = \sigma_*^2 = 4$, $c = c_* = 0.7$ and $\tau^2 = \tau_*^2 = 2$ are fixed, while the parameter μ is learned. The true value of this parameter is given by $\mu_* = 3$, and we consider two initial parameter estimates $\mu_0 = \{1, 7\}$. Meanwhile, the optimal sensor placement and the initial sensor placement is given by $\boldsymbol{o}_* = 4$, and we consider initial sensor placements $\boldsymbol{o}_0 = \{2, 6\}$. We remark that, as in any gradient based algorithm, the convergence of the proposed scheme may be sensitive to initialisation. In this case, however, it appears to be robust to this choice.

For simplicity, we choose to integrate all SDEs using a standard Euler-Maruyama discretisation, with $\Delta t = 0.01$, though similar results are obtained for other choices of Δt. We provide results for several choices of learning rates of the form $\gamma_t^{\mu} = \gamma_0^{\mu} t^{-\eta_{\mu}}$ and $\gamma_t^{o} = \gamma_0^{o} t^{-\eta_o}$, where $0 < \gamma_0^{\mu}, \gamma_0^{o} < \infty$, and $0 < \eta_{\mu} < \eta_o < 1$. We consider both learning rates for which the learning rate condition in Proposition 3.1 is satisfied (when $0.5 \leq \eta_{\mu}, \eta_o \leq 1$), and learning rates for which this condition is violated (when $0 < \eta_{\mu}, \eta_o < 0.5$). In this case, the online parameter estimates and optimal sensor converge to their true values, regardless of the choice of learning rate or the initialisation. We note, however, that the rate of convergence does depend on the choice of learning rate.

To obtain (optimal) convergence rates which are independent of the choice of learning rate, a standard approach in discrete time, including the two-timescale case [354], is to use Polyak-Ruppert averaging [13, 381, 401]. In the spirit of this scheme, in the continuous-time, two-timescale setting, we can consider new parameter estimates $\{\bar{\theta}_t\}_{t\geq 0}$ and $\{\bar{\boldsymbol{o}}_t\}_{t\geq 0}$ defined according to

$$\bar{\theta}_t = \frac{1}{t}\int_0^t \theta_s \mathrm{d}s \quad , \quad \bar{\boldsymbol{o}}_t = \frac{1}{t}\int_0^t \boldsymbol{o}_s \mathrm{d}s. \tag{3.80}$$

A detailed theoretical analysis of this approach, which extends the results in [354] to the continuous-time setting using the tools established in [371, 372, 420, 422], is beyond the scope of this chapter. We do provide tentative numerical evidence, however, to suggest that such results can also be expected to hold in continuous time. In particular, in Figure 3.2, we plot the sequence of averaged optimal sensor placements (Figure 3.2a) and the

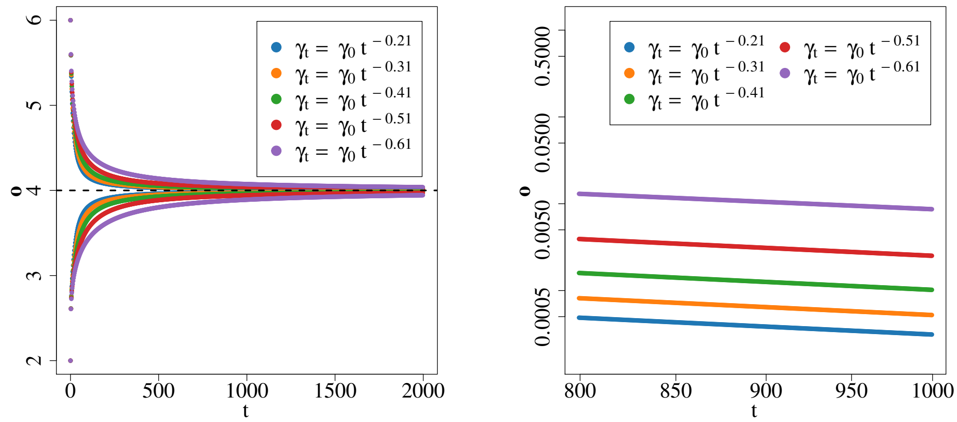


(a) Online parameter estimates.

(b) Optimal sensor placements.

Figure 3.1: The sequence of online parameter estimates & optimal sensor placements for the partially observed Beneš SDE, for several choices of the learning rates $\{\gamma_{\theta_t}\}_{t\geq0}$ and $\{\gamma_{\boldsymbol{o}_t}\}_{t\geq1}$.

corresponding $\mathbb{L}^1$ error for large times (Figure 3.2b), for several choices of the learning rate. The latter illustration, in particular, indicates that the convergence rate is now independent of the learning rate. One can obtain similar results for the corresponding sequence of averaged online parameter estimates (plots omitted).

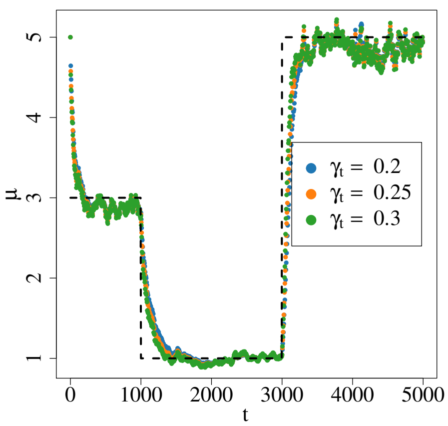


(a) Optimal sensor placements.

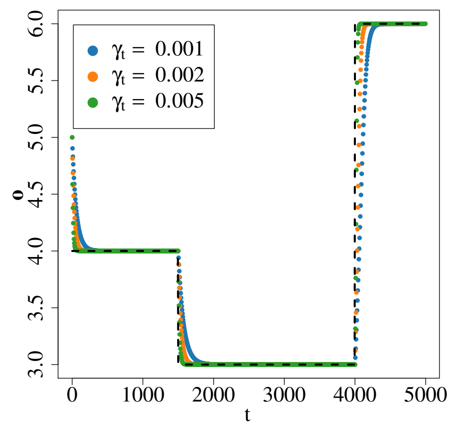
(b) $\mathbb{L}_1$ error (log-log plot).

Figure 3.2: The sequence of averaged optimal sensor placements for the partially observed Beneš SDE, for several choices of the learning rates $\{\gamma_t^{\theta}\}_{t\geq0}$ and $\{\gamma_t^{\boldsymbol{o}}\}_{t\geq1}$.

We conclude this section by investigating the performance of the stochastic gradient descent algorithm under the assumption that the true model parameters and the optimal sensor placements are no longer static, but now change in time. This is a scenario of particular practical interest. In this case, we must specify constant learning rates for both the parameter estimates and the sensor placements. While this violates the learning rate condition in Proposition 3.1, it is a standard choice when the model parameters are dynamic (e.g., [315]). In particular, although there is no longer any guarantee that that the algorithm iterates will converge to the stationary points of the two objective functions, they can be expected to oscillate around these points, with amplitude proportional to the learning rate. The performance of the algorithm is shown in Figure 3.3. As anticipated, the online parameter estimates (sensor placements) are able to track changes in the true model parameter (optimal sensor placement) in real time. It is worth noting that, while here we have considered the case in which the model parameters change discontinuously in time, we obtain similar results when the model parameters change continuously in time.



(a) Online parameter estimates.



(b) Optimal sensor placements.

Figure 3.3: The sequence of online parameter estimates & optimal sensor placements for the partially observed Beneš SDE, in the case of a time-varying parameter and optimal sensor location.

3.5 Conclusions

In this chapter, we have demonstrated in detail how a two-timescale stochastic gradient descent algorithm in continuous time can be applied to the problem of joint online parameter estimation and optimal sensor placement for a partially observed diffusion processes. Moreover, under suitable assumptions on the process consisting of the latent signal, the filter, and the filter derivatives, we have established a.s. convergence of this algorithm to the stationary points of the asymptotic log-likelihood and the asymptotic sensor placement objective function. Although we have focused on this specific application, it is important to emphasise that the proposed methodology is applicable to any problem involving the optimisation of two interdependent objective functions, either or both of which may depend on an ergodic diffusion process.

A natural extension of the work presented in this chapter is to consider the problem of optimal sensor scheduling (rather than optimal sensor placement). This, we recall, refers to determining the optimal subset of a finite number of measurement sensors at each instant in time. In this case, the picture is not quite as straightforward. In particular, the design variable $\boldsymbol{o}_t \in \{0,1\}^{n_o}$ is now discrete, being a vector of zeros and ones which represents which sensors are 'activated' at each time (e.g., [30, 211, 291]). As such, the gradient based methods proposed in this chapter are not immediately applicable. One can, however, recover a continuous version of this problem via the so-called control parametrisation enhancing transform (CPET) [178, 290, 291, 292, 300, 446], or via an appropriate relaxation (e.g., [233]). Thus, in principle, the joint online parameter estimation and optimal sensor scheduling problem should also be amenable to a two-timescale stochastic gradient descent scheme similar to the one proposed here.

Appendices

3.A Extensions to Proposition 3.1

In this Appendix, we discuss in more detail extensions to Proposition 3.1 in the case the Kushner-Stratonovich equation only admits a finite-dimensional recursive approximation. Under the stated assumption, Proposition 3.1 guarantees that[19]

$$\lim_{t\to\infty} \nabla_\theta \tilde{\mathcal{L}}^{(\text{filter})}(\theta_t, \boldsymbol{o}_t) = \lim_{t\to\infty} \nabla_\theta \tilde{\mathcal{J}}^{(\text{filter})}(\theta_t, \boldsymbol{o}_t) = 0 \ , \quad \text{a.s.} \tag{3.81}$$

where $\tilde{\mathcal{L}}^{(\text{filter})}(\theta, \boldsymbol{o})$ and $\tilde{\mathcal{J}}^{(\text{filter})}(\theta, \boldsymbol{o})$ are the representations of the asymptotic log-likelihood and the asymptotic sensor placement objective in terms of the approximate finite dimensional filter, and θ_t and $\boldsymbol{o}_t$ are the parameter estimates and optimal sensor placements generated by Algorithm (3.43a) - (3.43b). In the case that one only has access to an approximate filter, the functions $\tilde{\mathcal{L}}^{(\text{filter})}(\theta, \boldsymbol{o})$ and $\tilde{\mathcal{J}}^{(\text{filter})}(\theta, \boldsymbol{o})$ are only approximations of the true objective functions $\tilde{\mathcal{L}}(\theta, \boldsymbol{o})$ and $\tilde{\mathcal{J}}(\theta, \boldsymbol{o})$. As such, it would clearly be preferable to obtain a result of the form

$$\lim_{t\to\infty} \nabla_\theta \tilde{\mathcal{L}}(\theta_t, \boldsymbol{o}_t) = \lim_{t\to\infty} \nabla_\theta \tilde{\mathcal{J}}(\theta_t, \boldsymbol{o}_t) = 0 \ , \quad \text{a.s.} \tag{3.82}$$

For now, we will consider the slightly easier task of trying to obtain a result of the form

$$\lim_{t\to\infty} ||\nabla_\theta \tilde{\mathcal{L}}(\theta_t, \boldsymbol{o}_t)|| = \lim_{t\to\infty} ||\nabla_\theta \tilde{\mathcal{J}}(\theta_t, \boldsymbol{o}_t)|| = 0 \tag{3.83}$$

where the mode of convergence is to be specified. To make progress towards this goal, let us consider the simple decomposition

$$\begin{aligned} ||\nabla_\theta \tilde{\mathcal{L}}(\theta_t, \boldsymbol{o}_t)|| \leq\ & ||\nabla_\theta \tilde{\mathcal{L}}(\theta_t, \boldsymbol{o}_t) - \nabla_\theta \tilde{\mathcal{L}}^{(\text{filter})}(\theta_t, \boldsymbol{o}_t)|| \\ & + ||\nabla_\theta \tilde{\mathcal{L}}^{(\text{filter})}(\theta_t, \boldsymbol{o}_t)|| \end{aligned} \tag{3.84}$$

where, for the sake of brevity, we have now restricted our attention to the log-likelihood (an analogous decomposition holds for the asymptotic sensor placement objective function). Proposition 3.1 guarantees that the second term in this decomposition converges to zero a.s. as $t \to \infty$. It thus remains to bound the first term. Evidently this bound will depend on the properties of the filter, and vanishes if the filter is exact. To obtain such a bound,

[19] For the purpose of this discussion, we will ignore the projection device which ensures that the algorithm iterates remain within Θ and Ω, respectively.

we can write

$$||\nabla_\theta \tilde{\mathcal{L}}(\theta_t, \boldsymbol{o}_t) - \nabla_\theta \tilde{\mathcal{L}}^{(\text{filter})}(\theta_t, \boldsymbol{o}_t)|| \tag{3.85}$$

$$\leq ||\nabla_\theta \tilde{\mathcal{L}}(\theta_t, \boldsymbol{o}_t) - \frac{1}{t}\nabla_\theta \mathcal{L}_t(\theta_t, \boldsymbol{o}_t)|| \tag{3.86}$$
$$+ ||\frac{1}{t}\nabla_\theta \mathcal{L}_t(\theta_t, \boldsymbol{o}_t) - \frac{1}{t}\nabla_\theta \mathcal{L}_t^{(\text{filter})}(\theta_t, \boldsymbol{o}_t)||$$
$$+ ||\frac{1}{t}\nabla_\theta \mathcal{L}_t^{(\text{filter})}(\theta_t, \boldsymbol{o}_t) - \nabla_\theta \tilde{\mathcal{L}}^{(\text{filter})}(\theta_t, \boldsymbol{o}_t)||$$

where $\mathcal{L}_t(\theta, \boldsymbol{o})$ denotes the true log-likelihood at time t, and $\mathcal{L}_t^{(\text{filter})}(\theta, \boldsymbol{o})$ denotes the filter representation of the log-likelihood at time t. Under our assumptions, it is straightforward to show that the first term and the third term in this decomposition converge to zero a.s. as $t \to \infty$ (see also [430, Proposition 1]). We thus turn our attention to the central term. Using our previous expression for the log-likelihood function, c.f. (3.8), and its representation in terms of the approximate filter, c.f. (3.48), we have that

$$\frac{1}{t}\nabla_\theta \mathcal{L}_t(\theta, \boldsymbol{o}) - \frac{1}{t}\nabla_\theta \mathcal{L}_t^{(\text{filter})}(\theta, \boldsymbol{o}) \tag{3.87}$$

$$= \frac{1}{t}\nabla_\theta \left[\int_0^t R^{-1}(\boldsymbol{o})\hat{C}_s(\theta, \boldsymbol{o}) \cdot \mathrm{d}y_s - \frac{1}{2}\int_0^t ||R^{-\frac{1}{2}}(\boldsymbol{o})\hat{C}_s(\theta, \boldsymbol{o})||^2 \mathrm{d}s\right] \tag{3.88}$$

$$- \frac{1}{t}\nabla_\theta \left[\int_0^t R^{-1}(\boldsymbol{o})\psi_C(\theta, \boldsymbol{o}, M_s) \cdot \mathrm{d}y_s - \frac{1}{2}\int_0^t ||R^{-\frac{1}{2}}(\boldsymbol{o})\psi_C(\theta, \boldsymbol{o}, M_s)||^2 \mathrm{d}s\right]. \tag{3.89}$$

It follows, after some rearrangement of this expression, that given suitable bounds on the quantities $||\psi_C(\theta, \boldsymbol{o}, M_s) - \hat{C}_s(\theta, \boldsymbol{o})||$ and $||\psi_C^\theta(\theta, \boldsymbol{o}, M_s, M_s^\theta) - \hat{C}_s^\theta(\theta, \boldsymbol{o})||$, it will be possible to bound this term. In many cases (e.g., linear observations), this corresponds to bounds on $||\psi_x(\theta, \boldsymbol{o}, M_s) - \hat{x}_s(\theta, \boldsymbol{o})||$ and $||\psi_x^\theta(\theta, \boldsymbol{o}, M_s) - \hat{x}_s^\theta(\theta, \boldsymbol{o})||$, where, for example, ψ_x denotes the estimate of the conditional mean $\hat{x}$ in terms of the approximate filter. In general, it will be necessary to verify these bounds on a case by case basis. There are, however, some notable exceptions, including the Ensemble Kalman-Bucy Filter (EnKBF) (e.g., [141, 143]). Let us briefly demonstrate how existing results on this filter can be applied in our context. In order to simplify the presentation, in what follows we will assume that the observations are linear: that is, $C(\theta, \boldsymbol{o}, x) = C(\theta, \boldsymbol{o})x$.

Suppose that $(v_t^{i,N}, w_t^{i,N}, x_0^{i,N})_{i=1}^N$ are independent copies of (v_t, w_t, x_0). The EnKBF consists of N interacting particles $(x_t^{i,N}(\theta, \boldsymbol{o})_{i=1}^N$ which evolve according to following system of interacting stochastic differential equations

$$\mathrm{d}x_t^{i,N}(\theta, \boldsymbol{o}) = A(\theta, x_t^{i,N}(\theta, \boldsymbol{o}))\mathrm{d}t + B(\theta, x_t^{i,N}(\theta, \boldsymbol{o}))\mathrm{d}v_t^{i,N} \tag{3.90}$$
$$+ P_t^N(\theta, \boldsymbol{o})C^T(\theta, \boldsymbol{o})R^{-1}(\boldsymbol{o})(\mathrm{d}y_t - C(\theta, \boldsymbol{o})x_t^{i,N}(\theta, \boldsymbol{o})\mathrm{d}t - \mathrm{d}w_t^{i,N})$$

where

$$P_t^N(\theta, \boldsymbol{o}) = \frac{1}{N-1} \sum_{i=1}^{N} (x_t^{i,N}(\theta, \boldsymbol{o}) - m_t^N(\theta, \boldsymbol{o}))(x_t^{i,N}(\theta, \boldsymbol{o}) - m_t^N(\theta, \boldsymbol{o}))^T \tag{3.91}$$

$$m_t^N(\theta, \boldsymbol{o}) = \frac{1}{N} \sum_{i=1}^{N} x_t^{i,N}(\theta, \boldsymbol{o}).$$

represent the (empirical) filter estimates of the conditional covariance $\hat{\Sigma}_t(\theta, \boldsymbol{o})$ and the conditional mean $\hat{x}_t(\theta, \boldsymbol{o})$. Under additional assumptions (e.g., $A(\theta, \boldsymbol{o}, x)$ is linear), it is possible to show that, for all $p \geq 1$, and for sufficiently large N, (e.g., [143, Theorem 3.6])

$$\mathbb{E}\left[||m_t^N(\theta, \boldsymbol{o}) - \hat{x}_t(\theta, \boldsymbol{o})||^p\right]^{\frac{1}{p}} \leq \frac{K(p)}{N^{\frac{1}{2}}} \tag{3.92}$$

where $K(p) < \infty$ is a constant independent of N. Suppose, in addition, that one could establish a similar bound for the tangent EnKBF (this remains an open problem in the general case). Then, using these results, our existing assumptions (e.g., polynomial growth, uniformly bounded moments for the filter and the tangent filter), and the Hölder inequality, after some algebra one arrives at

$$\mathbb{E}\left[||\frac{1}{t}\nabla_\theta \mathcal{L}_t(\theta, \boldsymbol{o}) - \frac{1}{t}\nabla_\theta \mathcal{L}_t^{(\text{filter})}(\theta, \boldsymbol{o})||^p\right]^{\frac{1}{2p}} \leq \frac{K(p)}{N^{\frac{1}{2}}} \tag{3.93}$$

It follows, substituting this bound into (3.86), substituting (3.86) into (3.84), that

$$\lim_{t\to\infty} \mathbb{E}\left[||\nabla_\theta \tilde{\mathcal{L}}(\theta_t, \boldsymbol{o}_t)||^p\right]^{\frac{1}{2p}} \leq \frac{K(p)}{N^{\frac{1}{2}}} \tag{3.94}$$

One can follow the same argument to obtain an identical bound for $\nabla_\theta \tilde{\mathcal{J}}(\theta, \boldsymbol{o})$. It follows immediately from these bounds that the limit (3.83) holds in $\mathbb{L}^p$, for all $p \geq 1$, under the additional limit that $N \to \infty$ (i.e., as the number of particles goes to infinity). We remark that a rigorous result of this type has recently been established for (discrete-time) recursive maximum likelihood estimation in non-linear state-space models [435].

4

Online Parameter Estimation and Optimal Sensor Placement for the Partially Observed Stochastic Advection-Diffusion Equation

Summary. *In this chapter, we consider the problem of jointly performing online parameter estimation and optimal sensor placement for a partially observed infinite dimensional linear diffusion process. We present a principled solution to this problem in the form of a continuous-time, two-timescale stochastic gradient descent algorithm, which recursively seeks to maximise the log-likelihood with respect to the unknown model parameters, and to minimise the expected mean squared error of the hidden state estimate with respect to the sensor locations. This represents an extension of the algorithm introduced in the previous chapter to the case in which the hidden state is infinite dimensional. We also demonstrate in detail how to apply the proposed approach in the case that the hidden signal is governed by the two-dimensional stochastic advection-diffusion equation, and provide extensive numerical results demonstrating its efficacy in several cases of practical interest.*

4.1 Introduction

High dimensional spatio-temporal data are increasingly common across epidemiology, engineering, meteorology, environmental monitoring, and the applied sciences. In such applications, statistical spatio-temporal models are essential tools for performing inference and prediction. The phenomena of interest, even if defined continuously over space and in time, can typically only be measured at a limited number of spatial locations. Furthermore, they are often highly complex insofar as the dependence structure across time and space is non-trivial, non-separable, or non-stationary. It is often also the case that inference is desired at a very large number of spatial locations, that the data are obtained with significant observational uncertainty, and that there are missing observations at numerous spatial or temporal locations.

Various regimes have been proposed for modelling spatiotemporal processes (see also [125] for an excellent overview). Traditional approaches to this problem have focused on the geostatistical paradigm, which requires complete specification of the joint space-time covariance structure [122, 202, 326], or on the use of multivariate time series methods, which specify a set of spatially correlated time series [190, 280, 400, 469]. Alternatively, authors have considered time as an extra dimension, in which case standard spatial statistical techniques can be applied [124].

Each of these formulations is flawed, however. The first is limited by the relatively small known class of valid spatiotemporal covariance functions, despite some significant efforts at extension [186, 203, 394, 427], and by the inability of such covariance functions to realistically capture correlations in complex dynamical processes. The second approach doesn't provide for accurate predictions at unmonitored sites, due to the lack of a continuous spatial component in the model [469]. Perhaps more critically, it is difficult and costly to implement when the number of spatial locations is high. The final approach ignores the fundamental differences between space and time: in particular, time has a natural ordering, while space does not.

In the presence of complex temporal *and* spatial components, it is natural to combine the first two of these approaches, to obtain a statistical model that is both temporally dynamic and spatially descriptive. In the geostatistical setting, such models are traditionally referred to as space-time dynamic models [123, 469] or dynamical spatiotemporal models [125]. In such models, the temporal evolution of the spatiotemporal process of interest can be defined in various ways, common choices being simple random walk dynamics (e.g., [82, 227, 429]) or via a stochastic partial differential (or difference) equation (e.g., [72, 219, 220, 308, 417, 418, 428, 467, 468, 474].

In this chapter, we consider a dynamic spatio-temporal statistical model governed by a dissipative stochastic partial differential equation (SPDE), namely, the stochastic advection-diffusion equation. This equation, typically as part of a larger hierarchical model (see, e.g., [125]), is frequently used in environmental monitoring applications to model phenomena such as precipitation [71, 308, 418], air pollution [31, 309], chemical contamination of surface soil [353], groundwater flow [261, 455], and sediment transport [329].

In particular, we will consider the case in which the hidden state of interest is a space-time varying scalar field, $u_t(\boldsymbol{x})$, on some bounded two-dimensional domain $\Pi \subseteq \mathbb{R}^2$. This state is modelled using the stochastic advection-diffusion equation, which is given by

$$\frac{\partial u_t(\boldsymbol{x})}{\partial t} = -\boldsymbol{\mu}(\boldsymbol{x})^T \nabla u_t(\boldsymbol{x}) + \nabla \cdot \Sigma(\boldsymbol{x}) \nabla u_t(\boldsymbol{x}) - \zeta(\boldsymbol{x}) u_t(\boldsymbol{x}) + f_t(\boldsymbol{x}) + b(\boldsymbol{x}) \varepsilon_t(\boldsymbol{x}) \tag{4.1}$$

where $\boldsymbol{x} = (x_1, x_2)^T \in \Pi$, $\nabla = (\partial/\partial x_1, \partial/\partial x_2)^T$ is the gradient operator, $\nabla \cdot$ is the divergence operator, $f_t(\boldsymbol{x})$ is a deterministic forcing, and $\varepsilon_t(\boldsymbol{x})$ is a Gaussian noise process which is temporally white and spatially coloured. This might appear as a restrictive choice for the dynamics, but much of the subsequent methodology is generic, and could thus theoretically be applied to other models of interest (see Chapter 3 for a rigorous treatment). Moreover, this model results in a tractable but non-separable space-time covariance operator [418], and thus its spatiotemporal dynamics are interpretable for practitioners.

In addition, the terms in this equation can, if desired, be given a clear physical interpretation. In particular, the first term describes transport effects, also termed convection or advection, with $\boldsymbol{\mu}(\boldsymbol{x}) = (\mu_1(\boldsymbol{x}), \mu_2(\boldsymbol{x}))^T \in \mathbb{R}^2$ the drift or velocity field. The second term describes a possibly anisotropic diffusion, with $\Sigma(\boldsymbol{x}) = [\Sigma_{i,j}(\boldsymbol{x})]_{i,j=1,2} \in \mathbb{R}^{2\times 2}$ the diffusivity or diffusion matrix. This matrix can further be parametrised as (e.g., [418])

$$\Sigma^{-1}(\boldsymbol{x}) = \frac{1}{\rho_1^2(\boldsymbol{x})} \begin{pmatrix} \cos\alpha(\boldsymbol{x}) & \sin\alpha(\boldsymbol{x}) \\ -\gamma(\boldsymbol{x})\sin\alpha(\boldsymbol{x}) & \gamma\cos\alpha(\boldsymbol{x}) \end{pmatrix}^T \begin{pmatrix} \cos\alpha(\boldsymbol{x}) & \sin\alpha(\boldsymbol{x}) \\ -\gamma(\boldsymbol{x})\sin\alpha(\boldsymbol{x}) & \gamma\cos\alpha(\boldsymbol{x}) \end{pmatrix} \tag{4.2}$$

in which case $\rho_1(\boldsymbol{x}) \in \mathbb{R}_+$ can be viewed as the range, which determines the amount of diffusion; $\gamma(\boldsymbol{x}) \in \mathbb{R}_+$ as the anisotropic amplitude, which determines the amount of anisotropy; and $\alpha(\boldsymbol{x}) \in [0, \frac{\pi}{2}]$ as the anisotropic direction, which determines the direction of the anisotropy. In the case that $\gamma(\boldsymbol{x}) \equiv 1$, this matrix is symmetric, and the diffusion is isotropic. The third term describes damping, with $\zeta(\boldsymbol{x}) \in \mathbb{R}_+$ the damping rate, or damping coefficient. The fourth term $f_t(\boldsymbol{x})$ describes a deterministic forcing, i.e., a source or sink, while the final term $b(\boldsymbol{x})\varepsilon_t(\boldsymbol{x})$ describes a spatially weighted stochastic forcing.

We will assume, as in many typical applications, that the unknown state of the SPDE cannot be observed directly, but that it generates a continuous sequence of noisy observations via a finite set of measurement sensors. In the spirit of data assimilation and uncertainty quantification, we are then interested in determining the conditional distribution of the latent state, given the history of observations. Under the assumption that the model parameters are known, this distribution can be obtained directly via the infinite-dimensional Kalman-Bucy filter (see, e.g., [43, 127]). In practice, however, the model parameters are unknown, are must be estimated from the data (e.g., [18, 24, 63, 259]). Moreover, the locations of the measurement sensors are not fixed, and it may be possible to improve upon the state estimate by obtaining an optimal sensor placement (e.g., [41, 80, 128, 472]).

4.1.1 Literature Review

4.1.1.1 Parameter Estimation

The study of parameter estimation for stochastic partial differential equations, or distributed parameter systems, was initiated in the late 1960s (e.g., [259, 379]), and has since been the subject of numerous papers and several monographs (e.g., [28, 33, 113, 319] and references therein). Although the majority of literature on this subject has been written for fully observed processes, several authors have also consider the 'partially observed' case, in which observations of the infinite dimensional system are corrupted by some additional noise process [4, 5, 6, 7, 15, 18].

We will focus on parameter estimation methods based on the maximum likelihood principle. In the offline setting, the asymptotic properties (e.g., consistency, asymptotic normality) of such methods for partially observed linear distributed parameter systems have been relatively well studied (e.g., [7, 15, 18, 24, 247, 255, 352]). Perhaps somewhat surprisingly, a rigorous treatment of online maximum likelihood estimation in the infinite-dimensional setting remains an open problem.[1] While we will not attempt to review them here, it is worth noting that rigorous treatments of other approaches to online parameter estimation in infinite-dimensional systems have also been proposed (e.g., [33]).

4.1.1.2 Optimal Sensor Placement

In contrast to online parameter estimation, the problem of optimal sensor placement for state estimation in linear distributed parameter systems, has been studied by a large number of authors, and in a wide variety of contexts. The first comprehensive treatment of this problem was provided by Bensoussan [41, 43], who formulated it as an application of optimal control on the infinite dimensional Ricatti equation governing the covariance operator of the optimal filter. In particular, sensor locations were treated as control variables, and the performance index was taken as the sum of a measurement cost term and an accuracy cost term, the latter being the trace of the covariance operator at some terminal time. The solution was obtained in the form of a two-point boundary value

[1] We refer to Chapter 3 for a comprehensive survey of the relevant literature in the finite-dimensional setting.

problem, on the basis of which Bensoussan derived necessary and sufficient conditions for optimality by using the variational inequality. This extended the work of Athans [11], who obtained similar conditions in the finite dimensional case. Chen and Seinfeld [109], Omatu [368], and Curtain and Ichikawa [128] later also derived necessary and/or sufficient conditions for optimality from alternative perspectives. In particular, the first derived necessary conditions via an application of a minimum principle for distributed parameter systems. The second derived necessary and sufficient conditions, in the case of pointwise observations, using the existence and uniqueness theorem and the comparison theorem for partial differential equations of Riccati type. The third derived necessary conditions by introducing an appropriate Hilbert space formulation, and formulating the distributed parameter system in terms of mild evolution operators.

Following these early results, subsequent works focused largely on recursive computational methods which could be applied to determine the set of optimal sensor locations as the minima of the trace of the covariance operator of the optimal filter at some terminal time, or some variant thereof. Yu and Seinfeld [485] developed a sub-optimal technique for sequentially locating sensors in linear distributed parameter systems, which was extended by Chen and Seinfeld [109] to allow for the optimal simultaneous allocations of a finite number of a sensors among a set of locations given a priori. Subsequently, Aidarous et al. [2, 3] and Amouroux et al. [9] proposed gradient methods which could be applied for the optimal simultaneous allocation of a finite number of sensors to any set of locations in the domain of interest. Later computational efforts tended to focus on suboptimal, but more computationally tractable approaches [84, 262, 360, 368, 386], given the difficulty of solving the nonlinear infinite-dimensional Ricatti equation [408].

A common feature of all these computational methods was the requirement to transform the infinite dimensional system into a finite dimensional system. This was typically approached by the use of a truncated eigenvalue-eigenfunction expansion for either the state [89, 109, 485] or the state estimate [2, 3, 9] in its first K terms, sorted in increasing order of the eigenvalues.[2,3] The second of these approaches was largely motivated by the work of Curtain et al. [127, 128, 129]. By introducing an appropriate Hilbert space formu-

[2] In the former case, the idea was to represent the state variable as an infinite series of eigenfunctions of the relevant partial differential operator. This yielded an equivalent model described by an infinite sequence of ordinary differential equation in the time-varying coefficients of that expansion. This infinite sequence could then then approximated by an K-dimensional vector, by truncating in the first K terms. It was then possible to apply a finite dimensional filtering algorithm to this finite dimensional system, to obtain the a finite dimensional approximation of the Ricatti equation governing the filter covariance.

[3] In the latter case, the idea was to apply an infinite dimensional filtering algorithm to the original infinite dimensional system, to obtain the infinite dimensional state filtered estimate, and the infinite dimensional Ricatti equation governing the filter covariance. The filtered estimate could then be represented as an infinite series of eigenfunctions, yielding an equivalent estimate in terms of the infinite sequence of coefficients of that expansion. This infinite sequence could then be approximated by a K-dimensional vector, by truncating in the first K terms. This procedure could also be used to obtain a finite dimensional approximation of the covariance operator.

lation, Curtain showed that, for distributed parameter systems whose solutions can be expressed as a linear combination of eigenfunctions, one is justified in computing optimal locations using a suitable approximation of the covariance operator. In particular, it was proved that, under certain conditions, the trace of this finite dimensional approximation converges to the trace of the infinite dimensional covariance operator, and thus optimal sensor locations for the approximate finite dimensional system converge to the true optimal sensor locations for the finite dimensional system. We note that, using a slightly different operator formulation, Colantouni [114] also established conditions under which minimising the trace of the covariance operator is equivalent to minimising the trace of a finite dimensional approximate covariance matrix.

Recently, and in the spirit of Bensoussan's original approach, Burns et al. [77, 78, 79, 80, 216, 391] have provided a rigorous general framework for determining optimal location and trajectories of sensor networks for linear stochastic distributed parameter systems. The optimisation problem is precisely formulated as the minimisation of a functional involving the trace of a solution to the integral Ricatti equation, with constraints given by the trajectory of the sensor network [80]. The existence of Bochner integrable solutions to this equation, and thus the existence of optimal sensor locations, has been established [78, 79], and a Galerkin type numerical scheme for the approximation of these solutions developed, for which convergence is proved in $\mathbb{L}^p$ norm [80]. A suitable gradient descent scheme has also been proposed [80], and implemented in numerical examples for both stationary and moving sensor networks. In these authors' most recent contribution, the functional that penalises deviations with respect to the Kalman-Bucy filter is combined with a 'worst case scenario' functional, which involves a further optimisation problem for directional sensitivities over a set of admissible perturbations [216].

Morris et al. and Wu et al. [356, 445, 472, 488] have also recently revisited the problem of optimal sensor locations for linear stochastic, possibly time-varying, distributed parameter systems, and have developed a rigorous mathematical framework for this problem based on its duality with the problem of optimal actuator locations for linear control distributed parameter systems. In [472], Wu et al. consider the minimisation of the trace of the mild solution of the infinite dimensional Ricatti equation at some finite time as the sensor placement criterion, and prove the existence and convergence of optimal sensor locations for this problem. The obtained results were then applied to a simple advection-diffusion model. In [488], Zhang and Morris consider the minimisation of the trace of the mild solution of the infinite dimensional *algebraic* Ricatti equation as the sensor placement criterion. It is shown that the steady-state error variance is the trace of the solution of this equation, extending a well known result in the finite dimensional setting. The existence and convergence of optimal sensor locations for this problem is then derived, using results previously obtained for optimal actuator locations in [356]. The obtained results are applied to a number of simple examples, using the algorithm derived for the dual control problem in [135]. Tang and Morris have since extended this approach to the case that the form of the observation operator is not assumed fixed; that is, the shape as well as the location of the sensors is a design variable [445].

In recent years, significant contributions have also been made by Demetriou and coworkers (e.g.,[149, 150, 151, 152, 153, 154, 156]). These include the development of various schemes for the guidance of static and mobile sensors.

We conclude this section with the remark that the problem of optimal sensor placement for parameter estimation in distributed parameter systems has also been studied extensively. This is not the focus of the current study, and thus we will not review the literature on the topic in great depth. We do, however, remark that popular approaches to this problem originate in the classical theory of optimum experimental design [12, 121]. In this framework, the adopted optimisation criteria are essentially the same, i.e. the maximisation of various scalar measures of performance based on the Fisher information matrix (FIM) associated with the parameters to be estimated [454]. This approach dates back to the work of Rafajlowicz [385, 387]. For a comprehensive overview of this research area, we refer to [373, 453].

4.1.2 Contributions

In this chapter, we tackle the problem of *joint* online parameter estimation and optimal sensor placement for the partially observed stochastic advection-diffusion equation. There is clear motivation for this combined approach. In the vast majority of practical applications, both parameter estimation and optimal sensor placement are relevant. It would thus be highly convenient to solve them simultaneously, and, if possible, in an online fashion (i.e., in real time). Moreover, they are often interdependent, in the sense that the optimal sensor placement can vary significantly according to the current parameter estimate. Thus, tackling them together can result in significant performance improvements (see Figure 4.1).

To solve this problem, we propose a continuous-time, two-timescale stochastic gradient descent algorithm, formulated for the partially observed, infinite-dimensional linear diffusion process governed by the stochastic advection-diffusion equation. This algorithm can be seen as a formal extension of the algorithm in Chapter 3 to the setting in which the latent state is infinite-dimensional. We establish, using the theoretical results in Chapter 3, a.s. convergence of the online parameter estimates and recursive optimal sensor placements generated by a suitable finite dimensional approximation of this algorithm to the stationary points of the asymptotic log-likelihood and the asymptotic filter covariance, respectively. These results are obtained under a set of easy-to-verify sufficient conditions, specific to the linear Gaussian case.

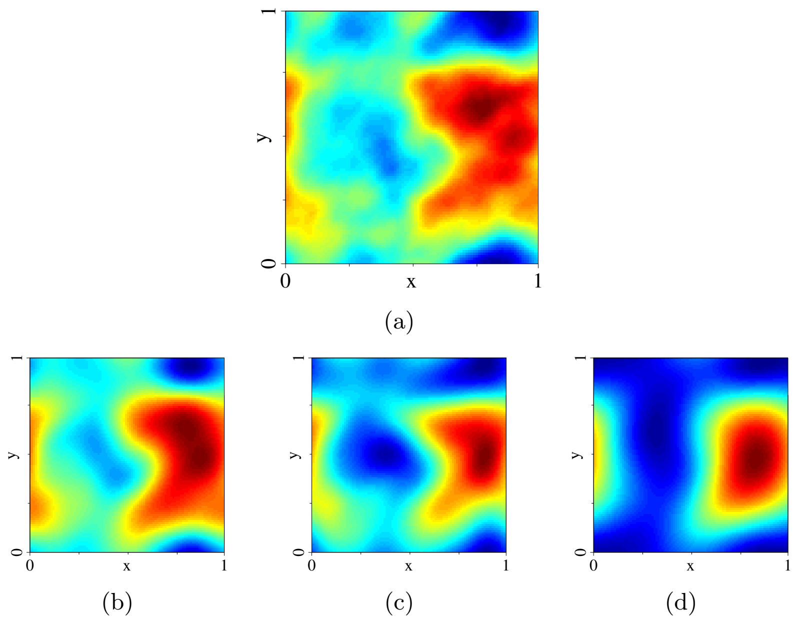


Figure 4.1: A comparison of the true, hidden state $u_t(\boldsymbol{x})$ (Fig. 4.1a) and the optimal state estimate $\hat{u}_t(\theta, \boldsymbol{o}, \boldsymbol{x})$ obtained in three possible scenarios: using the true parameters and an optimal sensor placement (Fig. 4.1b), using the true parameters but a sub-optimal sensor placement (Fig. 4.1c), and using an optimal sensor placement but the incorrect parameters (Fig. 4.1d). In this example, the hidden state is only accurately reconstructed in the first scenario, when the model parameters are successfully estimated **and** the sensors are optimally placed.

We then provide several detailed numerical case studies illustrating the performance of this method in different scenarios of practical interest. Our numerical results indicate that the algorithm is highly effective, and applicable to cases involving static and non-static model parameters, moving source terms, multiple noise and bias parameters, different specifications of the sensor placement objective function, and different specifications of the upper and lower-level objective functions.

4.1.3 Chapter Organisation

The remainder of this chapter is organised as follows. In Section 4.2, we precisely formulate the stochastic advection-diffusion equation as a functional stochastic differential equation on an appropriate separable Hilbert space. In Section 4.3, we present the two-timescale stochastic gradient descent algorithm for joint parameter estimation and optimal sensor placement. We present our methodology in a generic abstract framework, and thus in principle it could be applied to partially observed linear SPDEs other than the stochastic advection-diffusion equation. In Section 4.4, we provide several numerical examples illustrating the application of the proposed methodology. Finally, in Section 4.5, we provide some concluding remarks.

4.2 The Partially Observed Stochastic Advection-Diffusion Equation

In this section, we provide some background on the stochastic advection-diffusion partial differential equation. In particular, we now outline how this SPDE can be defined as a functional evolution equation on an appropriate separable Hilbert space (see also, e.g., [27, 80, 472]). We restrict our attention to the case of periodic boundary conditions following the treatment in [418]. This choice is largely motivated by expositional convenience, and the ability to efficiently perform numerical approximations using the Fast Fourier Transform (FFT). We should emphasise, however, that the joint online parameter estimation and optimal sensor placement algorithm subsequently introduced in Section 4.3 is generic, and does not rely on this assumption.

4.2.1 Preliminaries

We will suppose that the region of interest is the unit torus $\Pi := [0,1]^2$, with $\boldsymbol{x} = (x_1, x_2)^T \in \Pi$ a point on this space. We are interested in a space-time varying scalar field $u : \Pi \times [0,\infty) \rightarrow \mathbb{R}$, and will write $u_t(\boldsymbol{x})$ to denote the value of the field at spatial location $\boldsymbol{x} \in \Pi$ and time $t \in [0,\infty)$. We will assume that this field satisfies periodic boundary conditions. The function space of interest is thus given by $\mathcal{H} = L_2^{\text{per.}}(\Pi)$, the space of periodic square-integrable functions on $\Pi = [0,1]^2$.

It is natural to work with the Fourier characterisation of this space. In particular, suppose we write $\{\phi_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathbb{Z}^2}$ for the set of orthonormal Fourier basis functions for $\mathcal{H}$, namely, $\phi_{\boldsymbol{k}}(\boldsymbol{x}) = \exp(i\boldsymbol{k}^T\boldsymbol{x})$. We can then write

$$\mathcal{H} = \left\{\varphi : \varphi(\boldsymbol{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^2} \alpha_{\boldsymbol{k}}\phi_{\boldsymbol{k}}(\boldsymbol{x}) : \alpha_{-\boldsymbol{k}} = \overline{\alpha_{\boldsymbol{k}}}, \sum_{\mathbf{k}\in\mathbb{Z}^2} (\alpha_{\boldsymbol{k}})^2 < \infty\right\}. \tag{4.3}$$

4.2.2 The Signal Equation

Using standard results on infinite dimensional systems (e.g., [130, 369]), we can formulate the stochastic advection-diffusion partial differential equation (4.1) as a functional evolution equation on $\mathcal{H}$. Let $u_t = u_t(\cdot) = \{u_t(\boldsymbol{x}) : \boldsymbol{x} \in \Pi\} \in \mathcal{H}$ denote the state of the infinite-dimensional system. Then we can write

$$\mathrm{d}u_t = \mathcal{A}(\theta)u_t\mathrm{d}t + \mathcal{B}\mathrm{d}v_t^\theta, \quad u(0) = u_0 \in \mathcal{H} \tag{4.4}$$

where $\theta \in \Theta \subset \mathbb{R}^{n_\theta}$ is an n_θ-dimensional parameter, $\mathcal{A}(\theta)$ and $\mathcal{B}$ are abstract operators to be defined below, $v_t^\theta = v_t^\theta(\cdot) = \{v_t^\theta(\boldsymbol{x}) : \boldsymbol{x} \in \Pi\}$ is a space-time Brownian motion, and u_0

is a $\mathcal{H}$-valued Gaussian random variable, which is independent of v_θ. We are interested in weak solutions of this equation, to be understood path-wise on the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

The terms in this equation are defined explicitly as follows. Firstly, $\mathcal{A}(\theta) : \mathcal{D}(\mathcal{A}(\theta)) \rightarrow \mathcal{H}$ is the two-dimensional advection diffusion operator, defined according to

$$\mathcal{A}(\theta)\varphi = -\sum_{i=1}^{2} \mu_i(\boldsymbol{x}) \frac{\partial \varphi}{\partial x_i} + \sum_{i,j=1}^{2} \frac{\partial}{\partial x_i} \left(\Sigma_{i,j}(\boldsymbol{x}) \frac{\partial \varphi}{\partial x_j} \right) - \zeta(\boldsymbol{x})\varphi + f(\boldsymbol{x}) \ , \quad \varphi \in \mathcal{D}(\mathcal{A}(\theta)) \tag{4.5}$$

where $\mathcal{D}(\mathcal{A}(\theta)) = \{\varphi \in \mathcal{H} : \frac{\partial \varphi}{\partial x_i}, \frac{\partial^2 \varphi}{\partial x_i \partial x_j} \in L_2(\Pi) \ , i,j = 1,2\}$, and $\mu_i(\boldsymbol{x})$, $\Sigma_{i,j}(\boldsymbol{x})$, $\zeta(\boldsymbol{x})$, $f(\boldsymbol{x})$: $\Pi \rightarrow \mathbb{R}$, $i,j = 1,2$, are the real-valued, continuously differentiable functions defined in the introduction. For the chosen parametrisation of $\Sigma(\boldsymbol{x})$, the operator $-\mathcal{A}(\theta)$ is (strongly) elliptic of order 2 (e.g., [138, 376, 410]), bounded and coercive [156]. Moreover, $\mathcal{A}(\theta)$ generates an exponentially stable C_0-semigroup $\mathcal{S}(\theta, t) = e^{\mathcal{A}(\theta)t}$ over $L^2(\Pi)$ [80, 410].[4]

Meanwhile, $\mathcal{B} : \mathcal{H} \rightarrow \mathcal{H}$ is a spatial disturbance operator defined, for some spatial disturbance function $b \in \mathcal{H}$, via

$$\mathcal{B}\varphi = b(\boldsymbol{x})\varphi \ , \quad \varphi \in \mathcal{H}. \tag{4.6}$$

Finally, v_t^θ is an $\mathcal{H}$-valued Wiener process with incremental covariance operator $\mathcal{Q}(\theta)$: $\mathcal{H} \rightarrow \mathcal{H}$. We will assume that this covariance operator satisfies $\mathcal{Q}(\theta)\phi_{\boldsymbol{k}} = \eta_{\boldsymbol{k}}^2(\theta)\phi_{\boldsymbol{k}}$ for all $\boldsymbol{k} \in \mathbb{Z}^2$, for some bounded sequence of real numbers $\{\eta_{\boldsymbol{k}}^2(\theta)\}_{\boldsymbol{k} \in \mathbb{Z}^2}$ satisfying $\eta_{-\boldsymbol{k}}(\theta) = \eta_{\boldsymbol{k}}(\theta)$ and $\sum_{\boldsymbol{k}} \eta_{\boldsymbol{k}}^2(\theta) < \infty$.[5] We thus work with a diagonal covariance operator with respect to the Fourier basis, although other choices could easily be considered. It follows from standard results (see, e.g., [129, 131]) that

$$v_t^\theta := \sum_{\boldsymbol{k} \in \mathbb{Z}^2} \eta_{\boldsymbol{k}}(\theta)\phi_{\boldsymbol{k}}(\boldsymbol{x}) z_t^{\boldsymbol{k}}, \tag{4.7}$$

where $\{z_t^{\boldsymbol{k}}\}_{\boldsymbol{k} \in \mathbb{Z}}$ are a set of suitably defined independent Brownian motions (see, e.g., [316] for a precise definition). Following [302, 418], we will assume that $\{\eta_{\boldsymbol{k}}^2(\theta)\}_{\boldsymbol{k} \in \mathbb{Z}^2}$ are defined by

$$\eta_{\boldsymbol{k}}(\theta) = \frac{\sigma}{2\pi} \left(\boldsymbol{k}^T \boldsymbol{k} + \frac{1}{\rho_0^2} \right)^{-\nu}, \tag{4.8}$$

where $\sigma > 0$ is a marginal variance parameter, $\rho_0 > 0$ is a spatial range parameter, and $\nu > 0$ is a smoothness parameter.[6] This yields a noise process with the Matérn

[4] For ease of exposition, we have assumed here that $\mathcal{A}(\theta)$ is time-invariant. It is straightforward, however, to extend all of the results in this chapter to the case of a time-dependent operator $\mathcal{A}(\theta, t)$. In this case, we would also require some standard assumptions on the regularity of the map $t \rightarrow \mathcal{A}(\theta, t)$ (e.g., [303, 376, 444]).

[5] We will additionally require that $\eta_{\boldsymbol{k}}^2(\theta) = \mathcal{O}(|\boldsymbol{k}|^{-2(1+\varepsilon)})$ for some $\varepsilon > 0$. This ensures that $\sum_{\boldsymbol{k}} \eta_{\boldsymbol{k}}(\theta)^2 < \infty$, so that the operator $\mathcal{Q}(\theta)$ is of trace class.

[6] It is straightforward to verify that, in this case, $\eta_{\boldsymbol{k}}^2(\theta) = \mathcal{O}(|\boldsymbol{k}|^{-2(1+\varepsilon)})$ for some $\varepsilon > 0$.

covariance function in space, which is perhaps the most widely used covariance function in spatial statistics [124, 207, 426]. In many applications, the smoothness parameter is difficult to estimate [302, 418]. Thus, as in [418], we will henceforth assume that $\nu = 1$. This particular choice corresponds to the so-called Whittle covariance function in space, which can arguably be regarded as 'the elementary correlation in two dimensions' [466]. In principle, however, other values of ν could also be considered.

4.2.3 The Spectral Signal Equation

Using the Fourier characterisation, it is possible to write the solution of the signal equation as

$$u_t = \sum_{\mathbf{k}\in\mathbb{Z}^2} \alpha_t^{\boldsymbol{k}} \phi_{\boldsymbol{k}} \ , \quad \alpha_t^{\boldsymbol{k}} = \langle u_t, \phi_{\boldsymbol{k}} \rangle = \int_{\Pi} u_t(\boldsymbol{x}) \overline{\phi_{\boldsymbol{k}}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}. \tag{4.9}$$

It is thus equivalent to consider the parametrisation of u_t via the set of Fourier coefficients $\{\alpha_t^{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathbb{Z}^2}$. Taking the inner product of both sides of the signal equation with $\phi_{\boldsymbol{k}}$, we see that the $\alpha_t^{\boldsymbol{k}}$'s obey the following infinite dimensional stochastic differential equation

$$\mathrm{d}\alpha_t^{\boldsymbol{k}} = \sum_{\boldsymbol{j}\in\mathbb{Z}^2} \lambda_{\boldsymbol{j},\boldsymbol{k}}(\theta) \alpha_t^{\boldsymbol{j}} \mathrm{d}t + \sum_{\boldsymbol{j}\in\mathbb{Z}^2} \xi_{\boldsymbol{j},\boldsymbol{k}} \eta_{\boldsymbol{j}}(\theta) \mathrm{d}z_t^{\boldsymbol{j}} \ , \quad \boldsymbol{k} \in \mathbb{Z}^2, \tag{4.10}$$

where $\lambda_{\boldsymbol{j},\boldsymbol{k}}(\theta) = \langle \mathcal{A}(\theta)\phi_{\boldsymbol{j}}, \phi_{\boldsymbol{k}} \rangle$ and $\xi_{\boldsymbol{j},\boldsymbol{k}} = \langle \mathcal{B}\phi_{\boldsymbol{j}}, \phi_{\boldsymbol{k}} \rangle$. We will sometimes refer to this as the 'spectral' signal equation. This parametrisation of the signal process is highly convenient, as it allows us to perform inference on a vector whose coordinates evolve according to a SDE, even if this vector happens to have infinite length. In our numerical simulations, we will restrict attention to the case in which the advection-diffusion operator $\mathcal{A}(\theta)$ is spatially invariant, in which case we obtain a particularly simple form for the $\lambda_{\boldsymbol{j},\boldsymbol{k}}(\theta)$, namely,

$$\lambda_{\boldsymbol{j},\boldsymbol{k}}(\theta) = -\left(i\boldsymbol{j}^T\boldsymbol{\mu} + \boldsymbol{j}^T\Sigma\boldsymbol{j} + \zeta\right)\delta_{\boldsymbol{j},\boldsymbol{k}}, \tag{4.11}$$

where $\delta_{\boldsymbol{j},\boldsymbol{k}}$ denotes the standard Kronecker delta function. We will also assume, unless otherwise stated, that the spatial weighting operator $\mathcal{B}$ is the identity operator, in which case we also have $\xi_{\boldsymbol{j},\boldsymbol{k}} = \delta_{\boldsymbol{j},\boldsymbol{k}}$. Under these assumptions, the spectral signal equation (4.10) diagonalises completely; that is, the $\alpha_t^{\boldsymbol{k}}$'s evolve independently of one another.

4.2.4 The Observation Equation

We will assume, as in many typical applications, that the signal process cannot be observed directly, but that instead we obtain a continuous sequence of noisy observations $y = \{y_t\}_{t\geq 0}$, taking values in $\mathbb{R}^{n_y}$, via a set of n_y sensors located at $\boldsymbol{o} = \{\boldsymbol{o}_i\}_{i=1}^{n_y} \in \Pi^{n_y}$. In particular, we will assume that the observations are generated according to

$$\mathrm{d}y_t = \mathcal{C}(\theta, \boldsymbol{o})u_t \mathrm{d}t + \mathrm{d}w_t^{\boldsymbol{o}} \ , \quad y(0) = 0, \tag{4.12}$$

where $\mathcal{C}(\theta, \boldsymbol{o}) : \mathcal{H} \to \mathbb{R}^{n_y}$ is a bounded linear operator to be specified below, and $w_t^{\boldsymbol{o}}$ is a $\mathbb{R}^{n_y}$ valued Wiener process with incremental covariance $\mathcal{R}(\boldsymbol{o}) \in \mathbb{R}^{n_y \times n_y}$, which is independent of both v_θ and u_0. While the use of a linear observation equation is somewhat restrictive, it does encompass most typical observation schemes used in practice.

We suppose that each sensor provides a noisy, possibly biased, average of the latent signal around its current location, in which case the observation operator takes the form

$$\mathcal{C}(\theta, \boldsymbol{o})\varphi = \begin{pmatrix} \mathcal{C}_1(\theta, \boldsymbol{o}_1)\varphi \\ \vdots \\ \mathcal{C}_{n_y}(\theta, \boldsymbol{o}_{n_y})\varphi \end{pmatrix} \ , \quad \mathcal{C}_i(\theta, \boldsymbol{o}_i)\varphi = \frac{\int_\Pi \mathcal{K}_{\boldsymbol{o}_i}(\boldsymbol{x})\varphi(\boldsymbol{x})\mathrm{d}\boldsymbol{x}}{\int_\Pi \mathcal{K}_{\boldsymbol{o}_i}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}} + \beta_i \ , \quad \varphi \in \mathcal{H} \tag{4.13}$$

where $K_{\boldsymbol{o}_i} : \Pi \to \Pi$ are suitably chosen weighting functions, which decrease as $|\boldsymbol{x} - \boldsymbol{o}_i|$ increases, and $\beta_i \in \mathbb{R}$ are bias terms. In our numerics, we restrict our attention to the case in which each sensor provides an unweighted average of the latent signal process within a small fixed region of its current location (e.g., [80, 316]). This corresponds to the choice

$$\mathcal{K}_{\boldsymbol{o}_i}(\boldsymbol{x}) = \mathbb{1}\left\{\boldsymbol{x} \in \Pi : |\boldsymbol{x} - \boldsymbol{o}_i| \leq r\right\}, \quad r > 0. \tag{4.14}$$

For simplicity, we will also assume that the sensors are independent, in which case the covariance matrix $\mathcal{R}(\boldsymbol{o})$ reduces to a diagonal matrix; and that the sensors can be categorised into p_1 distinct 'noise classes', and into p_2 distinct 'bias' classes, where $1 \leq p_1, p_2 \leq n_y$. By this, we mean that all observations generated by sensors belonging to a particular class have the same variance (or the same bias).

4.3 Joint Online Parameter Estimation and Optimal Sensor Placement

In this section, we present the joint online parameter estimation and optimal sensor placement algorithm for a generic partially observed infinite dimensional linear diffusion process. The material presented in this section follows closely the material presented in Section 3.2, adapted appropriately to the infinite-dimensional setting.

4.3.1 Preliminaries

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space together with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ which satisfies the usual conditions. Let $\mathcal{G}$, $\mathcal{H}$ be separable Hilbert spaces. We will write $\mathcal{L}_1(\mathcal{G}, \mathcal{H})$ to denote the space of bounded linear operators from $\mathcal{G}$ to $\mathcal{H}$, and $\mathcal{L}_1(\mathcal{H})$ in the case $\mathcal{G} = \mathcal{H}$. We consider the following family of partially observed infinite dimensional linear diffusion processes:

$$\mathrm{d}u_t = \mathcal{A}(\theta)u_t \mathrm{d}t + \mathcal{B}\mathrm{d}v_t^{\theta} \ , \quad u_0 \in \mathcal{H}, \tag{4.15}$$

$$\mathrm{d}y_t = \mathcal{C}(\theta, \boldsymbol{o})u_t \mathrm{d}t + \mathrm{d}w_t^{\boldsymbol{o}} \ , \quad y_0 = 0, \tag{4.16}$$

where $\theta \in \Theta \subset \mathbb{R}^{n_\theta}$ is an n_θ-dimensional parameter, $\boldsymbol{o} = \{\boldsymbol{o}_i\}_{i=1}^{n_y} \in \Pi^{n_y} \subset \mathbb{R}^{2n_y}$ is a set of n_y sensor locations, with $\boldsymbol{o}_i \in \Pi \subseteq \mathbb{R}^2$ for $i = 1, \ldots, n_y$, $u = \{u_t\}_{t\geq 0}$ denotes the latent $\mathcal{H}$-valued signal process, $y = \{y_t\}_{t\geq 0}$ denotes the $\mathbb{R}^{n_y}$-valued observation process, and $v_\theta = \{v_t^\theta\}_{t\geq 0}$, $w_{\boldsymbol{o}} = \{w_t^{\boldsymbol{o}}\}_{t\geq 0}$ are independent Wiener processes, with incremental covariances $\mathcal{Q}(\theta) \in \mathcal{L}_1(\mathcal{H})$, $\mathcal{R}(\boldsymbol{o}) \in \mathcal{L}_1(\mathbb{R}^{n_y})$, which correspond to the signal noise and the measurement noise, respectively.

We assume that, for all $\theta \in \Theta$, $\mathcal{A}(\theta) : \mathcal{H} \to \mathcal{H}$ is the infinitesimal generator of a C_0-semigroup $\mathcal{S}_t(\theta)$ on $\mathcal{H}$. We also assume that, for all $\theta \in \Theta$, $\boldsymbol{o} \in \Omega^{n_y}$, $\mathcal{B} : \mathcal{H} \to \mathcal{H}$ and $\mathcal{C}(\theta, \boldsymbol{o}) : \mathcal{H} \to \mathbb{R}^{n_y}$ are bounded linear operators: $\mathcal{B} \in \mathcal{L}_1(\mathcal{H})$ and $\mathcal{C}(\theta, \boldsymbol{o}) \in \mathcal{L}_1(\mathcal{H}, \mathbb{R}^{n_y})$. Finally, we assume that the initial state u_0 is a $\mathcal{H}$-valued Gaussian random variable with mean $\hat{u}_0(\theta) \in \mathcal{H}$ and covariance $\Sigma_0(\theta) \in \mathcal{L}_1(\mathcal{H})$, which is independent of v_θ and $w_{\boldsymbol{o}}$ for all $\theta \in \Theta$, $\boldsymbol{o} \in \Omega^{n_y}$. Clearly, this abstract framework includes the partially observed stochastic advection-diffusion equation defined in Section 4.2.

4.3.2 The Infinite-Dimensional Kalman-Bucy Filter

We begin with a brief review of the infinite dimensional linear filtering problem. That is, the problem of determining the conditional law of the latent signal process, given the history of observations $\mathcal{F}_t^Y = \sigma\{y(s) : 0 \leq s \leq t\}$. In the linear Gaussian case, it is well known that that conditional distribution of the latent signal process is Gaussian, and thus determined uniquely by its mean and covariance. These quantities can be obtained explicitly via the infinite-dimensional Kalman-Bucy filter (e.g., [43, 127, 129]).

In particular, suppose we write $\hat{u}(\theta, \boldsymbol{o}) = \{\hat{u}_t(\theta, \boldsymbol{o})\}_{t\geq 0}$ to denote the conditional mean of the signal given $\mathcal{F}_t^Y$, and $\Sigma(\theta, \boldsymbol{o}) = \{\Sigma_t(\theta, \boldsymbol{o})\}_{t\geq 0}$ its conditional covariance. Then $\Sigma_t(\theta, \boldsymbol{o})$ is a weak solution of the operator Ricatti equation [129, Theorem 6.10]

$$\begin{aligned}\dot{\Sigma}_t(\theta, \boldsymbol{o}) &= \mathcal{A}(\theta)\Sigma_t(\theta, \boldsymbol{o}) + \Sigma_t(\theta, \boldsymbol{o})\mathcal{A}^*(\theta) \\ &\quad + \mathcal{B}\mathcal{Q}(\theta)\mathcal{B}^* - \Sigma_t(\theta, \boldsymbol{o})\mathcal{C}^*(\theta, \boldsymbol{o})\mathcal{R}^{-1}(\boldsymbol{o})\mathcal{C}(\theta, \boldsymbol{o})\Sigma_t(\theta, \boldsymbol{o}),\end{aligned} \tag{4.17}$$

and $\hat{u}_t(\theta, \boldsymbol{o})$ is a mild solution of the stochastic evolution equation [129, Theorem 6.21]

$$\begin{aligned} \mathrm{d}\hat{u}_t(\theta,\boldsymbol{o}) &= \mathcal{A}(\theta)\hat{u}_t(\theta,\boldsymbol{o})\mathrm{d}t \\ &\quad + \Sigma_t(\theta,\boldsymbol{o})\mathcal{C}^*(\theta,\boldsymbol{o})\mathcal{R}^{-1}(\boldsymbol{o})\big(\mathrm{d}y_t - \mathcal{C}(\theta,\boldsymbol{o})\hat{u}_t(\theta,\boldsymbol{o})\mathrm{d}t\big) \end{aligned} \tag{4.18}$$

We refer to [42, 129] for a precise definition of 'weak' and 'mild' solutions. It is worth noting that the differential form for $\hat{u}_t(\theta, \boldsymbol{o})$ only holds under certain additional technical assumptions on the noise processes (see [128, Theorem 6.21] for details). More generally, the optimal estimator $\hat{u}_t(\theta, \boldsymbol{o})$ is obtained as the solution of the following integral equation [128, Corollary 6.11])

$$\hat{u}_t(\theta,\boldsymbol{o}) = \mathcal{S}_t(\theta)\hat{u}_0(\theta,\boldsymbol{o}) + \int_0^t \mathcal{S}_t(\theta)\Sigma_s(\theta,\boldsymbol{o})\mathcal{C}^*(\theta,\boldsymbol{o})\mathcal{R}^{-1}(\boldsymbol{o})\mathrm{d}\rho_s(\theta,\boldsymbol{o}), \tag{4.19}$$

where $\rho(\theta, \boldsymbol{o}) = \{\rho_t(\theta, \boldsymbol{o})\}_{t\geq 0}$ is the so-called innovations process, defined according to $\rho_t(\theta, \boldsymbol{o}) = y_t - \int_0^t C(\boldsymbol{o})\hat{u}_s(\theta, \boldsymbol{o})\mathrm{d}s$.

In what follows, it will also be useful to define $\hat{u}^{\theta}(\theta, \boldsymbol{o}) = \{\hat{u}^{\theta}_t(\theta, \boldsymbol{o})\}_{t\geq 0}$ and $\hat{u}^{\boldsymbol{o}}(\theta, \boldsymbol{o}) = \{\hat{u}^{\boldsymbol{o}}_t(\theta, \boldsymbol{o})\}_{t\geq 0}$ as the 'filter derivatives' of the conditional mean, and $\Sigma^{\theta}(\theta, \boldsymbol{o}) = \{\Sigma^{\theta}_t(\theta, \boldsymbol{o})\}_{t\geq 0}$ and $\Sigma^{\boldsymbol{o}}(\theta, \boldsymbol{o}) = \{\Sigma^{\boldsymbol{o}}_t(\theta, \boldsymbol{o})\}_{t\geq 0}$, to denote the 'filter derivatives' of the conditional covariance, respectively. By this we mean, for example, that $\hat{u}^{\theta}(\theta, \boldsymbol{o}) = \{\hat{u}^{\theta}_t(\theta, \boldsymbol{o})\}_{t\geq 0}$ is the process defined, for all $t \geq 0$, according to $\hat{u}^{\theta}_t(\theta, \boldsymbol{o}) = \nabla_{\theta}\hat{u}_t(\theta, \boldsymbol{o})$. By definition, these quantities are the solutions, interpreted in the appropriate sense, of the equations obtained upon formal differentiation of equations (4.17) - (4.18). For brevity, the explicit forms of these equations are omitted.

Using the infinite-dimensional Kalman-Bucy filter, we can now obtain the log-likelihood of the observations (e.g., [5, 18, 22, 23, 304]), and define the optimal sensor placement objective function (e.g., [80, 109, 211, 254]) as

$$\mathcal{L}_t(\theta,\boldsymbol{o}) = \int_0^t \big\langle \mathcal{R}^{-1}(\boldsymbol{o})\mathcal{C}(\theta,\boldsymbol{o})\hat{u}_s(\theta,\boldsymbol{o}), \mathrm{d}y(s)\big\rangle - \frac{1}{2}\int_0^t ||\mathcal{R}^{-\frac{1}{2}}(\boldsymbol{o})\mathcal{C}(\theta,\boldsymbol{o})\hat{u}_s(\theta,\boldsymbol{o})||^2\mathrm{d}s, \tag{4.20a}$$

$$\mathcal{J}_t(\theta,\boldsymbol{o}) = \int_0^t \mathrm{Tr}\left[\mathcal{M}_s\Sigma_s(\theta,\boldsymbol{o})\right]\mathrm{d}s \tag{4.20b}$$

where $\mathcal{M}_s \in \mathcal{L}_1(\mathcal{H})$ is a bounded, possibly time-varying linear operator designed to weight significant parts of the state estimate. In the online setting, as discussed in Chapter 3, we are interested in optimising the asymptotic variants of these two functions, viz

$$\tilde{\mathcal{L}}(\theta,\boldsymbol{o}) = \lim_{t\to\infty}\frac{1}{t}\mathcal{L}_t(\theta,\boldsymbol{o}) \quad , \quad \tilde{\mathcal{J}}(\theta,\boldsymbol{o}) = \lim_{t\to\infty}\frac{1}{t}\mathcal{J}_t(\theta,\boldsymbol{o}) \tag{4.21}$$

4.3.3 Joint Online Parameter Estimation and Optimal Sensor Placement

We will achieve this, as in the finite-dimensional case, using a continuous-time, two-timescale stochastic gradient descent algorithm. In particular, suppose some initialisation at $\theta_0 \in \Theta$, $\boldsymbol{o}_0 \in \Omega^{n_y}$. Then, simultaneously, the parameter estimates $\{\theta_t\}_{t\geq 0}$ and the sensor locations $\{\boldsymbol{o}_t\}_{t\geq 0}$ are generated according to

$$\mathrm{d}\theta_t = \begin{cases} -\gamma_t^{\theta}\big[\mathcal{C}(\theta,\boldsymbol{o})\hat{u}_t^{\theta}(\theta,\boldsymbol{o})\big]^T\mathcal{R}^{-1}(\boldsymbol{o})\big[\mathcal{C}(\theta,\boldsymbol{o})\hat{u}_t(\theta,\boldsymbol{o})\mathrm{d}t - \mathrm{d}y_t\big]\big|_{\substack{\theta=\theta_t \\ \boldsymbol{o}=\boldsymbol{o}_t}} & , \quad \theta_t \in \Theta, \\ 0 & , \quad \theta_t \notin \Theta, \end{cases} \tag{4.22a}$$

$$\mathrm{d}\boldsymbol{o}_t = \begin{cases} -\gamma_t^{\boldsymbol{o}}\mathrm{Tr}^{\boldsymbol{o}}\big[\mathcal{M}(t)\Sigma_t(\theta,\boldsymbol{o})\big]^T\mathrm{d}t\big|_{\substack{\theta=\theta_t \\ \boldsymbol{o}=\boldsymbol{o}_t}} & , \quad \boldsymbol{o}_t \in \Omega^{n_y}, \\ 0 & , \quad \boldsymbol{o}_t \notin \Omega^{n_y}. \end{cases} \tag{4.22b}$$

where $\{\gamma_t^{\theta}\}_{t\geq 0}$ and $\{\gamma_t^{\boldsymbol{o}}\}_{t\geq 0}$ are non-negative, non-increasing real functions such that $\lim_{t\to\infty}\gamma_t^{\theta}/\gamma_t^{\boldsymbol{o}} = 0$ or $\lim_{t\to\infty}\gamma_t^{\boldsymbol{o}}/\gamma_t^{\theta} = 0$. The choice between these two conditions on the learning rates determines which of the algorithm iterates moves on a slower timescale. The first choice implies that the parameter estimates move on a slower timescale than the sensor placements, and is generally preferred if parameter estimation is the primary objective. The second implies that the sensor placements move on a slower timescale than the parameter estimates, and is generally preferred if optimal sensor placement is the primary objective.

This algorithm represents a formal extension of Algorithm (3.43a) - (3.43b) in Chapter 3 to the infinite-dimensional linear Gaussian setting. Let us briefly provide some justification for this extension. Firstly, under weak conditions, the log-likelihood for a partially observed infinite-dimensional linear diffusion process is both well-defined and consistent (e.g., [22, Theorem 8.4] or [15, Theorem 2] in the general case, and [7, Corollary 4.1] for hyperbolic systems). Meanwhile, assuming that the observation operator is continuous with respect to the sensor locations, the sensor placement objective function admits an infimum, i.e., an optimal sensor placement ([80, Theorem 5.3] and [488, Theorem 4.1]).

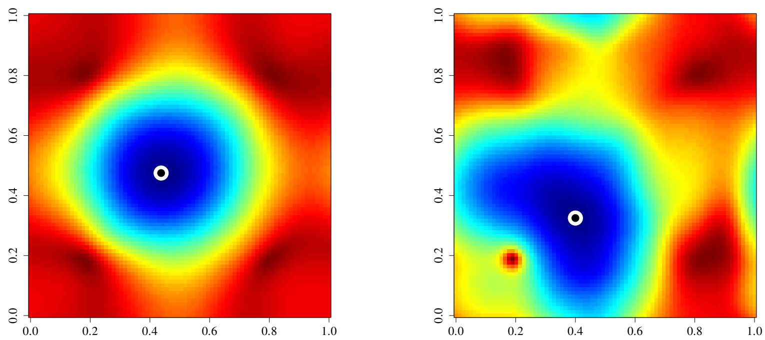
We should remark that, if the true model parameters were known, then one could compute the asymptotic sensor placement objective function (and its gradient) prior to receiving any observations by solving the algebraic Ricatti equation

$$\begin{aligned} 0 = \mathcal{A}(\theta)\Sigma_\infty(\theta,\boldsymbol{o}) &+ \Sigma_\infty(\theta,\boldsymbol{o})\mathcal{A}^\dagger(\theta) \\ &+ \mathcal{B}\mathcal{Q}(\theta)\mathcal{B}^\dagger - \Sigma_\infty(\theta,\boldsymbol{o})\mathcal{C}^\dagger(\theta,\boldsymbol{o})\mathcal{R}^{-1}(\boldsymbol{o})\mathcal{C}(\theta,\boldsymbol{o})\Sigma_\infty(\theta,\boldsymbol{o}). \end{aligned} \tag{4.23}$$

In particular, provided $\mathcal{M}_t \rightarrow \mathcal{M}_\infty$ as $t \rightarrow \infty$, in this case it is possible to show that

$$\tilde{\mathcal{J}}(\theta, \boldsymbol{o}) := \lim_{t\rightarrow\infty} \frac{1}{t}\mathcal{J}_t(\theta, \boldsymbol{o}) = \mathrm{Tr}\left[M_\infty \Sigma_\infty(\theta, \boldsymbol{o})\right]. \tag{4.24}$$

In this case, it would arguably be preferable to use a (non-stochastic) gradient descent algorithm on the asymptotic objective function directly in order to obtain the optimal sensor placement. Indeed, this approach is rather more standard in the literature (e.g., [3, 9, 488]). If, however, the true model parameters are unknown, then one can no longer compute the true asymptotic sensor placement objective function (or its gradient) prior to receiving any observations, since the true solution of the algebraic Ricatti equation (and thus the true asymptotic objective function and the true optimal sensor placement) depends on knowledge of the true parameters (see Figure 4.2). This observation highlights the need to tackle the parameter estimation and optimal sensor placement problems together using Algorithm (4.22a) - (4.22b).



(a) (b)

Figure 4.2: The 'optimisation landscape'. Plots of the asymptotic sensor placement objective function, $\tilde{\mathcal{J}}(\theta, \boldsymbol{o})$, and the corresponding optimal sensor placement, $\hat{\boldsymbol{o}} = \arg\min_{\boldsymbol{o}\in\Omega} \tilde{\mathcal{J}}(\theta, \boldsymbol{o})$, for two possible specifications of the model parameters θ.

In practice, we cannot implement this algorithm directly, as it depends on the infinite-dimensional solutions of the Kalman-Bucy filtering equations. We are thus required to use a Galerkin discretisation, and project onto a finite-dimensional Hilbert space (e.g., [198, 398]). Under standard assumptions, the approximate, finite-dimensional solutions of the Kalman-Bucy filtering equations (4.17) - (4.18), and the algebraic Ricatti equation (4.23), converge to the true, infinite-dimensional solutions as the order of the projection is increased, uniformly in time (e.g., [27, 80, 140, 198, 472, 488]). We remark that these conditions are satisfied by the partially observed stochastic advection-diffusion equation introduced in Section 4.2 (see, for example, [80, Section 6.1]).

On this basis, it is reasonable to expect that, under similar conditions, the finite-dimensional approximations of the asymptotic log-likelihood and the asymptotic objective function, namely, $\tilde{\mathcal{L}}_n(\theta, \boldsymbol{o})$ and $\tilde{\mathcal{J}}_n(\theta, \boldsymbol{o})$, will converge to their true, infinite-dimensional counterparts; as will the corresponding approximations of the maximum likelihood estimate and the optimal sensor placement, namely,

$$\hat{\theta}_n := \arg\max_{\theta \in \Theta} \tilde{\mathcal{L}}_n(\theta, \boldsymbol{o}) \quad \text{and} \quad \hat{\boldsymbol{o}}_n := \arg\min_{\boldsymbol{o} \in \Omega^{n_y}} \tilde{\mathcal{J}}_n(\theta, \boldsymbol{o}). \tag{4.25}$$

In fact, rigorous convergence results of this type have already been obtained in the case of the sensor placement objective function, and the optimal sensor placement. In particular, under precisely those conditions required for convergence of the approximate, finite-dimensional filter, the finite-dimensional approximation of the sensor placement objective function, and the corresponding approximation of the optimal sensor placement, do indeed converge to their true values (e.g., [80, Theorem 6.3], [472, Theorem 4.1.2], and [488, Theorem 4.3]). While corresponding results have not explicitly been derived for the asymptotic log-likelihood and the maximum likelihood estimate, very similar arguments could be applied in this setting (see [52, 113, 222, 223, 224, 225, 226, 320, 321] for some relevant results in the fully observed case).

In this context, we have strong justification for implementing a finite-dimensional version of Algorithm (4.22a) - (4.22b), in which the filter and filter derivatives are replaced by their finite-dimensional approximations. As noted above, the resulting algorithm is a particular case of the joint online parameter estimation and optimal sensor placement algorithm analysed in Proposition 3.1 in Chapter 3. Thus, under suitable conditions on the process consisting of the latent state, the optimal filter, and the filter derivatives, the parameter estimates and the optimal sensor placements generated by this algorithm are guaranteed to converge to the stationary points of the (finite-dimensional approximation) of the asymptotic log-likelihood and the asymptotic sensor placement objective function, respectively. That is,

$$\lim_{t \to \infty} \nabla_{\theta} \tilde{\mathcal{L}}_n(\theta_t, \boldsymbol{o}_t) = \lim_{t \to \infty} \nabla_{\boldsymbol{o}} \tilde{\mathcal{J}}_n(\theta_t, \boldsymbol{o}_t) = 0. \tag{4.26}$$

These conditions can be found in full in [412, Appendix A], as well as sufficient conditions in the linear Gaussian case. The required conditions coincide with standard conditions required for stability of the Kalman-Bucy filter, and are thus in some sense the weakest conditions under which an asymptotic result of this type can be established.

4.4 Numerical Results

In this section, we provide extensive numerical examples illustrating the performance of the joint online parameter estimation and optimal sensor placement algorithm for the partially observed stochastic advection-diffusion equation. The `R` code is available at https://github.com/louissharrock/RML-ROSP. All simulations are performed on a 2012 MacBook Pro with 2.7 GHz Intel Core i7 processor and 16GB RAM.

4.4.1 Numerical Considerations

For numerical purposes, we are required to project the infinite-dimensional solution of the signal equation onto a finite dimensional Hilbert space. In particular, we will consider the finite dimensional subspace $\mathcal{H}_n \subset \mathcal{H}$ spanned by the truncated set of Fourier basis functions $\{\phi_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\Lambda_n}$, where $\Lambda_n \subset \mathbb{Z}^2$ is the following set of wave-numbers

$$\Lambda_n = \left\{\boldsymbol{k} \in \mathbb{Z}^2 : -\left(\frac{n}{2}-1\right) \leq k_1, k_2 \leq \frac{n}{2}\right\} \ , \quad n \in 2\mathbb{N} \ , \quad |\Lambda_n| = n^2. \tag{4.27}$$

We should emphasise that this choice of basis is not unique. Indeed, in principle, one can consider any finite dimensional basis (e.g., Chebyshev polynomials, finite-elements, etc.), provided that the resulting projection converges in an appropriate sense as its dimension is increased (see, e.g., Theorem 4.2 in [488] for some precise conditions). Indeed, other choices of the finite dimensional basis may be more appropriate in the case of non-periodic boundary conditions or other more complex geometries. We highlight, in particular, the Gaussian Markov Random Field approach introduced in [302], which makes use of piece-wise linear basis functions on a triangulation of the domain.

For $n \geq 1$, let $\Pi_n : \mathcal{H} \to \mathcal{H}_n$ denote the orthogonal projection onto this space, defined in the usual fashion. The Galerkin projection of u_t is then given by $u_t^n = u_t^n(\cdot) = \{u_t^n(\boldsymbol{x}) : \boldsymbol{x} \in \Pi\} \in \mathcal{H}_n$, where

$$u_t^n(\boldsymbol{x}) = \Pi_n u_t(\boldsymbol{x}) = \sum_{\mathbf{k}\in\Lambda_n} \alpha_t^{\boldsymbol{k}} \phi_{\boldsymbol{k}}(\boldsymbol{x}) \ , \quad \alpha_t^{\boldsymbol{k}} = \langle u_t, \phi_{\boldsymbol{k}}\rangle = \int_\Pi u_t(\boldsymbol{x})\overline{\phi_{\boldsymbol{k}}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \tag{4.28}$$

and the vector of Fourier coefficients $\{\alpha_t^{\boldsymbol{k}}\}_{\boldsymbol{k}\in\Lambda_n}$ now obey the finite dimensional SDE

$$\mathrm{d}\alpha_t^{\boldsymbol{k}} = \sum_{\boldsymbol{j}\in\Lambda_n} \lambda_{\boldsymbol{j},\boldsymbol{k}}(\theta)\alpha_t^{\boldsymbol{j}}\mathrm{d}t + \sum_{\boldsymbol{j}\in\Lambda_n} \xi_{\boldsymbol{j},\boldsymbol{k}}\eta_{\boldsymbol{j}}(\theta)\mathrm{d}z_t^{\boldsymbol{j}} \ , \quad \boldsymbol{k} \in \Lambda_n, \tag{4.29}$$

with $\lambda_{\boldsymbol{j},\boldsymbol{k}}(\theta)$ and $\xi_{\boldsymbol{j},\boldsymbol{k}}(\theta)$ defined as previously. This high dimensional SDE will provide an approximation for the original, infinite dimensional SPDE. Indeed, as $n \to \infty$, one can show that the finite-dimensional approximation u_t^n does indeed converge in law to the true solution u_t (see, e.g., [418]). It is convenient to rewrite this equation, as well as the corresponding observation equation, in vector form. Let $\alpha_t^n = \{\alpha_t^{\boldsymbol{k}}\}_{\boldsymbol{k}\in\Lambda_n}$ denote the n^2-dimensional vector of Fourier coefficients. We can then write

$$\mathrm{d}\alpha_t^n = A_n(\theta)\alpha_t^n\mathrm{d}t + B_n\mathrm{d}v_t^{n,\theta} \ , \quad \alpha_0^n \in \mathcal{H}_n \tag{4.30}$$

$$\mathrm{d}y_t^n = C_n(\theta, \boldsymbol{o})\alpha_t^n\mathrm{d}t + \mathrm{d}w_t^{\boldsymbol{o}} \ , \quad y_0^n = 0, \tag{4.31}$$

where $A_n(\theta) \in \mathbb{R}^{n^2 \times n^2}$, $B_n \in \mathbb{R}^{n^2 \times n^2}$, and $C_n(\theta, \boldsymbol{o}) = \mathcal{C}(\theta, \boldsymbol{o})|_{\mathbb{R}^{n^2}} \in \mathbb{R}^{n_y \times n^2}$ are the matrices

$$[A_n(\theta)]_{j,k} = \lambda_{\boldsymbol{j},\boldsymbol{k}}(\theta) \quad , \quad [B_n]_{j,k} = \xi_{\boldsymbol{j},\boldsymbol{k}} \quad , \quad [C_n(\theta, \boldsymbol{o})]_{j,k} = \mathcal{C}_j(\theta, \boldsymbol{o})\phi_{\boldsymbol{k}}, \tag{4.32}$$

and where $v_{n,\theta}(t)$ is the $\mathbb{R}^{n^2}$-valued Wiener process with incremental covariance matrix $Q_n(\theta) = \text{diag}\{\eta^2_{\boldsymbol{k}}(\theta)\}_{\boldsymbol{k} \in \Lambda_n}\}$. In our numerical simulations, we will typically set $n = 50$, so that the simulated observations correspond to noisy realisations of the projection of the true infinite dimensional signal onto an $n^2 = 2500$ dimensional subspace. We observe numerically that coefficients for larger wave-numbers (i.e., the higher frequency components) are very close to zero. This agrees with the results reported by Sigrist et al. [418] and Liu et al. [308], and suggests that the model dynamics are dominated by the low-frequency components.

Given this, it is permissible to apply the finite-dimensional Kalman-Bucy filter and tangent filter, and implement the joint online parameter estimation and optimal sensor placement algorithm, using a reduced Fourier basis of $K \ll n^2$ basis functions. This is typical in similar applications (e.g., [125, 418]). In particular, following [418], we used a reduced Fourier basis given by $\{\phi_{\boldsymbol{k}}\}_{\boldsymbol{k} \in \Gamma_{m,n}}$, where $\Gamma_{m,n} \subseteq \Lambda_n \subset \mathbb{Z}^2$ is the following set of wave-numbers

$$\Gamma_{m,n} = \{\boldsymbol{k} \in \mathbb{Z}^2 : k_1^2 + k_2^2 \leq m\} \cap \Lambda_n \ , \quad m \in \mathbb{N}_0 \ , \quad K = |\Gamma_{m,n}|. \tag{4.33}$$

In our simulations, we will set $m = 5$, which yields a Fourier truncation with $K = 21$ basis functions. Numerical results indicate that, for our purposes, this choice represents a reasonable trade-off between accuracy - both of the optimal sensor placement (see Table 4.1) and the optimal state estimate (see Figure 4.3a) - and computational cost. This choice is also comparable with other related works (e.g., [80, 418, 488]). It is worth noting that, even if high-frequency components were more dominant, it may not be possible to capture these when using a small number of sensors, spaced at large intervals. A similar observation was previously made in [418]. Regarding the time discretisation, we use an exponential Euler scheme for the finite-dimensional approximation of the partially observed diffusion process (4.30) - (4.31), and implement the discrete-time analogue of the stochastic gradient descent algorithm (4.22a) - (4.22b).

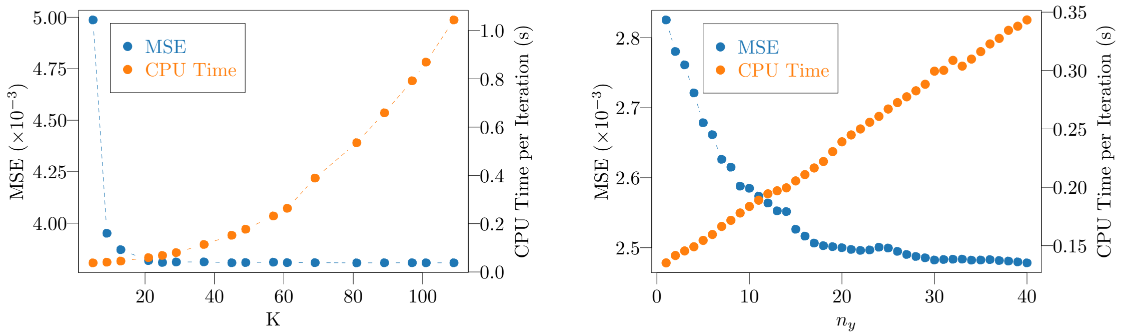
4.4.2 Numerical Experiments

4.4.2.1 Simulation I

We first investigate the convergence of the parameter estimates and the optimal sensor placements under conditions which guarantee convergence to the stationary points of the asymptotic log-likelihood and the asymptotic sensor placement objective function, respec-

	K	5	13	*21*	37	57	81	101
Case I	MSE ($\times 10^{-2}$)	5.75	4.28	2.23	2.67	1.43	0.28	0.13
	CPU Time per Iteration (s)	0.03	0.05	0.06	0.11	0.23	0.54	0.87
Case II	MSE ($\times 10^{-2}$)	12.49	9.07	0.32	0.16	0.05	0.02	0.01
	CPU Time per Iteration (s)	0.04	0.05	0.07	0.12	0.24	0.52	0.83

Table 4.1: The mean squared error (MSE) of the approximate optimal sensor placement, and the CPU time per iteration, for different choices of the number of basis functions K, in two cases of interest. In Case I, there are 25 sensors, all of whose positions are to be optimised. In Case II, there are 40 sensors, 5 of whose positions are to be optimised. In both cases, the sensors are initially uniformly placed at random over $\Pi = [0,1]^2$, and the algorithm is run for $T = 1 \times 10^4$ iterations. While increasing the number of basis functions beyond $K = 21$ can further decrease the error in the approximate optimal sensor placement, this choice is adequate to obtain a relatively accurate approximation at a relatively low computational cost, particularly when there are relatively few movable sensors (Case II).



(a) Number of basis functions.

(b) Number of movable sensors.

Figure 4.3: The MSE of the optimal state estimate, and the CPU time (per iteration), for different choices of (a) the number of basis functions K and (b) the number of movable sensors. In (a) we use 25 sensors, all of which are movable. In (b) we use 40 sensors, varying the number of movable sensors, and use $K = 45$ basis functions. In both cases, the sensors are initially uniformly placed at random over $\Pi = [0,1]^2$, and the algorithm is run for $T = 1 \times 10^4$ iterations.

tively (see [412, Appendix A]). We assume that the true model parameters and the initial parameter estimates are given respectively by

$$\theta^* = (\rho_0 = 0.50, \sigma^2 = 0.2, \zeta = 0.5, \rho_1 = 0.1, \gamma = 2.0, \alpha = \tfrac{\pi}{4}, \mu_x = 0.30, \mu_y = -0.3, \tau^2 = 0.01),$$

$$\theta_0 = (\rho_0 = 0.25, \sigma^2 = 0.8, \zeta = 0.1, \rho_1 = 0.2, \gamma = 1.2, \alpha = \tfrac{\pi}{3}, \mu_x = 0.1, \mu_y = -0.15, \tau^2 = 0.10).$$

We also assume that we have $n_y = 8$ sensors in $\Pi = [0,1]^2$. We suppose that the sensors are independent, have zero bias, and generate noisy measurements with variance τ^2. Thus, in the observation equation, we have $\beta = (0, \ldots, 0)^T$ and $\mathcal{R} = \mathrm{diag}(\tau^2)$. In this test simulation, in order to verify the convergence of our algorithm, we suppose that our objective is to obtain the optimal sensor placement with respect to the state estimate at a set of known 'target' locations. This is achieved by choosing an operator $\mathcal{M}$ in the sensor placement objective function $\tilde{\mathcal{J}}(\theta, \boldsymbol{o})$, c.f. (4.20b), which places an emphasis on minimising the uncertainty in the state estimate at these target locations. We provide an explicit definition of this operator in Appendix 4.A. In particular, we assume that the target sensor locations and the initial sensor locations are given, respectively, by

$$\boldsymbol{o}^* = \left\{ \begin{pmatrix} 0.00 \\ 0.59 \end{pmatrix}, \begin{pmatrix} 0.50 \\ 0.66 \end{pmatrix}, \begin{pmatrix} 0.33 \\ 0.33 \end{pmatrix}, \begin{pmatrix} 0.75 \\ 0.50 \end{pmatrix}, \begin{pmatrix} 0.08 \\ 0.08 \end{pmatrix}, \begin{pmatrix} 0.58 \\ 0.83 \end{pmatrix}, \begin{pmatrix} 0.83 \\ 0.92 \end{pmatrix}, \begin{pmatrix} 0.25 \\ 0.83 \end{pmatrix} \right\}, \tag{4.34}$$

$$\boldsymbol{o}_0 = \left\{ \begin{pmatrix} 0.84 \\ 0.65 \end{pmatrix}, \begin{pmatrix} 0.34 \\ 0.50 \end{pmatrix}, \begin{pmatrix} 0.43 \\ 0.31 \end{pmatrix}, \begin{pmatrix} 0.60 \\ 0.34 \end{pmatrix}, \begin{pmatrix} 0.27 \\ 0.26 \end{pmatrix}, \begin{pmatrix} 0.51 \\ 0.18 \end{pmatrix}, \begin{pmatrix} 0.08 \\ 0.23 \end{pmatrix}, \begin{pmatrix} 0.25 \\ 0.08 \end{pmatrix} \right\}. \tag{4.35}$$

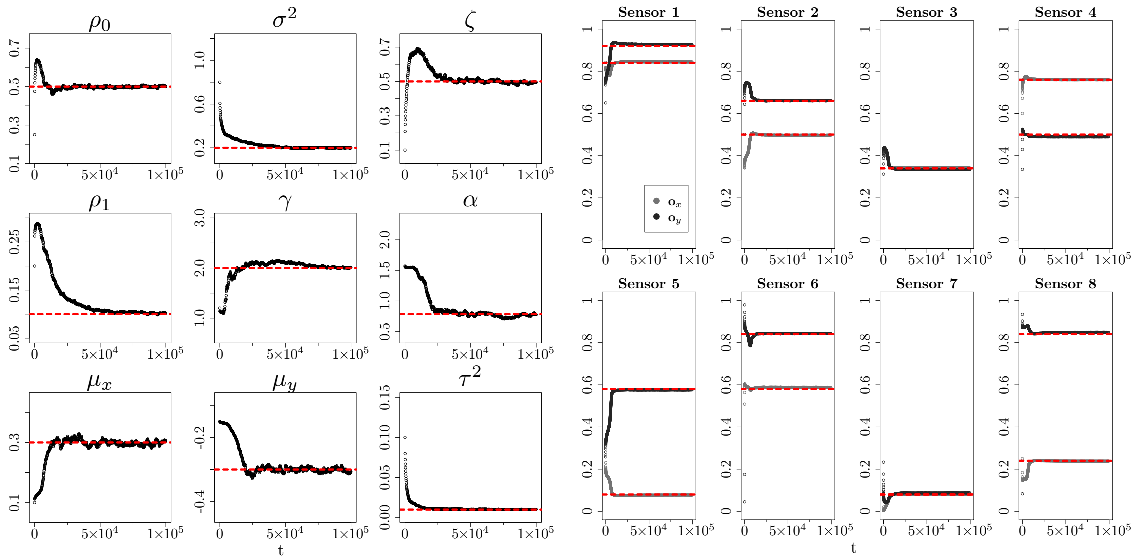
It remains to specify the learning rates $\{\gamma_t^{i,\theta}\}_{t\geq 0}^{i=1,\dots,9}$ and $\{\gamma_t^{j,\boldsymbol{o}}\}_{t\geq 0}^{j=1,\dots,8}$, where the indices i, j now make explicit the fact that the step sizes are permitted to vary between parameters, and between sensors. In this simulation, we assume that our primary objective is to estimate the true model parameters, and our secondary objective is to optimally place the measurement sensors. We thus set $\gamma_t^{i,\theta} = \gamma_0^{\theta,i} t^{-\varepsilon^{i,\theta}}$ and $\gamma_t^{j,\boldsymbol{o}} = \gamma_0^{j,\boldsymbol{o}} t^{-\varepsilon^{j,\boldsymbol{o}}}$, where $\gamma_0^{i,\theta}, \gamma_0^{j,\boldsymbol{o}} > 0$ and $0.5 < \varepsilon_{\boldsymbol{o}}^j < \varepsilon_\theta^i \leq 1$ for all $i = 1, \dots, 9$ and $j = 1, \dots, 8$, with the values of $\gamma_0^{i,\theta}$, $\varepsilon^{i,\theta}$, $\gamma_0^{j,\boldsymbol{o}}$, and $\varepsilon^{j,\boldsymbol{o}}$ tuned individually. In our numerics, the specific values of the learning rates are chosen on the basis of initial experiments. In principle, however, one can use any one of a number of adaptive learning rate methods to automate this choice, including backtracking line search, Adagrad [168], Adadelta [486], Adam [248], AMSgrad [393], and others. This choice of learning rate satisfies all of the conditions of Proposition 3.1 in Chapter 3. In particular, it guarantees that

$$\lim_{t\to\infty} \frac{\gamma_t^{i,\theta}}{\gamma_t^{j,\boldsymbol{o}}} = 0 \quad \forall i = 1, \dots, 9 \ , \ j = 1, \dots, 8. \tag{4.36}$$

This implies that the parameter estimates $\{\theta_t\}_{t\geq 0}$ move on a slower timescale than the sensor placements $\{\boldsymbol{o}_t\}_{t\geq 0}$. Thus, the sensor placements see the parameter estimates as quasi-static, while the parameter estimates see the sensor placements as almost equilibrated. In practice, this means that $\boldsymbol{o}_t$ will asymptotically track the sensor placements which are optimal with respect to the *current* parameter estimates. This is particularly advantageous when the optimal sensor placement depends significantly on the parameters (see Section 4.4.2.2).

The performance of the two-timescale stochastic gradient descent algorithm is visualised in Figure 4.4, in which we plot the sequence of online parameter estimates and optimal sensor placements, Figure 4.5, in which we plot a single component of the optimal state estimate, and in Figure 4.6, in which we plot the time evolution of the mean squared error

(MSE) for the corresponding filter. As expected, all of the parameter estimates converge to within a small neighbourhood of their true values (Figure 4.4a), and all of the sensors converge to one of the target locations (Figure 4.4b). As a result, the performance of the filter is improved to near-optimal after approximately $T = 2 \times 10^4$ iterations (black line in Figure 4.6). This number is largely determined by the initial magnitudes of the learning rates $\{\gamma_t^{i,\theta}\}_{t\geq 0}^{i=1,\ldots,9}$ and $\{\gamma_t^{j,\boldsymbol{o}}\}_{t\geq 0}^{j=1,\ldots,8}$. In particular, increasing one or more of these values will often decrease the time taken for the algorithm iterates to converge.



(a) Online parameter estimates.

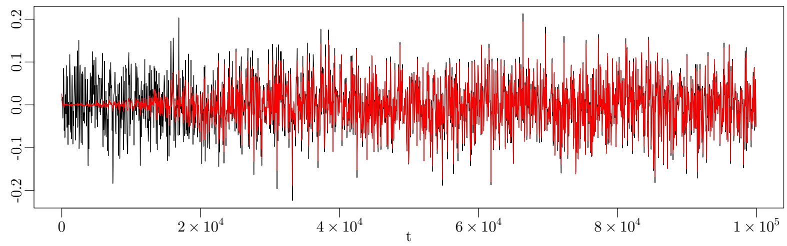
(b) Optimal sensor placements.

Figure 4.4: Simulation Ia. The online parameter estimates and optimal sensor placements (black); and the true parameters and optimal sensor placements (red, dashed). In this simulation, the parameter estimates move on the slower timescale, and the sensor placements move on the faster timescale. The total CPU time required for this simulation was 2368 seconds (0.02368 seconds per iteration).

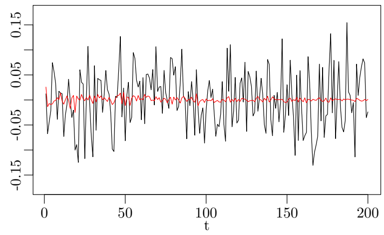
It is also possible to apply our algorithm when the primary objective is to obtain the optimal sensor placement, and the secondary objective is to estimate the true model parameters. That is, the order of the two optimisation problems is reversed. In particular, this is achieved by choosing learning rates which no longer satisfy (4.36), but instead satisfy

$$\lim_{t\to\infty} \frac{\gamma_t^{j,\boldsymbol{o}}}{\gamma_t^{i,\theta}} = 0 \quad \forall i = 1,\ldots,9\ ,\ j = 1,\ldots,8. \tag{4.37}$$

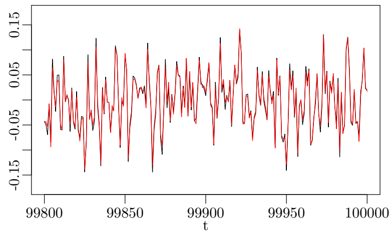
This implies, of course, that the sensor placements $\{\boldsymbol{o}_t\}_{t\geq 0}$ now move on a slower timescale than the parameter estimates $\{\theta_t\}_{t\geq 0}$. The performance of the two-timescale stochastic gradient descent algorithm in this scenario, with all other assumptions unchanged from



(a) The entire learning period.



(b) The first 200 iterations.



(c) The final 200 iterations.

Figure 4.5: Simulation Ia. A single component of the optimal state estimate $\hat{\alpha}_n(t)$ obtained using the true parameters and the target sensor locations (black), and using the sequence of online parameters and optimal sensor placements (red).

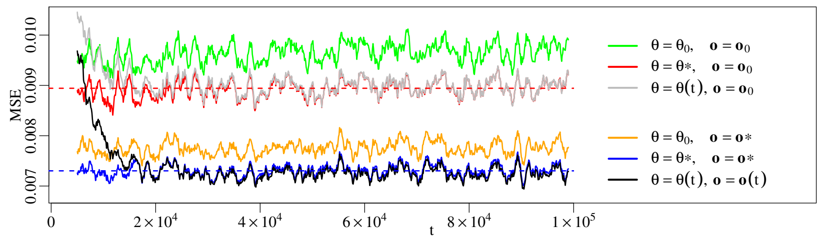
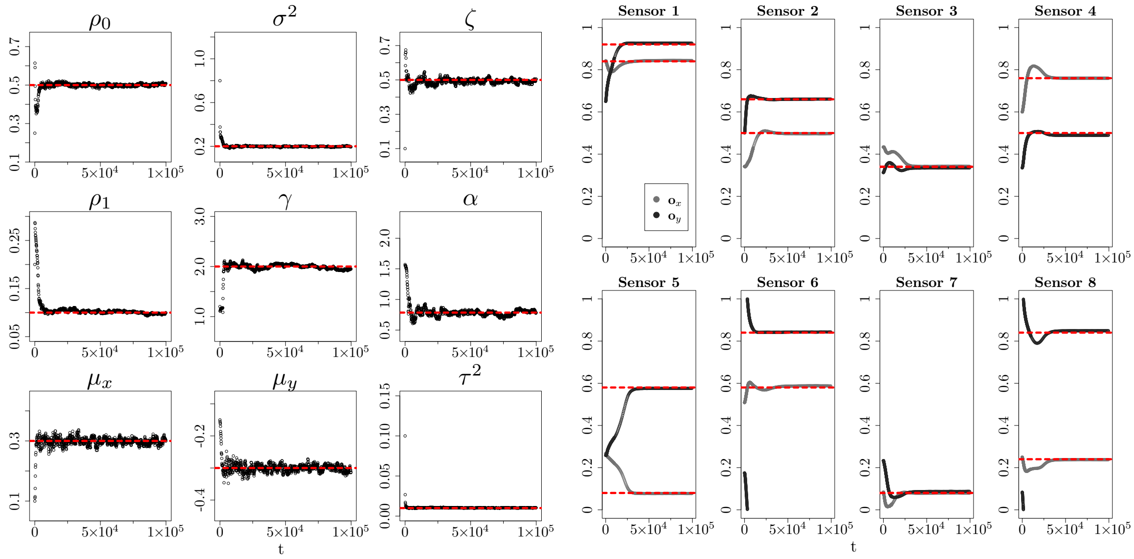


Figure 4.6: Simulation Ia. The moving average of the MSE of the optimal state estimate under different learning scenarios (various colours). We also plot the average of the MSE for the true parameters and the initial sensor placement (red, dashed), and the average of the MSE for the true parameters and the optimal sensor placement (blue, dashed). The MSE is calculated at $n^2 = 50^2$ uniformly spaced grid points on $\Pi = [0,1]^2$.

the first simulation, is illustrated in Figure 4.7. Once more, we observe that all of the parameter estimates converge to within a small neighbourhood of their true values, and that the sensors converge to the target locations. Unsurprisingly, given the alternative learning rates, the convergence of the parameter estimates is somewhat faster than before, while the convergence of the optimal sensor placements is somewhat slower.

It is worth re-emphasising, at this stage, that the convergence of our algorithm does *not* depend on whether the learning rates satisfy (4.36) or (4.37). That is to say, *a priori*, the algorithm has no preference over which of the parameter estimates or the optimal sensor placements moves on the faster time scale, and which moves on the slower time scale. This is a clear advantage of the two-timescale approach.



(a) Online parameter estimates.

(b) Optimal sensor placements.

Figure 4.7: Simulation Ib. The online parameter estimates and optimal sensor placements (black); and the true parameters and optimal sensor placements (red, dashed). In this simulation, the parameter estimates move on the faster timescale, and the sensor placements move on the slower timescale.

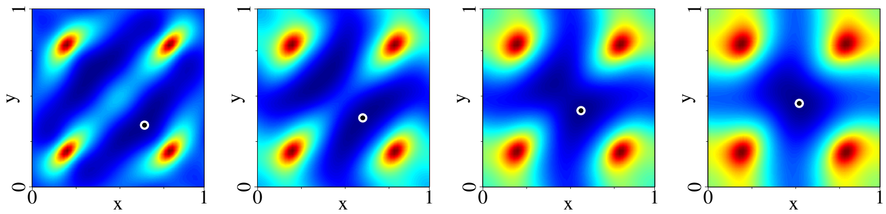
4.4.2.2 Simulation II

In our second numerical experiment, we investigate the performance of our algorithm in a scenario where the optimal sensor placement depends to a significant extent on the value of one of the model parameters. In this simulation, we will assume that the values of $\theta_{2:9} = (\sigma^2, \zeta, \rho_1, \gamma, \alpha, \mu_x, \mu_y, \tau^2)$ are known, and fixed equal to their true values, while the value of $\theta_1 = \rho_0$ is unknown. The true value and the initial value of the unknown parameter are given by $\rho_0^* = 0.3$ and $\rho_0 = 0.01$, respectively.

We now assume that we have $n_y = 5$ sensors in $\Pi = [0,1]^2$, all of which are independent, have zero bias, and the same variance. The locations of the first 4 sensors are fixed, while the location of the final sensor is to be optimised. In contrast to the previous simulation, we now suppose our objective is to obtain the optimal sensor placement with respect to the state estimate over the entire spatial domain (i.e., not weighted towards a set of target locations). The locations of the fixed sensors, and the initial location of the sensor whose location is to be optimised, are shown in Figure 4.9b.

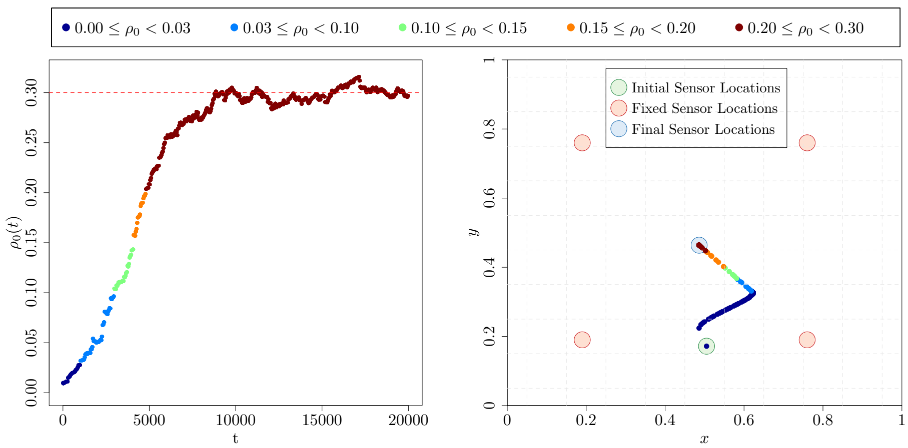
It remains to specify the learning rates $\{\gamma_t^{\rho_0}\}_{t\geq 0}$ and $\{\gamma_t^{\boldsymbol{o}_5}\}_{t\geq 0}$. In this case, we set $\gamma_t^{\rho_0} = 0.1t^{-0.55}$ and $\gamma_t^{\boldsymbol{o}_5} = 0.1t^{-0.51}$, implying that the sensor placements move on a faster timescale than the parameter estimates. Thus, as outlined previously, $\boldsymbol{o}_t$ should asymptotically track $\boldsymbol{o}^*(\rho_0(t))$, the sensor placement which is optimal with respect to the current parameter estimate. This is clearly advantageous in the current scenario, in which the optimal sensor placement is known to depend on the unknown model parameter. This is clearly visualised in Figure 4.8, which contains plots of the asymptotic sensor placement objective function, and the corresponding optimal sensor placement, for several different values of the unknown model parameter. For this configuration of fixed sensors, the optimal location of the additional sensor is to the south-east (or north-west) of centre for small ρ_0 (Figure 4.8a), and converges to the centre as ρ_0 increases (Figures 4.8b - 4.8d).

The performance of the two-timescale gradient descent algorithm is illustrated in Figure 4.9, in which we have plotted the sequence of online parameter estimates $\{(\rho_0)_t\}_{t\geq 0}$ and optimal sensor placements $\{(\boldsymbol{o}_5)_t\}_{t\geq 0}$. As expected, the online parameter estimate, on the slow-timescale, is seen to converge to the true value of $\rho_0^* = 0.3$ over the course of the entire learning period. Meanwhile, the optimal sensor placement, on the fast-timescale, begins by moving rapidly from its initial position to a location to the south-east of centre. It then moves slowly towards the centre of the domain as the online parameter estimate of ρ_0 increases towards its true value. Thus, the optimal sensor placement does indeed track the local optimum of the sensor placement objective function, while the online parameter estimate converges to its true value.



(a) $\rho_0 = 0.03$ (b) $\rho_0 = 0.10$ (c) $\rho_0 = 0.15$ (d) $\rho_0 = 0.20$

Figure 4.8: Simulation II. Heat maps of the sensor placement objective function, and the optimal sensor placement, for different values of ρ_0.



(a) Online parameter estimates. (b) Optimal sensor placements.

Figure 4.9: Simulation II. The online parameter estimates and optimal sensor placements (various colours); and the true parameter (red, dashed). The total CPU time required for this simulation was 185s (0.00925s per iteration).

We should note, at this point, that the asymptotic sensor placement objective function (and the asymptotic log-likelihood function) can admit multiple local optima (see Figure 4.8). There is thus no guarantee that the sensor placements (or the parameter estimates) generated by our algorithm will always converge to the global optimum (i.e., the true optimal sensor placements or the true parameter values). On this point, let us make several remarks. Firstly, this is a necessary feature of any gradient based method; such methods are only guaranteed to converge to a global optimum under the rather restrictive assumption of global convexity (see [413]). Secondly, we use a stochastic gradient descent method, updating the sensor placements and the parameter estimates in the directions of noisy estimates of the gradients of the asymptotic sensor placement objective and the asymptotic log-likelihood. In comparison to a (non-stochastic) gradient descent scheme, this approach is significantly more likely to avoid local minima and saddle points (e.g., [65, 189]). One can also use momentum [384] or additional random noise to help to escape local minima. Finally, one could run the algorithm multiple times using random restarts.

4.4.2.3 Simulation III

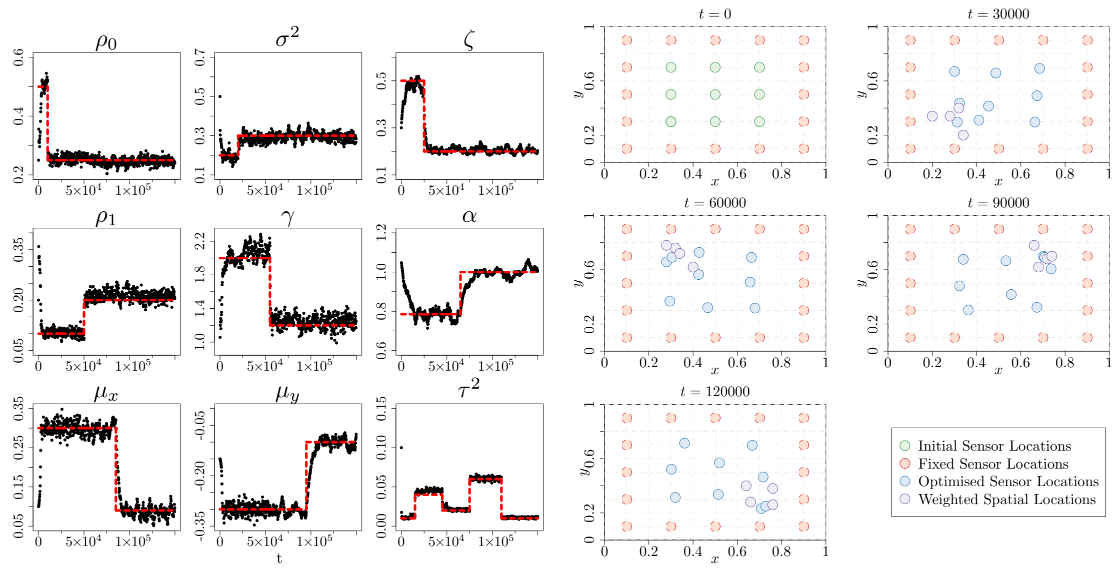
In our third numerical experiment, we investigate the performance of our algorithm under the assumption that the true model parameters $\theta^* = \theta_t^*$ are no longer static, and contain change-points at certain points in time. The values of these parameters are shown in Figure 4.10. Meanwhile, the initial parameter estimates are now given by

$$\theta_0 = (\rho_0 = 0.25, \sigma^2 = 0.5, \zeta = 0.3, \rho_1 = 0.2, \gamma = 1.5, \alpha = \tfrac{\pi}{3}, \mu_x = 0.1, \mu_y = -0.15, \tau^2 = 0.1).$$

We also now suppose that the optimal sensor locations $\boldsymbol{o}^* = \boldsymbol{o}_t^*$ vary in time. In particular, we now consider a scenario in which our objective is to obtain the sensor placement which minimises the uncertainty in the state estimate over the entire spatial domain, but now weighted slightly towards a set of four time-varying spatial locations. Once again, this is achieved by a suitable choice of spatial weighting operator in the sensor placement objective function (see Appendix 4.A). On this occasion, we assume that we have $n_y = 25$ sensors in $\Pi = [0,1]^2$, each with zero bias and equal variance. The first 16 sensors are distributed evenly towards the boundary of the spatial domain, with their locations fixed. Meanwhile, the locations of the final 9 sensors are to be optimised. We show the locations of the fixed sensors (red), the initial sensor locations of the nine sensors to be optimised (green) and the weighted spatial locations at four time points (purple) in Figure 4.10b.

It remains, once more, to specify the learning rates $\{\gamma_t^{i,\theta}\}_{t\geq 0}^{i=1,\dots,9}$, $\{\gamma_t^{j,\boldsymbol{o}}\}_{t\geq 0}^{j=17,\dots,20}$. In this simulation, we set the learning rates for the parameter estimates and the sensors placements as constant. That is, $\gamma_t^{i,\theta} = \gamma_0^{i,\theta}$ and $\gamma_t^{j,\boldsymbol{o}} = \gamma_0^{j,\boldsymbol{o}}$, with the specific values of $\gamma_0^{i,\theta}$, $\gamma_0^{j,\boldsymbol{o}}$ tuned individually. This is a standard choice when the true parameters are no longer static (e.g., [315]). The choice of constant learning rates violates one of the conditions required for convergence of the parameter estimates and the optimal sensor placements, namely, that $\int_0^\infty (\gamma_t^\theta)^2 \mathrm{d}t < \infty$ and $\int_0^\infty (\gamma_t^o)^2 \mathrm{d}t < \infty$. There is thus no longer any guarantee that the algorithm iterates will converge to the stationary points of the two objective functions. They are, however, expected to oscillate around the optimal points. The advantage of constant learning rates is that the algorithm iterates can now adapt rapidly to changes in the true model parameters and the optimal sensor placements.

In practice, the two-timescale stochastic gradient algorithm still performs remarkably well in this scenario (Figure 4.10). The online parameter estimates generated by the algorithm are able to track the changes in the dynamic model parameters in real time (Figure 4.10a), while the sensor placements update in response to changes in the time-varying weighted spatial locations (Figure 4.10b).



(a) Online parameter estimates.

(b) Optimal sensor placements

Figure 4.10: Simulation III. The online parameter estimates and optimal sensor placements; and the true parameters (red, dashed). The total CPU time required for this simulation was 9345s (0.0623s per iteration).

Let us make some brief remarks regarding the optimal sensor placements shown in Figure 4.10b. In general, we see that, at any given time instant, the sensors tend to be positioned closer to the current locations of the four weighted points than they would be in a completely uniform configuration. For example, at $t = 30000$, all of the sensors, to a greater or lesser extent, have moved towards the south-west of the domain (top right hand panel in Figure 4.10b). At the same time, the sensors also maintain a relatively even distribution across the entire centre of the domain. This should not come as a surprise; indeed, for the chosen sensor placement objective, this does indeed represent the optimal placement of the available measurement sensors. In particular, this configuration represents a trade-off between attempting to minimise the uncertainty of the state estimate over the entire spatial domain (which favours a uniform placement of sensors), while also placing a slightly greater emphasis on the accuracy of the state estimate at the four time-varying weighted locations (which favours a placement of sensors close to these locations).

4.4.2.4 Simulation IV

In our fourth numerical experiment, we investigate the ability of our algorithm to estimate multiple unknown bias and variance parameters. We thus relax our previous assumption that the sensors all have zero bias, and the same variance. This scenario is of significant practical interest: in real-data applications, it is often necessary to calibrate the bias and variance of many measurement sensors simultaneously, and in real time.

In this simulation, we assume that we have $n_y = 11$ sensors in $\Pi = [0,1]^2$, six of which have unknown bias and variance. The true values, and initial estimates, of these parameters, are given respectively by

$$\boldsymbol{\tau}^{*2} = (\tau_1^2 = 0.01, \tau_2^2 = 0.01, \tau_3^2 = 0.05, \tau_4^2 = 0.05, \tau_5^2 = 0.10, \tau_6^2 = 0.10) \tag{4.38}$$

$$\boldsymbol{\tau}_0^2 = (\tau_1^2 = 0.05, \tau_2^2 = 0.03, \tau_3^2 = 0.15, \tau_3^2 = 0.20, \tau_5^2 = 0.02, \tau_6^2 = 0.25) \tag{4.39}$$

and

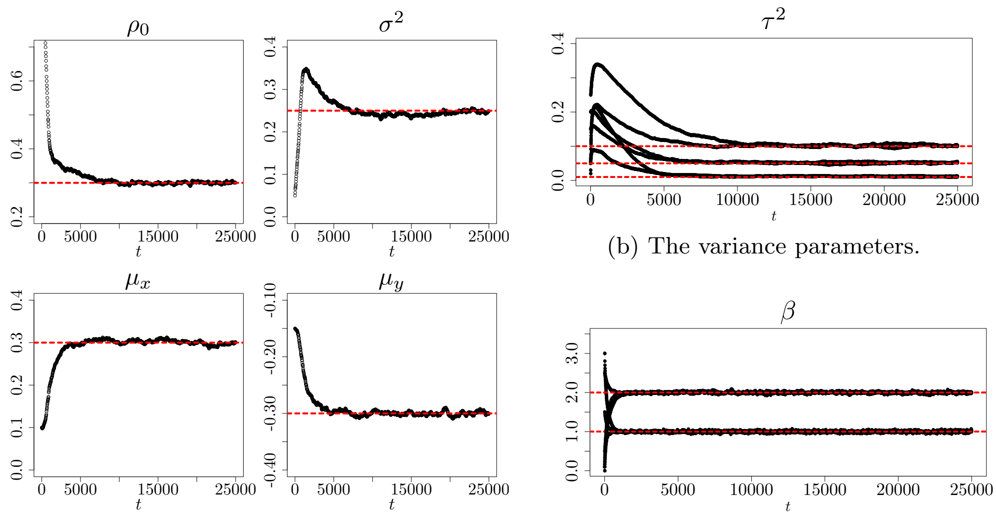
$$\boldsymbol{\beta}^* = (\beta_1 = 1.00, \beta_2 = 1.00, \beta_3 = 1.00, \beta_4 = 2.00, \beta_5 = 2.00, \beta_6 = 2.00) \tag{4.40}$$

$$\boldsymbol{\beta}_0 = (\beta_1 = 0.10, \beta_2 = 3.00, \beta_3 = 1.50, \beta_4 = 2.50, \beta_5 = 0.00, \beta_6 = 0.50). \tag{4.41}$$

We estimate the bias and variance of each of these sensors independently. In terms of the parameters in the signal equation, we now assume that the values of $\theta_{3:6} = (\zeta, \rho_1, \gamma, \alpha)$ are known, while the values of $\theta_{1,2,7:8} = (\rho_0, \sigma^2, \mu_x, \mu_y)$ are to be estimated. The true values and initial values of these parameters are shown in Figure 4.11a.

Regarding the sensor placement, we assume that the locations of the six sensors whose biases and variances are unknown are to be optimised, while the locations of the remaining five sensors are fixed. The objective is to minimise the uncertainty in the state estimate over the entire spatial domain, as in Simulation II. The locations of the fixed sensors are distributed non-uniformly, close to the boundary of the domain, while the initial locations of the movable are distributed non-uniformly, close to the centre of the domain. Finally, the step-sizes are of the same form as those in the Simulation Ib.

The performance of the two-timescale algorithm is shown in Figure 4.11, in which we have plotted the sequence of online parameter estimates for the unknown parameters. As previously, the parameter estimates are all seen to converge to their true values. Thus, in particular, the algorithm correctly identifies the biases and variances of each of the measurement sensors. Meanwhile, the final locations of the movable measurement sensors are distributed more evenly throughout the domain (plot omitted), leading to a 27% reduction in the error in the optimal state estimate (0.026 to 0.019).



(a) The signal parameters. (b) The variance parameters. (c) The bias parameters.

Figure 4.11: Simulation IV. The online parameter estimates (black); and the true parameters (red, dashed). The total CPU time required for this simulation was 888s (0.0355s per iteration).

4.4.2.5 Simulation V

In our final numerical simulation, we investigate the performance of the two-timescale stochastic gradient descent algorithm in the presence of a spatially weighted disturbance in the signal noise. We will assume, in this simulation, that $\theta_{1:6} = (\rho_0, \sigma^2, \zeta, \rho_1, \gamma, \alpha)$ are known, while $\theta_{7:9} = (\mu_x, \mu_y, \tau^2)$ are to be estimated. The true values and initial estimates of these parameters are given respectively by

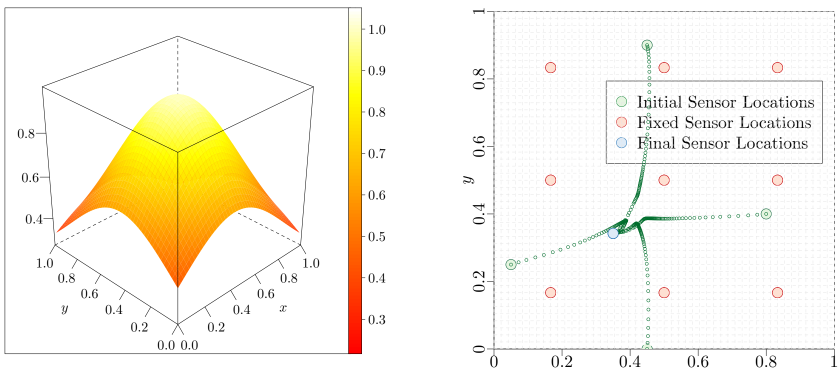
$$\theta^* = (\mu_x = 0.10, \mu_y = -0.10, \tau^2 = 0.01), \tag{4.42}$$

$$\theta_0 = (\mu_x = 0.39, \mu_y = -0.41, \tau^2 = 0.50). \tag{4.43}$$

We now assume that we have $n_y = 10$ sensors $\Pi = [0,1]^2$, each with zero bias and equal variance. The locations of nine of these sensors are fixed, while the location of the final sensor is to be optimised. The locations of the fixed sensors, and the initial location of the sensor whose location is to be optimised, are shown in Figure 4.12b. As in the second numerical experiment, we will suppose that the objective is to obtain the optimal state estimate over the entire spatial domain (i.e., not only at a set of target locations). We also now suppose that there is a localised disturbance in the signal noise around the point $(\frac{5}{12}, \frac{5}{12})$. Thus, in the signal equation, we now specify the spatial weighting function

$$b(\boldsymbol{x}) = b(x, y) = \operatorname{sech}\left[\left(\frac{(x - \frac{5}{12})^2}{0.2^2} + \frac{(y - \frac{5}{12})^2}{0.2^2}\right)^{\frac{1}{2}}\right]. \tag{4.44}$$

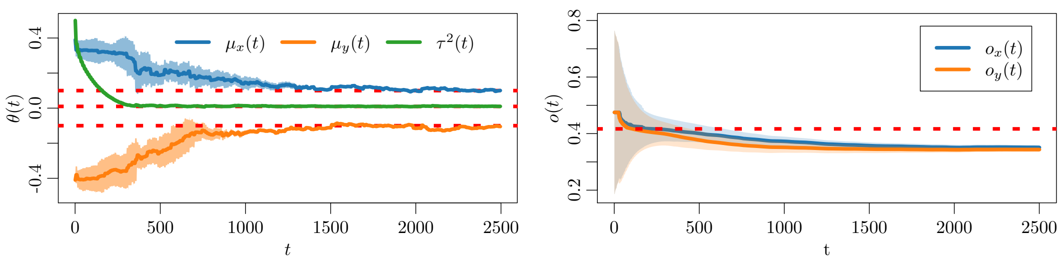
The performance of the two-timescale algorithm is illustrated in Figures 4.12 and 4.13, in which we have plotted trial averaged sequences of the optimal sensor placements, and the online parameter estimates, respectively. As previously, the online parameter estimates all converge to a small neighbourhood of their true values. Meanwhile, the movable sensor is seen to converge to a location close to, but not directly at, the centre of the local disturbance. The slight offset to the south-west of the centre of this disturbance is explained by the presence of the fixed sensor at $(0.5, 0.5)$, which is just to the north-east of the centre of the disturbance. These numerical results corroborate those also obtained in, e.g., [80, 488].



(a) The spatial weighting $b(\boldsymbol{x})$.

(b) Sensor placements on $\Pi = [0, 1]^2$.

Figure 4.12: Simulation V. The optimal sensor placements for four different initial conditions. The average CPU time required for this simulation was 41s (0.016s per iteration).



(a) Online parameter estimates.

(b) Optimal sensor placements.

Figure 4.13: Simulation V. The online parameter estimates and optimal sensor placements (various colours, solid); and the true parameters and centre of the signal noise disturbance (red, dashed). The plots are averaged over 400 trials which use different initialisations of the movable sensor.

4.5 Conclusions

In this chapter, we have considered the problem of joint online parameter estimation and optimal sensor placement for a partially observed, infinite-dimensional linear diffusion process. We have presented a solution to this problem in the form of a two-timescale stochastic gradient descent algorithm, and shown in detail how this algorithm can be successfully applied to a partially observed stochastic advection-diffusion equation, which depends in a highly non-linear fashion on a set of nine or more unknown model parameters. Our numerical results have illustrated the effectiveness of the proposed approach in a number of scenarios of practical interest. Moreover, they have highlighted the advantages of tackling the problems of online parameter estimation and optimal sensor placement together.

Appendices

4.A The Spatial Weighting Operator

In this appendix, we provide an explicit definition of the spatial weighting operator $\mathcal{M}_t : \mathcal{H} \to \mathcal{H}$. Let $m_t(\cdot) \in \mathcal{H}$ be a spatial weighting function (to be defined below). Then we define the operator $\mathcal{M}_t$ according to

$$(\mathcal{M}_t\varphi)(\boldsymbol{x}) = m_t(\boldsymbol{x})\varphi(\boldsymbol{x}) \ , \quad \varphi \in \mathcal{H}. \tag{4.45}$$

Let us now motivate a suitable definition for the spatial weighting function. We first note that, using Mercer's Theorem [345], it is possible to show that (e.g., [109, 114, 262])

$$\mathrm{Tr}\left[\mathcal{M}_t\Sigma_t(\theta,\boldsymbol{o})\right] = \int\limits_{\Pi} m_t(\boldsymbol{x})\tilde{\Sigma}_t(\theta,\boldsymbol{o},\boldsymbol{x},\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \tag{4.46}$$

where $\tilde{\Sigma}_t(\theta,\boldsymbol{o},\cdot,\cdot) : \Pi \times \Pi \to \mathbb{R}$ is the kernel operator (or covariance function) associated with the covariance operator $\Sigma_t(\theta,\boldsymbol{o})$. It follows, in particular, that the sensor placement objective function can be written in the form

$$\mathcal{J}_t(\theta,\boldsymbol{o}) = \int\limits_0^t \left[\int\limits_{\Pi} m_s(\boldsymbol{x})\tilde{\Sigma}_s(\theta,\boldsymbol{o},\boldsymbol{x},\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right]\mathrm{d}s. \tag{4.47}$$

On the basis of this expression, one appropriate choice for the spatial weighting function is given by

$$m_t(\boldsymbol{x}) = c_0\mathbb{1}_{\boldsymbol{x}\in\Pi_t^w} + c_1\mathbb{1}_{\boldsymbol{x}\in\Pi\setminus\Pi_t^w} \ , \quad \boldsymbol{x}\in\Pi, \tag{4.48}$$

where $\Pi_t^w \subseteq \Pi$ denotes a 'weighted' or 'target' spatial region, which corresponds to the region in which we are most interested in minimising the uncertainty in the optimal state estimate, and $0 \leq c_1 \leq c_0 \leq 1$ are positive constants. The choice of the constants c_0 and c_1, or equivalently the ratio $\frac{c_0}{c_1} \in [1,\infty)$, determines the extent to which the objective function will prioritise minimising the uncertainty in the state estimate in the region Π_t^w, relative to the region $\Pi \setminus \Pi_t^w$. In our numerics, we use the following specific definitions of c_0, c_1, and Π_t^w.

Simulation I

In this simulation, we set $c_0 = 1$, $c_1 = 0$, and $\Pi_t^w = \Pi_w = \bigcup_{i=1}^{8}\{\boldsymbol{x} \in \Pi : |\boldsymbol{x} - \boldsymbol{x}_i| \leq r\}$, where $\{\boldsymbol{x}_i\}_{i=1}^{8}$ are the 8 'target' locations defined in (4.34), and $r > 0$ is a small positive constant. This yields $m_t(\boldsymbol{x}) = \sum_{i=1}^{8} \mathbb{1}_{\{\boldsymbol{x}\in\Pi:|\boldsymbol{x}-\boldsymbol{x}_i|\leq r\}}$, and thus

$$\mathrm{Tr}\left[\mathcal{M}_t\Sigma_t(\theta,\boldsymbol{o})\right] = \sum_{i=1}^{8} \int\limits_{\{\boldsymbol{x}\in\Pi:|\boldsymbol{x}-\boldsymbol{x}_i|<r\}} \tilde{\Sigma}_t(\theta,\boldsymbol{o},\boldsymbol{x},\boldsymbol{x})\mathrm{d}\boldsymbol{x} \tag{4.49}$$

so that the objective function only seeks to minimise the uncertainty in the state estimate close to the target locations $\{\boldsymbol{x}_i\}_{i=1}^{8}$. We remark that similar results are obtained if one sets $c_1 = \varepsilon$, for some $\varepsilon \ll 1$. In this case, the weighted trace of the covariance is given by a similar expression to (4.51) (see below).

Simulation II, IV, V

In these simulations, we set $c_0 = c_1 = 1$, or equivalently $\frac{c_0}{c_1} = 1$. In this case, the spatial weighting function reduces to the identity, since $m_t(\boldsymbol{x}) = \mathbb{1}_{\boldsymbol{x}\in\Pi_w} + \mathbb{1}_{\boldsymbol{x}\in\Pi\setminus\Pi_w} = \mathbb{1}_{\boldsymbol{x}\in\Pi}$, and we have

$$\mathrm{Tr}\left[\mathcal{M}_t\Sigma_t(\theta,\boldsymbol{o})\right] = \int\limits_{\Pi} \tilde{\Sigma}_t(\theta,\boldsymbol{o},\boldsymbol{x},\boldsymbol{x})\mathrm{d}\boldsymbol{x} \tag{4.50}$$

so that the objective function equally weights the uncertainty in the state estimate at all spatial locations.

Simulation III

In this simulation we set $c_0 = 1$, $c_1 = 0.01$, and $\Pi_t^w = \bigcup_{i=1}^{4}\{\boldsymbol{x} \in \Pi : |\boldsymbol{x} - \boldsymbol{x}_t^i| \leq r\}$, where $\{\boldsymbol{x}_t^i\}_{i=1}^{4}$ are the 4 time-varying locations shown in purple in Figure 4.10b. This yields $m_t(\boldsymbol{x}) = \sum_{i=1}^{4} \mathbb{1}_{\{\boldsymbol{x}\in\Pi:|\boldsymbol{x}-\boldsymbol{x}_t^i|\leq r\}} + 0.01\mathbb{1}_{\{\boldsymbol{x}\in\Pi:\cap_{i=1}^{4}|\boldsymbol{x}-\boldsymbol{x}_t^i|>r\}}$ and

$$\mathrm{Tr}\left[\mathcal{M}_t\Sigma_t(\theta,\boldsymbol{o})\right] = \sum_{i=1}^{4} \int\limits_{\{\boldsymbol{x}\in\Pi:|\boldsymbol{x}-\boldsymbol{x}_t^i|<r\}} \tilde{\Sigma}_t(\theta,\boldsymbol{o},\boldsymbol{x},\boldsymbol{x})\mathrm{d}\boldsymbol{x} + 0.01 \int\limits_{\{\boldsymbol{x}\in\Pi:\cap_{i=1}^{4}|\boldsymbol{x}-\boldsymbol{x}_t^i|>r\}} \tilde{\Sigma}_t(\theta,\boldsymbol{o},\boldsymbol{x},\boldsymbol{x})\mathrm{d}\boldsymbol{x} \tag{4.51}$$

so that the objective function strongly weights the uncertainty in the state estimate in the regions close to the time-varying locations $\{\boldsymbol{x}_t^i\}_{i=1}^{4}$, but also contains a contribution from the uncertainty in the state estimate at all other locations.

5

Parameter Estimation for the McKean-Vlasov Stochastic Differential Equation

Summary. *In this chapter, we consider the problem of parameter estimation for a stochastic McKean-Vlasov equation, and the associated system of weakly interacting particles. We first establish consistency and asymptotic normality of the offline maximum likelihood estimator for the interacting particle system in the limit as the number of particles $N \to \infty$. We then turn our attention to online parameter estimation. We propose a recursive estimator for the parameters of the McKean-Vlasov SDE, which evolves according to a continuous-time stochastic gradient descent algorithm on the asymptotic log-likelihood of the interacting particle system. We prove that this estimator converges in $\mathbb{L}^1$ to the stationary points of the asymptotic log-likelihood of the McKean-Vlasov SDE in the joint limit as $N \to \infty$ and $t \to \infty$, under suitable conditions which guarantee ergodicity and uniform-in-time propagation of chaos. We then demonstrate, under the additional condition of global strong concavity, that our estimator converges in $\mathbb{L}^2$ to the unique maximiser of this asymptotic log-likelihood function, and establish an $\mathbb{L}^2$ convergence rate. We also obtain analogous results under the condition that, rather than observing multiple trajectories of the interacting particle system, we instead observe multiple independent replicates of the McKean-Vlasov SDE itself or, less realistically, a single sample path of the McKean-Vlasov SDE and its law. Our theoretical results are demonstrated via several numerical examples, including a linear mean field model and a stochastic opinion dynamics model.*

5.1 Introduction

In this chapter, we consider a family of McKean-Vlasov stochastic differential equations (SDEs) on $\mathbb{R}^d$, parametrised by $\theta \in \mathbb{R}^p$, of the form

$$\mathrm{d}x_t^\theta = B(\theta, x_t^\theta, \mu_t^\theta)\mathrm{d}t + \sigma(x_t^\theta)\mathrm{d}w_t, \quad t \geq 0 \tag{5.1}$$

$$\mu_t^\theta = \mathcal{L}(x_t^\theta), \tag{5.2}$$

where $B : \mathbb{R}^p \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ are Borel measurable functions, $(w_t)_{t\geq 0}$ is a $\mathbb{R}^d$-valued standard Brownian motion, and $\mathcal{L}(x_t^\theta)$ denotes the law of x_t^θ. We assume that $x_0 \in \mathbb{R}^d$, or that x_0 is a $\mathbb{R}^d$-valued random variable with law μ_0, independent of $(w_t)_{t\geq 0}$. This equation is non-linear in the sense of McKean [334, 335, 434]; in particular, the coefficients depend on the law of the solution, in addition to the solution itself. We will restrict our attention to the case in which the dependence on the law only enters linearly in the drift, namely, that

$$B(\theta, x, \mu) = b(\theta, x) + \int_{\mathbb{R}^d} \phi(\theta, x, y)\mu(\mathrm{d}y), \tag{5.3}$$

for some Borel measurable functions $b : \mathbb{R}^p \times \mathbb{R}^d \to \mathbb{R}^d$ and $\phi : \mathbb{R}^p \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$. This choice of dynamics, while not the most general possible, is sufficiently broad for many applications of interest. Moreover, it includes the popular case in which b and ϕ both have gradient forms, that is, $b(\theta, x) = \nabla V_\theta(x)$ and $\phi(\theta, x, y) = \nabla W_\theta(x - y)$, in which case V_θ and W_θ are referred to as the confinement potential and the interaction potential, respectively (e.g., [170, 327]).

The McKean-Vlasov SDE arises naturally as the hydrodynamical limit ($N \to \infty$) of the mean-field interacting particle system (IPS)

$$\mathrm{d}x_t^{\theta,i,N} = B(\theta, x_t^{\theta,i,N}, \mu_t^{\theta,N})\mathrm{d}t + \sigma(x_t^{\theta,i,N})\mathrm{d}w_t^i\ , \quad i = 1, \ldots, N \tag{5.4}$$

where $(w_t^i)_{t\geq 0}$ are N independent $\mathbb{R}^d$-valued independent standard Brownian motions, x_0^i are a family of i.i.d. $\mathbb{R}^d$-valued random variables with common law μ_0, independent of $(w_t^i)_{t\geq 0}$, and $\mu_t^{\theta,N} = \frac{1}{N}\sum_{i=1}^N \delta_{x_t^{\theta,i,N}}$ is the empirical law of the interacting particles. In particular, under relatively weak assumptions, it is well known that the empirical law $\mu_t^{\theta,N} \to \mu_t^\theta$ weakly as $N \to \infty$ (e.g., [367]). This phenomenon is commonly known as the propagation of chaos [434].

The McKean-Vlasov SDE also has a natural connection to a non-linear, non-local partial differential equation on the space of probability measures (e.g., [95]). In particular, under some regularity conditions on b and ϕ, one can show that $\mathcal{L}(x_t^\theta)$ is absolutely continuous with respect to the Lebesgue measure for all $t \geq 0$ [335, 443] and its density, which we will denote by u_t^θ, satisfies a non-linear partial differential equation of the form

$$\frac{\partial u_t^\theta(x)}{\partial t} = \nabla\left[\frac{1}{2}\sigma(x)\sigma^T(x)\nabla u_t^\theta(x) + u_t^\theta(x)\left[b(\theta,x) + \int_{\mathbb{R}^d}\phi(\theta,x,y)u_t^\theta(y)\mathrm{d}y\right]\right]. \tag{5.5}$$

In the particular case that $b(x) = \nabla V(x)$ and $\phi(x,y) = \nabla W(x-y)$, this is commonly referred to as the granular media equation or the kinetic Fokker-Planck equation (e.g., [34, 95]).

5.1.1 Literature Review

The systematic study of McKean-Vlasov SDEs was first initiated by McKean [334] in the 1960s, inspired by Kac's programme in Kinetic Theory [235]. We refer to [187, 338, 434, 459] for some other classical references. In the last two decades, the study of non-linear diffusions has continued to receive considerable attention, with extensive results on well-posedness (e.g., [100, 221]), existence and uniqueness (e.g., [32, 234, 351]), ergodicity (e.g., [57, 93, 95, 171, 212, 327, 449]), and propagation of chaos (e.g., [34, 81, 170, 327, 328]). This has no doubt been motivated, at least in part, by the increasing number of applications for McKean-Vlasov SDEs, including in statistical physics [38], multi-agent systems [34], mean-field games [91], stochastic control [74], filtering [126], mathematical biology (including neuroscience [21] and structured models of population dynamics [76]), epidemic dynamics [25], social sciences (including opinion dynamics [101] and cooperative behaviours [85]), financial mathematics [201], and, perhaps most recently, high dimensional sampling [307] and neural networks [421].

Despite the recent renewed interest in the study of McKean-Vlasov SDEs, however, the problem of parameter estimation for this class of equations has received relatively little attention. This is contrast to the wealth of literature on parameter inference in linear (i.e., not measure dependent) diffusion processes (e.g., [52, 63, 273, 304]). Recently, Wen et al. [465] established the asymptotic consistency and asymptotic normality of the (offline) maximum likelihood estimator (MLE) for a broad class of McKean-Vlasov SDEs, based on continuous observation of $(x_t)_{t\in[0,T]}$. These results have since been extended by Liu et al. to the path-dependent case [306]. We also mention the work of Catalot and Laredo [194, 195], who have studied parametric inference for a particular class of one-dimensional nonlinear self-stabilising SDEs using an approximate log-likelihood function, again based on continuous observation of the non-linear diffusion process, and established the asymptotic properties (consistency, normality, convergence rates) of the resulting estimators in several asymptotic regimes (e.g., small noise and long time limit). More recently, Gomes et al. [205] have considered parameter estimation for a McKean-Vlasov PDE, based on independent realisations of the associated non-linear SDE, in the context of models for pedestrian dynamics.

In a slightly different framework, Maestra and Hoffmann [146] consider non-parametric estimation of the drift-term in a McKean-Vlasov SDE, and the solution of the corresponding non-linear Fokker-Planck equation, based on continuous observation of the associated IPS over a fixed time horizon, namely $(x_t^i)_{t\in[0,T]}^{i=1,\dots,N}$, in the limit as $N \to \infty$. The authors obtain adaptive estimators based on the solution map of the Fokker-Planck equation, and prove their optimality in a minimax sense. Moreover, in the case of the so-called Vlasov model, which in our notation corresponds to the case in which $b(x) = \nabla V(x)$ and $\phi(x, y) = \nabla W(x - y)$, the authors derive an estimator of the interaction potential, and establish its consistency. We also refer to [322, 323, 324] for some other recent contributions on non-parametric inference for IPSs. While these approaches are interesting and potentially very useful, we should emphasise that they are tangential and very different to this contribution.

Despite these recent contributions, however, to our knowledge there are no existing works which tackle the problem of online parameter estimation for McKean-Vlasov SDEs. The main purpose of this chapter is to address this gap. There is significant motivation for this approach. Indeed, in comparison to classical (offline) methods, which process observations in a batch fashion, online methods perform inference in real time, can track changes in parameters over time, are more computationally efficient, and have significantly smaller storage requirements. Even for standard diffusion processes, literature on online parameter estimation is somewhat sparse, with some notable recent exceptions [420, 422, 430]. The problem of recursive estimation in continuous time stochastic processes was first rigorously analysed by Levanony et al. [296], who proposed an online MLE which, irrespective of initial conditions, was shown to be consistent and asymptotically efficient. This estimator, however, involves computing gradients of a Girsanov log-likelihood, $\mathcal{L}_t(\theta)$, every time a new observation arrives; as a result, it is computationally expensive, and cannot be implemented in a truly online fashion, since $\nabla_\theta \mathcal{L}_t(\theta)$ depends on the entire trajectory of the process x_t. This problem has recently been revisited by Sirignano and Spiliopoulos [420, 422], who propose an online statistical learning algorithm - 'stochastic gradient descent in continuous time' - for the estimation of the parameters in a fully observed ergodic diffusion process. These authors establish the a.s. convergence of this estimator in the sense that $||\nabla_\theta g(\theta_t)|| \to 0$ as $t \to \infty$ a.s., for some suitably defined objective function $g(\theta)$ [420], and, under additional assumptions, also obtain an $\mathbb{L}^p$ convergence rate and a central limit theorem [422]. These results have since also been extended to partially observed diffusion processes [430] and jump-diffusion processes [50].

There also exists relatively little previous literature on statistical inference for IPSs, in the limit as the number of particles $N \to \infty$. Let us briefly review the main existing results on this subject. In the context of parameter estimation, the mean field regime was first analysed by Kasonga [242], who considered a system of interacting diffusion processes, depending linearly on some unknown parameter, and established that the MLE based on continuous observations over a fixed time interval $[0, T]$ is consistent and asymptotically normal in the limit as $N \to \infty$. Bishwal [53] later extended these results to the case in which the parameter to be estimated is a function of time, proving consistency and asymptotic normality of the sieve estimator (in the case of continuous observations) and an approximate MLE (in the case of discrete observations). In this chapter, we extend the results in [242] in another direction, establishing consistency and asymptotic normality of the offline MLE when the parametrisation is not linear.

More recently, Giesecke et al. [201] have established the asymptotic properties (consistency, asymptotic normality, and asymptotic efficiency) of an approximate MLE for a much broader class of dynamic interacting stochastic systems, widely applicable in financial mathematics, which additionally allow for discontinuous (i.e., jump) dynamics. In addition, Chen [110] has established the optimal convergence rate for the MLE in an interacting parameter system with linear interaction for ϕ, simultaneously in the large N (mean-field limit) and large T (long-time dynamics) regimes. None of the these works, however, considers parameter estimation for the IPS in the online setting.

5.1.2 Contributions

The main contributions of this chapter relate to both the methodology and the theory of parameter estimation for the McKean-Vlasov SDE (5.1) - (5.2). Regarding methodology:

- We discuss how one can formulate an appropriate approximation to the true likelihood function in this problem, under various modelling assumptions.
- We distinguish between cases in which the data consists of multiple paths of the IPS (Case I), multiple independent samples of the McKean-Vlasov SDE (Case II), or, less realistically, a single sample path of the McKean-Vlasov SDE and its law (Case III).

In each of these cases, we perform a rigorous asymptotic analysis of the MLE, with a focus on online parameter estimation. Our main theoretical contributions can be summarised as follows:

- In Case I, we establish asymptotic consistency and asymptotic normality of the offline MLE, in the limit as the number of particles $N \to \infty$. Our results generalise those in [242] to the case in which b and ϕ depend non-linearly on the parameter.

- In all three cases, we propose online estimators for the parameters of the McKean-Vlasov SDE, which evolve according to continuous-time stochastic gradient descent algorithms with respect to appropriate asymptotic log-likelihood functions.

- We prove that each of these estimators converges in $\mathbb{L}^1$ to the stationary points of the asymptotic log-likelihood of the McKean-Vlasov SDE, under suitable conditions which guarantee ergodicity and uniform-in-time propagation of chaos. In Cases I - II, this convergence holds in the joint limit as $N \to \infty$ and $t \to \infty$. In Case III, it holds solely in the limit as $t \to \infty$. These proofs combine ideas from [81, 171, 327, 420].

- We prove, under the additional condition that the asymptotic log-likelihood of the McKean Vlasov SDE is strongly concave, that these estimators converge in $\mathbb{L}^2$ to its unique global maximiser, in the same limits outlined above. In each case, we also obtain explicit convergence rates.

Finally, we provide numerical examples to illustrate the application of these results to several cases of interest, namely, a linear mean-field model, and a stochastic opinion dynamics model. It is worth emphasising that, given the connection between the McKean-Vlasov SDE (5.1) - (5.2) and the non-linear, non-local PDE (5.5), the results of this chapter are also applicable when one is primarily interested in parameter estimation for the non-linear PDE (5.5).

5.1.3 Chapter Organisation

The remainder of this chapter is organised as follows. In Section 5.2, we formulate the estimation problem, and propose a recursive estimator for the McKean-Vlasov SDE. In Section 5.3, we state our conditions and our main results regarding the asymptotic properties of the offline and online MLEs. In Section 5.4, we provide the proofs of these results. In Section 5.5, we provide several numerical examples illustrating the performance of the proposed algorithm. Finally, in Section 5.6, we provide some concluding remarks.

5.1.4 Additional Notation

We will assume throughout this chapter that all stochastic processes are defined canonically on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, equipped with filtration $(\mathcal{F}_t)_{t\geq 0}$. We will use $\langle \cdot, \cdot \rangle$ and $||\cdot||$ to denote, respectively, the Euclidean inner product and the corresponding norm on $\mathbb{R}^d$. We write $\mathcal{P}(\mathbb{R}^d)$ and $\mathcal{P}_p(\mathbb{R}^d)$, $p > 0$, for the collection of all probability measures on $\mathbb{R}^d$, and the collection of all probability measures on $\mathbb{R}^d$ with finite p^{th} moment. In a slight abuse of notation, we will frequently write $\mu(||\cdot||^p)$ for the p^{th} moment of μ; that is, $\mu(||\cdot||^p) = \int_{\mathbb{R}^d} ||x||^p \mu(\mathrm{d}x)$. For $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$, we write $\mathbb{W}_p(\mu, \nu)$ to denote the Wasserstein distance between μ and ν, viz

$$\mathbb{W}_p(\mu,\nu) = \inf_{\pi \in \Pi(\mu,\nu)} \left[\int\limits_{\mathbb{R}^d \times \mathbb{R}^d} ||x-y||^p \pi(\mathrm{d}x, \mathrm{d}y) \right]^{\frac{1}{\max\{1,p\}}}. \tag{5.6}$$

where $\Pi(\mu,\nu)$ for the set of all couplings of μ,ν. That is, if $\pi \in \Pi(\mu,\nu)$, then $\pi(A \times \mathbb{R}^d) = \mu(A)$ and $\pi(\mathbb{R}^d \times A) = \nu(A)$ for all $A \in \mathcal{B}(\mathbb{R}^d)$. Finally, if $(x_t)_{t\geq 0}$ is a solution of the McKean-Vlasov SDE with $x_0 = x \in \mathbb{R}^d$, we will occasionally make explicit the dependence on the initial condition by writing $\mu_t^x = \mathcal{L}(x_t)$ for the law of x_t. We can also then write $\mathbb{E}_x[f(x_t)] = \int_{\mathbb{R}^d} f(y)\mu_t^x(\mathrm{d}y)$.

5.2 Parameter Estimation for the McKean-Vlasov SDE

We will assume, throughout this chapter, that there exists a true (static) parameter $\theta_0 \in \mathbb{R}^p$ which generates observations $(x_t)_{t\geq 0} := (x_t^{\theta_0})_{t\geq 0}$ of the McKean-Vlasov SDE (5.1). Thus, we operate under the standard well specified regime, and in our notation will suppress the dependence of the observed path on the true parameter θ_0. We will assume the same condition when instead we observe trajectories of the IPS (5.4), in which case the observations are given by $(x_t^{i,N})_{t\geq 0}^{i=1,\ldots,N} = (x_t^{\theta_0,i,N})_{t\geq 0}^{i=1,\ldots,N}$.

5.2.1 The Likelihood Function

Let $\mathbb{P}_t^{\theta}$ denote the probability measure induced by a path $(x_s^{\theta})_{s\in[0,t]}$ of the McKean-Vlasov SDE (5.1). Then, under certain regularity conditions, to be specified below, one can use the Girsanov formula to obtain a likelihood function as (e.g., [465])

$$\begin{aligned}\mathcal{L}_t(\theta) = \log \frac{\mathrm{d}\mathbb{P}_t^{\theta}}{\mathrm{d}\mathbb{P}_t^{\theta_0}} = \int\limits_0^t \left\langle \left[B(\theta,x_s,\mu_s) - B(\theta_0,x_s,\mu_s)\right], (\sigma(x_s)\sigma^T(x_s))^{-1}\mathrm{d}x_s \right\rangle \\ - \frac{1}{2}\int\limits_0^t \left[\left|\left|\sigma^{-1}(x_s)B(\theta,x_s,\mu_s)\right|\right|^2 - \left|\left|\sigma^{-1}(x_s)B(\theta_0,x_s,\mu_s)\right|\right|^2 \right] \mathrm{d}s.\end{aligned} \tag{5.7}$$

Suppose, for a moment, that the diffusion coefficient σ also depended on the unknown parameter. In this case, the measures $\{\mathbb{P}_{\theta}\}$ would be mutually singular, and the likelihood function would not be well defined. We thus adopt the standard condition of parameter independence for the diffusion coefficient, and for convenience set $\sigma = 1$ (e.g., [63, 296, 465]). In the case that σ is an unknown constant, it can be estimated separately using standard methods (e.g., [193]). In fact, there are various different approaches in this case, including those based on a quasi log-likelihood function [215], or on a least squares type function for the diffusion coefficient [420]. The methods outlined in this chapter can be extended to either of these cases, as well as to parameter estimation under other criteria.

In order to proceed, it will be convenient to define the functions $G : \mathbb{R}^p \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ and $L : \mathbb{R}^p \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ according to

$$G(\theta, x, \mu) := B(\theta, x, \mu) - B(\theta_0, x, \mu) \tag{5.8}$$

$$L(\theta, x, \mu) := -\frac{1}{2}||G(\theta, x, \mu)||^2. \tag{5.9}$$

We are now ready to state our first basic assumption. This is a Novikov-type condition which ensures that $\frac{\mathrm{d}\mathbb{P}_t^{\theta}}{\mathrm{d}\mathbb{P}_t^{\theta_0}}$ exists and is a martingale. We note that several slightly weaker versions of this condition are also possible (e.g., [306, 465]).

Assumption A.1. *For all $\theta \in \mathbb{R}^p$, $t \geq 0$, the function $G : \mathbb{R}^p \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ satisfies*

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^t ||G(\theta, x_s, \mu_s)||^2 \mathrm{d}s\right)\right] < \infty. \tag{5.10}$$

Under this assumption, it follows from Girsanov's Theorem that $\mathbb{P}_t^{\theta}$ is absolutely continuous with respect to $\mathbb{P}_t^{\theta_0}$ for all $\theta \in \mathbb{R}^p$, $t > 0$ (e.g. [304, Theorem 7.19], [306, 465]), and that the log-likelihood for an observed path of the McKean-Vlasov SDE (5.1) - (5.2) is given by

$$\mathcal{L}_t(\theta) = \int_0^t L(\theta, x_s, \mu_s)\mathrm{d}s + \int_0^t \langle G(\theta, x_s, \mu_s), \mathrm{d}w_s\rangle. \tag{5.11}$$

While, in general, it is possible to observe a sample path $(x_t)_{t\geq 0}$ of a (McKean-Vlasov) SDE, in general one does not have direct access to its law $(\mu_t)_{t\geq 0}$. As such, it is generally not possible to compute the likelihood function $\mathcal{L}_t(\theta)$ in (5.11) directly. On this basis, even if one is interested in fitting data to the McKean-Vlasov SDE, it will typically be necessary to approximate the corresponding likelihood function.

In order to make such an approximation, we will henceforth assume that we can simultaneously observe multiple continuous sample paths, which is much more typical of the data that we observe in practice. There are now two possibilities. The first is to assume that the observed paths correspond to the trajectories of N particles $(x_t^{i,N})_{t\geq 0}^{i=1,\dots,N}$ from the IPS (5.4). In this case, we can approximate $\mathcal{L}_t(\theta)$ by the Girsanov log-likelihood for

Case	Data-Generating Model	Observation(s)	Likelihood Function *Approximate*	*Ideal*
Case I	IPS (5.4)	$(x_t^{i,N})_{t\geq 0}^{i=1,\dots,N}$	$\mathcal{L}_t^N(\theta)$ in (5.12)	$\mathcal{L}_t(\theta)$ in (5.11)
Case II	MVSDE (5.1) - (5.2)	$(x_t^{i})_{t\geq 0}^{i=1,\dots,N}$	$\mathcal{L}_t^{[N]}(\theta)$ in (5.13)	$\mathcal{L}_t(\theta)$ in (5.11)
Case III	MVSDE (5.1) - (5.2)	$(x_t,\mu_t)_{t\geq 0}$	-	$\mathcal{L}_t(\theta)$ in (5.11)

Table 5.1: Parameter Estimation: Summary of Different Cases

the IPS, which is given by (e.g., [53, 110, 242])

$$\mathcal{L}_t^N(\theta) := \frac{1}{N}\sum_{i=1}^N \mathcal{L}_t^{i,N}(\theta) = \frac{1}{N}\sum_{i=1}^N \left[\int_0^t L(\theta, x_s^{i,N},\mu_s^N)\mathrm{d}s + \int_0^t \langle G(\theta, x_s^{i,N},\mu_s^N), \mathrm{d}w_s^i\rangle\right], \tag{5.12}$$

where $\mu_t^N = \frac{1}{N}\sum_{j=1}^N \delta_{x_t^{j,N}}$ denotes the empirical measure of the IPS, and we have included $\frac{1}{N}$ as a normalisation factor. We will refer to this as **Case I**. The second possibility is to instead assume the observed paths are N independent instances $(x_t^i)_{t\geq 0}^{i=1,\dots,N}$ of the McKean-Vlasov SDE (5.1). In this case, we can approximate $\mathcal{L}_t(\theta)$ by

$$\mathcal{L}_t^{[N]}(\theta) := \frac{1}{N}\sum_{i=1}^N \mathcal{L}_t^{[i,N]}(\theta) = \frac{1}{N}\sum_{i=1}^N \left[\int_0^t L(\theta, x_s^{i},\mu_s^{[N]})\mathrm{d}s + \int_0^t \langle G(\theta, x_s^{i},\mu_s^{[N]}), \mathrm{d}w_s^i\rangle\right], \tag{5.13}$$

where $\mu_t^{[N]} = \frac{1}{N}\sum_{i=1}^N \delta_{x_t^i}$ denotes the empirical measure of the sample paths. In this approximation, the functions $\mathcal{L}_t^{[i,N]}(\theta)$, $i = 1,\dots,N$, correspond to N Monte Carlo approximations of $\mathcal{L}_t(\theta)$, obtained by substituting $\mu_t^{[N]}$ for μ_t. The approximation $\mathcal{L}_t^{[N]}(\theta)$ then follows by independence. We will refer to this case as **Case II**. Finally, we will refer to the rather unrealistic case in which we directly observe a single path $(x_t)_{t\geq 0}$ of the McKean-Vlasov SDE (5.1), as well as its law $(\mu_t)_{t\geq 0}$, as **Case III**. These cases are summarised in Table 5.1.

In what follows, our exposition will primarily focus on Case I, which provides the most interesting and challenging case in which to perform asymptotic analysis in both N and t. One can consider Case I and Case II as approximations to Case III that are amenable to implementation. In the limit as $N \to \infty$, standard propagation-of-chaos results (e.g., [327]) show that the dynamics of the observations in Cases I and II will coincide. In our results, we will establish rigorously that this also holds for the different implied likelihood functions, $\mathcal{L}_t^N(\theta)$ and $\mathcal{L}_t^{[N]}(\theta)$. This should not be a surprise given the similarities between these two functions: in particular, they are identical as functions of x and μ^N. We will also demonstrate that, as $N \to \infty$, these two 'approximations' also coincide with $\mathcal{L}_t(\theta)$, the 'ideal' likelihood function implied by the less realistic Case III. Moreover, we show that the same is true of the resulting parameter estimates.

5.2.2 Offline Parameter Estimation

In the offline setting, the objective is to estimate the true parameter θ_0 after receiving a batch of data over a fixed time interval $[0, t]$. Let us first consider the 'idealised' framework of Case III, in which one directly observes both $(x_s)_{s\in[0,t]}$ and $(\mu_s)_{s\in[0,t]}$ from the McKean-Vlasov SDE (5.1) - (5.2). In this case, one can achieve this objective directly by seeking to maximise the value of $\mathcal{L}_t(\theta)$ in order to obtain the MLE

$$\hat{\theta}_t = \underset{\theta\in\mathbb{R}^p}{\arg\sup}\, \mathcal{L}_t(\theta). \tag{5.14}$$

The asymptotic properties (i.e., consistency, asymptotic normality) of this estimator in the limit as $t \to \infty$, under similar conditions to our own (see Section 5.3), have recently been established [306, 465]. In this chapter, we are more interested in Case I, in which we assume that we observe N sample paths $(x_t^{i,N})_{s\in[0,t]}^{i=1,\dots,N}$ following the dynamics of the IPS (5.4). In this case, we aim instead to maximise the value of $\mathcal{L}_t^N(\theta)$, and are thus interested in the asymptotic properties of the following MLE

$$\hat{\theta}_t^N = \underset{\theta\in\mathbb{R}^p}{\arg\sup}\, \mathcal{L}_t^N(\theta). \tag{5.15}$$

The asymptotic properties of this estimator as $t \to \infty$, for fixed N, are covered by well established results for parameter estimation in standard SDEs (e.g., [52, 296, 304]). Conversely, there are very few results on the properties of this MLE in the limit as $N \to \infty$, aside from in the case of a linear parametrisation [53, 242]. We thus find it instructive to revisit this problem. In Theorems 5.1 - 5.2, we extend previous results to the more general and possible non-linear setting (in the sense of parametrisation), establishing consistency and asymptotic normality of this estimator as $N \to \infty$, for fixed t.

5.2.3 Online Parameter Estimation

In the online setting, our objective is to estimate the true parameter θ_0 in real time, using the continuous stream of observations. Once more, let us begin in the 'idealised' framework of Case III. In this case, a standard approach to this task would be to seek to recursively maximise the asymptotic log-likelihood function $\tilde{\mathcal{L}}(\theta)$ of the McKean-Vlasov

SDE, which, provided the limit exists, could be defined according to

$$\tilde{\mathcal{L}}(\theta) = \lim_{t\to\infty} \frac{1}{t}\mathcal{L}_t(\theta). \tag{5.16}$$

In the spirit of [50, 420, 430], this could be achieved using stochastic approximation by defining an estimator $\theta = (\theta_t)_{t\geq 0}$ which follows the gradient of the integrand of the log-likelihood in (5.11), evaluated with the current parameter estimate. Thus, initialised at $\theta_{\text{init}} \in \mathbb{R}^p$, θ_t evolves according to a McKean-Vlasov SDE of the form

$$\mathrm{d}\theta_t = \gamma_t \Big(\underbrace{\nabla_\theta L(\theta_t, x_t, \mu_t)\mathrm{d}t}_{\text{(noisy) ascent term}} + \underbrace{\nabla_\theta B(\theta_t, x_t, \mu_t)\mathrm{d}w_t}_{\text{noise term}} \Big) \tag{5.17}$$

where $\gamma_t : \mathbb{R}_+ \to \mathbb{R}_+^p$ is a positive, non-increasing function known as the learning rate.[1] This evolution equation represents a continuous-time stochastic gradient ascent scheme on the asymptotic log-likelihood function. To see this, let us rewrite the parameter update equation (5.17) in the form

$$\mathrm{d}\theta_t = \gamma_t \Big(\underbrace{\nabla_\theta \tilde{\mathcal{L}}(\theta_t)\mathrm{d}t}_{\text{(true) ascent term}} + \underbrace{(\nabla_\theta L(\theta_t, x_t, \mu_t) - \nabla_\theta \tilde{\mathcal{L}}(\theta_t))\mathrm{d}t}_{\text{fluctuations term}} + \underbrace{\nabla_\theta B(\theta_t, x_t, \mu_t)\mathrm{d}w_t}_{\text{noise term}} \Big) \tag{5.18}$$

The first term in this decomposition represents the true ascent direction $\nabla_\theta \tilde{\mathcal{L}}(\theta_t)$, the second term the deviation between the stochastic gradient ascent direction $\nabla_\theta L(\theta_t, x_t, \mu_t)$ and the true (deterministic) gradient ascent direction $\nabla_\theta \tilde{\mathcal{L}}(\theta_t)$, while the third term is a zero-mean noise term. Heuristically, we might expect that, provided the learning rate γ_t decreases (sufficiently quickly) with time, the ascent term will dominate the fluctuations term and the noise term when t is sufficiently large. If this is the case, we could then reasonably expect that θ_t will converge to a local maximum of $\tilde{\mathcal{L}}(\theta)$.

Similarly to the offline case, the 'ideal' online estimator (5.17) cannot typically be implemented in practice, since we do not have access to the law $(\mu_t)_{t\geq 0}$. Instead, as remarked previously, we will typically observe multiple continuous sample paths. Once again, let us first consider the case in which the N sample paths $(x_t)_{s\in[0,t]}^{i=1,\dots,N}$ are assumed to correspond to the trajectories of the IPS (5.4) (Case I). In this case, it is natural to consider the 'approximate' update equation

$$\mathrm{d}\theta_t^{i,N} = \gamma_t \left[\nabla_\theta L(\theta_t, x_t^{i,N}, \mu_t^N)\mathrm{d}t + \nabla_\theta B(\theta_t, x_t^{i,N}, \mu_t^N)\mathrm{d}w_t^i \right], \tag{5.19}$$

[1] One also arrives at this estimator by considering a 'least-squares' type objective, i.e., minimisation of the function $||G(\theta, x, \mu)||^2$ [420].

for some $i = 1, \ldots, N$, or, averaging over all of the interacting particles,

$$\mathrm{d}\theta_t^N = \gamma_t \frac{1}{N} \sum_{i=1}^N \left[\nabla_\theta L(\theta_t, x_t^{i,N}, \mu_t^N)\mathrm{d}t + \nabla_\theta B(\theta_t, x_t^{i,N}, \mu_t^N)\mathrm{d}w_t^i \right]. \tag{5.20}$$

We can also use these update equations in Case II, in which we instead assume that the N sample paths $(x_t^i)_{t\geq 0}^{i=1,\ldots,N}$ correspond to independent replicates of the McKean-Vlasov SDE (5.1). This should not be surprising on the basis of our previous remarks: in particular, the likelihood functions in Cases I and II are identical up to specification of the data. In Case II, we must simply replace $x_t^{i,N}$ by x_t^i, and μ_t^N by $\mu_t^{[N]}$ in (5.19) and (5.20). We will denote the resulting estimates by $(\theta_t^{[i,N]})_{t\geq 0}$ and $(\theta_t^{[N]})_{t\geq 0}$.

Let us briefly remark on these two schemes. The advantage of (5.19) is that the computation can be performed locally at each particle, following a message passing step for retrieving μ_t^N. It is thus convenient for a distributed implementation. On the other hand, (5.20) will typically be more accurate, as we will later demonstrate (see Theorems 5.4 and 5.4*). In Case I, these two schemes can be seen as stochastic gradient descent algorithms for maximising the 'partial' asymptotic log-likelihood of the i^{th} particle in the IPS, or the 'complete' asymptotic log-likelihood of all of the particles, respectively. That is,

$$\tilde{\mathcal{L}}^{i,N}(\theta) = \lim_{t\to\infty} \frac{1}{t}\mathcal{L}_t^{i,N}(\theta) \quad \text{or} \quad \tilde{\mathcal{L}}^{N}(\theta) = \lim_{t\to\infty} \frac{1}{t}\mathcal{L}_t^{N}(\theta). \tag{5.21}$$

Reasoning as before, we expect that, under suitable conditions on the learning rate, θ_t^N and $\theta_t^{i,N}$ will converge to local maxima of $\tilde{\mathcal{L}}^N(\theta)$ and $\tilde{\mathcal{L}}^{i,N}(\theta)$ as $t \to \infty$. Moreover, assuming uniform-in-time propagation of chaos, we can also now expect that $\tilde{\mathcal{L}}^N(\theta)$ and $\tilde{\mathcal{L}}^{i,N}(\theta)$ will converge to $\tilde{\mathcal{L}}(\theta)$ as $N \to \infty$. Thus, in the joint limit as $t \to \infty$ and $N \to \infty$ it seems reasonable to hypothesise that θ_t^N and $\theta_t^{i,N}$ will in fact converge to local maxima of $\tilde{\mathcal{L}}(\theta)$, the asymptotic log-likelihood of the original McKean-Vlasov SDE. In Theorems 5.3 - 5.4, we will establish rigorously that this is indeed the case.

5.3 Main Results

In this section, we present our main results on the asymptotic properties of the offline and online MLEs, as well as our assumptions.

5.3.1 Assumptions

Let us begin by stating our basic assumptions.

Assumption B.1. *For all $\theta \in \mathbb{R}^p$, $b(\theta, \cdot) : \mathbb{R}^d \to \mathbb{R}^d$ has the following properties.*

(i) $b(\theta,\cdot)$ *is locally Lipschitz. That is, for all* $x,x'\in\mathbb{R}^d$ *such that* $||x||,||x'||<R$*, there exists* $0<L_1<\infty$ *such that*

$$||b(\theta,x)-b(\theta,x')||\leq L_1||x-x'||. \tag{5.22}$$

(ii) $b(\theta,\cdot)$ *is 'monotonic'. That is, for all* $x,x'\in\mathbb{R}^d$*, there exists* $\alpha>0$ *such that*

$$\big\langle x-x',b(\theta,x)-b(\theta,x')\big\rangle\leq-\alpha||x-x'||^2. \tag{5.23}$$

Assumption B.2. *For all* $\theta\in\mathbb{R}^p$*,* $\phi(\theta,\cdot,\cdot):\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}^d$ *has the following properties.*

(i) $\phi(\theta,\cdot,\cdot)\in C^2(\mathbb{R}^d,\mathbb{R}^d)$*. That is,* ϕ *is twice continuously differentiable with respect to both of its arguments.*

(ii) $\phi(\theta,\cdot,\cdot)$ *is globally Lipschitz. In particular, there exists* $0<2L_2<\alpha$ *such that, for all* $x,y,x',y'\in\mathbb{R}^d$*,*

$$||\phi(\theta,x,y)-\phi(\theta,x',y')||\leq L_2(||x-x'||+||y-y'||). \tag{5.24}$$

or, in place of (ii),

(ii)' $\phi(\theta,\cdot,\cdot)$ *is 'anti-symmetric'. That is, for all* $x,y\in\mathbb{R}^d$*,* $\phi(\theta,x,y)=-\phi(\theta,y,x)$*.*

(ii)" $\phi(\theta,\cdot,\cdot)$ *increases as a function of the distance between its arguments. That is, for all* $x,y,x',y'\in\mathbb{R}^d$*,*

$$\langle(x-y)-(x'-y'),\phi(x,y)-\phi(x',y')\rangle\leq 0. \tag{5.25}$$

These two conditions are used to establish existence and uniqueness of the strong solution to the McKean-Vlasov SDE, uniform moment bounds, uniform-in-time propagation of chaos, and the existence of, and exponential convergence to, a unique invariant measure (e.g., [81, 458]). We provide a precise statement of these well known results in Appendix 5.A, which we will frequently make use of to prove the main results in this chapter.

In the literature on non-linear diffusions, it is typical, as noted previously, to consider the case in which $b(\theta,x)=-\nabla V(\theta,x)$ for some confinement potential V, and $\phi(\theta,x,y)=-\nabla W(\theta,x-y)$ for some interaction potential W. In this context, Condition B.1(ii) is equivalent to the condition that V is strongly convex with parameter α, and Conditions B.2(ii)'-(ii)" are equivalent to the conditions that W is even and convex (see [327]). These are perhaps the simplest and most well established conditions under which the results listed above (uniform-in-time propagation of chaos, exponential convergence to a unique invariant measure) can be obtained; we have thus adopted them here for ease of exposition.

This being said, let us remark briefly upon some weaker conditions under which these results still hold, and therefore under which the main results of our chapter will also still hold (albeit with some additional technical overhead). In the case that there is no confinement potential (i.e. $V \equiv 0$), and the interaction potential is uniformly convex with gradient that is locally Lipschitz with polynomial growth, Malrieu established uniform-in-time propagation of chaos and exponential convergence to equilibrium [328]. Cattiaux et al. [95] later established the same results in the case that the interaction potential is degenerately convex. Meanwhile, in [93, 94], the authors establish exponential convergence to equilibrium under the strict convexity condition $\text{Hess}(V + 2W) \geq \beta I_d$, for some $\beta > 0$.

In the case that $V + 2W$ is not convex, far fewer results are available; indeed, without additional conditions on V and W, even the existence of a unique stationary distribution is not guaranteed (see, e.g., [212, 213, 449, 450]). This being said, Bolley et al. [57] proved uniform exponential convergence to equilibrium in both degenerately convex, and weakly non-convex cases. More recently, [170, 171] have established uniform-in-time propagation of chaos and exponential convergence to equilibrium in the non-convex case, provided the confinement potential V is strictly convex outside a ball, and the interaction potential is globally Lipschitz with sufficiently small Lipschitz constant. For a recent extension of these results, see also [308].

We will also require the following regularity condition.

Assumption C.1. *The functions $b : \mathbb{R}^p \times \mathbb{R}^d \to \mathbb{R}^d$ and $\phi : \mathbb{R}^p \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ have the following properties.*

(i) $\nabla_\theta b(\cdot, x), \nabla_\theta \phi(\cdot, x, y) \in \mathcal{C}^2(\mathbb{R}^p)$ *for all* $x, y \in \mathbb{R}^d$, $\frac{\partial^2}{\partial x^2}\nabla_\theta b \in \mathcal{C}(\mathbb{R}^p, \mathbb{R}^d)$, $\frac{\partial^2}{\partial x^2}\nabla_\theta \phi \in \mathcal{C}(\mathbb{R}^p, \mathbb{R}^d, \mathbb{R}^d)$, *and* $\nabla_\theta^i b(\theta, \cdot) \in \mathcal{C}^{1+\alpha}(\mathbb{R}^d)$, $\nabla_\theta^i \phi(\theta, \cdot, \cdot) \in \mathcal{C}^{1+\alpha}(\mathbb{R}^d, \mathbb{R}^d)$, $i = 1, 2$, *uniformly in* $\theta \in \mathbb{R}^p$ *for some* $\alpha \in (0, 1)$.[2]

(ii) The functions $\nabla_\theta^i b(\theta, \cdot)$ *and* $\nabla_\theta^i \phi(\theta, \cdot, \cdot)$ *are locally Lipschitz with polynomial growth. That is, there exist constants* $q, K < \infty$ *such that, for* $i = 0, 1, 2, 3$,

$$\left|\left|\nabla_\theta^i b(\theta, x) - \nabla_\theta^i b(\theta, x')\right|\right| \leq K||x - x'|| \left[1 + ||x||^q + ||x'||^q\right] \tag{5.26}$$

$$\left|\left|\nabla_\theta^i \phi(\theta, x, y) - \nabla_\theta^i \phi(\theta, x', y')\right|\right| \leq K \left[||x - x'|| + ||y - y'||\right] \tag{5.27}$$

$$\cdot \left[1 + ||x||^q + ||x'||^q + ||y||^q + ||y'||^q\right]. \tag{5.28}$$

(iii) $b(\theta_0, \cdot) \in C^{2+\alpha}(\mathbb{R}^d)$, $\phi(\theta_0, \cdot, \cdot) \in C^{2+\alpha}(\mathbb{R}^d, \mathbb{R}^d)$ *with* $\alpha \in (0, 1)$. *Namely, these functions have two derivatives, with all partial derivatives Hölder continuous with exponent* α.

[2]In fact, we only require that these properties hold for the function $L(\theta, x, \mu)$, as defined in (5.8) - (5.9). We find it more convenient, however, to specify this condition in terms of the functions $b(\theta, x)$ and $\phi(\theta, x, y)$.

In the offline setting, these conditions are required in order to control the growth of the log-likelihood function and its derivatives. In the online setting, they are required in order to control the ergodic behaviour of the solution of the McKean-Vlasov SDE (and the associated IPS), which is central to establishing convergence of the online MLE. In particular, they ensure that fluctuations terms of the form $\int_0^t \gamma_s(\nabla_\theta L(\theta_s, x_s, \mu_s) - \nabla_\theta \tilde{\mathcal{L}}(\theta_s))\mathrm{d}s$, which arise due to the noisy online estimate of the gradient of the asymptotic log-likelihood function $\nabla_\theta \tilde{\mathcal{L}}(\theta_s)$, c.f. (5.18), tend to zero sufficiently quickly as $t \to \infty$. Using an approach which is now well established in the analysis of stochastic approximation algorithms in continuous time (e.g., [50, 413, 420, 422, 430]), we control such terms by rewriting them in terms of the solutions of some related Poisson equations. Condition C.1 ensures that these solutions are unique, and that they grow at most polynomially in a suitable sense (see Lemma 5.16 in Appendix 5.C).

We should remark that, for the sake of convenience and to remain in line with much of the recent literature, we have restricted our attention to the case in which the measure enters only linearly in the drift coefficient $B(\theta, x, \mu)$. As such, our main conditions are stated in terms of the functions $b : \mathbb{R}^p \times \mathbb{R}^d \times \mathbb{R}^d$ and $\phi : \mathbb{R}^p \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$. Our main results, however, can be extended straightforwardly to more general choices of interaction function, under suitable conditions on $B : \mathbb{R}^p \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$. In particular, in the online setting, we simple require conditions which guarantee the existence of a unique invariant measure, and uniform-in-time propagation of chaos. As an example, we can replace Condition C.1(ii) by $||\nabla_\theta B(\theta, x, \mu)|| \leq K[1 + ||x||^q + \mu(|| \cdot ||^q)]$. Finally, we will require the following assumption on the initial condition.

Assumption D.1. *The initial law satisfies $\mu_0 \in \mathcal{P}_k(\mathbb{R}^d)$ for all $k \in \mathbb{N}$.*

This condition guarantees that the solutions of the McKean-Vlasov SDE and the associated IPS have bounded moments of all orders (see Proposition 5.2), and so do their invariant measures (see Lemma 5.3). In turn, this ensures that one can control the polynomial growth of the log-likelihood and its derivatives (in the offline case), and the polynomial growth of the solutions of the relevant Poisson equations (in the online case). We should remark that, in the offline case, we can significantly weaken this assumption: in particular, we only require that $\mu_0 \in \mathcal{P}_q(\mathbb{R}^d)$, where q is the order of the polynomial growth of the functions $b(\theta, \cdot)$ and $\phi(\theta, \cdot, \cdot)$ (see Condition C.1). One can also slightly relax this condition in the online case, though in a much more cumbersome fashion.[3] We note that this condition is trivially satisfied in the case that $x_0 \in \mathbb{R}^d$.

[3] In particular, in the online case, one requires $\mu_0 \in \mathcal{P}_k(\mathbb{R}^d)$, where k is the maximum order of polynomial growth of a solution of any of the relevant Poisson equations appearing in the proofs of Theorem 5.3 and Theorem 5.4.

5.3.2 Offline Parameter Estimation

In the case of offline parameter estimation, we will require the following additional assumptions.

Assumption E.1. *For all $t > 0$, and for all $\theta \in \mathbb{R}^p$, the function $m_t : \mathbb{R}^p \to \mathbb{R}$, defined according to*

$$m_t(\theta) = \int_0^t \int_{\mathbb{R}^d} L(\theta, x, \mu_s)\mu_s(\mathrm{d}x)\mathrm{d}s \tag{5.29}$$

satisfies $\sup_{||\theta-\theta_0||>\delta} m_t(\theta) < 0$ *a.s.* $\forall \delta > 0$.

Assumption E.2. *For all $t > 0$, the matrix $I_t(\theta_0) = [I_t(\theta_0)]_{k,l=1,\dots,p} \in \mathbb{R}^{p\times p}$, defined according to*

$$[I_t(\theta_0)]_{kl} = \int_0^t \int_{\mathbb{R}^d} [\nabla_\theta B(\theta_0, x, \mu_s)]_k [\nabla_\theta B(\theta_0, x, \mu_s)]_l \mu_s(\mathrm{d}x)\mathrm{d}s \tag{5.30}$$

is positive-definite, with $\lambda^T I_t(\theta_0)\lambda$ increasing for all $\lambda \in \mathbb{R}^p$, and $I_0(\theta_0) = 0$.

The first of these two conditions relates to parameter identifiability, guaranteeing the uniqueness of θ_0 as the optimal parameter in the sense of some asymptotic cost, and is necessary in order to establish consistency of the MLE as $N \to \infty$. It can be seen, in some sense, as an analogue of the classical condition used to obtain consistency in the long-time regime (e.g., [63], [312, pp. 137-139], [296, pp. 252 - 253] [390, Condition A_5]). It is also closely related to the so-called 'coercivity condition', introduced in [58], which appears in the study of non-parametric inference for IPSs (see also [301, 322, 323, 324]). Notably, this condition holds if the parametrisation is linear. Meanwhile, the second condition is necessary in order to establish asymptotic normality, and can be seen as a generalisation of a similar condition introduced in [242] (see also [53]).

We are now ready to state our two main results in the offline case.

Theorem 5.1. *Assume that Conditions A.1, B.1 - B.2, C.1, D.1, and E.1 hold. Let $\Theta \subseteq \mathbb{R}^p$ be a compact set, and suppose $\theta_0 \in \Theta$. Then, for all $t > 0$, $\hat{\theta}_t^N$ is a weakly consistent estimator of θ_0 as $N \to \infty$. That is, as $N \to \infty$,*

$$\hat{\theta}_t^N \xrightarrow{\mathbb{P}} \theta_0. \tag{5.31}$$

Proof. See Section 5.4.1. □

Theorem 5.2. *Assume that Conditions A.1, B.1 - B.2, C.1, D.1, and E.1 - E.2 hold. Let $\Theta \subseteq \mathbb{R}^p$ be a compact set, and suppose $\theta_0 \in \Theta$. Then, for all $t > 0$, $N^{\frac{1}{2}}(\hat{\theta}_t^N - \theta_0)$ is asymptotically normal with mean zero and variance $I_t^{-1}(\theta_0)$. That is, as $N \to \infty$,*

$$N^{\frac{1}{2}}(\hat{\theta}_t^N - \theta_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0, I_t^{-1}(\theta_0)). \tag{5.32}$$

Proof. See Section 5.4.2. □

5.3.3 Online Parameter Estimation

In the online case, we will first require the following standard condition on the learning rate.

Assumption F.1. *The learning rate $\gamma_t : \mathbb{R}_+ \to \mathbb{R}_+$ is a positive, non-increasing function such that $\int_0^\infty \gamma_t \mathrm{d}t = \infty$, $\int_0^\infty \gamma_t^2 \mathrm{d}t < \infty$, $\int_0^\infty \gamma_t' \mathrm{d}t < \infty$. Moreover, there exists $p > 0$ such that $\lim_{t\to\infty} \gamma_t^2 t^{2p+\frac{1}{2}} = 0$.*

This condition can be seen as the continuous-time analogue of the standard step-size condition used in the convergence analysis of stochastic approximation algorithms in discrete time (e.g., [396, 420]).

We now proceed with some additional assumptions, which will only be required for our $\mathbb{L}^2$ convergence results (Theorems 5.4, 5.4*, 5.4†, 5.4‡).

Assumption F.2. *Let $\Phi_{s,t} = \exp(-2\eta \int_s^t \gamma_u \mathrm{d}u)$, for the constant η defined below in Condition H.1. The learning rate $\gamma_t : \mathbb{R}_+ \to \mathbb{R}_+$ satisfies $\int_0^t \gamma_s^2 \Phi_{s,t} \mathrm{d}s = O(\gamma_t)$, $\int_0^t \gamma_t' \Phi_{s,t} \mathrm{d}s = O(\gamma_t)$, $\int_0^t \gamma_s \Phi_{s,t} \mathrm{d}s = O(1)$, $\int_0^t \gamma_s \Phi_{s,t} e^{-\lambda s} \mathrm{d}s = O(\gamma_t)$, and $\Phi_{1,t} = O(\gamma_t)$.*

This is another condition on the learning rate, first introduced in [422], and is specific to stochastic gradient descent in continuous time. A standard choice of learning rate which satisfies both of these conditions is $\gamma_t = C_\gamma (C_0 + t)^{-1}$, where $C_\gamma, C_0 > 0$ are positive constants such that $C_\gamma \eta > 1$.

In addition, we introduce the following two assumptions.

Assumption G.1. *There exists a positive constant $R < \infty$, and an almost everywhere positive function $\kappa : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$, such that, for all $||\theta|| \geq R$,*

$$\langle \nabla_\theta L(\theta, x, \mu), \theta \rangle \leq -\kappa(x, \mu) ||\theta||^2. \tag{5.33}$$

Assumption G.2. *Define the function $\tau : \mathbb{R}^p \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ according to*

$$\tau(\theta, x, \mu) = \Big\langle \nabla_\theta B(\theta, x, \mu) \nabla_\theta B^T(\theta, x, \mu) \frac{\theta}{||\theta||}, \frac{\theta}{||\theta||} \Big\rangle^{\frac{1}{2}} \tag{5.34}$$

Then, there exists $0 < q, K < \infty$ such that, for all $\theta, \theta' \in \mathbb{R}^p$, for all $x, y \in \mathbb{R}^d$,

$$|\tau(\theta, x) - \tau(\theta', x)|| \leq K ||\theta - \theta'|| (1 + ||x||^q + ||\mu(|| \cdot ||^2)||^{\frac{q}{2}}) \tag{5.35}$$

These two conditions ensure, via the comparison theorem (e.g., [228, 478]), that the online parameter estimates generated by the McKean-Vlasov SDE and the IPS, namely $(\theta_t)_{t\geq 0}$ and $(\theta_t^{i,N})_{t\geq 0}$, have uniformly bounded moments (see Lemma 5.1). We refer to [243] for some more general conditions under which this result still holds. The first condition relates to the drift terms in the two parameter update equations, and can be seen as a recurrence condition; the second condition relates to the diffusion terms, and can be seen as an extension of Condition B.2(ii). In particular, in the case that $\theta \in \mathbb{R}$, Condition G.2 essentially reduces to Condition B.2(ii). This condition was introduced in [420], and has since also appeared in [50].

Finally, to establish consistency, we will require the following assumptions on the concavity of the log-likelihood.

Assumption H.1. *The function $\tilde{\mathcal{L}}(\theta)$ is strongly concave. That is, there exists $\eta > 0$ such that, for all $\theta, \theta' \in \mathbb{R}^p$,*

$$\tilde{\mathcal{L}}(\theta') \leq \tilde{\mathcal{L}}(\theta) + \nabla\tilde{\mathcal{L}}(\theta)^T(\theta' - \theta) - \frac{\eta}{2}||\theta' - \theta||^2. \tag{5.36}$$

Assumption H.1'. *The function $\tilde{\mathcal{L}}^{i,N}(\theta)$ is strongly concave, for all $N \in \mathbb{N}$, for all $i = 1, \ldots, N$. That is, there exists $\eta^{i,N} > 0$ such that, for all $\theta, \theta' \in \mathbb{R}^p$,*

$$\tilde{\mathcal{L}}^{i,N}(\theta') \leq \tilde{\mathcal{L}}^{i,N}(\theta) + \nabla\tilde{\mathcal{L}}^{i,N}(\theta)^T(\theta' - \theta) - \frac{\eta^{i,N}}{2}||\theta' - \theta||^2. \tag{5.37}$$

These conditions relate to the properties of the asymptotic log-likelihoods of the McKean Vlasov SDE and the IPS, respectively. They imply, in particular, that $\tilde{\mathcal{L}}(\theta)$ and $\tilde{\mathcal{L}}^{i,N}(\theta)$ have unique maximisers, say θ_* and θ_*^N. Under certain identifiability assumptions, these must in fact be equal to the true parameter θ_0 (e.g., [296]). We note that the first of these conditions is slightly weaker than the second. Indeed, under the first assumption, we only establish that $\theta_t^{i,N} \xrightarrow{\mathbb{L}^2} \theta_0$ as $t \to \infty$ and $N \to \infty$ (Theorem 5.4), while under the second assumption, we establish that $\theta_t^{i,N} \xrightarrow{\mathbb{L}^2} \theta_0$ as $t \to \infty$ for all $N \in \mathbb{N}$ (Theorem 5.4†). Thus, under the second assumption, there is no requirement to take the limit as $N \to \infty$. We also obtain a sharper $\mathbb{L}^2$ convergence rate.

We are now ready to state our main results in the online case. These results, categorised according to different cases introduced in Section 5.2.1, are summarised in Table 5.2. We begin by considering Case I.

Case I

In this case, we assume that we observe the trajectories of N particles $(x_t^i)_{t\geq 0}^{i=1,\ldots,N}$ of the IPS (5.4). We can thus generate online parameter estimates according to (5.19) or (5.20). We here show that, in the limit as $N \to \infty$ and $t \to \infty$, these parameter estimates can maximise $\tilde{\mathcal{L}}(\theta)$, the asymptotic log-likelihood of the McKean-Vlasov SDE. Thus, the proposed approach with finite N can be thought of as a principled approximate method for estimating the unknown parameter θ of the McKean-Vlasov SDE in an online fashion. In our first result, we establish $\mathbb{L}^1$ convergence of (5.19) and (5.20) to the stationary points of $\tilde{\mathcal{L}}(\theta)$.

Theorem 5.3. *Assume that Conditions A.1, B.1 - B.2, C.1, D.1, and F.1 hold. Then, in $\mathbb{L}^1$, it holds that*

$$\lim_{N\to\infty}\lim_{t\to\infty}||\nabla_\theta\tilde{\mathcal{L}}(\theta_t^{i,N})|| = \lim_{t\to\infty}\lim_{N\to\infty}||\nabla_\theta\tilde{\mathcal{L}}(\theta_t^{i,N})|| = 0, \tag{5.38}$$

$$\lim_{N\to\infty}\lim_{t\to\infty}||\nabla_\theta\tilde{\mathcal{L}}(\theta_t^{N})|| = \lim_{t\to\infty}\lim_{N\to\infty}||\nabla_\theta\tilde{\mathcal{L}}(\theta_t^{N})|| = 0. \tag{5.39}$$

Proof. See Section 5.4.3. □

In our second result, under additional assumptions, we establish $\mathbb{L}^2$ convergence to the unique maximiser of $\tilde{\mathcal{L}}(\theta)$.

Theorem 5.4. *Assume that Conditions A.1, B.1 - B.2, C.1, D.1, F.1 - F.2, G.1 - G.2, and H.1 hold. Then, for sufficiently large t, $N \geq 1$, $1 \leq i \leq N$, there exist positive constants K_1, K_2, K_3 such that*

$$\mathbb{E}\left[||\theta_t^{i,N} - \theta_0||^2\right] \leq (K_1 + K_2)\,\gamma_t + \frac{K_3}{N^{\frac{1}{2}}}, \tag{5.40}$$

$$\mathbb{E}\left[||\theta_t^{N} - \theta_0||^2\right] \leq (K_1 + \frac{K_2}{N})\gamma_t + \frac{K_3}{N^{\frac{1}{2}}}, \tag{5.41}$$

Proof. See Section 5.4.4. □

We can also obtain similar results in Cases II and III. Indeed, having established these two convergence results in Case I, analogous results in the remaining cases follow via very similar arguments. With this mind, and in the interest of brevity, we have chosen to omit detailed proofs of the main results in Cases II and III. These proofs can be found in [414].

Case II

In this case, we assume that we observe independent sample paths $(x_t^i)_{t\geq 0}^{i=1,\dots,N}$ of the McKean-Vlasov SDE (5.1). We thus generate online parameter estimates according to (5.19) or (5.20), replacing $x_t^{i,N}$ by x_t^i, and μ_t^N by $\mu_t^{[N]}$ where appropriate. In this case, we obtain the following statement of our results, similarly to Case I.

Case	**Theorems**	**Parameter Estimates**	**Objective Function**		**Convergence Rate**
Case I	5.3 - 5.4	$\theta_t^{i,N}$ from (5.19)	$\tilde{\mathcal{L}}(\theta)$	MSVDE	$(K_1+K_2)\gamma_t + \frac{K_3}{N^{\frac{1}{2}}}$
		θ_t^{N} from (5.20)			$(K_1+\frac{K_2}{N})\gamma_t + \frac{K_3}{N^{\frac{1}{2}}}$
Case II	5.3^* - 5.4^*	$\theta_t^{[i,N]}$ from (5.19)	$\tilde{\mathcal{L}}(\theta)$	MSVDE	$(K_1^*+K_2^*)\,\gamma_t$
		$\theta_t^{[N]}$ from (5.20)			$(K_1^*+\frac{K_2^*}{N})\gamma_t$
Case III	$5.3^\dagger$ - $5.4^\dagger$	θ_t from (5.17)	$\tilde{\mathcal{L}}(\theta)$	MSVDE	$(K_1^\dagger+K_2^\dagger)\gamma_t$
Case I (*finite N*)	$5.3^\ddagger$ - $5.4^\ddagger$	$\theta_t^{i,N}$ from (5.19)	$\tilde{\mathcal{L}}^{i,N}(\theta)$	IPS (Partial)	$(K_1^\ddagger+K_2^\ddagger)\gamma_t$
		θ_t^{N} from (5.20)	$\tilde{\mathcal{L}}^{N}(\theta)$	IPS (Complete)	$(K_1^\ddagger+\frac{K_2^\ddagger}{N})\gamma_t$

Table 5.2: Online Parameter Estimation: Summary of Main Results

Theorem 5.3*. *Assume that Conditions A.1, B.1 - B.2, C.1, D.1, and F.1 hold. Then, in $\mathbb{L}^1$, it holds that*

$$\lim_{N\to\infty}\lim_{t\to\infty}||\nabla_\theta\tilde{\mathcal{L}}(\theta_t^{[i,N]})|| = \lim_{t\to\infty}\lim_{N\to\infty}||\nabla_\theta\tilde{\mathcal{L}}(\theta_t^{[i,N]})|| = 0, \tag{5.42}$$
$$\lim_{N\to\infty}\lim_{t\to\infty}||\nabla_\theta\tilde{\mathcal{L}}(\theta_t^{[N]})|| = \lim_{t\to\infty}\lim_{N\to\infty}||\nabla_\theta\tilde{\mathcal{L}}(\theta_t^{[N]})|| = 0. \tag{5.43}$$

Theorem 5.4*. *Assume that Conditions A.1, B.1 - B.2, C.1, D.1, F.1 - F.2, G.1 - G.2, and H.1 hold. Then, for sufficiently large t, there exist positive constants K_1^*, K_2^*, such that*

$$\mathbb{E}\left[||\theta_t^{[i,N]}-\theta_0||^2\right] \leq (K_1^*+K_2^*)\,\gamma_t, \tag{5.44}$$
$$\mathbb{E}\left[||\theta_t^{[N]}-\theta_0||^2\right] \leq (K_1^*+\frac{K_2^*}{N})\gamma_t. \tag{5.45}$$

Proof. See [414, Appendix F]. □

Let us briefly compare the results obtained in Case I (Theorems 5.3 - 5.4) and in Case II (Theorems 5.3^* - 5.4^*). As remarked previously, the online parameter estimates in both of these cases follow the same parameter update equations; the only difference is the assumed form of the data-generating model. It is thus expected that the results obtained in these two cases will be similar, if not identical. In Theorems 5.3 and 5.3^*, this is indeed seen to be the case. These results establish that, regardless of the assumed form of the data-generating mechanism, the online parameter estimates generated via (5.19) or (5.20) converge to the stationary points of $\tilde{\mathcal{L}}(\theta)$ as $N\to\infty$ and $t\to\infty$. On the other hand, in Theorems 5.4 and 5.4^*, a difference does arise between the two $\mathbb{L}^2$ convergence rates. In particular, in Case I (Theorem 5.4) there is an additional $\mathcal{O}(\frac{1}{N^{\frac{1}{2}}})$ term. We can interpret this term as a penalty for the mismatch between the likelihood implied by the assumed data-generating model in Case I, namely the IPS (5.4), and the likelihood implied by the McKean-Vlasov SDE (5.1) - (5.2), which is the function that we are seeking to optimise.

Case III

In this case, we assume that we can observe not only a sample path $(x_t)_{t\geq 0}$ of the non-linear SDE, but also its law $(\mu_t)_{t\geq 0}$. We can thus generate online parameter estimates according to (5.17). In this case, we obtain the following statement of our results.

Theorem 5.3†. *Assume that Conditions A.1, B.1 - B.2, C.1, D.1, F.1 hold. Then, in $\mathbb{L}^1$, it holds that*

$$\lim_{t\to\infty} ||\nabla_\theta \tilde{\mathcal{L}}(\theta_t)|| = 0. \tag{5.46}$$

Theorem 5.4†. *Assume that Conditions A.1, B.1 - B.2, C.1, D.1, F.1 - F.2, G.1 - G.2 and H.1 hold. Then, for sufficiently large t, there exist positive constants $K_1^{\dagger}, K_2^{\dagger}$, such that*

$$\mathbb{E}\left[||\theta_t - \theta_0||^2\right] \leq (K_1^{\dagger} + K_2^{\dagger})\gamma_t. \tag{5.47}$$

Proof. See [414, Appendix G]. □

These results represent an extension of [420, Theorem 2.4] and [422, Proposition 2.13], respectively, to the McKean-Vlasov case. We should remark that a rather more direct proof of these results may be possible, which does not require us to pass between the McKean-Vlasov SDE and the IPS, but rather which works directly with the non-linear equation. This would require a significant extension of the recent results obtained in [397] regarding the regularity of the solutions of a non-linear, non-local Poisson equation.

In some sense, this scenario is mainly of theoretical interest, since in practice we do not measure the law of the non-linear process. In principle, one could circumvent this by integrating the McKean-Vlasov PDE (5.5) in parallel with the parameter update equation (5.17). In particular, starting from some initial law μ_0 and some initial parameter estimate θ_0, one would simultaneously update $(\mu_t)_{t\geq 0}$ according to (5.5), now integrated along the path of the online parameter estimates, and update $(\theta_t)_{t\geq 0}$ according to (5.17), now integrated along the path of the approximate laws. This yields an estimator which only requires us to observe a single sample path $(x_t)_{t\geq 0}$ of the McKean-Vlasov SDE, but at the cost of having to solve numerically the non-linear, non-local PDE (5.5), and analyse the resulting numerical error. We leave a rigorous analysis of this approach to future work.

Case I (*finite* N)

For the sake of completeness, we conclude this section by revisiting Case I, now under the condition that the number of particles N if fixed and finite, and that we are only interested in long-time asymptotics. In particular, our objective is now simply to maximise the asymptotic log-likelihood of the IPS, $\tilde{\mathcal{L}}^{i,N}(\theta)$. In this case, we have the following.

Theorem 5.3‡. *Assume that Conditions A.1, B.1 - B.2, C.1, D.1, and F.1 hold. Then, in $\mathbb{L}^1$, it holds that*

$$\lim_{t\to\infty} ||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N})|| = \lim_{t\to\infty} ||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N})|| = 0. \tag{5.48}$$

$$\lim_{t\to\infty} ||\nabla_\theta \tilde{\mathcal{L}}^{N}(\theta_t^{N})|| = \lim_{t\to\infty} ||\nabla_\theta \tilde{\mathcal{L}}^{N}(\theta_t^{N})|| = 0. \tag{5.49}$$

Theorem 5.4‡. *Assume that Conditions B.1 - B.2, C.1, D.1, F.1 - F.2, G.1 - G.2, and H.1' hold. Then, for sufficiently large t, there exist positive constants $K_1^{\ddagger}, K_2^{\ddagger}$, such that*

$$\mathbb{E}\left[||\theta_t^{i,N} - \theta_0||^2\right] \leq \left(K_1^{\ddagger} + K_2^{\ddagger}\right)\gamma_t. \tag{5.50}$$

$$\mathbb{E}\left[||\theta_t^{N} - \theta_0||^2\right] \leq (K_1^{\ddagger} + \frac{K_2^{\ddagger}}{N})\gamma_t. \tag{5.51}$$

Proof. See [414, Appendix H]. □

Theorem 5.4‡ demonstrates that, if the asymptotic log-likelihood of the IPS is sufficiently well-behaved (i.e., strongly concave) for finite values of $N \in \mathbb{N}$, then the parameter estimate generated using the IPS is guaranteed to converge to the true parameter value as $t \to \infty$ for all values of $N \in \mathbb{N}$. In particular, it is no longer necessary to take the limit as $N \to \infty$. This is clear upon comparison of the convergence rates (5.40) - (5.41) obtained in Theorem 5.4 and the convergence rates (5.50) - (5.51) obtained in Theorem 5.4‡.

It is worth emphasising that the differences between (5.40) - (5.41) in Theorem 5.4 and (5.50) - (5.51) in Theorem 5.4‡, arise solely due to the different assumptions imposed, namely Assumption H.1 and Assumption H.1'. In particular, the additional terms appearing in (5.40) - (5.41) can loosely be regarded as upper bounds on the difference between the maxima of $\tilde{\mathcal{L}}(\cdot)$ and $\tilde{\mathcal{L}}^{i,N}(\cdot)$, which only arise when $\tilde{\mathcal{L}}(\cdot)$ if strongly concave, but $\tilde{\mathcal{L}}^{i,N}(\cdot)$ is not. That is, if Assumption H.1 is satisfied but Assumption H.1' is not. Meanwhile, if Assumption H.1' is satisfied, then the global maxima of $\tilde{\mathcal{L}}^{i,N}(\cdot)$ and $\tilde{\mathcal{L}}(\cdot)$ will both coincide with the true parameter, and thus this difference vanishes.

5.4 Proof of Main Results

In this section, we provide proofs of our main results. Many of these proofs will rely on additional auxiliary lemmas; in the interest of brevity, the statements and proofs of these lemmas have been deferred to the appendices.

5.4.1 Proof of Theorem 5.1

We begin by establishing consistency of the offline MLE as $N \to \infty$. We should emphasise that, throughout this proof, the value of t will be fixed and finite. This being said, our method of proof will broadly follow the classical approach for establishing strong consistency of the MLE in a different asymptotic regime, namely, in the limit as $t \to \infty$ (e.g., [63]). Since we consider an entirely different asymptotic regime, however, at times we will need to rely on slightly different arguments (e.g., Lemma 5.2), and, of course, different conditions (e.g., Condition E.1).

Proof. Let $\mathbb{P}^{\theta}_{t,N}$ denote the probability measure induced by $(x_s^{\theta,i,N})_{s\in[0,t]}^{i=1,\dots,N}$. We begin with the observation that, since $\Theta \subseteq \mathbb{R}^p$ is compact, for all $t \geq 0$, and for all $N \in \mathbb{N}$, there exists $\hat{\theta}_t^N \in \Theta$ such that

$$\frac{\mathrm{d}\mathbb{P}^{\theta}_{t,N}}{\mathrm{d}\mathbb{P}^{\theta_0}_{t,N}}\bigg|_{\theta=\hat{\theta}_t^N} \geq \frac{\mathrm{d}\mathbb{P}^{\tilde{\theta}}_{t,N}}{\mathrm{d}\mathbb{P}^{\theta_0}_{N,t}} \quad \text{a.s.} \tag{5.52}$$

for all $\tilde{\theta} \in \Theta$. We thus have, setting $\tilde{\theta} = \theta_0$ in the above, that $\frac{\mathrm{d}\mathbb{P}^{\theta}_{t,N}}{\mathrm{d}\mathbb{P}^{\theta_0}_{t,N}}|_{\theta=\hat{\theta}_t^N} \geq 1$ a.s. from which it follows that $\mathcal{L}_t^N(\hat{\theta}_t^N) = \log[\frac{\mathrm{d}\mathbb{P}_{\theta_t^N}}{\mathrm{d}\mathbb{P}_{\theta_{0,t}^N}}]_{\theta=\hat{\theta}_t^N} \geq 0$ a.s., It follows, using the definition of the log-likelihood, that, a.s. ,

$$\frac{1}{N}\sum_{i=1}^N \int_0^t \langle G(\theta, x_s^{i,N}, \mu_s^N), \mathrm{d}w_s^i\rangle_{\theta=\hat{\theta}_t^N} \geq \frac{1}{2N}\sum_{i=1}^N \int_0^t \left|\left|G(\hat{\theta}_t^N, x_s^{i,N}, \mu_s^N)\right|\right|^2 \mathrm{d}s \geq 0. \tag{5.53}$$

In addition, by Lemma 5.2, we have that $\frac{1}{N}\sum_{i=1}^N \int_0^t \langle G(\theta, x_s^{i,N}, \mu_s^N), \mathrm{d}w_s^i\rangle_{\theta=\hat{\theta}_t^N} \xrightarrow{\mathbb{P}} 0$ as $N \to \infty$. It follows straightforwardly that, as $N \to \infty$,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N \int_0^t \left|\left|G(\hat{\theta}_t^N, x_s^{i,N}, \mu_s^N)\right|\right|^2 \mathrm{d}s \xrightarrow{\mathbb{P}} 0. \tag{5.54}$$

We next observe, making use of the Cauchy-Schwarz inequality, that

$$\left|\frac{1}{N}\sum_{i=1}^N \int_0^t ||G(\theta, x_s^{i,N}, \mu_s^N)||^2 \mathrm{d}s - \frac{1}{N}\sum_{i=1}^N \int_0^t ||G(\theta', x_s^{i,N}, \mu_s^N)||^2 \mathrm{d}s\right| \tag{5.55}$$

$$\leq \left[\frac{1}{N}\sum_{i=1}^{N}\int_0^t \big|\big|G(\theta,x_s^{i,N},\mu_s^N) - G(\theta',x_s^{i,N},\mu_s^N)\big|\big|^2\mathrm{d}s\right]^{\frac{1}{2}} \tag{5.56}$$

$$\cdot\left[\frac{1}{N}\sum_{i=1}^{N}\int_0^t \big|\big|G(\theta,x_s^{i,N},\mu_s^N) + G(\theta',x_s^{i,N},\mu_s^N)\big|\big|^2\mathrm{d}s\right]^{\frac{1}{2}}$$

$$\leq K||\theta-\theta'||\left[\frac{1}{N}\sum_{i=1}^{N}\int_0^t \Big|\Big|\frac{1}{N}\sum_{j=1}^{N}(1+||x_s^{i,N}||^q+||x_s^{j,N}||^q)\Big|\Big|^2\mathrm{d}s\right]^{\frac{1}{2}} \tag{5.57}$$

$$\cdot\left[\frac{2}{N}\sum_{i=1}^{N}\left[\int_0^t \big|\big|G(\theta,x_s^{i,N},\mu_s^N)\big|\big|^2\mathrm{d}s + \int_0^t \big|\big|G(\theta',x_s^{i,N},\mu_s^N)\big|\big|^2\mathrm{d}s\right]\right]^{\frac{1}{2}}$$

where in the final line we have used Conditions C.1(i) - C.1(ii). In addition, the uniform moment bounds on the IPS (Proposition 5.2), which follow from Condition D.1, together with Condition C.1(ii), imply that all terms on the RHS of this inequality are bounded. It follows immediately that the function $\frac{1}{N}\sum_{i=1}^N\int_0^t\big|\big|G(\theta,x_s^{i,N},\mu_s^N)\big|\big|^2\mathrm{d}s$ is Lipschitz continuous in θ, uniformly in N. Combining this with (5.54), we have that, as $N\rightarrow\infty$,

$$\hat{\theta}_t^N \xrightarrow{\mathbb{P}} \mathcal{D}_t^N = \left\{\theta\in\Theta: \lim_{N\rightarrow\infty}\frac{1}{N}\sum_{i=1}^{N}\int_0^t \big|\big|G(\theta,x_s^{i,N},\mu_s^N)\big|\big|^2\mathrm{d}s = 0\right\} \tag{5.58}$$

by which we we mean more precisely that $\inf_{\theta\in\mathcal{D}_t}||\hat{\theta}_t^N-\theta||\xrightarrow{\mathbb{P}}0$ as $N\rightarrow\infty$. It remains to observe that, by a repeated application of the McKean-Vlasov Law of Large Numbers (Proposition 5.6), as $N\rightarrow\infty$, and for all $t>0$, we have

$$\mathcal{D}_t^N \xrightarrow{\mathbb{P}} \mathcal{D}_t = \left\{\theta\in\Theta: \int_0^t\left[\int_{\mathbb{R}^d}||G(\theta,x,\mu_s)||^2\,\mu_s(\mathrm{d}x)\right]\mathrm{d}s = 0\right\} = \{\theta_0\}, \tag{5.59}$$

where in the second equality we have also made use of the identifiability condition in Condition E.1. It follows immediately, combining (5.58) and (5.59) that, for all fixed $t>0$, as $N\rightarrow\infty$, $\hat{\theta}_t^N\xrightarrow{\mathbb{P}}\theta_0$. □

5.4.2 Proof of Theorem 5.2

The proof of this theorem, similarly to the previous proof, combines well known techniques used to establishing strong consistency of the MLE as $t\rightarrow\infty$ (e.g., [296]) with ideas relevant to the asymptotic regime as $N\rightarrow\infty$ (e.g., [242]). Once again, we emphasise that throughout this proof the value of t will be fixed and finite, and we will consider the limit only as $N\rightarrow\infty$.

Proof. We begin by considering a Taylor expansion of $\nabla_\theta \mathcal{L}_t^N(\theta)$ around the true value of the parameter $\theta = \theta_0$, viz,

$$0 = \nabla_\theta \mathcal{L}_t^N(\hat{\theta}_t^N) = \nabla_\theta \mathcal{L}_t^N(\theta_0) + (\theta_t^N - \theta_0)\nabla_\theta^2 \mathcal{L}_t(\bar{\theta}_t^N) \tag{5.60}$$

where $\bar{\theta}_t^N$ is point in the segment connecting $\hat{\theta}_t^N$ and θ_0. The validity of this expansion is based on the sample path continuity of the log-likelihood and its derivatives. It follows that

$$N^{\frac{1}{2}}(\hat{\theta}_t^N - \theta_0)\nabla_\theta^2 \mathcal{L}_t^N(\bar{\theta}_t^N) = -N^{\frac{1}{2}}\nabla_\theta \mathcal{L}_t^N(\theta_0) \tag{5.61}$$

To deal with the terms in this equation, we will rely extensively on a multivariate version of Rebolledo's Central Limit Theorem [392], as stated in [242, Corollary to Theorem 2]. Let us begin by considering the RHS. First observe that

$$\begin{aligned} N^{\frac{1}{2}}\nabla_\theta \mathcal{L}_t^N(\theta_0) &= N^{-\frac{1}{2}}\sum_{i=1}^N \int_0^t \left\langle \nabla_\theta B(\theta_0, x_s^{i,N}, \mu_s^N), \mathrm{d}w_s^i \right\rangle \\ &\quad + N^{-\frac{1}{2}}\sum_{i=1}^N \int_0^t \nabla_\theta B(\theta_0, x_s^{i,N}, \mu_s^N) G(\theta_0, x_s^{i,N}, \mu_s^N)\mathrm{d}s \\ &= N^{-\frac{1}{2}}\sum_{i=1}^N \int_0^t \left\langle \nabla_\theta B(\theta_0, x_s^{i,N}, \mu_s^N), \mathrm{d}w_s^i \right\rangle \end{aligned} \tag{5.62--5.63}$$

where in the second line we have used the fact that, by definition, $G(\theta_0, \cdot, \cdot) = 0$ is identically equal to zero. It follows, using also Condition C.1(ii) (the polynomial growth property) and Proposition 5.2 (uniform moment bounds for the solutions of the IPS), that for all $t \geq 0$, $(N^{\frac{1}{2}}\nabla_\theta \mathcal{L}_t^N(\theta_0))_{N\in\mathbb{N}}$ is a sequence of local square integrable martingales, which implies that the first condition of [242, Corollary to Theorem 2] is satisfied.

Next, observe that the process $(N^{\frac{1}{2}}\nabla_\theta \mathcal{L}_t^N(\theta))_{t\geq 0}$ is continuous (in time), and thus the second condition of [242, Corollary to Theorem 2] (the Lindenberg condition) is satisfied. Finally, we have that, for all $k, l = 1, \ldots, p$, as $N \to \infty$,

$$\begin{aligned} &\left\langle \left[N^{\frac{1}{2}}\nabla_\theta \mathcal{L}_t^N(\theta_0)\right]_k, \left[N^{\frac{1}{2}}\nabla_\theta \mathcal{L}_t^N(\theta_0)\right]_l \right\rangle \\ &= \frac{1}{N}\sum_{i=1}^N \int_0^t [\nabla_\theta B(\theta_0, x_s^{i,N}, \mu_s^N)]_k [\nabla_\theta B(\theta_0, x_s^{i,N}, \mu_s^N)]_l \mathrm{d}s \\ &\xrightarrow{\mathbb{P}} \int_0^t \left[\int_{\mathbb{R}^d} [\nabla_\theta B(\theta_0, x, \mu_s)]_k \, [\nabla_\theta B(\theta_0, x, \mu_s)]_l \, \mu_s(\mathrm{d}x)\right] \mathrm{d}s = [I_t(\theta_0)]_{kl}\,, \end{aligned} \tag{5.64--5.66}$$

where in the final line, we have used a repeated application of the weak law of large numbers for the empirical distribution of the IPS (Proposition 5.6), and the definition of

$I_t(\theta)$ (see Condition E.2). Thus, the final condition in [242, Corollary to Theorem 2] is satisfied. It follows from this result that

$$-N^{\frac{1}{2}}\nabla_\theta\mathcal{L}_t^N(\theta_0) \xrightarrow{\mathcal{D}} \mathcal{N}_p(0, I_t(\theta_0)). \tag{5.67}$$

It remains to prove that $\nabla^2_\theta\mathcal{L}_t^N(\bar{\theta}_t^N) \xrightarrow{\mathbb{P}} -I_t(\theta_0)$. In fact, since $\hat{\theta}_t^N \xrightarrow{\mathbb{P}} \theta_0$ as $N\to\infty$ by Theorem 5.1, the continuity of $\{\nabla^2_\theta\mathcal{L}_t^N(\cdot)\}_{N\in\mathbb{N}}$ in θ implies that this limit holds provided we can establish that $\nabla^2_\theta\mathcal{L}_t^N(\theta_0) \xrightarrow{\mathbb{P}} -I_t(\theta_0)$. To do so, let us begin with the observation, via a simple calculation, we have that

$$\begin{aligned}\left[\nabla^2_\theta\mathcal{L}_t^N(\theta_0)\right]_{kl} &= \frac{1}{N}\sum_{i=1}^N\int_0^t\left[\nabla^2_\theta B(\theta_0, x_s^{i,N}, \mu_s^N)\right]_{kl}\mathrm{d}w_s^i \\ &- \frac{1}{N}\sum_{i=1}^N\int_0^t[\nabla_\theta B(\theta_0, x_s^{i,N}, \mu_s^N)]_k[\nabla_\theta B(\theta_0, x_s^{i,N}, \mu_s^N)]_l\mathrm{d}s\end{aligned} \tag{5.68}$$

Arguing as in the proof of Lemma 5.2 (see Appendix 5.B), we can show that, as $N\to\infty$, we have

$$\frac{1}{N}\sum_{i=1}^N\int_0^t\nabla^2_\theta B(\theta_0, x_s^{i,N}, \mu_s^N)\mathrm{d}w_s^i \xrightarrow{\mathbb{P}} 0. \tag{5.69}$$

Moreover, we have already established, c.f. (5.66), that, as $N\to\infty$, we have

$$\frac{1}{N}\sum_{i=1}^N\int_0^t[\nabla_\theta B(\theta_0, x_s^{i,N}, \mu_s^N)]_k[\nabla_\theta B(\theta_0, x_s^{i,N}, \mu_s^N)]_l\mathrm{d}s \xrightarrow{\mathbb{P}} [I_t(\theta_0)]_{kl}. \tag{5.70}$$

It follows, substituting (5.69) - (5.70) into (5.68), that $\nabla^2_\theta\mathcal{L}_t^N(\theta_0) \xrightarrow{\mathbb{P}} -I_t(\theta_0)$ as $N\to\infty$. By our previous remarks, this completes the proof. □

5.4.3 Proof of Theorem 5.3

We will prove Theorem 5.3 via a sequence of intermediate Lemmas. In fact, once these lemmas have been established, the proof itself follows straightforwardly.

Before we present this proof, it will first be necessary to introduce some additional notation. Recall from Section 5.2.1 (e.g., Table 5.1) that $(x_t^i)_{t\geq0}$ denotes a solution of the McKean-Vlasov SDE (5.1), where the Brownian motion $(w_t)_{t\geq0}$ is replaced by $(w_t^i)_{t\geq0}$. We will now also write $(\mu_t^i)_{t\geq0}$ to denote the law of this solution,[4] and, for the corresponding

[4] We remark that $(\mu_t)_{t\geq0}^i = (\mu_t)_{t\geq0}$ for all $i = 1,\ldots,N$. Nonetheless, we will use this notation to emphasise that we are considering solution of the McKean-Vlasov SDE with Brownian motion $(w_t^i)_{t\geq0}$.

log-likelihood function, write

$$\mathcal{L}_t^i(\theta) = \int_0^t L(\theta, x_s^i, \mu_s^i)\mathrm{d}s + \int_0^t \langle G(\theta, x_s^i, \mu_s^i), \mathrm{d}w_s^i\rangle. \tag{5.71}$$

We can now proceed to the proof of Theorem 5.3.

Proof. Using the triangle inequality, we can decompose the asymptotic log-likelihood of interest as

$$\begin{aligned} ||\nabla_\theta \tilde{\mathcal{L}}(\theta_t^{i,N})|| \leq & \underbrace{||\nabla_\theta \tilde{\mathcal{L}}(\theta_t^{i,N}) - \tfrac{1}{t}\nabla_\theta \mathcal{L}_t^i(\theta_t^{i,N})||}_{\to 0 \text{ as } t\to\infty \ \ \forall N\in\mathbb{N} \text{ by Lemma 5.4.A}} + \underbrace{||\tfrac{1}{t}\nabla_\theta \mathcal{L}_t^i(\theta_t^{i,N}) - \tfrac{1}{t}\nabla_\theta \mathcal{L}_t^{i,N}(\theta_t^{i,N})||}_{\to 0 \text{ as } N\to\infty \ \ \forall t\in\mathbb{R}_+ \text{ by Lemma 5.4.C}} \\ & + \underbrace{||\tfrac{1}{t}\nabla_\theta \mathcal{L}_t^{i,N}(\theta_t^{i,N}) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N})||}_{\to 0 \text{ as } t\to\infty \ \ \forall N\in\mathbb{N} \text{ by Lemma 5.4.B}} + \underbrace{||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N})||}_{\to 0 \text{ as } t\to\infty \ \ \forall N\in\mathbb{N} \text{ by Lemma 5.4.D}}. \end{aligned} \tag{5.72}$$

or, almost identically,

$$\begin{aligned} ||\nabla_\theta \tilde{\mathcal{L}}(\theta_t^{N})|| \leq & \underbrace{||\nabla_\theta \tilde{\mathcal{L}}(\theta_t^{N}) - \tfrac{1}{t}\nabla_\theta \mathcal{L}_t^i(\theta_t^{N})||}_{\to 0 \text{ as } t\to\infty \ \ \forall N\in\mathbb{N} \text{ by Lemma 5.4.A}} + \underbrace{||\tfrac{1}{t}\nabla_\theta \mathcal{L}_t^i(\theta_t^{N}) - \tfrac{1}{t}\nabla_\theta \mathcal{L}_t^{N}(\theta_t^{N})||}_{\to 0 \text{ as } N\to\infty \ \ \forall t\in\mathbb{R}_+ \text{ by Lemma 5.4.C}} \\ & + \underbrace{||\tfrac{1}{t}\nabla_\theta \mathcal{L}_t^{N}(\theta_t^{N}) - \nabla_\theta \tilde{\mathcal{L}}^{N}(\theta_t^{N})||}_{\to 0 \text{ as } t\to\infty \ \ \forall N\in\mathbb{N} \text{ by Lemma 5.4.B}} + \underbrace{||\nabla_\theta \tilde{\mathcal{L}}^{N}(\theta_t^{N})||}_{\to 0 \text{ as } t\to\infty \ \ \forall N\in\mathbb{N} \text{ by Lemma 5.4.D}}. \end{aligned} \tag{5.73}$$

where $\tilde{\mathcal{L}}^{i,N}(\theta)$ and $\tilde{\mathcal{L}}^{N}(\theta)$ are defined in Lemma 5.4.B. In both of these inequalities, all of the stated limits hold in $\mathbb{L}^1$. This completes the proof. □

Let us make brief two remarks regarding this result. Firstly, the first, third, and fourth limits all hold a.s (see Lemmas 5.4.A, 5.4.B and 5.4.D). Thus, if we could extend Lemma 5.4.C to include a.s. convergence, Theorem 5.3 would also hold a.s. Secondly, we actually obtain $\mathbb{L}^1$ rates for the first, second, and third terms (see Lemmas 5.4.A, 5.4.B and 5.4.C). Thus, if we could extend the results of Lemma 5.4.D to include an $\mathbb{L}^1$ convergence rate (e.g., on the infimum), possibly under additional assumptions, then Theorem 5.3 would also include an $\mathbb{L}^1$ convergence rate.

Before we proceed to the proofs of the intermediate Lemmas 5.4.A - 5.4.D, it is instructive to provide a brief high level overview.

(i) In Lemma 5.4.A, we establish the existence of $\tilde{\mathcal{L}}(\theta)$, the asymptotic log-likelihood of the McKean-Vlasov SDE, as well as its derivatives. We provide explicit expressions for these functions in terms of the unique invariant measure of the McKean-Vlasov SDE, prove an appropriate convergence result as $t \to \infty$ (both a.s. and in $\mathbb{L}^1$), and establish convergence rates.

(ii) In Lemma 5.4.B, we establish the existence of $\tilde{\mathcal{L}}^{i,N}(\theta)$ and $\tilde{\mathcal{L}}^{N}(\theta)$, the 'marginal' and 'joint' asymptotic log-likelihoods of the IPS, as well as their derivatives. As above, we provide explicit expressions for these functions in terms of the unique invariant measure of the IPS, prove an appropriate convergence result as $t \to \infty$ (both a.s. and in $\mathbb{L}^1$), and establish convergence rates.

(iii) In Lemma 5.4.C, we prove that, for all $t \geq 0$, the gradient of the asymptotic log-likelihood(s) of the IPS converges to the gradient of the asymptotic log-likelihood of the McKean-Vlasov SDE as $N \to \infty$ (in $\mathbb{L}^1$). We also provide $\mathbb{L}^1$ convergence rates. The proof of this result relies on classical uniform-in-time propagation of chaos results.

(iv) In Lemma 5.4.D, we establish that, for all $N \in \mathbb{N}$, the gradient of the asymptotic log-likelihood(s) of the IPS, evaluated at the relevant online parameter updates generated by the IPS, converges to zero as $t \to \infty$ (both a.s. and in $\mathbb{L}^1$). This result can be seen as a generalisation of [420, Theorem 2.4].

5.4.3.1 Proof of Lemma 5.4.A

Lemma 5.4.A. *Assume that Conditions B.1 - B.2, C.1, and D.1 hold. Then the processes* $\frac{1}{t}\nabla_\theta^m \mathcal{L}_t^i(\theta)$, $m = 0, 1, 2$, *converge, both a.s. and in* $\mathbb{L}^1$, *to the functions*

$$\nabla_\theta^m \tilde{\mathcal{L}}(\theta) = \int_{\mathbb{R}^d} \nabla_\theta^m L(\theta, x, \mu_\infty)\mu_\infty(\mathrm{d}x). \tag{5.74}$$

In addition, there exist positive constants K_m^1, K_m^2 *such that*

$$\left|\mathbb{E}\left[\frac{1}{t}\nabla_\theta^m \mathcal{L}_t^i(\theta) - \nabla_\theta^m \tilde{\mathcal{L}}(\theta)\right]\right| \leq \frac{K_m^1(1-e^{-\lambda t})}{\lambda t} + \frac{K_m^2(1+\sqrt{t})^{\frac{1}{2}}}{t^{\frac{1}{2}}}. \tag{5.75}$$

Proof. We will prove Lemma 5.4.A for $m = 0$, with $m = 1, 2$ proved similarly. For $m = 1, 2$, we remark only that the processes $\frac{1}{t}\nabla_\theta^m \mathcal{L}_t^i(\theta)$, and hence also $\nabla_\theta^m \tilde{\mathcal{L}}(\theta)$, exist due to Condition C.1. With this established, the proof when $m = 1, 2$ is essentially identical to the proof when $m = 0$. Let us begin by recalling the definition of $\frac{1}{t}\mathcal{L}_t^i(\theta)$, viz

$$\frac{1}{t}\mathcal{L}_t^i(\theta) = \underbrace{\frac{1}{t}\int_0^t L(\theta, x_s^i, \mu_s^i)\mathrm{d}s}_{I_1^N(\theta,t)} + \underbrace{\frac{1}{t}\int_0^t \langle G(\theta, x_s^i, \mu_s^i), \mathrm{d}w_s^i\rangle}_{I_2^N(\theta,t)} \tag{5.76}$$

We first consider the first term on the RHS. We will characterise the asymptotic behaviour of this term via the following decomposition

$$\underbrace{\frac{1}{t}\int_0^t L(\theta,x_s^i,\mu_s^i)\mathrm{d}s}_{I_1^N(\theta,t)} = \underbrace{\frac{1}{t}\int_0^t \left[L(\theta,x_s^i,\mu_s^i) - L(\theta,x_s^i,\mu_\infty)\right]\mathrm{d}s}_{I_{1,1}^N(\theta,t)} + \underbrace{\frac{1}{t}\int_0^t L(\theta,x_s^i,\mu_\infty)\mathrm{d}s}_{I_{1,2}^N(\theta,t)} \quad (5.77)$$

where μ_∞ is the unique invariant measure of the McKean-Vlasov SDE, which exists via Proposition 5.3 (see Appendix 5.A). We begin with the observation that, as $t\to\infty$,

$$\frac{1}{t}\int_0^t [L(\theta,x_s^i,\mu_s^i) - L(\theta,x_s^i,\mu_\infty)]\mathrm{d}s \xrightarrow{\text{a.s.}} 0, \quad (5.78)$$

$$\frac{1}{t}\int_0^t L(\theta,x_s^i,\mu_\infty)\mathrm{d}s \xrightarrow[\mathbb{L}^1]{\text{a.s.}} \tilde{\mathcal{L}}(\theta), \quad (5.79)$$

the former by Proposition 5.3, and the latter by an appropriate version of the ergodic theorem (e.g., [395, Chapter X]). Let us now demonstrate that $I_{1,1}^N(\theta,t)$ also converges to zero in $\mathbb{L}^1$. Using Lemma 5.7, we can write

$$\big|\big|L(\theta,x_s^i,\mu_s^i) - L(\theta,x_s^i,\mu_\infty)\big|\big| \leq K\mathbb{W}_2(\mu_s^i,\mu_\infty)\left[1+||x_s^i|||^q + \mu_\infty(||\cdot||^q) + \mu_s^i(||\cdot||^q)\right] \quad (5.80)$$

$$\leq K\big[1+||x_s^i||^q\big]e^{-\lambda s} \quad (5.81)$$

where in the second line we have additionally made use of Proposition 5.2 (moment bounds for the McKean-Vlasov SDE), Proposition 5.3 (the exponential contractivity of the McKean-Vlasov SDE), and Lemma 5.3 (moment bounds for the invariant measure of the McKean-Vlasov SDE). It follows straightforwardly, making use once more of Proposition 5.2, and allowing the value of K to change from line to line, that

$$\mathbb{E}[|I_{1,1}^N(t)|] \leq \frac{1}{t}\int_0^t K\big(1+\mathbb{E}\big[||x_s^i||^q\big]\big)e^{-\lambda s}\mathrm{d}s \leq \frac{K}{t}\int_0^t e^{-\lambda s}\mathrm{d}s \leq \frac{K(1-e^{-\lambda t})}{\lambda t}, \quad (5.82)$$

so that the convergence of $I_{1,1}^N(\theta,t)$ to zero does also hold in $\mathbb{L}^1$. We thus have, substituting (5.78) - (5.79) into (5.77), that $I_1^N(\theta,t)\to\tilde{\mathcal{L}}(\theta)$, both a.s. and in $\mathbb{L}^1$.

Let us now try to establish the convergence rate of this term. We have already established a (non-asymptotic) bound for $I_{1,1}^N(\theta,t)$, so it remains to consider $I_{1,2}^N(\theta,t)$. We can bound the deviation between this term and the asymptotic log-likelihood using arguments similar to those found in, for example, [171]. First note that, using Lemma 5.7 and Lemma 5.3

(moment bounds for the invariant measure of the McKean-Vlasov SDE), we have

$$|L(\theta, x, \mu_\infty) - L(\theta, y, \mu_\infty)| \leq K||x-y||[1 + ||x||^q + ||y||^q]. \tag{5.83}$$

We can thus utilise Lemma 5.5 to obtain

$$\left|\mathbb{E}_{x_0^i}[L(\theta, x, \mu_\infty)] - \int_{\mathbb{R}^d} L(\theta, y, \mu_\infty)\mu_\infty(\mathrm{d}y)\right| \leq K\left[1 + ||x_0^i||^q\right]e^{-\lambda s} \tag{5.84}$$

from which, in particular, it follows that

$$|\mathbb{E}[I_{1,2}(\theta, t)]| \leq \left|\mathbb{E}_{x_0^i}\left[\frac{1}{t}\int_0^t L(\theta, x_s^i, \mu_\infty)\mathrm{d}s - \int_{\mathbb{R}^d} L(\theta, y, \mu_\infty)\mu_\infty(\mathrm{d}y)\right]\right| \tag{5.85}$$

$$\leq \frac{1}{t}\int_0^t \left|\mathbb{E}_{x_0^i}[L(\theta, x_s^i, \mu_\infty)] - \int_{\mathbb{R}^d} L(\theta, y, \mu_\infty)\mu_\infty(\mathrm{d}y)\right|\mathrm{d}s \tag{5.86}$$

$$\leq \frac{K(1 - e^{-\lambda t})}{\lambda t}\left[1 + ||x_0^i||^q\right] \leq \frac{K(1 - e^{-\lambda t})}{\lambda t}. \tag{5.87}$$

where, as previously, we have allowed the value of the constant K to change from line to line. Substituting (5.82) and (5.87) into (5.77), we thus have that, for some $K_0^1 > 0$,

$$\left|\mathbb{E}\left[\frac{1}{t}\int_0^t L(\theta, x_s^i, \mu_s^i)\mathrm{d}s - \tilde{\mathcal{L}}(\theta)\right]\right| \leq \frac{K_0^1(1 - e^{-\lambda t})}{\lambda t}. \tag{5.88}$$

We now turn our attention $I_2^N(\theta, t)$, the second term in (5.76). We begin with the observation that, by the Itô's isometry, Condition C.1(ii) (the polynomial growth of G), Proposition 5.2 (the bounded moments of the McKean-Vlasov SDE), and Lemma 5.4 (the asymptotic growth rate of the moments of the McKean-Vlasov SDE), we have that

$$\mathbb{E}\left[\left|\int_0^t \langle G(\theta, x_s^i, \mu_s^i), \mathrm{d}w_s^i\rangle\right|^2\right] = \mathbb{E}\left[\int_0^t ||G(\theta, x_s^i, \mu_s^i)||^2\mathrm{d}s\right] \tag{5.89}$$

$$\leq \mathbb{E}\left[\int_0^t K\left(1 + ||x_s^i||^q + \mathbb{E}\left[||x_s^i||^q\right]\right)\mathrm{d}s\right] \tag{5.90}$$

$$\leq Kt\left[1 + \mathbb{E}[\sup_{0\leq s\leq t}||x_s^i||^q]\right] \leq Kt\left[1 + \sqrt{t}\right] \tag{5.91}$$

where the value of the constant K is allowed to change from line to line. It follows, making use of the triangle inequality and the Hölder inequality that, for some $K_0^2 > 0$, we have

$$\left|\mathbb{E}\left[\frac{1}{t}\langle G(\theta, x_s^i, \mu_s^i), \mathrm{d}w_s^i\rangle\right]\right| \leq \frac{K_0^2(1+\sqrt{t})^{\frac{1}{2}}}{t^{\frac{1}{2}}}, \tag{5.92}$$

so that this term converges in $\mathbb{L}^1$ to zero, and we have the required rate. It remains only to demonstrate a.s. convergence of this term to zero. To do so, consider the local martingale

$$M_t = \int_0^t \frac{1}{s}\langle G(\theta, x_s^i, \mu_s^i), \mathrm{d}w_s^i\rangle = \frac{1}{t}\int_0^t \langle G(\theta, x_s^i, \mu_s^i), \mathrm{d}w_s^i\rangle + \int_0^t \frac{1}{s^2}\left[\int_0^s \langle G(\theta, x_u^i, \mu_u^i), \mathrm{d}w_u^i\rangle\right]\mathrm{d}s, \tag{5.93}$$

where the second line follows from Itô's Lemma. Using the Itô isometry, Condition C.1(ii) (the polynomial growth of G), and Proposition 5.2 (the bounded moments of the McKean-Vlasov SDE), and arguing similarly to above, we have

$$\sup_{t>0}\mathbb{E}\left[|M_t|^2\right] = \mathbb{E}\left[\int_0^\infty \frac{1}{s^2}\mathbb{E}\left[||G(\theta, x_s^i, \mu_s^i)||^2\right]\mathrm{d}s\right] \leq K\left[\int_0^t \frac{1}{s^2}\left(1+\mathbb{E}\left[||x_s^i||^q\right]\right)\mathrm{d}s\right] < \infty. \tag{5.94}$$

By Doob's martingale convergence theorem [164], there thus exists a finite random variable M_∞ such that $M_t \to M_\infty$ a.s. It follows immediately that $\frac{1}{t}\int_0^t \langle G(\theta, x_s^i, \mu_s^i), \mathrm{d}w_s^i\rangle$ also converges to zero a.s., as claimed. Putting everything together, we thus have that $\frac{1}{t}\mathcal{L}_t^i(\theta)$ converges to $\tilde{\mathcal{L}}(\theta)$ both a.s. and in $\mathbb{L}^1$, and, combining (5.76), (5.88) and (5.92), that

$$\begin{aligned}\left|\mathbb{E}\left[\frac{1}{t}\mathcal{L}_t^i(\theta) - \tilde{\mathcal{L}}(\theta)\right]\right| &\leq \left|\mathbb{E}\left[\frac{1}{t}\int_0^t L(\theta, x_s^i, \mu_s^i)\mathrm{d}s - \tilde{\mathcal{L}}(\theta)\right]\right| + \left|\mathbb{E}\left[\frac{1}{t}\int_0^t \langle G(\theta, x_s^i, \mu_s^i), \mathrm{d}w_s^i\rangle\right]\right| \\ &\leq \frac{K_0^1(1-e^{-\lambda t})}{\lambda t} + \frac{K_0^2(1+\sqrt{t})^{\frac{1}{2}}}{t^{\frac{1}{2}}}.\end{aligned} \tag{5.95}$$

□

5.4.3.2 Proof of Lemma 5.4.B

5.4.3.2.1 Additional Notation

In order to state and prove the next Lemma, it will be useful to introduce some additional notation. Let $\hat{x}_t^N \in (\mathbb{R}^d)^N$ be the process consisting of the concatenation of the N solutions of the IPS (5.4), viz, $\hat{x}_t^N = (x_t^{1,N}, \ldots, x_t^{N,N})^T$. Observe that this process is the solution of the following SDE on $(\mathbb{R}^d)^N$

$$\mathrm{d}\hat{x}_t^N = \hat{B}(\theta, \hat{x}_t^N)\mathrm{d}t + \mathrm{d}\hat{w}_t^N, \tag{5.96}$$

where $\hat{w}_t^N$ is a $(\mathbb{R}^d)^N$-valued Brownian motion, and the function $\hat{B}(\theta,\cdot) : (\mathbb{R}^d)^N \to (\mathbb{R}^d)^N$ is of the form $\hat{B}(\theta, \hat{x}^N) = (\hat{B}^{1,N}(\theta,\hat{x}^N), \dots, \hat{B}^{N,N}(\theta,\hat{x}^N))^T$, where, for $i = 1, \dots, N$, $\hat{B}^{i,N}(\theta,\cdot) : (\mathbb{R}^d)^N \to \mathbb{R}^d$ is defined according to

$$\hat{B}^{i,N}(\theta,\hat{x}^N) = b(\theta, x^{i,N}) + \frac{1}{N}\sum_{j=1}^N \phi(\theta, x^{i,N}, x^{j,N}). \tag{5.97}$$

It will also be useful to define, for $i = 1, \dots, N$, the functions $\hat{G}^{i,N}(\theta,\cdot) : (\mathbb{R}^d)^N \to \mathbb{R}^d$ and $\hat{L}^{i,N}(\theta,\cdot) : (\mathbb{R}^d)^N \to \mathbb{R}$ according to

$$\hat{G}^{i,N}(\theta,\hat{x}^N) = \hat{B}^{i,N}(\theta,\hat{x}^N) - \hat{B}^{i,N}(\theta_0,\hat{x}^N) \tag{5.98}$$

$$\hat{L}^{i,N}(\theta,\hat{x}) = -\frac{1}{2}||\hat{G}^{i,N}(\theta,\hat{x}^N)||^2. \tag{5.99}$$

Finally, we will write $\hat{\mu}_t^N = \mathcal{L}(\hat{x}_t^N)$ to denote the law of $\hat{x}_t^N = (x_t^{1,N}, \dots, x_t^{N,N})$. We should be careful not to confuse this with $\mu_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{x_t^{i,N}}$, the empirical measure of the IPS.

Lemma 5.4.B. *Assume that Conditions B.1 - B.2, C.1, and D.1 hold. Then, for all $N \in \mathbb{N}$, the processes $\frac{1}{t}\nabla_\theta^m \mathcal{L}_t^{i,N}(\theta)$ and $\frac{1}{t}\nabla_\theta^m \mathcal{L}_t^N(\theta)$, $m = 0,1,2$, converge, both a.s. and in $\mathbb{L}^1$, to the functions*

$$\nabla_\theta^m \tilde{\mathcal{L}}^{i,N}(\theta) = \int_{(\mathbb{R}^d)^N} \nabla_\theta^m \hat{L}^{i,N}(\theta,\hat{x}^N)\hat{\mu}_\infty^N(\mathrm{d}\hat{x}^N) \quad , \quad \nabla_\theta^m \tilde{\mathcal{L}}^N(\theta) = \frac{1}{N}\sum_{i=1}^N \nabla_\theta^m \tilde{\mathcal{L}}^{i,N}(\theta). \tag{5.100}$$

In addition, there exist positive constants K_m^1, K_m^2, independent of N, such that

$$\left|\mathbb{E}\left[\frac{1}{t}\nabla_\theta^m \mathcal{L}_t^{i,N}(\theta) - \nabla_\theta^m \tilde{\mathcal{L}}^{i,N}(\theta)\right]\right| \leq \frac{K_m^1(1-e^{-\lambda t})}{\lambda t} + \frac{K_m^2(1+\sqrt{t})^{\frac{1}{2}}}{t^{\frac{1}{2}}} \tag{5.101}$$

and this bound also holds if $\mathcal{L}_t^{i,N}(\cdot)$ and $\tilde{\mathcal{L}}^{i,N}(\cdot)$ are replaced with $\mathcal{L}_t^N(\cdot)$ and $\tilde{\mathcal{L}}^N(\cdot)$.

Proof of Lemma 5.4.B. We will begin by proving that the two statements hold for the function $\mathcal{L}_t^{i,N}(\theta)$. The proof, in this case, is very similar to the proof of Lemma 5.4.A, with some simplifications. We will provide a sketch of the proof, signposting differences with the previous proof where necessary. As previously, we will only consider the case $m = 0$, with the results for $m = 1,2$ proved analogously. We begin by recalling the definition of the function $\frac{1}{t}\mathcal{L}_t^{i,N}(\theta)$ from (5.12), which we now write in the form

$$\frac{1}{t}\mathcal{L}_t^{i,N}(\theta) = \frac{1}{t}\int_0^t \hat{L}^{i,N}(\theta,\hat{x}_s^N)\mathrm{d}s + \frac{1}{t}\int_0^t \langle \hat{G}^{i,N}(\theta,\hat{x}_s^N),\mathrm{d}w_s^i\rangle \tag{5.102}$$

We begin with the first term on the RHS. By Proposition 5.4, the IPS admits a unique invariant measure $\hat{\mu}_\infty^N \in \mathcal{P}((\mathbb{R}^d)^N)$. Thus, for all $N\in\mathbb{N}$, by the ergodic theorem (e.g., [395, Chapter X]) we have that as $t\to\infty$,

$$\frac{1}{t}\int_0^t \hat{L}^{i,N}\left(\theta,\hat{x}_s^N\right)\mathrm{d}s \xrightarrow[\mathbb{L}^1]{\text{a.s.}} \int_{(\mathbb{R}^d)^N} \hat{L}^{i,N}\left(\theta,\hat{x}^N\right)\hat{\mu}_\infty(\mathrm{d}\hat{x}^N) = \tilde{\mathcal{L}}^{i,N}(\theta), \tag{5.103}$$

Let us now obtain the required convergence rate. By the remark after Lemma 5.7, the function $\hat{L}^{i,N}(\theta,\hat{x}^N)$ satisfies the conditions of Lemma 5.5. Thus, we can apply Lemma 5.5 to obtain

$$\left|\mathbb{E}_{\hat{x}_0}[\hat{L}^{i,N}(\theta,\hat{x}_t^N)] - \int_{(\mathbb{R}^d)^N}\hat{L}^{i,N}(\theta,\hat{y})\hat{\mu}_\infty^N(\mathrm{d}\hat{y}^N)\right| \leq K\left[1+||x_0^{i,N}||^q + \frac{1}{N}\sum_{j=1}^N ||x_0^{j,N}||^q\right]e^{-\lambda t} \tag{5.104}$$

and so, arguing as in (5.85) - (5.87) in the proof of Lemma 5.4.A, we have

$$\left|\mathbb{E}_{\hat{x}_0}\left[\frac{1}{t}\int_0^t \hat{L}^{i,N}(\theta,\hat{x}_s^N)\mathrm{d}s - \int_{(\mathbb{R}^d)^N}\hat{L}^{i,N}(\theta,\hat{y}^N)\hat{\mu}_\infty^N(\mathrm{d}\hat{y}^N)\right]\right| \tag{5.105}$$

$$\leq \frac{K(1-e^{-\lambda t})}{\lambda t}\left[1+||x_0^{i,N}||^q + \frac{1}{N}\sum_{j=1}^N ||x_0^{j,N}||^q\right] \leq \frac{K_0^1(1-e^{-\lambda t})}{\lambda t}. \tag{5.106}$$

It remains to bound the second term on the RHS of (5.102). We show that this term converges to zero a.s. and in $\mathbb{L}^1$, and satisfies the required convergence rate, using essentially identical arguments to those used in the proof of Lemma 5.4.A, c.f. (5.89) - (5.94). This concludes the proof.

We now turn our attention to the function $\mathcal{L}_t^N(\theta)$. The proof of the statements regarding this function now follows easily. In particular, using the definition of $\mathcal{L}_t^N(\theta)$, c.f. (5.12), and the results above, we have (once more restricting attention to the case $m=0$)

$$\frac{1}{t}\mathcal{L}_t^N(\theta) = \frac{1}{t}\left[\frac{1}{N}\sum_{i=1}^N \mathcal{L}_t^{i,N}(\theta)\right] = \frac{1}{N}\sum_{i=1}^N\left[\frac{1}{t}\mathcal{L}_t^{i,N}(\theta)\right] \xrightarrow[\mathbb{L}^1]{\text{a.s.}} \frac{1}{N}\sum_{i=1}^N \tilde{\mathcal{L}}^{i,N}(\theta) = \tilde{\mathcal{L}}^N(\theta), \tag{5.107}$$

and, for the required bound,

$$\left|\left|\mathbb{E}\left[\frac{1}{t}\nabla_\theta^m\mathcal{L}_t^N(\theta)-\nabla_\theta^m\tilde{\mathcal{L}}^N(\theta)\right]\right|\right| = \left|\left|\frac{1}{N}\sum_{i=1}^N\mathbb{E}\left[\frac{1}{t}\nabla_\theta^m\mathcal{L}_t^{i,N}(\theta)-\nabla_\theta^m\tilde{\mathcal{L}}^{i,N}(\theta)\right]\right|\right| \tag{5.108}$$

$$\leq \frac{1}{N}\sum_{i=1}^N\left|\left|\mathbb{E}\left[\frac{1}{t}\nabla_\theta^m\mathcal{L}_t^{i,N}(\theta)-\nabla_\theta^m\tilde{\mathcal{L}}^{i,N}(\theta)\right]\right|\right| \tag{5.109}$$

$$\leq \frac{1}{N}\sum_{i=1}^N\left[\frac{K_m(1-e^{-\lambda t})}{\lambda t}+\frac{K_m(1+\sqrt{t})^{\frac{1}{2}}}{t^{\frac{1}{2}}}\right] \tag{5.110}$$

$$= \frac{K_m(1-e^{-\lambda t})}{\lambda t}+\frac{K_m(1+\sqrt{t})^{\frac{1}{2}}}{t^{\frac{1}{2}}}. \tag{5.111}$$

□

5.4.3.3 Proof of Lemma 5.4.C

Lemma 5.4.C. *Assume that Conditions B.1 - B.2, C.1, and D.1 hold. Then, for all $\theta\in\mathbb{R}^p$, for all $t\geq 0$, for all $i=1,\dots,N$, we have, in $\mathbb{L}^1$, that*

$$\lim_{N\to\infty}||\tfrac{1}{t}\nabla_\theta\mathcal{L}_t^{i,N}(\theta)|| = ||\tfrac{1}{t}\nabla_\theta\mathcal{L}_t^{i}(\theta)||, \tag{5.112}$$

$$\lim_{N\to\infty}||\tfrac{1}{t}\nabla_\theta\mathcal{L}_t^{N}(\theta)|| = ||\tfrac{1}{t}\nabla_\theta\mathcal{L}_t^{i}(\theta)||. \tag{5.113}$$

In addition, there exists a positive constant K such that, for all $\theta\in\mathbb{R}^p$, for all $N\in\mathbb{N}$,

$$\mathbb{E}\left[\left|\left|\frac{1}{t}\nabla_\theta\mathcal{L}_t^i(\theta)-\frac{1}{t}\nabla_\theta\mathcal{L}_t^{i,N}(\theta)\right|\right|\right]\leq\frac{K}{\sqrt{N}}\left(1+\frac{1}{\sqrt{t}}\right), \tag{5.114}$$

and this bound also holds if $\mathcal{L}_t^{i,N}(\cdot)$ is replaced by $\mathcal{L}_t^N(\cdot)$.

Proof. We begin by proving that the two statements hold for $\mathcal{L}_t^{i,N}(\theta)$. First recall that

$$\frac{1}{t}\nabla_\theta\mathcal{L}_t^i(\theta) = \underbrace{-\frac{1}{t}\int_0^t\nabla_\theta G(\theta,x_s^i,\mu_s^i)G(\theta,x_s^i,\mu_s^i)\mathrm{d}s}_{I_1^i(\theta,t)}+\underbrace{\frac{1}{t}\int_0^t\left\langle\nabla_\theta G(\theta,x_s^i,\mu_s^i),\mathrm{d}w_s^i\right\rangle}_{I_2^i(\theta,t)} \tag{5.115}$$

$$\frac{1}{t}\nabla_\theta\mathcal{L}_t^{i,N}(\theta) = \underbrace{-\frac{1}{t}\int_0^t\left[\frac{1}{N}\sum_{j=1}^N\nabla_\theta G(\theta,x_s^{i,N},x_s^{j,N})\right]\left[\frac{1}{N}\sum_{j=1}^N G(\theta,x_s^{i,N},x_s^{j,N})\right]\mathrm{d}s}_{I_1^{i,N}(\theta,t)} \tag{5.116}$$

$$+\underbrace{\frac{1}{t}\int_0^t\left\langle\frac{1}{N}\sum_{j=1}^N\nabla_\theta G(\theta,x_s^{i,N},x_s^{j,N}),\mathrm{d}w_s^i\right\rangle}_{I_2^{i,N}(\theta,t)}$$

Let us seek bounds for $\mathbb{E}||I_1^i(\theta,t)-I_1^{i,N}(\theta,t)||$ and $\mathbb{E}||I_2^i(\theta,t)-I_2^{i,N}(\theta,t)||$, starting with the latter. By Lemma 5.10 (see Appendix 5.C), for all $s\geq 0$, there exists a positive constant K such that

$$\mathbb{E}\Big[\big|\big|\nabla_\theta G(\theta,x_s^i,\mu_s^i)-\frac{1}{N}\sum_{j=1}^N\nabla_\theta G(\theta,x_s^{i,N},x_s^{j,N})\big|\big|^2\Big]\leq\frac{K}{N}. \tag{5.117}$$

Thus, making using of the triangle inequality, the Itô isometry, and Fubini's Theorem, we have that

$$\mathbb{E}\left[||I_2^i(\theta,t)-I_2^{i,N}(\theta,t)||^2\right]\leq\frac{1}{t^2}\mathbb{E}\Bigg[\int_0^t\left|\left|\nabla_\theta G(\theta,x_s^i,\mu_s^i)-\frac{1}{N}\sum_{j=1}^N\nabla_\theta G(\theta,x_s^{i,N},x_s^{j,N})\right|\right|^2\mathrm{d}s\Bigg] \tag{5.118}$$

$$\leq\frac{K}{Nt}. \tag{5.119}$$

and thus, by the Hölder inequality,

$$\mathbb{E}\big[||I_2^i(\theta,t)-I_2^{i,N}(\theta,t)||\big]\leq\frac{K}{\sqrt{Nt}} \tag{5.120}$$

We will now obtain, in much the same fashion, a bound for $\mathbb{E}[||I_1^i(\theta,t)-I_1^{i,N}(\theta,t)||]$. Once again, by Lemma 5.10, for all $s\geq 0$, we have that

$$\mathbb{E}\Big[\big|\big|\nabla_\theta G(\theta,x_s^i,\mu_s^i)-\frac{1}{N}\sum_{j=1}^N\nabla_\theta G(\theta,x_s^{i,N},x_s^{j,N})\big|\big|^2\Big]\leq\frac{K}{N}, \tag{5.121}$$

$$\mathbb{E}\Big[\big|\big|G(\theta,x_s^i,\mu_s^i)-\frac{1}{N}\sum_{j=1}^N G(\theta,x_s^{i,N},x_s^{j,N})\big|\big|^2\Big]\leq\frac{K'}{N}. \tag{5.122}$$

To proceed, consider the following inequality, which follows straightforwardly from the triangle inequality and the Cauchy-Schwarz inequality,

$$\mathbb{E}\left[||YZ-Y_NZ_N||\right]\leq\mathbb{E}\left[||Y-Y_N||^2\right]^{\frac{1}{2}}\mathbb{E}\left[||Z||^2\right]^{\frac{1}{2}}+\mathbb{E}\left[||Y_N||^2\right]^{\frac{1}{2}}\mathbb{E}\left[||Z-Z_N||^2\right]^{\frac{1}{2}} \tag{5.123}$$

Suppose we let $Y=\nabla_\theta G(\theta,x_s^i,\mu_s^i)$, $Y_N=N^{-1}\sum_{j=1}^N\nabla_\theta G(\theta,x_s^{i,N},x_s^{j,N})$, $Z=G(\theta,x_s^i,\mu_s^i)$, and $Z_N=N^{-1}\sum_{j=1}^N G(\theta,x_s^{i,N},x_s^{j,N})$. Then, once more allowing the value of the constant

K to change from line to line, this inequality yields

$$\mathbb{E}\Big[\big|\big|\nabla_\theta G(\theta,x_s^i,\mu_s^i)G(\theta,x_s^i,\mu_s^i)-\frac{1}{N}\sum_{j=1}^N\nabla_\theta G(\theta,x_s^{i,N},x_s^{j,N})\cdot\frac{1}{N}\sum_{j=1}^N G(\theta,x_s^{i,N},x_s^{j,N})\big|\big|\Big] \tag{5.124}$$

$$\leq \underbrace{\mathbb{E}\Big[\big|\big|\nabla_\theta G(\theta,x_s^i,\mu_s^i)-\frac{1}{N}\sum_{j=1}^N\nabla_\theta G(\theta,x_s^{i,N},x_s^{j,N})\big|\big|^2\Big]^{\frac{1}{2}}}_{\leq\frac{K}{\sqrt{N}}\text{ by (5.121)}}\cdot\underbrace{\mathbb{E}\Big[\big|\big|G(\theta,x_s^i,\mu_s^i)\big|\big|^2\Big]^{\frac{1}{2}}}_{\leq K} \tag{5.125}$$

$$+\underbrace{\mathbb{E}\Big[\big|\big|G(\theta,x_s^i,\mu_s^i)-\frac{1}{N}\sum_{j=1}^N G(\theta,x_s^{i,N},x_s^{j,N})\big|\big|^2\Big]^{\frac{1}{2}}}_{\leq\frac{K'}{\sqrt{N}}\text{ by (5.122)}}\cdot\underbrace{\mathbb{E}\Big[\big|\big|\frac{1}{N}\sum_{j=1}^N G(\theta,x_s^{i,N},x_s^{j,N})\big|\big|^2\Big]^{\frac{1}{2}}}_{\leq K'}$$

$$\leq\frac{K}{\sqrt{N}}, \tag{5.126}$$

where in the penultimate line we have used Condition C.1 (the polynomial growth of G) and Proposition 5.2 (the moment bounds for the IPS), to conclude that each of the expectations are bounded above by some positive constants. That is, for example,

$$\mathbb{E}\left[\big|\big|G(\theta,x_s^i,\mu_s^i)\big|\big|^2\right]\leq\mathbb{E}\left[K\left(1+||x_s^i||^q+\int_{\mathbb{R}^d}||y||^q\mu_s^i(\mathrm{d}y)\right)\right]\leq K\left(1+\mathbb{E}\left[||x_s^i||^q\right]\right)\leq K^2. \tag{5.127}$$

It follows straightforwardly that

$$\mathbb{E}\left[||I_1^i(\theta,t)-I_1^{i,N}(\theta,t)||\right]\leq\frac{1}{t}\int_0^t\frac{K}{\sqrt{N}}\mathrm{d}s=\frac{K}{\sqrt{N}}. \tag{5.128}$$

Combining inequalities (5.120) and (5.128), and making use of the triangle inequality one final time, we have that

$$\mathbb{E}\left[\left|\left|\frac{1}{t}\nabla_\theta\mathcal{L}_t^i(\theta)-\frac{1}{t}\nabla_\theta\mathcal{L}_t^{i,N}(\theta)\right|\right|\right]\leq\frac{K}{\sqrt{N}}\left(1+\frac{1}{\sqrt{t}}\right) \tag{5.129}$$

This establishes convergence in $\mathbb{L}^1$ as $N\to\infty$, for all $t\geq 0$. It remains only to establish that the statements of the lemma also hold for $\mathcal{L}_t^N(\theta)$. This is straightforward. Indeed, we omit the calculations, which are essentially identical to those used at the end of the proof of Lemma 5.4.B, c.f. (5.108) - (5.111). □

5.4.3.4 Proof of Lemma 5.4.D

Lemma 5.4.D. *Assume that Conditions B.1 - B.2, C.1, D.1 and F.1 hold. Then, for all $N \in \mathbb{N}$, we have, both a.s. and in $\mathbb{L}^1$, that*

$$\lim_{t\to\infty} ||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N})|| = 0, \tag{5.130}$$

$$\lim_{t\to\infty} ||\nabla_\theta \tilde{\mathcal{L}}^{N}(\theta_t^{N})|| = 0. \tag{5.131}$$

Proof. We will prove the first statement of the lemma, with the second proved identically.[5] In particular, we will use a modified version of the approach in [420], which itself is a continuous-time version of the approach first introduced in [46]. In the interest of completeness, we will include the proof in full here, adapted appropriately to the current setting and our notation.

We will require the following additional notation. Define an arbitrary constant $\kappa > 0$, with $\lambda = \lambda(\kappa) > 0$ to be determined. Set $\sigma = 0$, and define the cycle of random stopping times

$$0 = \sigma_0 \leq \tau_1 \leq \sigma_1 \leq \tau_2 \leq \sigma_2 \leq \dots \tag{5.132}$$

according to

$$\tau_k = \inf\left\{t > \sigma_{k-1} : ||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N})|| \geq \kappa\right\} \tag{5.133}$$

$$\sigma_k = \sup\Big\{t > \tau_k : \tfrac{1}{2}||\nabla \tilde{\mathcal{L}}^{i,N}(\theta_{\tau_k}^{i,N})|| \leq ||\nabla \tilde{\mathcal{L}}^{N}(\theta_s^{i,N})|| \leq 2||\nabla \tilde{\mathcal{L}}^{i,N}(\theta_{\tau_k}^{i,N})|| \quad \forall s \in [\tau_k, t], \tag{5.134}$$

$$\int_{\tau_k}^{t} \gamma(s)\mathrm{d}s \leq \rho\Big\}$$

The purpose of these stopping times is to control the periods of time for which $||\nabla \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N})||$ is close to zero, and those for which it is away from zero. In addition, let $\eta > 0$, and set $\sigma_{k,\eta} = \sigma_k + \eta$. First consider the case in which there are a finite number of stopping times τ_k. In this case, there exists finite t_0 such that, for all $t \geq t_0$, $[||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N})||] < \kappa$. Now consider the case in which there are an infinite number of stopping times τ_k. Then, using Lemmas 5.13 - 5.14 (see Appendix 5.C), there exist $0 < \beta_1 < \beta$, and $k_0 \in \mathbb{N}$, such that for all $k \geq k_0$, a.s. ,

$$\tilde{\mathcal{L}}^{i,N}(\theta_{\sigma_k}^{i,N}) - \tilde{\mathcal{L}}^{i,N}(\theta_{\tau_k}^{i,N}) \geq \beta \tag{5.135}$$

$$\tilde{\mathcal{L}}^{i,N}(\theta_{\tau_k}^{i,N}) - \tilde{\mathcal{L}}^{i,N}(\theta_{\sigma_{k-1}}^{i,N}) \geq -\beta_1. \tag{5.136}$$

[5] We remark that Lemmas 5.11 - 5.14, which are essential to this proof, all apply to both $\mathcal{L}_t^{i,N}(\theta)$ and $\mathcal{L}_t^N(\theta)$, and thus can still be used to establish the second statement.

It follows straightforwardly that

$$\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_{n+1}}) - \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_{k_0}}) = \sum_{k=k_0}^{n} \left[\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\sigma_k}) - \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k}) + \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_{k+1}}) - \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\sigma_k})\right] \tag{5.137}$$

$$\geq \sum_{k=k_0}^{n} (\beta - \beta_1) = (n+1-k_0)(\beta - \beta_1) \tag{5.138}$$

Since $\beta - \beta_1 > 0$, this implies that $\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_{n+1}}) \to \infty$ as $n \to \infty$. But this is in contradiction with Lemma 5.8, which states that $\tilde{\mathcal{L}}^{i,N}(\theta)$ is bounded from above. Thus, there must exist a finite time t_0 such that, for all $t \geq t_0$, $||\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_t)|| < \kappa$. Since our original choice of κ was arbitrary, this completes the proof that, for all $N \in \mathbb{N}$, a.s. ,

$$\lim_{t\to\infty} ||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_t)|| = 0. \tag{5.139}$$

Finally, we observe that, by Lemma 5.8, $||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta)||$ is bounded above for all $\theta \in \mathbb{R}^p$. Thus, we also have convergence in $\mathbb{L}^1$ via Lebesgue's dominated convergence theorem (e.g., [464, Chapter 5]). □

5.4.4 Proof of Theorem 5.4

Before we proceed to the proof of Theorem 5.4, we state the following Lemma, which provides uniform moment bounds for the online parameter estimate, and will be used frequently in this proof.

Lemma 5.1. *Assume that Conditions B.1 - B.2, C.1, D.1, F.1, and G.1 - G.2 hold. Then, for all $q \geq 1$, for all $i = 1, \ldots, N$, for all $N \in \mathbb{N}$, there exists K such that*

$$\sup_{t>0} \mathbb{E}\left[||\theta_t||^q\right] \leq K \tag{5.140}$$

$$\sup_{t>0} \mathbb{E}\left[||\theta^{i,N}_t||^q\right] \leq K. \tag{5.141}$$

Proof. This Lemma follows straightforwardly as an extension of [422, Lemma A.1], making use of the appropriate bounds in Conditions G.1 - G.2. □

Proof of Theorem 5.4. The proof of this result closely follows the proof of [422, Theorem 2.7], adapted appropriately to our particular case. We will begin by proving the first statement of the theorem. To begin, let us recall the following form of the parameter update equation (5.19):

$$\mathrm{d}\theta^{i,N}_t = \gamma_t \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_t)\mathrm{d}t + \gamma_t\big(\nabla_\theta L(\theta^{i,N}_t, x^{i,N}_t, \mu^N_t) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_t)\big)\mathrm{d}t \tag{5.142}$$

$$
\begin{aligned}
&+ \gamma_t \nabla_\theta B(\theta_t^{i,N}, x_t^{i,N}, \mu_t^N)\mathrm{d}w_t^i \\
&= \gamma_t \nabla_\theta \tilde{\mathcal{L}}(\theta_t^{i,N})\mathrm{d}t + \gamma_t\big(\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N}) - \nabla_\theta \tilde{\mathcal{L}}(\theta_t^{i,N})\big)\mathrm{d}t \qquad (5.143)\\
&+ \gamma_t\big(\nabla_\theta L(\theta_t^{i,N}, x_t^{i,N}, \mu_t^N) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N})\big)\mathrm{d}t + \gamma_t \nabla_\theta B(\theta_t^{i,N}, x_t^{i,N}, \mu_t^N)\mathrm{d}w_t^i.
\end{aligned}
$$

Using a first order Taylor expansion, we have that

$$
\nabla_\theta \tilde{\mathcal{L}}(\theta_t^{i,N}) = \nabla_\theta \tilde{\mathcal{L}}(\theta_0) + \nabla^2 \tilde{\mathcal{L}}(\tilde{\theta}_t^{i,N})(\theta_t^{i,N} - \theta_0) = \nabla^2 \tilde{\mathcal{L}}(\tilde{\theta}_t^{i,N})(\theta_t^{i,N} - \theta_0) \qquad (5.144)
$$

where $\tilde{\theta}_t^{i,N}$ is point in the segment connecting $\theta_t^{i,N}$ and θ_0. Substituting this into (5.143), we obtain the following equation for $Z_t^{i,N} = \theta_t^{i,N} - \theta_0$

$$
\begin{aligned}
\mathrm{d}Z_t^{i,N} &= \gamma_t \nabla_\theta^2 \tilde{\mathcal{L}}(\tilde{\theta}_t^{i,N}) Z_t^{i,N}\mathrm{d}t + \gamma_t\big(\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N}) - \nabla_\theta \tilde{\mathcal{L}}(\theta_t^{i,N})\big)\mathrm{d}t \qquad (5.145)\\
&+ \gamma_t\big(\nabla_\theta L(\theta_t^{i,N}, x_t^{i,N}, \mu_t^N) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N})\big)\mathrm{d}t + \gamma_t \nabla_\theta B(\theta_t^{i,N}, x_t^{i,N}, \mu_t^N)\mathrm{d}w_t^i.
\end{aligned}
$$

Applying Itô's formula to the function $||\cdot||^2$, we obtain

$$
\begin{aligned}
\mathrm{d}||Z_t^{i,N}||^2 &= 2\gamma_t \big\langle Z_t^{i,N}, \nabla_\theta^2 \tilde{\mathcal{L}}(\tilde{\theta}_t^{i,N}) Z_t^{i,N}\big\rangle \mathrm{d}t + \gamma_t\big\langle Z_t^{i,N}, \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N}) - \nabla_\theta \tilde{\mathcal{L}}(\theta_t^{i,N})\big\rangle \mathrm{d}t \qquad (5.146)\\
&+ \gamma_t\big\langle Z_t^{i,N}, \nabla_\theta L(\theta_t^{i,N}, x_t^{i,N}, \mu_t^N) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N})\big\rangle \mathrm{d}t\\
&+ \gamma_t\big\langle Z_t^{i,N}, \nabla_\theta B(\theta_t^{i,N}, x_t^{i,N}, \mu_t^N)\mathrm{d}w_t^i\big\rangle + \gamma_t^2 \big|\big|\nabla_\theta B(\theta_t^{i,N}, x_t^{i,N}, \mu_t^N)\big|\big|_F^2 \mathrm{d}t
\end{aligned}
$$

Due to the strong concavity of $\tilde{\mathcal{L}}(\theta)$ (Condition H.1), it then follows that

$$
\begin{aligned}
\mathrm{d}||Z_t^{i,N}||^2 + 2\eta\gamma_t ||Z_t^{i,N}||^2 \mathrm{d}t &\leq \gamma_t\big\langle Z_t^{i,N}, \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N}) - \nabla_\theta \tilde{\mathcal{L}}(\theta_t^{i,N})\big\rangle \mathrm{d}t \qquad (5.147)\\
&+ \gamma_t\big\langle Z_t^{i,N}, \nabla_\theta L(\theta_t^{i,N}, x_t^{i,N}, \mu_t^N) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N})\big\rangle \mathrm{d}t\\
&+ \gamma_t\big\langle Z_t^{i,N}, \nabla_\theta B(\theta_t^{i,N}, x_t^{i,N}, \mu_t^N)\mathrm{d}w_t^i\big\rangle\\
&+ \gamma_t^2 \big|\big|\nabla_\theta B(\theta_t^{i,N}, x_t^{i,N}, \mu_t^N)\big|\big|_F^2 \mathrm{d}t
\end{aligned}
$$

where $||\cdot||_F$ is the Frobenius norm. Now, let us define the function $\Phi_{t,t'} = \exp[-2\eta \int_t^{t'} \gamma_u \mathrm{d}u]$, with $\partial_t \Phi_{t,t'} = 2\eta\gamma_t \Phi_{t,t'}$. Using the product rule, and (5.147), we obtain

$$
\begin{aligned}
\mathrm{d}\left[\Phi_{t,t'} ||Z_t^{i,N}||^2\right] &= \Phi_{t,t'}\left[\mathrm{d}||Z_t^{i,N}||^2 + 2\eta\gamma_t ||Z_t^{i,N}||^2 \mathrm{d}t\right] \qquad (5.148)\\
&\leq \gamma_t \Phi_{t,t'} \big\langle Z_t^{i,N}, \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N}) - \nabla_\theta \tilde{\mathcal{L}}(\theta_t^{i,N})\big\rangle \mathrm{d}t \qquad (5.149)\\
&+ \gamma_t \Phi_{t,t'} \big\langle Z_t^{i,N}, \nabla_\theta L(\theta_t^{i,N}, x_t^{i,N}, \mu_t^N) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N})\big\rangle \mathrm{d}t\\
&+ \gamma_t \Phi_{t,t'} \big\langle Z_t^{i,N}, \nabla_\theta B(\theta_t^{i,N}, x_t^{i,N}, \mu_t^N)\mathrm{d}w_t^i\big\rangle\\
&+ \gamma_t^2 \Phi_{t,t'} \big|\big|\nabla_\theta B(\theta_t^{i,N}, x_t^{i,N}, \mu_t^N)\big|\big|_F^2 \mathrm{d}t
\end{aligned}
$$

Rewriting this in integral form, setting $t' = t$, and taking expectations, we obtain

$$\begin{aligned}
\mathbb{E}\left[||Z_t^{i,N}||^2\right] &\leq \mathbb{E}\left[\Phi_{1,t}||Z_1^{i,N}||^2\right] && (5.150)\\
&\quad + \mathbb{E}\left[\int_1^t \gamma_s \Phi_{s,t} \big\langle Z_s^{i,N}, \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_s^{i,N}) - \nabla_\theta \tilde{\mathcal{L}}(\theta_s^{i,N})\big\rangle \mathrm{d}s\right]\\
&\quad + \mathbb{E}\left[\int_1^t \gamma_s \Phi_{s,t} \big\langle Z_s^{i,N}, \nabla_\theta L(\theta_s^{i,N}, x_s^{i,N}, \mu_s^N) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_s^{i,N})\big\rangle \mathrm{d}s\right]\\
&\quad + \mathbb{E}\left[\int_1^t \gamma_s^2 \Phi_{s,t} \big|\big|\nabla_\theta B(\theta_s^{i,N}, x_s^{i,N}, \mu_s^N)\big|\big|_F^2 \mathrm{d}s\right]\\
&= \mathbb{E}\left[\Omega_{t,i,N}^{(1)}\right] + \mathbb{E}\left[\Omega_{t,i,N}^{(2)}\right] + \mathbb{E}\left[\Omega_{t,i,N}^{(3)}\right] + \mathbb{E}\left[\Omega_{t,i,N}^{(4)}\right] && (5.151)
\end{aligned}$$

We will deal with each of these terms separately, beginning with $\Omega_{t,i,N}^{(1)}$. For this term, we have that, for sufficiently large t,

$$\mathbb{E}\left[\Omega_{t,i,N}^{(1)}\right] = \Phi_{1,t}\mathbb{E}\left[||Z_1^{i,N}||^2\right] \leq K^{(1)}\gamma_t \tag{5.152}$$

which follows from Lemma 5.1 (the moment bounds for $\theta_s^{i,N}$), and Condition F.1 (the conditions on the learning rate).

We now turn our attention to $\Omega_{t,i,N}^{(2)}$. For this term, substituting the bound in Lemma 5.19 into (5.150), we immediately obtain

$$\begin{aligned}
\mathbb{E}\left[\Omega_{t,i,N}^{(2)}\right] &\leq \int_1^t \gamma_s \Phi_{s,t} \mathbb{E}\left[||Z_s^{i,N}|| \sup_{\theta_s^{i,N}} ||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_s^{i,N}) - \nabla_\theta \mathcal{L}(\theta_s^{i,N})||\right] \mathrm{d}s && (5.153)\\
&\leq K\left[\frac{1}{N^{\frac{1}{2}}}\right] \int_1^t \gamma_s \Phi_{s,t} \mathrm{d}s \leq K^{(2)}\left[\frac{1}{N^{\frac{1}{2}}}\right]. && (5.154)
\end{aligned}$$

where in the last line we have used Condition F.1 (the conditions on the learning rate) to bound the integral.

We now turn our attention to $\Omega_{t,i,N}^{(3)}$. We will analyse this term by constructing an appropriate Poisson equation. Let us define

$$R^{i,N}(\theta, \hat{x}^N) = \left\langle \theta - \theta_0, \nabla_\theta \hat{L}^{i,N}(\theta, \hat{x}^N) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta) \right\rangle, \tag{5.155}$$

where, as previously, $\hat{x}^N = (x^{1,N}, \ldots, x^{N,N})$. It is straightforward to verify that this function satisfies all of the conditions of Lemma 5.15. Thus, by Lemma 5.15, the Poisson equation

$$\mathcal{A}_x v^{i,N}(\theta,\hat{x}^N) = R^{i,N}(\theta,\hat{x}^N) \quad , \quad \int_{\mathbb{R}^d} v^{i,N}(\theta,\hat{x}^N)\hat{\mu}^N_\infty(\mathrm{d}\hat{x}^N) = 0 \tag{5.156}$$

has a unique twice differentiable solution which satisfies

$$\sum_{j=0}^{2}\left|\frac{\partial^j v^{i,N}}{\partial\theta^i}(\theta,\hat{x}^N)\right| + \left|\frac{\partial^2 v^{i,N}}{\partial\theta\partial x}(\theta,\hat{x}^N)\right| \leq K\bigg(1+||x^{i,N}||^q + \frac{1}{N}\sum_{j=1}^{N}||x^{j,N}||^q\bigg). \tag{5.157}$$

Now, by Itô's formula, we have that

$$\begin{aligned} v^{i,N}(\theta^{i,N}_t,\hat{x}^N_t) - v^{i,N}(\theta^{i,N}_s,\hat{x}^N_s) &= \int_s^t \mathcal{A}_\theta v^{i,N}(\theta^{i,N}_u,\hat{x}^N_u)\mathrm{d}u + \int_s^t \mathcal{A}_{\hat{x}^N} v^{i,N}(\theta^{i,N}_u,\hat{x}^N_u)\mathrm{d}u \qquad (5.158)\\ &+ \int_s^t \gamma_u \partial_\theta v^{i,N}(\theta^{i,N}_u,\hat{x}^N_u)\nabla_\theta \hat{B}^{i,N}(\theta_u,\hat{x}^N_u)\mathrm{d}w^i_u \\ &+ \int_s^t \partial_x v^{i,N}(\theta^{i,N}_u,\hat{x}^N_u)\mathrm{d}\hat{w}^N_u \\ &+ \int_s^t \gamma_u\left[\partial_\theta\partial_{\hat{x}} v^{i,N}(\theta^{i,N}_u,\hat{x}^N_u)\nabla_\theta\hat{B}^{i,N}(\theta^{i,N}_u,\hat{x}^N_u)\right]\mathrm{d}u \end{aligned}$$

where $\hat{w}^N_u$ was defined in (5.96). It follows, now writing $v^{i,N}_t := v^{i,N}(\theta^{i,N}_t,\hat{x}^N_t)$, that

$$\begin{aligned} R^{i,N}(\theta^{i,N}_t,\hat{x}^N_t)\mathrm{d}t &= \mathcal{A}_{\hat{x}^N} v^{i,N}(\theta^{i,N}_t,\hat{x}^N_t)\mathrm{d}t \qquad (5.159)\\ &= \mathrm{d}v^{i,N}_t - \mathcal{A}_\theta v^{i,N}(\theta^{i,N}_t,\hat{x}^N_t)\mathrm{d}t \qquad (5.160)\\ &\quad - \gamma_t\partial_\theta v^{i,N}(\theta^{i,N}_t,\hat{x}^N_t)\nabla_\theta\hat{B}^{i,N}(\theta^{i,N}_t,\hat{x}^N_t)\mathrm{d}w^i_t \\ &\quad - \partial_{\hat{x}} v^{i,N}(\theta^{i,N}_t,\hat{x}^N_t)\mathrm{d}\hat{w}^N_t \\ &\quad - \gamma_t\left[\partial_\theta\partial_{\hat{x}} v^{i,N}(\theta^{i,N}_t,\hat{x}^N_t)\nabla_\theta\hat{B}^{i,N}(\theta^{i,N}_t,\hat{x}^N_t)\right]\mathrm{d}t \end{aligned}$$

Thus, we can rewrite $\Omega^{(3)}_{t,i,N}$ as

$$\begin{aligned} \Omega^{(3)}_{t,i,N} &= \int_1^t \gamma_s\Phi_{s,t}\underbrace{\left\langle\theta^{i,N}_s-\theta_0,\nabla_\theta\hat{L}^{i,N}(\theta^{i,N}_s,\hat{x}^N_s) - \nabla_\theta\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_s)\right\rangle\mathrm{d}s}_{R^{i,N}(\theta^{i,N}_s,\hat{x}^N_s)\mathrm{d}s} \qquad (5.161)\\ &= \int_1^t \gamma_s\Phi_{s,t}\mathrm{d}v^{i,N}_s - \int_1^t \gamma_s\Phi_{s,t}\mathcal{A}_\theta v^{i,N}(\theta^{i,N}_s,\hat{x}^N_s)\mathrm{d}s \qquad (5.162)\\ &\quad - \int_1^t \gamma_s^2\Phi_{s,t}\partial_\theta v^{i,N}(\theta^{i,N}_s,\hat{x}^N_s)\nabla_\theta\hat{B}^{i,N}(\theta^{i,N}_s,\hat{x}^N_s)\mathrm{d}w^i_s \end{aligned}$$

$$
-\int_1^t \gamma_s \Phi_{s,t} \partial_{\hat{x}} v^{i,N}(\theta_s^{i,N}, \hat{x}_s^N) \mathrm{d}\hat{w}_s^N
$$
$$
-\int_1^t \gamma_s^2 \Phi_{s,t} \partial_\theta \partial_x v^{i,N}(\theta_s, \hat{x}_s^N) \nabla_\theta \hat{B}^{i,N}(\theta_s^{i,N}, \hat{x}_s^N) \mathrm{d}s
$$

We can rewrite the first term in this expression by applying Itô's formula to the function $f(s, v_s) = \gamma_s \Phi_{s,t} v_s$. This yields

$$
\gamma_t \Phi_{t,t} v_t^{i,N} - \gamma_1 \Phi_{1,t} v_1^{i,N} = \int_1^t \gamma_s \Phi_{s,t} \mathrm{d}v_s^{i,N} + \int_1^t \dot{\gamma}_s \Phi_{s,t} v_s^{i,N} \mathrm{d}s + \int_1^t 2\eta \gamma_s^2 \Phi_{s,t} v_s^{i,N} \mathrm{d}s. \tag{5.163}
$$

Substituting the resulting expression for $\int_1^t \gamma_s \Phi_{s,t} \mathrm{d}v_s^{i,N}$ into (5.161), and taking expectations, we obtain

$$
\begin{aligned}
\mathbb{E}\left[\Omega_{t,i,N}^{(3)}\right] &= \mathbb{E}\left[\gamma_t \Phi_{t,t} v^{i,N}(\theta_t^{i,N}, \hat{x}_t^N)\right] - \mathbb{E}\left[\gamma_1 \Phi_{1,t} v^{i,N}(\theta_1^{i,N}, \hat{x}_1^N)\right] && (5.164)\\
&\quad - \mathbb{E}\left[\int_1^t \dot{\gamma}_s \Phi_{s,t} v^{i,N}(\theta_s^{i,N}, \hat{x}_s^N) \mathrm{d}s\right] - \mathbb{E}\left[\int_1^t 2\eta\gamma_s^2 \Phi_{s,t} v^{i,N}(\theta_s^{i,N}, \hat{x}_s^N) \mathrm{d}s\right]\\
&\quad - \mathbb{E}\left[\int_1^t \gamma_s \Phi_{s,t} \mathcal{A}_\theta v^{i,N}(\theta_s^{i,N}, \hat{x}_s^N) \mathrm{d}s\right]\\
&\quad - \mathbb{E}\left[\int_1^t \gamma_s^2 \Phi_{s,t} \partial_\theta \partial_x v^{i,N}(\theta_s^{i,N}, \hat{x}_s^N) \nabla_\theta \hat{B}^{i,N}(\theta_s^{i,N}, \hat{x}_s^N) \mathrm{d}s\right]\\
&\leq K\left[\gamma_t + \int_1^t \left(\dot{\gamma}_s + \gamma_s^2\right) \Phi_{s,t} \mathrm{d}s\right] \leq K^{(3)} \gamma_t, && (5.165)
\end{aligned}
$$

where in the penultimate inequality we have used the polynomial growth of $v^{i,N}(\theta, \hat{x}^N)$ and $\partial_\theta \partial_x v^{i,N}(\theta, \hat{x}^N)$, Condition C.1(ii) (which implies the polynomial growth of $\nabla_\theta \hat{B}^{i,N}(\theta, \hat{x}^N)$), Proposition 5.2 (the moment bounds for $\hat{x}_t^N$), Lemma 5.1 (the moment bounds for $\theta_s^{i,N}$), and in the final inequality we have used Condition F.1 (the conditions on the learning rate). It remains only to bound $\Omega_{t,i,N}^{(4)}$. For this term, once more making use of the above assumptions, we obtain

$$
\mathbb{E}\left[\Omega_{t,i,N}^{(4)}\right] = \mathbb{E}\left[\int_1^t \gamma_s^2 \Phi_{s,t} \big|\big|\nabla_\theta B(\theta_s, x_s, \mu_s)\big|\big|_F^2 \mathrm{d}s\right] \leq K \int_1^t \gamma_s^2 \Phi_{s,t} \mathrm{d}s \leq K^{(4)} \gamma_t. \tag{5.166}
$$

Combining inequalities (5.152), (5.153), (5.165), and (5.166), and setting $K_1 = \max\{K^{(1)}, K^{(3)}\}$,

$K_2 = K^{(4)}$, and $K_3 = K^{(2)}$, we thus have that

$$\mathbb{E}\left[||\theta_t^{i,N} - \theta_0||^2\right] \leq \mathbb{E}\left[\Omega_{t,i,N}^{(1)}\right] + \mathbb{E}\left[\Omega_{t,i,N}^{(2)}\right] + \mathbb{E}\left[\Omega_{t,i,N}^{(3)}\right] + \mathbb{E}\left[\Omega_{t,i,N}^{(4)}\right] \tag{5.167}$$

$$\leq (K_1 + K_2)\gamma_t + \frac{K_3}{N^{\frac{1}{2}}}, \tag{5.168}$$

which completes the proof of the first statement of the theorem.

Let us now turn our attention to the second statement. The proof of this bound goes through almost verbatim. Let us briefly highlight the main points of difference. To begin, we now have the following decomposition of the parameter update equation

$$\begin{aligned}\mathrm{d}\theta_t^N = {} & \gamma_t \nabla_\theta \tilde{\mathcal{L}}(\theta_t^N)\mathrm{d}t + \gamma_t \frac{1}{N}\sum_{i=1}^N \big(\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^N) - \nabla_\theta \tilde{\mathcal{L}}(\theta_t^N)\big)\mathrm{d}t \\ & + \gamma_t \frac{1}{N}\sum_{i=1}^N \big(\nabla_\theta L(\theta_t^N, x_t^{i,N}, \mu_t^N) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^N)\big)\mathrm{d}t \\ & + \gamma_t \frac{1}{N}\sum_{i=1}^N \nabla_\theta B(\theta_t^{i,N}, x_t^{i,N}, \mu_t^N)\mathrm{d}w_t^i.\end{aligned} \tag{5.169}$$

Using a Taylor expansion around θ_0, defining $Z_t^N = \theta_t^N - \theta_0$, applying Itô's formula to the function $||Z_t^N||^2$, and using the strong concavity of $\tilde{\mathcal{L}}(\theta)$, as in (5.144) - (5.147), we obtain

$$\begin{aligned}\mathrm{d}||Z_t^N||^2 + 2\eta\gamma_t||Z_t^N||^2\mathrm{d}t \leq {} & \gamma_t \frac{1}{N}\sum_{i=1}^N \big\langle Z_t^N, \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^{i,N}) - \nabla_\theta \tilde{\mathcal{L}}(\theta_t^N)\big\rangle \mathrm{d}t \\ & + \gamma_t \frac{1}{N}\sum_{i=1}^N \big\langle Z_t^N, \nabla_\theta L(\theta_t^{i,N}, x_t^{i,N}, \mu_t^N) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_t^N)\big\rangle \mathrm{d}t \\ & + \gamma_t \frac{1}{N}\sum_{i=1}^N \big\langle Z_t^N, \nabla_\theta B(\theta_t^N, x_t^{i,N}, \mu_t^N)\mathrm{d}w_t^i\big\rangle \\ & + \gamma_t^2 \frac{1}{N^2}\sum_{i=1}^N \big|\big|\nabla_\theta B(\theta_t, x_t, \mu_t)\big|\big|_F^2 \mathrm{d}t\end{aligned} \tag{5.170}$$

Continuing to follow our previous arguments, c.f. (5.148) - (5.151), we finally arrive at

$$\mathbb{E}\left[||Z_t^N||^2\right] \leq \frac{1}{N}\sum_{i=1}^N \left[\mathbb{E}\left[\tilde{\Omega}_{t,i,N}^{(1)}\right] + \mathbb{E}\left[\tilde{\Omega}_{t,i,N}^{(2)}\right] + \mathbb{E}\left[\tilde{\Omega}_{t,i,N}^{(3)}\right]\right] + \frac{1}{N^2}\sum_{i=1}^N \mathbb{E}\left[\tilde{\Omega}_{t,i,N}^{(4)}\right] \tag{5.171}$$

where, up to minor modifications, $\tilde{\Omega}_{t,i,N}^{(1)}, \ldots, \tilde{\Omega}_{t,i,N}^{(4)}$ are identical to $\Omega_{t,i,N}^{(1)}, \ldots, \Omega_{t,i,N}^{(4)}$ as defined in (5.150) - (5.151). In particular, all instances of $\theta_s^{i,N}$ in $\Omega_{t,i,N}^{(1)}, \ldots, \Omega_{t,i,N}^{(4)}$ have been replaced by θ_s^N in $\tilde{\Omega}_{t,i,N}^{(1)}, \ldots, \tilde{\Omega}_{t,i,N}^{(4)}$. We thus have, using the bounds established previously, c.f. (5.152), (5.153), (5.165), and (5.166), that

$$\mathbb{E}\left[||Z_t^N||^2\right] \leq \frac{1}{N}\sum_{i=1}^N \left(K_1\gamma_t + K_3\frac{1}{N^{\frac{1}{2}}}\right) + \frac{1}{N^2}\sum_{i=1}^N K_2\gamma_t = (K_1 + \frac{K_2}{N})\gamma_t + \frac{K_3}{N^{\frac{1}{2}}}. \quad (5.172)$$

□

5.5 Numerical Examples

To illustrate the results of Section 5.3, we now provide several illustrative examples of parameter estimation in McKean-Vlasov SDEs, and the associated systems of interacting particles. In particular, we consider a linear mean-field model, a model with bistable potential, a stochastic Kuramoto model, and a stochastic opinion dynamics model. In all cases, we simulate sample paths and implement the recursive MLE using a standard Euler-Maruyama scheme with $\Delta t = 0.1$.

5.5.1 Linear Mean Field Dynamics

We first consider a one-dimensional linear mean field model, parametrised by $\theta = (\theta_1, \theta_2)^T \in \mathbb{R}^2$, given by

$$\mathrm{d}x_t = -\left[\theta_1 x_t + \theta_2 \int_{\mathbb{R}} (x_t - y)\mu_t(\mathrm{d}y)\right] \mathrm{d}t + \sigma \mathrm{d}w_t, \quad (5.173)$$

$$\mu_t = \mathcal{L}(x_t). \quad (5.174)$$

where $\sigma > 0$ and $w = (w_t)_{t\geq 0}$ is a standard Brownian motion. We will assume that $x_0 \in \mathbb{R}$. This is clearly of the form of the McKean-Vlasov SDE (5.1) - (5.2) with $b(\theta, x) = -\theta_1 x$ and $\phi(\theta, x, y) = -\theta_2(x - y)$. The corresponding system of interacting particles is given by

$$\mathrm{d}x_t^{i,N} = -\left[\theta_1 x_t^{i,N} + \theta_2 \frac{1}{N}\sum_{j=1}^N (x_t^{i,N} - x_t^{j,N})\right] \mathrm{d}t + \sigma \mathrm{d}w_t^{i,N}, \quad i = 1, \ldots, N. \quad (5.175)$$

In this model, the parameter θ_1 controls the strength of attraction of the non-linear process (or, in the IPS, of each individual particle) towards zero, while the strength of the parameter θ_2 controls the strength of the attraction of the non-linear process (of each individual particle) towards its mean (the empirical mean). We remark that, in the case $\theta_2 = 0$, the non-linear process reduces to a one-dimensional Orstein-Uhlenbeck process, and the system of interacting particles reduces to N independent samples of this process. It is straightforward to show that this model satisfies all of the conditions specified in

Section 5.3.1. The full details can be found in [413, Appendix I].

5.5.1.1 Offline Parameter Estimation

We begin by illustrating the performance of the offline MLE. Since this model is linear in both of the parameters, in this case it is possible to obtain the maximum likelihood in closed form as (see also [242])

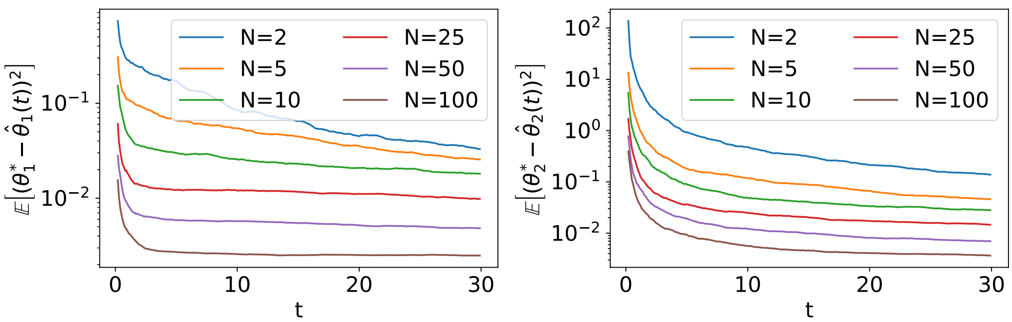
$$\hat{\theta}^N_{1,t} = \frac{A^N_t - B^N_t}{C^N_t - D^N_t} \quad , \quad \hat{\theta}^N_{2,t} = \frac{D^N_t A^N_t - C^N_t B^N_t}{(C^N_t)^2 - C^N_t D^N_t} \tag{5.176}$$

where we have defined, writing $\bar{x}^N_s = \frac{1}{N}\sum_{j=1}^N x^{j,N}_s$,

$$A^N_t = \int_0^t \sum_{i=1}^N (x^{i,N}_s - \bar{x}^N_s)\mathrm{d}x^{i,N}_s \ , \quad B^N_t = \int_0^t \sum_{i=1}^N x^{i,N}_s \mathrm{d}x^{i,N}_s \tag{5.177}$$

$$C^N_t = \int_0^t \sum_{i=1}^N (x^{i,N}_s - \bar{x}^N_s)^2\mathrm{d}s \quad , \quad D^N_t = \int_0^t \sum_{i=1}^N (x^{i,N}_s)^2\mathrm{d}s. \tag{5.178}$$

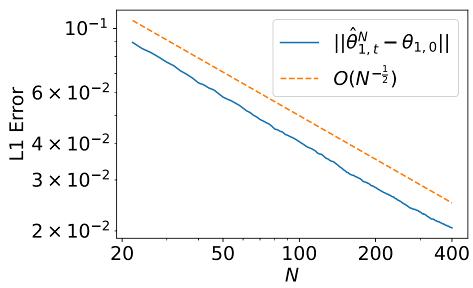
For our first simulation, we assume that the true parameter is given by $\theta^* = (1, 0.5)^T$, and that the diffusion coefficient is equal to the identity, $\sigma = 1$. The performance of the MLE is visualised in Figure 5.1, in which we plot the mean squared error (MSE) of the offline parameter estimate for $t \in [0, 30]$, and $N \in \{2, 5, 10, 25, 50, 100\}$, averaged over 500 random trials. As expected, the parameter estimates converge to the true parameter values (that is, the MSE converges to zero) as N increases with t fixed (see Theorem 5.1), and also as t increases with N fixed (see, e.g., [63, 296]).



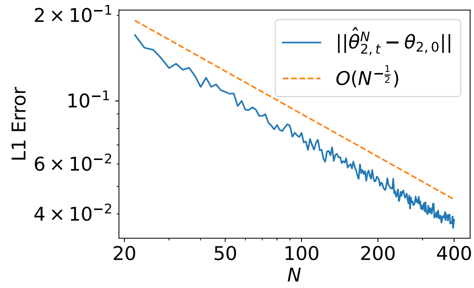
(a) $\hat{\theta}^N_{1,t}$. (b) $\hat{\theta}^N_{2,t}$.

Figure 5.1: $\mathbb{L}^2$ error of the offline MLE for $t \in [0, 30]$ and $N = \{2, 5, 10, 25, 50, 100\}$. The $\mathbb{L}^2$ error is plotted on a log-scale.

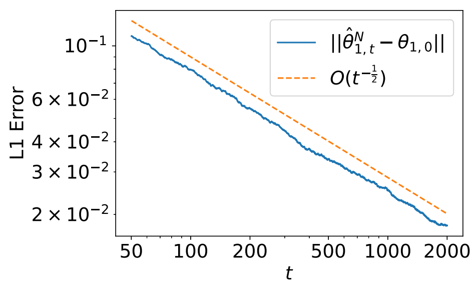
We investigate the convergence rate of the offline MLE further in Figure 5.2, in which we plot the mean absolute error (MAE) of the parameter estimate for $N \in \{20, 21, \ldots, 400\}$ with $t = 5$, and also for $t \in [50, 2000]$ with $N = 2$, averaged over 500 random trials. Our results suggest that the offline MLE for this model has an $\mathbb{L}^1$ convergence rate of order $O((Nt)^{-\frac{1}{2}})$. This is rather unsurprising: such a rate was recently established by Chen [110] for a linear mean field model (of arbitrary dimension) in the absence of the global confinement term.



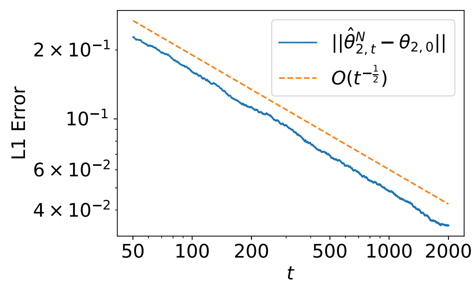
(a) $\hat{\theta}^N_{1,t=5}$.



(b) $\hat{\theta}^N_{2,t=5}$.



(c) $\hat{\theta}^{N=2}_{1,t}$.



(d) $\hat{\theta}^{N=2}_{2,t}$.

Figure 5.2: Log-log plot of the $\mathbb{L}^1$ error of the offline MLE for $t = 5$ and $N \in \{20, \ldots, 400\}$ (top panel), and for $t \in [50, 2000]$ and $N = 2$ (bottom panel).

To conclude this section, we provide numerical confirmation of the asymptotic normality of the MLE (Theorem 5.2). For the linear mean field model of interest, it is in fact possible to obtain the asymptotic information matrix in closed form (see also [242]). In particular, it is given by

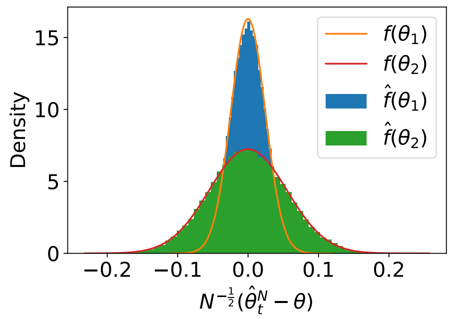
$$I_t(\theta) = \begin{pmatrix} D_t(\theta) & C_t(\theta) \\ C_t(\theta) & C_t(\theta) \end{pmatrix}, \tag{5.179}$$

where, with $\gamma(\theta) = -2(\theta_1 + \theta_2)$,

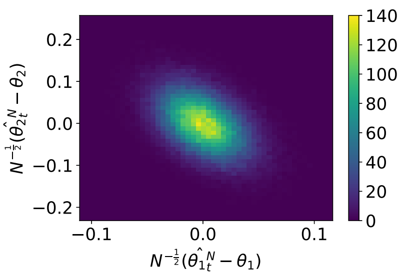
$$C_t(\theta) = \frac{1}{\gamma^2(\theta)}(e^{\gamma(\theta)t} - 1) - \frac{t}{\gamma(\theta)} + \frac{\sigma_0^2}{\gamma}(e^{\gamma(\theta)t} - 1), \tag{5.180}$$

$$D_t(\theta) = \frac{1}{\gamma^2(\theta)}(e^{\gamma(\theta)t} - 1) - \frac{t}{\gamma(\theta)} + \frac{\sigma_0^2}{\gamma(\theta)}(e^{\gamma(\theta)t} - 1) - \frac{\mu_0^2}{2\theta_1}(e^{-2\theta_1 t} - 1). \tag{5.181}$$

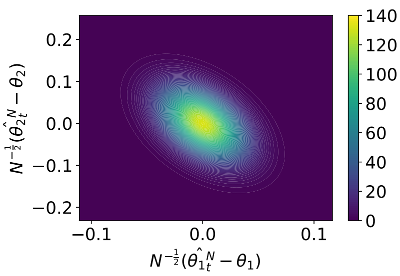
In Figure 5.3, we are thus able to provide a direct comparison of the asymptotic distribution of the MLE, and the approximate distribution obtained with a finite number of particles. Our results are computed using $T = 2$, $N = 500$, and $n = 10^5$ sample paths.



(a) Asymptotic & approximate marginals.



(b) Approximate bivariate density.



(c) Asymptotic bivariate density.

Figure 5.3: A comparison between the asymptotic normal distribution and the approximate normal distribution of the MLE for $N = 500$ particles. The histograms were obtained using $n = 10^5$ random samples.

5.5.1.2 Online Parameter Estimation

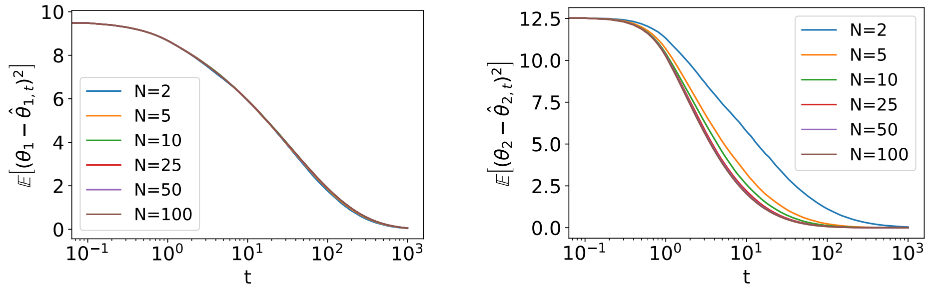
We now turn our attention to the online MLE, which for this model evolves according to

$$\mathrm{d}\theta_{1,t}^N = \frac{\gamma_{1,t}}{N\sigma^2}\sum_{i=1}^N \left[-x_t^{i,N}\mathrm{d}x_t^{i,N} - x_t^{i,N}(\theta_{1,t}^N x_t^{i,N} + \theta_{2,t}^N(x_t^{i,N} - \bar{x}_t^N))\mathrm{d}t\right], \tag{5.182}$$

$$\mathrm{d}\theta_{2,t}^N = \frac{\gamma_{2,t}}{N\sigma^2}\sum_{i=1}^N \left[-(x_t^{i,N} - \bar{x}_t^N)\mathrm{d}x_t^{i,N} - (x_t^{i,N} - \bar{x}_t^N)(\theta_{1,t}^N x_t^{i,N} + \theta_{2,t}^N(x_t^{i,N} - \bar{x}_t^N))\mathrm{d}t\right]. \tag{5.183}$$

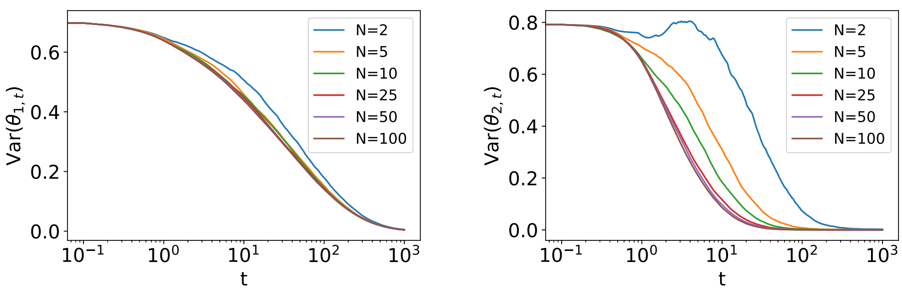
We will initially assume that one of the parameters is fixed (and equal to the true value), while the other parameter is to be estimated. The true parameters are given by $\theta_1^* = 0.5$ and $\theta_2^* = 0.1$. Meanwhile, the initial parameter estimates are randomly generated according to $\theta_1^0, \theta_2^0 \sim \mathcal{U}([2,5])$. Finally, the learning rates are given by $\gamma_{i,t} = \min\{\gamma_i^0, \gamma_i^0 t^{-\alpha}\}$, $i =$

$1, 2$, where $\gamma_1^0 = 0.05$, $\gamma_2^0 = 0.30$, and $\alpha = 0.51$. The performance of the stochastic gradient descent algorithm is visualised in Figures 5.4 and 5.5, in which we plot the MSE and the variance of the online parameter estimates for $t \in [0, 1000]$ and $N = \{2, 5, 10, 25, 50, 100\}$. The results are computed over 500 independent random trials. Interestingly, increasing the number of particles can result in a relatively significant reduction in the MSE of the interaction parameter θ_2, but has little consequence for the error of the confinement parameter θ_1. Meanwhile, there is a relatively significant reduction in the variance of both estimates.



(a) $\theta_{1,t}^N$.   (b) $\theta_{2,t}^N$.

Figure 5.4: $\mathbb{L}^2$ error of the online parameter estimates for $t \in [0, 1000]$ and $N = \{2, 5, 10, 25, 50, 100\}$. The time is plotted on a log-scale.



(a) $\theta_{1,t}^N$.   (b) $\theta_{2,t}^N$.

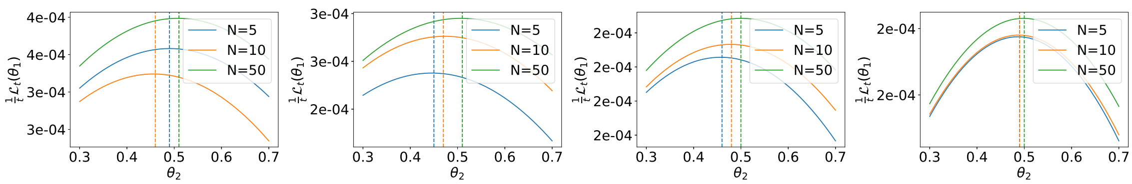
Figure 5.5: Variance of the online parameter estimates for $t \in [0, 1000]$ and $N = \{2, 5, 10, 25, 50, 100\}$. The time is plotted on a log-scale.

We should remark that, in the linear mean field model, with one parameter fixed, the online parameter estimates generated via the system of interacting particles will converge to the true value of the parameter (which coincides with the global minimum of the asymptotic log-likelihood of the McKean-Vlasov SDE) for all values of N. Indeed, for this model, the (asymptotic) log-likelihood of the IPS is strongly concave for all values of N, with unique global maximum at the true parameter values. This is visualised in Figures 5.6d and 5.7d, in which we have plotted approximations of profile asymptotic log-likelihood of the IPS for several values of N. We are thus in the regime of Case I with finite N, meaning θ_t^N will converge to the true parameter as $t \to \infty$, regardless of the value of N.

Figures 5.6 and 5.7 also provide a numerical illustration of why the finite-time performance of the online estimator improves with the number of particles (see Theorem 5.4‡), and why this improvement is more pronounced for the interaction parameter θ_2. As N increases, we observe that the time weighted average of the log-likelihood $\mathcal{L}_t^N(\theta)$ (the noisy objective function) much more closely resembles the asymptotic log-likelihood $\tilde{\mathcal{L}}^N(\theta)$ (the true objective function), even for small time values. This means, in particular, that the fluctuations terms appearing in the proof of Theorems 5.3‡ - 5.4‡ of the form

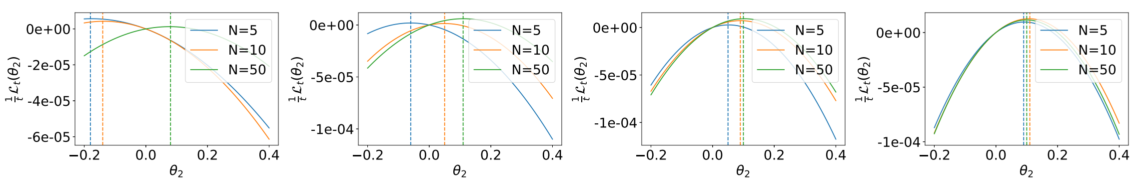
$$\int_0^t \gamma_s(\nabla_\theta \tilde{\mathcal{L}}^N(\theta_s^N) - \frac{1}{N}\sum_{i=1}^N \nabla_\theta L(\theta_s^N, x_s^N, \mu_s^N))\mathrm{d}s, \tag{5.184}$$

converge more rapidly to zero (as a function of time), for larger values of N. This disparity in the convergence rate of the log-likelihood (as a function of the time), for different values of N, appears to be much more significant for the interaction parameter θ_2 (Figure 5.7) than it is for the confinement parameter θ_1 (Figure 5.6). Consequently, the online parameter estimate $\theta_{2,t}^N$ converges more rapidly as N increases, while there is little difference in the convergence rate of $\theta_{1,t}^N$.



(a) $T = 1.0$. (b) $T = 2.5$. (c) $T = 5.0$. (d) $T = 7.5$.

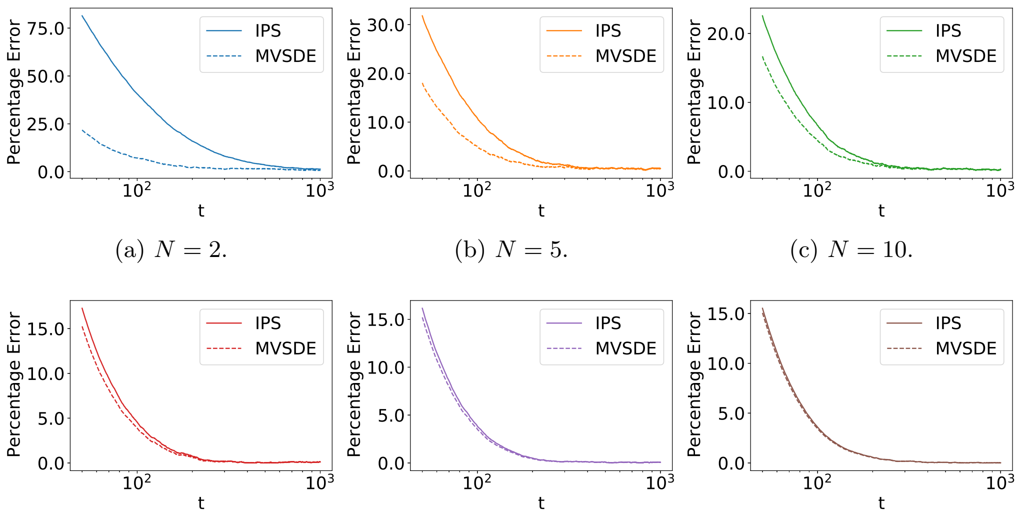
Figure 5.6: Plots of the average log-likelihood, $\frac{1}{T}\mathcal{L}_T^N(\theta_1)$, for $T = \{1, 2.5, 5.7.5\}$ and $N = \{5, 10, 50\}$.



(a) $T = 1.0$. (b) $T = 2.5$. (c) $T = 5.0$. (d) $T = 7.5$.

Figure 5.7: Plots of the average log-likelihood, $\frac{1}{T}\mathcal{L}_T^N(\theta_2)$, for $T = \{1, 2.5, 5.7.5\}$ and $N = \{5, 10, 50\}$.

We conclude this discussion with a comparison between the online parameter estimates generated using N particles from the IPS, and those generated using a single sample path of McKean-Vlasov SDE, and its law. We should emphasise that the latter is only possible when the solution of the non-linear equation is available. Illustrative results are provided in Figure 5.8, in which we plot the percentage error of the online parameter estimates for the interaction parameter, for several values of N. In each case, the estimate based on the McKean-Vlasov SDE converges more rapidly to the true parameter value. We also note, perhaps unsurprisingly, that this disparity becomes less apparent as the number of particles increases, reflecting the fact that the dynamics of the interacting particles increasingly resemble the dynamics of the solutions of the non-linear equation. Consistent with our previous observations, this disparity is also less apparent for the online estimates of the confinement parameter (results omitted).



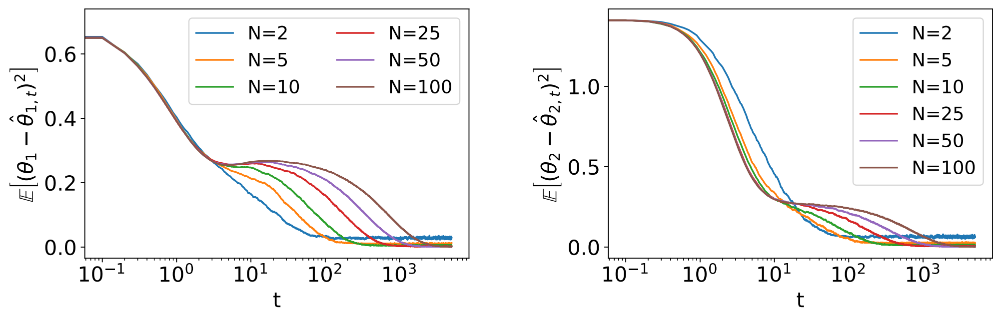
(a) $N = 2$. (b) $N = 5$. (c) $N = 10$.

(d) $N = 25$. (e) $N = 50$. (f) $N = 100$.

Figure 5.8: Percentage error of the online maximum likelihood estimates of the interaction parameter θ_2 for $T \in [0, 1000]$ and $N = \{2, 5, 10, 25, 50, 100\}$, generated using the IPS and the McKean-Vlasov SDE.

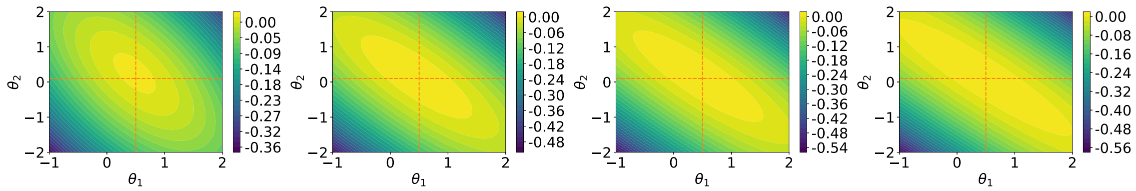
Let us now turn our attention to the case in which both parameters are unknown, and to be estimated from the data. For the sake of comparison, we will once more assume that that the true parameter is given by $\theta^* = (\theta_1^*, \theta_2^*) = (0.5, 0.1)$. The initial parameter estimates are now generated according to $\theta_1^0 \sim \mathcal{U}([-1, 2])$ and $\theta_2^0 \sim \mathcal{U}([-2, 2])$. Finally, we use constant learning rates, with $\gamma_{1,t} = 0.1$ and $\gamma_{2,t} = 0.2$. The performance of the stochastic gradient descent algorithm is illustrated in Figure 5.9, in which we plot the MSE of the online parameter estimates for both of the unknown parameters, averaged over 500 random trials.

In this case, the evolution of the MSE appears to indicate three distinct learning phases. In the initial phase, the performance of the online estimator improves as a function of the number of particles, with this improvement being more noticeable for the interaction parameter θ_2, as observed previously. Conversely, in the middle phase, the online estimator performs (significantly) better for smaller values of N. These observations are readily explained with reference to the asymptotic log-likelihood of the IPS for different values of N, as shown in Figure 5.10. In particular, far from the global maximum at $\theta^* = (0.5, 0.1)$, the asymptotic log-likelihood decreases more steeply as the value of N increases. Broadly speaking, we can think of this region of the optimisation landscape as responsible for the initial learning phase, hence the improved performance of the estimator for larger values of N. On the other hand, close to the global maximum, the asymptotic log-likelihood exhibits an increasingly large plateau as the value of N increases (i.e., an increasingly flat maximum). This region of the optimisation landscape is largely responsible for the middle learning phase, which explains the slower convergence of the estimator for larger values of N. In the final learning phase, the steady-state error of the recursive MLE appears to decrease as a function of the number of particles. This is unsurprising, given the $\mathcal{O}(\frac{1}{N^{\frac{1}{2}}})$ term appearing in Theorem 5.4.



(a) $\theta_{1,t}^N$. (b) $\theta_{2,t}^N$.

Figure 5.9: $\mathbb{L}^2$ error of the online MLEs for $T \in [0, 5000]$ and $N = \{2, 5, 10, 25, 50, 100\}$. The time is plotted on a log-scale.



(a) $N = 2$. (b) $N = 5$. (c) $N = 10$. (d) $N = 100$.

Figure 5.10: Contour plots of the asymptotic log-likelihood $\tilde{\mathcal{L}}^N(\theta)$ for $N = \{2, 5, 10, 100\}$.

5.5.2 A Bistable Potential

In this section, we consider a one-dimensional mean-field model, parametrised by $\theta \in \mathbb{R}$, of the form

$$\mathrm{d}x_t = -\left[\nabla_x V(x) + \int_{\mathbb{R}} \nabla_x W(\theta, x_t - y)\mu_t(\mathrm{d}y)\right]\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}w_t, \tag{5.185}$$

$$\mu_t = \mathcal{L}(x_t). \tag{5.186}$$

where $\beta > 0$ is inverse temperature, $w = (w_t)_{t\geq 0}$ is a standard Brownian motion, $V : \mathbb{R} \to \mathbb{R}$ is the bistable confinement potential

$$V(x) = \left[\frac{1}{4}x^4 - \frac{1}{2}x^2\right] \tag{5.187}$$

and $W(\theta, \cdot) : \mathbb{R} \to \mathbb{R}$ is the Curie-Weiss (i.e., quadratic) interaction potential

$$W(\theta, x - y) = \frac{1}{2}\theta\left[x - y\right]^2 \tag{5.188}$$

As in the previous example, the parameter θ represents the interaction strength. The corresponding system of interacting particles for this system is given by

$$\mathrm{d}x_t^{i,N} = -\left[\nabla_x V(x_t^{i,N}) + \frac{1}{N}\sum_{j=1}^{N} \nabla_x W(\theta, x_t^{i,N} - x_t^{j,N})\right]\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}w_t^{i,N}. \tag{5.189}$$

This model is interesting as, while the IPS always has a unique invariant measure (see [374, Chapter 4]), for sufficiently high interaction strengths, the McKean-Vlasov SDE (5.185) - (5.186) admits multiple invariant measures (e.g., [139, 204, 443]). Thus, in particular, the assumptions of Theorems 5.3 - 5.4 are no longer satisfied. Nonetheless, it is still of interest to investigate (recursive) parameter estimation for this model numerically.

For the quadratic interaction potential, a one-parameter family of invariant densities for the McKean-Vlasov equation can be obtained as (e.g., [204])

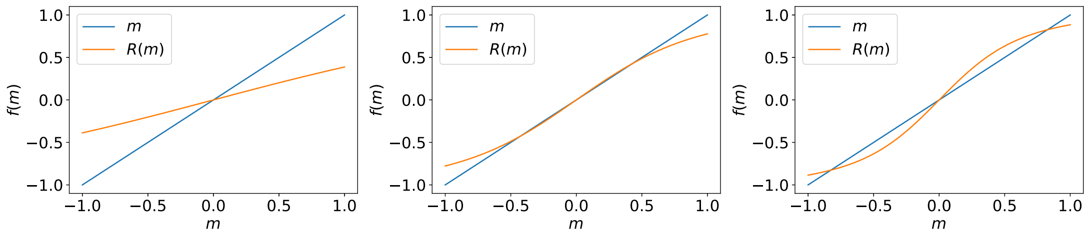
$$p_{\infty}(x; \theta, \beta, m) = \frac{e^{-\beta\left(\left[\frac{1}{4}x^4 - \frac{1}{2}x^2\right] + \theta\left(\frac{1}{2}x^2 - xm\right)\right)}}{\int_{\mathbb{R}} e^{-\beta\left(\left[\frac{1}{4}x^4 - \frac{1}{2}x^2\right] + \theta\left(\frac{1}{2}x^2 - xm\right)\right)}\mathrm{d}x} \tag{5.190}$$

These solutions are subject to the constraint that they provide the correct formula for the first moment, viz

$$m = \int_{\mathbb{R}} x p_{\infty}(x; \theta, \beta, m)\mathrm{d}x = R(m; \theta, \beta). \tag{5.191}$$

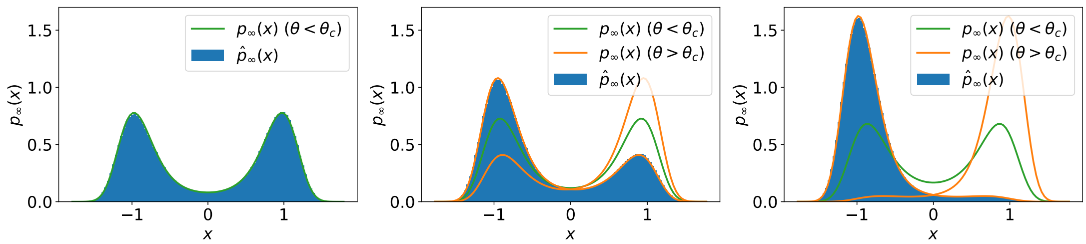
This is sometimes referred to as the self-consistency equation. Once a solution to this equation has been obtained, re-substituting into (5.190) yields the invariant density p_∞. Thus, the number of solutions to this equation determines the number of invariant measures of the McKean-Vlasov equation.

For the bistable potential (5.187), the self-consistency equation admits precisely one solution ($m = 0$) for sufficiently low interaction strengths. Meanwhile, there are an additional two solutions ($m = m_\pm$) above a critical interaction strength θ_c (e.g., [139, Theorem 3.3.2], [443, Theorem 4.1,4.2], [415]).[6] This is shown in Figure 5.11. It follows that there exists a unique invariant measure for $\theta < \theta_c$, and multiple invariant measures for $\theta > \theta_c$.[7] This is shown in Figure 5.12.



(a) $\theta = 0.05$. (b) $\theta = 0.15$. (c) $\theta = 0.25$.

Figure 5.11: Plots of $f(m) = m$ and $f(m) = R(m; \theta, \beta)$ for several values of θ, and fixed $\beta = 10$. The intersection points correspond to solutions of the self-consistency equation.



(a) $\theta = 0.05$. (b) $\theta = 0.15$. (c) $\theta = 0.25$.

Figure 5.12: Plots of the empirical invariant density $\hat{p}_\infty(x; \theta, \beta, m)$ (blue), and the true invariant density (densities) $p_\infty(x; \theta, \beta, m)$ (green, orange) for several values of θ, and fixed $\beta = 10$. We distinguish between the invariant density which exists for $\theta < \theta_c$ (green) and the two invariant densities which only exist for $\theta > \theta_c$ (orange).

[6] We note that one can compute θ_c analytically by solving the equation (see [183, 415])

$$\mathrm{Var}_{p_\infty}(x)|_{m=0} = \frac{1}{\beta\theta_c}. \tag{5.192}$$

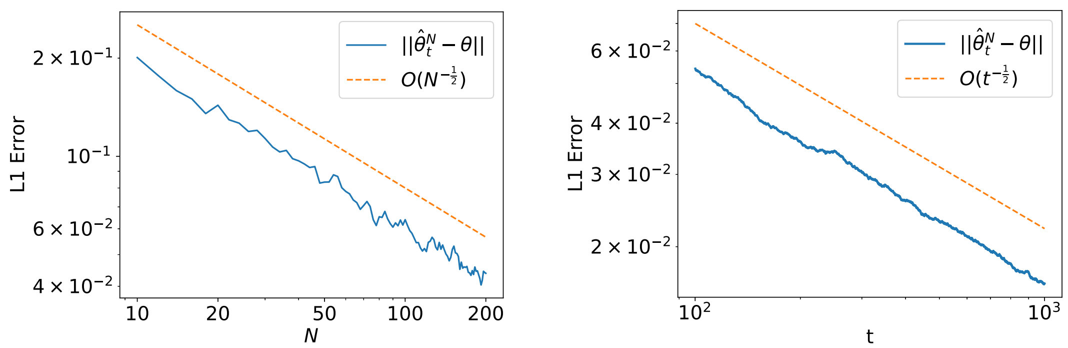
[7] This statement is typically given in terms of the temperature β^{-1}. In particular, a unique invariant measure exists at sufficiently high temperatures, while multiple invariant measures exist above a critical temperature β_c^{-1}.

5.5.2.1 Offline Parameter Estimation

We begin by considering offline parameter estimation of the unknown interaction parameter $\theta \in \mathbb{R}$. Since the model is linear in this parameter, in this case we can explicitly obtain

$$\hat{\theta}_t^N = \frac{-\sum_{i=1}^N \int_0^t \left(x_s^{i,N} - \frac{1}{N}\sum_{j=1}^N x_s^{j,N}\right) \mathrm{d}x_s^{i,N} - \int_0^t \sum_{i=1}^N \nabla_x V(x_s^{i,N}) \left(x_s^{i,N} - \frac{1}{N}\sum_{j=1}^N x_s^{j,N}\right) \mathrm{d}s}{\sum_{i=1}^N \int_0^t \left(x_s^{i,N} - \frac{1}{N}\sum_{j=1}^N x_s^{j,N}\right)^2 \mathrm{d}s}. \tag{5.193}$$

We plot the $\mathbb{L}^1$ error of the MLE, for several values of N and t, in Figure 5.13 below. The true value of the interaction parameter is $\theta_0 = 0.1$ (Figure 5.13a) and $\theta_0 = 0.2$ (Figure 5.13b). Our results illustrate a.s. convergence of the MLE in both large N and large T, and indicate that the MLE converges in N at a rate $\mathcal{O}((Nt)^{-\frac{1}{2}})$. This provides further evidence to suggest that the convergence rate recently obtained by Chen [110] also holds in the presence of a global confinement term.



(a) $\hat{\theta}^N_{1,t=5}$.

(b) $\hat{\theta}^{N=2}_{2,t}$.

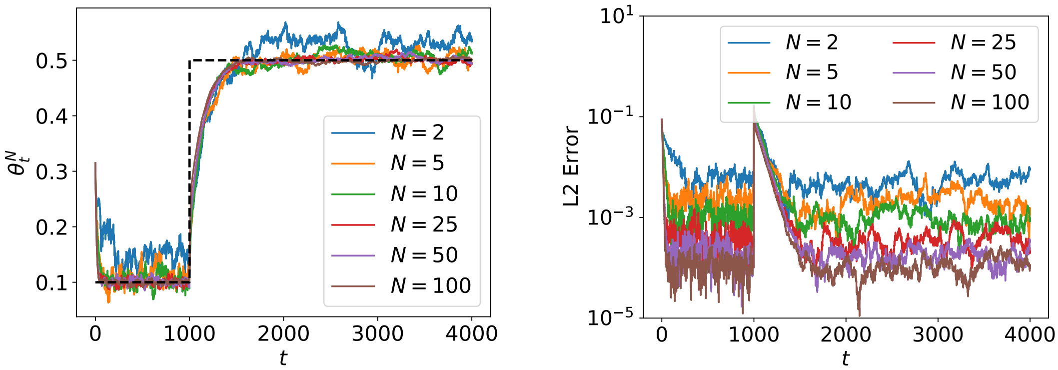
Figure 5.13: Log-log plot of the $\mathbb{L}^1$ error of the offline MLE for $t = 0.5$ and $N \in \{20, \ldots, 200\}$ (left hand panel), and for $t \in [100, 1000]$ and $N = 2$ (right hand panel).

5.5.2.2 Online Parameter Estimation

We now turn our attention to the online case. For this model, the recursive maximum likelihood estimator evolves according to

$$\mathrm{d}\theta_t^N = \frac{\beta}{2N}\gamma_t \sum_{i=1}^N \left[-(x_t^{i,N} - \bar{x}_t^N)\mathrm{d}x_t^{i,N} - (x_t^{i,N} - \bar{x}_t^N)(\nabla_x V(x_t^{i,N}) + \theta_t^N (x_t^{i,N} - \bar{x}_t^N))\mathrm{d}t\right]. \tag{5.194}$$

The performance of the online estimator is shown in Figure 5.14, in which we plot an illustrative sequence of online parameter estimates (Figure 5.14a), and the $\mathbb{L}^2$ error of the online parameter estimates, averaged over 500 independent runs (Figure 5.14b). In this case, we consider a time-varying parameter, which takes the values $\theta_0 = \{\theta_0^1, \theta_0^2\} = \{0.1, 0.5\}$.



(a) Sequence of online parameter estimates.

(b) $\mathbb{L}^2$ error.

Figure 5.14: Performance of the online MLE for $N \in \{2, 5, 10, 25, 50, 100\}$. The true value(s) of the time varying parameter are shown in black (dashed).

Let us make two brief observations. Firstly, and unsurprisingly, there is clear improvement in the performance of the online estimator as the number of particles increases. Secondly, as we have seen at various times throughout this thesis, the online estimator is able to track changes in the true parameter in real time. One interesting feature of this simulation is that initial true parameter value is below the critical threshold ($\theta_0^1 < \theta_c$), while the final true parameter value is above the critical threshold ($\theta_0^2 > \theta_c$). This is evident from the plots of self-consistency equation (Figure 5.11), or the plots of the invariant densities (Figure 5.12). Thus, the corresponding mean-field system exhibits a phase transition. We shall return to this point at the end of the next numerical example.

5.5.3 The Stochastic Kuramoto Model

We next consider a non-linear SDE on the one-dimensional torus $\mathbb{T}$, parametrised by $\theta \in \mathbb{R}$, of the form

$$\mathrm{d}x_t = \left[\theta \int_{\mathbb{R}} \sin(x_t - y)\mu_t(\mathrm{d}y)\right] \mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}w_t, \tag{5.195}$$

$$\mu_t = \mathcal{L}(x_t). \tag{5.196}$$

where $\beta > 0$ and $w = (w_t)_{t\geq 0}$ is a $\mathbb{T}$-valued Brownian motion, and all other terms are as defined previously. This equation represents the mean-field limit of the stochastic Kuramoto model (e.g., [45, 92]), or Kuramoto-Shinomoto-Sakaguchi model (e.g., [1, 264, 404]), which is given by

$$\mathrm{d}x_t^{i,N} = \frac{\theta}{N}\sum_{j=1}^{N}\sin(x_t^{i,N} - x_t^{j,N})\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}w_t^{i,N}, \tag{5.197}$$

This system of interacting particles models the synchronisation of noisy oscillators interacting through their phases. Thus, the parameter θ can be interpreted as the coupling strength between oscillators.

Aside from its wide applicability in various fields such as physics, chemistry, and biology (e.g., [1] and references therein), our interest in the stochastic Kuramoto model stems from the fact that its mean-field limit (5.195) - (5.196) exhibits a phase transition. Similarly to the previous example, when the coupling strength θ is smaller than a critical value θ_c the noise dominates, a uniform state is the only equilibrium, and the population always tends towards this incoherent state. On the other hand, when $\theta > \theta_c$ the coupling dominates, and a family of non-trivial coherent (or synchronised) equilibria exists, and the population tends to synchronise.

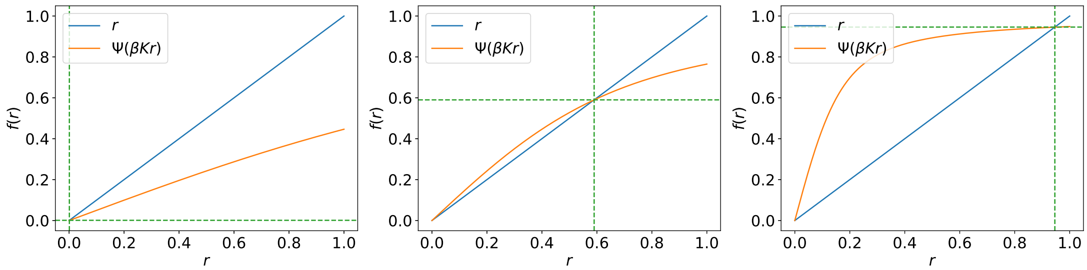
Let us now make this more precise. We first note that, if $p_\infty(x)$ is a stationary solution, then so too is $p_\infty(x + x_0)$, for arbitrary x_0. This is due to the invariance of (5.195) under rotations. In general, every stationary solutions can be written as $p_\infty(x + x_0)$ for some $x_0 \in [0, 2\pi)$, where (e.g., [45])

$$p_\infty(x;\theta,\beta,r) = \frac{e^{\beta\theta r\cos x}}{2\pi\int_{\mathbb{S}} e^{\beta\theta r\cos x}\mathrm{d}x}. \tag{5.198}$$

with r being a non-negative solution to the equation

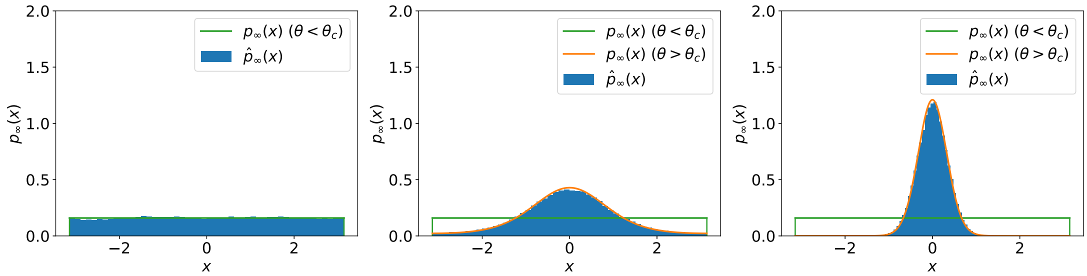
$$r := \Psi(\beta\theta r) = \frac{\int_{\mathbb{S}}\cos x\exp(\beta\theta r\cos x)\mathrm{d}x}{\int_{\mathbb{S}}\exp(\beta\theta r\cos x)\mathrm{d}x} \tag{5.199}$$

There are precisely one or two solutions to this equation. The first is the trivial solution $r = 0$, which holds for all values of θ. The second is a solution $0 < r < 1$, which only exists for $\theta > \theta_c = 2\beta^{-1}$. This is shown in Figure 5.15. In terms of stationary solutions, this means that for $\theta \leq \theta_c$ only the flat (incoherent) invariant measure $p_\infty(x) = \frac{1}{2\pi}$ is stationary. Meanwhile, for $\theta > \theta_c$, $\{p_\infty(\cdot + x_0)\}_{x_0\in\mathbb{S}}$ is a family of stationary solutions. These are the solutions which exhibit coherence or synchronisation. This is shown in Figure 5.16.



(a) $\theta = 0.10$. (b) $\theta = 0.25$. (c) $\theta = 1.00$.

Figure 5.15: Plots of $f(r) = r$ and $f(r) = \Psi(\beta\theta r)$ for several values of θ, and fixed $\beta = 10$. The intersection points correspond to solutions of the fixed point equation.



(a) $\theta = 0.10$. (b) $\theta = 0.25$. (c) $\theta = 1.00$.

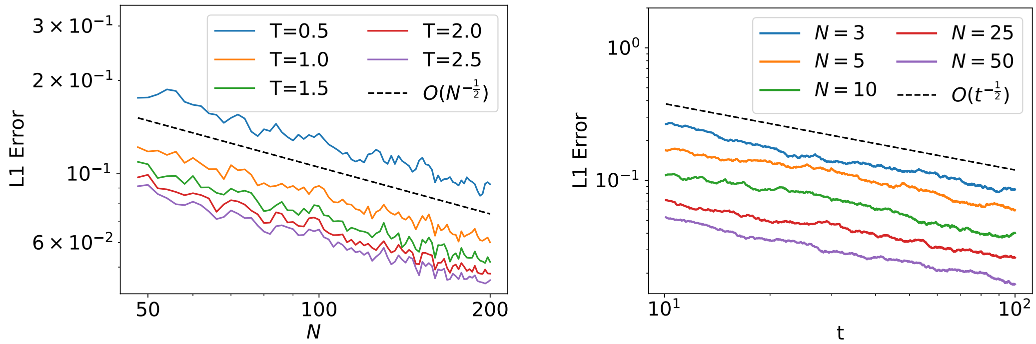
Figure 5.16: Plots of the empirical invariant density $\hat{p}_\infty(x;\theta,\beta,m)$ (blue), and the true invariant density (densities) $p_\infty(x;\theta,\beta,m)$ (green, orange) for several values of θ, and fixed $\beta = 10$. We distinguish between the invariant density which exists for $\theta < \theta_c$ (green) and the invariant densities which only exist for $\theta > \theta_c$ (orange).

5.5.3.1 Offline Parameter Estimation

We now turn our attention to offline estimation of the coupling strength $\theta \in \mathbb{R}$. Once more, since the parametrisation is linear in this parameter, we can obtain the maximum likelihood estimator in closed form as

$$\hat{\theta}_{\mathrm{MLE}} = \frac{-\sum_{i=1}^N \int_0^t \left(\frac{1}{N}\sum_{j=1}^N \sin(x_s^{i,N} - x_s^{j,N})\right) \mathrm{d}x_s^{i,N}}{\sum_{i=1}^N \int_0^t \left(\frac{1}{N}\sum_{j=1}^N \sin(x_s^{i,N} - x_s^{j,N})\right)^2 \mathrm{d}s}. \tag{5.200}$$

The performance of the MLE is illustrated in Figure 5.17. Our numerics indicate that the interaction parameter is also successfully estimated in this model; and tentatively suggest that the $\mathbb{L}^1$ error of the MLE once more converges in both N and t at a rate $\mathcal{O}(Nt)^{-\frac{1}{2}}$).



(a) $\mathbb{L}^1$ error as a function of N.

(b) $\mathbb{L}^1$ error as a function of t.

Figure 5.17: Log-log plot of the $\mathbb{L}^1$ error of the offline MLE (a) as a function of N, for several values of t, and (b) as a function of t, for several values of N (b).

5.5.3.2 Online Parameter Estimation

We now consider the online case. For this model, the recursive maximum likelihood estimator evolves according to the following SDE

$$\mathrm{d}\theta_t^N = \frac{\beta}{2N}\gamma_t \sum_{i=1}^N \left[-\frac{1}{N}\sum_{j=1}^N \sin(x_t^{i,N} - x_t^{j,N})\mathrm{d}x_t^{i,N} \right. \tag{5.201}$$

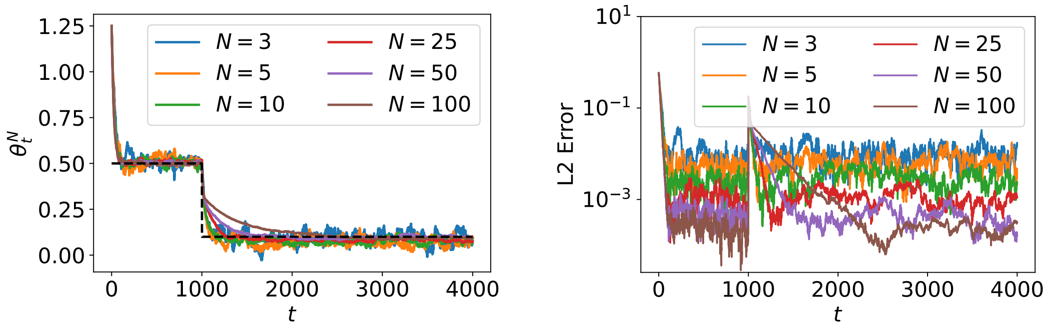
$$\left. -\frac{1}{N}\sum_{j=1}^N \sin(x_t^{i,N} - x_t^{j,N}) \left(\frac{\theta_t^N}{N}\sum_{j=1}^N \sin(x_t^{i,N} - x_t^{j,N}) \right) \mathrm{d}t \right]. \tag{5.202}$$

We illustrate the performance of the online estimator, for several values of N, in Figure 5.18. In this case, the results are averaged over 10 independent runs. We suppose, once more, that the true parameter is time-varying, with value

$$\theta_0 = \begin{cases} 0.5 & , \quad t \in [0, 1000] \\ 0.1 & , \quad t \in [1000, 4000] \end{cases} \tag{5.203}$$

Once again, the online estimator quickly reacts to the change point in the interaction parameter, closely tracking its true value. Interestingly, the estimator appears to react more slowly to changes in this parameter as N increases. On the other hand, the asymptotic performance of the estimator (in the sense of $\mathbb{L}^2$ error) improves as a function of N, as is to be expected on the basis of our theoretical results.

While the ability to infer the unknown interaction parameter (in an online fashion) is important in its own right, in this case it has some added significance. In particular, given the relationship between this parameter and the invariant measure of the stochastic Karumoto model, our results suggest that one could predict the asymptotic behaviour



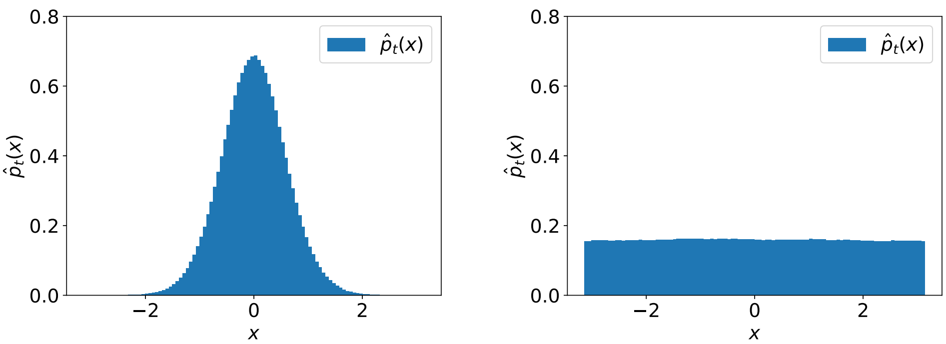
(a) Sequence of online parameter estimates.

(b) $\mathbb{L}^2$ error.

Figure 5.18: Performance of the online MLE for $N \in \{3, 5, 10, 25, 50, 100\}$. The true value(s) of the time varying parameter are shown in black (dashed).

(i.e., synchronisation or disorder) of this system in real time using the online parameter estimates. These remarks could equally apply to any mean-field model which admits different asymptotic regimes on the basis one or more unknown parameters (e.g., the model with bistable potential studied in the previous section).

Returning to the simulation at hand, we see in this case that the true parameter value at small times (and, thus, the parameter estimates at small times) are above the critical threshold ($0.5 = \theta_0^1 > \theta_c = 0.2$). Thus, at this stage, one would likely predict on the basis of the online parameter estimates that the population will synchronise. On the other hand, the true value of the parameter are large times (and, thus, the parameter estimates at large times) are below the critical threshold ($0.1 = \theta_0^2 < \theta_c = 0.2$). Thus, once more on the basis of the online parameter estimates, one would likely update the original prediction to expect convergence towards an incoherent state. These predictions are confirmed numerically in Figure 5.19.



(a) $t = 1000$.

(b) $t = 4000$.

Figure 5.19: Plots of the empirical density $\hat{p}_t(x)$ at two times.

Clearly, our confidence in predicting which asymptotic regime will occur (i.e., in predicting whether the true parameter is greater than or less than the critical parameter) is dependent on a number of factors. Principle among these are (i) the number of particles, (ii) the number of iterations, and (iii) the distance between the true parameter and the critical parameter. In particular, the larger the number of particles or number of iterations, the smaller the expected error of the parameter estimate, and thus the greater confidence in the prediction. Meanwhile, the larger the distance $|\theta_c - \theta_0|$, the greater the permissible error in the parameter estimate, and thus the greater the confidence in the prediction. While beyond the scope of this work, we believe that a precise probabilistic characterisation of these intuitions is a fruitful direction for future work.

5.5.4 Stochastic Opinion Dynamics

Finally, we consider a one-dimensional stochastic opinion dynamics model, parametrised by $\theta = (\theta_1, \theta_2)^T \in \mathbb{R}^2$, of the form

$$\mathrm{d}x_t = -\left[\int_{\mathbb{R}} \varphi_\theta(||x_t - y||)(x_t - y)\mu_t(\mathrm{d}y)\right]\mathrm{d}t + \sigma \mathrm{d}w_t, \tag{5.204}$$

where the *interaction kernel* $\varphi_\theta : \mathbb{R}_+ \to \mathbb{R}_+$ is defined according to

$$\varphi_\theta(r) = \begin{cases} \theta_1 \exp\left[-\dfrac{0.01}{1-(r-\theta_2)^2}\right] & , \quad r > 0 \\ 0 & , \quad r \leq 0 \end{cases} \tag{5.205}$$

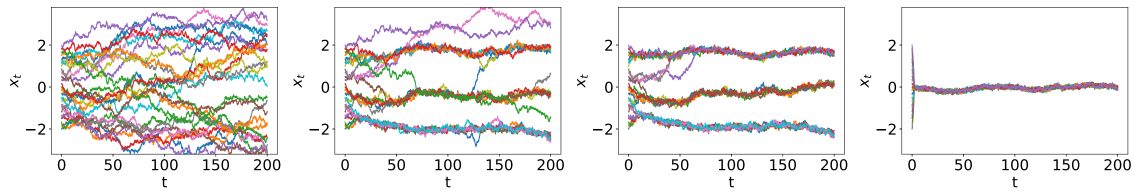
This function provides an approximation, infinitely differentiable on $\mathbb{R}_+$, to a scaled indicator function with magnitude θ_1 and support $[0, \theta_2 + 1]$. That is, writing $\tilde{\theta}_2 = \theta_2 + 1$,

$$\tilde{\varphi}_\theta(r) = \begin{cases} \theta_1 & , \quad 0 \leq r \leq \tilde{\theta}_2 \\ 0 & , \quad \tilde{\theta}_2 < r < \infty. \end{cases} \tag{5.206}$$

The system of interacting particles corresponding to (5.204) is given by

$$\mathrm{d}x_t^{i,N} = -\frac{1}{N}\sum_{j=1}^N \varphi_\theta(||x_t^{i,N} - x_t^{j,N}||)(x_t^{i,N} - x_t^{j,N})\mathrm{d}t + \sigma \mathrm{d}w_t. \tag{5.207}$$

Models of this form arise in various applications, from biology to the social sciences, in which φ_θ determines how the dynamics of one particle (e.g., the opinions of one person) may influence the dynamics of other particles (e.g., the opinions of other people). In this setting, one can interpret θ_1 as a scale parameter, which controls the strength of the attraction between particles, and $\tilde{\theta}_2$ as a range parameter, which determines the distance within which particles must be of one another to interact. For a more detailed account of such models, we refer to [54, 73, 120, 209, 257, 318, 357] and references therein. For deterministic models of this type, it is well known that, asymptotically, the particles merge into clusters, the number of which depends both on the interaction kernel (i.e., the range and strength of the interaction between particles) and the initialisation. In the stochastic setting, the random noise prohibits the formation of exact clusters; instead, the particles merge into metastable 'soft clusters' (see also [322]). This is shown in Figure 5.20.

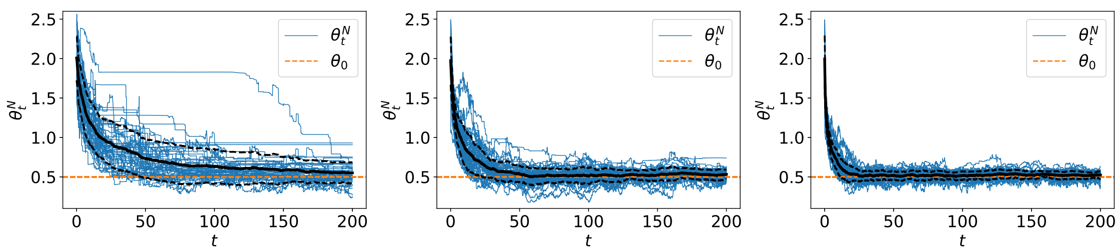


(a) $\theta_2 = 0.0$. (b) $\theta_2 = 0.3$. (c) $\theta_2 = 0.5$. (d) $\theta_2 = 1.0$.

Figure 5.20: Sample trajectories of the IPS for $\tilde{\theta}_2 = \{0.0, 0.3, 0.5, 1.0\}$.

We will focus on online parameter estimation in the case in which the scale parameter θ_1 is fixed, and the range parameter $\tilde{\theta}_2$ is to be estimated. We assume that $\theta_1 = \theta_1^* = 2$, and that $\tilde{\theta}_2^* = 0.5$. This corresponds to an interaction kernel with compact support on $[0, 0.5]$. The initial parameter estimates are generated according to $\tilde{\theta}_2^0 \sim \mathcal{U}([1.5, 2.5])$. Meanwhile, the initial particles are uniformly distributed over the interval $[-2, 2]$. Finally, we use constant learning rates with $\gamma_t = 0.002$.

The performance of the recursive MLE is illustrated in Figure 5.21, in which we plot the sequence of online parameter estimates for $\tilde{\theta}_2$. We provide results for several values of N, and for 50 different random initialisations. Encouragingly, (almost) all of the online parameter estimates converge to within a small neighbourhood of the true value of the parameter, suggesting that it is indeed possible to estimate the range of the interaction kernel in an online fashion. As in our previous simulations, the performance of the online estimator improves as the number of particles is increased.



(a) $N = 10$. (b) $N = 20$. (c) $N = 50$.

Figure 5.21: Sequence on online parameter estimates (blue) for the range parameter $\tilde{\theta}_2$, for 50 different random initialisation $\theta_2^0 \sim \mathcal{U}([1.5, 2.5])$, and $N = \{10, 20, 50\}$. We also plot the true parameter value (orange), the mean online parameter estimate plus/minus one standard deviation (black: solid, dashed).

We should remark that the performance of the online estimator is highly dependent on the initial conditions of the particles. This should not come as a surprise; indeed, if the distance between particles is greater than then support of the interaction kernel, then the interaction kernel (and its gradient) are identically zero, and thus so too are all of the terms in the parameter update equation. Thus, the value of the parameter estimate will remain unchanged. We see this phenomenon in Figure 5.21, particularly when there are fewer particles.

One can, of course, simulate even more interesting dynamics by considering sums of interaction kernels of the form (5.205), namely

$$\varphi_\theta(r) = \sum_{i=1}^{p} \varphi_\theta^i(r) := \begin{cases} \sum_{i=1}^{p} \theta_{1,i} \exp\left[-\frac{0.01}{1-(r-\theta_{2,i})^2}\right] & , \quad r \geq 0 \\ 0 & , \quad r < 0. \end{cases} \tag{5.208}$$

which provides an approximation for sums of indicator functions of the form (5.206). In Figure 5.22, we show illustrative results in the case $p = 2$. Once again, we suppose that the two scale parameters are fixed, and that the two range parameters are to be estimated. The true values of the scale parameters are $\theta_{1,1}^* = \theta_{1,2}^* = 1$, while the true values of the range parameters are $\tilde{\theta}_{2,1}^* = 0.6$, $\tilde{\theta}_{2,2}^* = 1.0$.[8] This specifies an interaction kernel with compact support on $[0, 1.0]$, with interactions twice as strong in $[0, 0.6]$, compared to in $[0.6, 1.0]$. In this simulation, the initial parameter estimates are fixed, and given by $\theta_{2,1}^0 = 0.1$ and $\theta_{2,2}^0 = 1.5$. Meanwhile, the initial particles are uniformly distributed over $[0, 5]$. Finally, we use a constant learning rate of $\gamma_{2,t} = 0.003$.

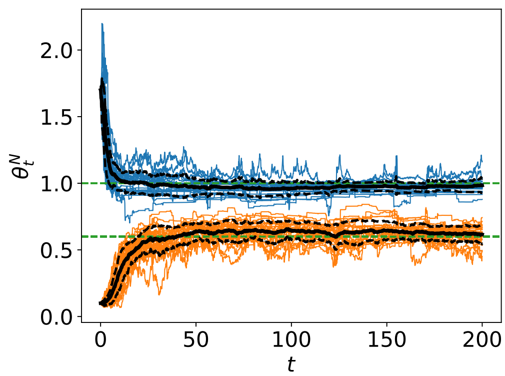


Figure 5.22: Sequence on online parameter estimates for the range parameters $\tilde{\theta}_{2,1}$ (blue) and $\tilde{\theta}_{2,2}$ (orange), for 25 different independent runs. We also plot the true parameter value (green), the mean online parameter estimate plus/minus one standard deviation (black: solid, dashed).

[8] Similarly to before, we have defined $\tilde{\theta}_{2,i} = \theta_{2,i} + 1$. Clearly, this re-scaled parameter represents the range of the i^{th} indicator function in the interaction kernel.

5.6 Conclusions

In this chapter, we have considered the problem of parameter estimation for a stochastic McKean-Vlasov equation and the associated system of weakly interacting particles. We established consistency and asymptotic normality of the offline MLE for the IPS as the number of particles $N \to \infty$, extending classical results in [242]. We also proposed an online estimator for the parameters of the McKean-Vlasov SDE, under various modelling assumptions. We demonstrated $\mathbb{L}^1$ convergence of this estimator to the stationary points of the asymptotic log-likelihood of the McKean-Vlasov SDE as $N \to \infty$ and $t \to \infty$ and, under additional assumptions, obtained an $\mathbb{L}^2$ convergence rate. Finally, we presented four numerical examples as a proof of concept: a toy model with linear dependence on a one-dimensional parameter in both the confinement potential and the interaction potential, a model with bistable potential and unknown interaction potential, the stochastic Kuramoto model, and a model commonly arising in the study of opinion dynamics.

Appendices

5.A Existing Results on the McKean-Vlasov SDE

Proposition 5.1 (Existence and Uniqueness, [81, Theorem 2.2.3]). *Assume that Conditions B.1(i) - B.2(i) hold. If $\mu_0^\theta \in \mathcal{P}_2(\mathbb{R}^d)$, the McKean-Vlasov SDE* (5.1) *has a unique strong solution $x^\theta = (x_t^\theta)_{t\geq 0}$ for all $t \geq 0$. In addition, the IPS* (5.4) *has a unique strong solution $x^{\theta,N} = (x_t^{\theta,N})_{t\geq 0}$ for all $t \geq 0$.*

Proposition 5.2 (Moment Bounds, [81, Lemma 2.3.1]). *Assume that Conditions B.1(i) - B.2(i) and D.1 hold. Then, for all $k \geq 0$, there exists $C_k > 0$ such that for all $\theta \in \mathbb{R}^p$, and for all $N \in \mathbb{N}$,*

$$\sup_{t\geq 0} \mathbb{E}||x_t^{\theta,i,N}||^k \leq C_k \left(\int_{\mathbb{R}^d} x^k \mu_0(\mathrm{d}x) + 1 \right) \tag{5.209}$$

$$\sup_{t\geq 0} \mathbb{E}||x_t^{\theta}||^k \leq C_k \left(\int_{\mathbb{R}^d} x^k \mu_0(\mathrm{d}x) + 1 \right) \tag{5.210}$$

Proposition 5.3 (Unique Invariant Measure of the MVSDE [81, Theorem 2.3.3]). *Assume that Conditions B.1 - B.2 hold, and $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$. Then the McKean-Vlasov SDE admits a unique equilibrium measure μ_∞ which is independent of the initial condition μ_0. Moreover, with $\lambda = \alpha - 2L_2$, the following contraction rate holds*

$$\mathbb{W}_2(\mu_t, \mu_\infty) \leq e^{-\lambda t} \mathbb{W}_2(\mu_0, \mu_\infty) \tag{5.211}$$

Proposition 5.4 (Unique Invariant Measure of the IPS, [414, Proposition A.4]). *Assume that Conditions B.1 - B.2 hold, and $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$. Then the IPS admits a unique equilibrium measure $\hat{\mu}_\infty^N$ which is independent of the initial condition $\hat{\mu}_0^N$. Moreover, with $\lambda = \alpha - 2L_2$, and writing $\hat{\mu}_t^{(k),N}$ for the law of a subset of $1 \leq k \leq N$ interacting particles, the following contraction rate holds*

$$\mathbb{W}_2(\hat{\mu}_t^{(k),N}, \hat{\mu}_\infty^{(k),N}) \leq e^{-\lambda t} \mathbb{W}_2(\mu_0^{\otimes k}, \hat{\mu}_\infty^{(k),N}). \tag{5.212}$$

Proposition 5.5 (Propagation of Chaos, [81, Lemma 2.4.1]). *Let $x^i = (x_t^i)_{t\geq 0}$ be N independent copies of the solutions of* (5.1) - (5.2) *driven by independent Brownian motions w^i. Assume that Conditions B.1 - B.2 and D.1 hold. Then there exist $0 < C < \infty$,* *independent of time, such that*

$$\sup_{t\geq 0}\mathbb{E}\left[||x_t^{i,N}-x_t^i||^2\right]\leq \frac{C}{N}. \tag{5.213}$$

Proposition 5.6 (A Law of Large Numbers, [117, Theorem 1.2], [367])**.** *Assume that Conditions B.1(i) - B.2(i) hold. If $(\mu_0^N)_{N\in\mathbb{N}}$ converge weakly to μ_0, then for all $g\in\mathcal{C}(\mathbb{R}^d)$ and for all $t\geq 0$, as $N\to\infty$,*

$$\lim_{N\to\infty}\left[\frac{1}{N}\sum_{i=1}^N g(x_t^{i,N})\right]\overset{\mathbb{P}}{=}\int_{\mathbb{R}^d} g(x)\mu_t(\mathrm{d}x). \tag{5.214}$$

5.B Proof of Lemma for Theorem 5.1 and Theorem 5.2

Lemma 5.2. *For all $T\geq 0$, for all $\theta\in\Theta\subseteq\mathbb{R}^p$,*

$$\lim_{N\to\infty}\sup_{0\leq t\leq T}\frac{1}{N}\sum_{i=1}^N\int_0^t\langle G(\theta,x_s^{i,N},\mu_s^N),\mathrm{d}w_s^i\rangle=0 \tag{5.215}$$

Proof. For ease of notation, let us define

$$M_t^N(\theta):=\frac{1}{N}\sum_{i=1}^N\int_0^t\langle G(\theta,x_s^{i,N},\mu_s^N),\mathrm{d}w_s^i\rangle. \tag{5.216}$$

Now, for all $N\in\mathbb{N}$, and for all $\theta\in\mathbb{R}^p$, $(M_t^N(\theta))_{t\geq 0}$ is a zero mean continuous square integrable martingale, with quadratic variation

$$\left[M^N(\theta)\right]_t=\frac{1}{N^2}\sum_{i=1}^N\int_0^t||G(\theta,x_s^{i,N},\mu_s^N)||^2\mathrm{d}s. \tag{5.217}$$

It follow, using the elementary fact that $\sup_x[f(x)-g(x)]\geq\sup_x f(x)-\sup_x g(x)$, and the martingale inequality [336, page 25], that

$$\mathbb{P}\bigg(\sup_{0\leq t\leq T}M_t^N(\theta)-\sup_{0\leq t\leq T}\frac{\alpha}{2}\left[M^N(\theta)\right]_t>\beta\bigg)\leq\mathbb{P}\bigg(\sup_{0\leq t\leq T}\left\{M_t^N(\theta)-\frac{\alpha}{2}\left[M^N(\theta)\right]_t\right\}>\beta\bigg) \tag{5.218}$$

$$<e^{-\alpha\beta}. \tag{5.219}$$

Thus, substituting (5.217) and using symmetry, we have that

$$\mathbb{P}\left(\sup_{0\leq t\leq T}\left|M_t^N(\theta)\right| > \beta + \frac{\alpha}{2N^2}\sum_{i=1}^N\int_0^T ||G(\theta, x_s^{i,N}, \mu_s^N)||^2\mathrm{d}s\right) < 2e^{-\alpha\beta}. \tag{5.220}$$

Let $\alpha = N^a$, $\beta = N^{-b}$, for some $0 < a < b < 1$. Then

$$\mathbb{P}\left(\sup_{0\leq t\leq T}|M_t^N(\theta)| > \frac{1}{N^b} + \frac{1}{2N^{1-a}}\frac{1}{N}\sum_{i=1}^N\int_0^T ||G(\theta, x_s^{i,N}, \mu_s^N)||^2\mathrm{d}s\right) < 2e^{-N^{a-b}}. \tag{5.221}$$

By a repeated application of Proposition 5.6 (the McKean-Vlasov Law of Large Numbers), we have that, as $N\rightarrow\infty$,

$$\frac{1}{N}\sum_{i=1}^N\int_0^T ||G(\theta, x_s^{i,N}, \mu_s^N)||^2\mathrm{d}s \xrightarrow{\mathbb{P}} \int_0^T\left[\int_{\mathbb{R}^d}||G(\theta, x, \mu_s)||^2\mu_s(\mathrm{d}x)\right]\mathrm{d}s. \tag{5.222}$$

By definition, Condition C.1(ii), and Proposition 5.2 the limiting function on the RHS is finite and non-random. Moreover, we have that $\sum_{N=1}^\infty e^{-N^{a-b}} < \infty$. The Borel-Cantelli Lemma thus implies

$$\lim_{N\rightarrow\infty}\sup_{0\leq t\leq T} M_t^N(\theta) = 0. \tag{5.223}$$

□

5.C Proof of Lemmas for Theorem 5.3

5.C.1 Additional Lemmas for Lemmas 5.4.A and 5.4.B

Lemma 5.3. *Assume that Conditions B.1 - B.2 and D.1 hold. Then, for all $k\in\mathbb{N}$, there exists a positive constant $K>0$ such that, for all $i = 1,\dots,N$, and for all $N\in\mathbb{N}$,*

$$\int_{\mathbb{R}^d}||x||^k\mu_\infty(\mathrm{d}x) \leq K, \tag{5.224}$$

$$\int_{(\mathbb{R}^d)^N}||x_i||^k\hat{\mu}_\infty^N(\mathrm{d}\hat{x}^N) \leq K. \tag{5.225}$$

Proof. By Proposition 5.3, the McKean-Vlasov SDE (5.1) - (5.2) admits a unique equilibrium measure μ_∞ which is independent of the initial condition μ_0. By the ergodic theorem (e.g., [395, Chapter X]), we thus have, for all $k \in \mathbb{N}$, that

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t ||x_s||^k \mathrm{d}s = \int_{\mathbb{R}^d} ||x||^k \mu_\infty(\mathrm{d}x) \ , \quad \text{a.s.} \tag{5.226}$$

Using Jensen's inequality and Proposition 5.2 (uniform moment bounds for the McKean-Vlasov SDE), we obtain uniform integrability of the family $\{\frac{1}{t}\int_0^t ||x_s||^k \mathrm{d}s\}_{t>0}$. In particular, for all $1 \leq k' < k$, for all $t > 0$, we have, for some $\varepsilon > 0$,

$$\mathbb{E}\left[\frac{1}{t}\int_0^t ||x_s||^{k'}\mathrm{d}s\right]^{1+\varepsilon} \leq \frac{1}{t}\int_0^t \mathbb{E}\left[||x_s||^{k'(1+\varepsilon)}\right]\mathrm{d}s \leq C_k\left(\int_{\mathbb{R}^d} x^k \mu_0(\mathrm{d}x)+1\right) < \infty \tag{5.227}$$

where in the final line we have used Condition D.1. It follows, taking expectations of (5.226), using uniform integrability in order to interchange the limit and the expectation, and once more making use of Proposition 5.2, that

$$\int_{\mathbb{R}^d} ||x||^k \mu_\infty(\mathrm{d}x) = \lim_{t\to\infty}\left[\frac{1}{t}\int_0^t \mathbb{E}\left[||x_s||^k\right]\mathrm{d}s\right] < \infty. \tag{5.228}$$

The proof of the bound for the IPS is identical, noting that all of the relevant results (Propositions 5.2 and 5.3) have analogues for for the IPS (Propositions 5.2 and 5.4). $\square$

Lemma 5.4. *Assume that Conditions B.1 - B.2 and D.1 hold. Then, for all $k \geq 1$, and for all $t \geq 0$, there exists $K > 0$ such that*

$$\mathbb{E}\left[\sup_{0\leq s\leq t}||x_s||^k\right] \leq Kt^{\frac{1}{2}}, \tag{5.229}$$

$$\mathbb{E}\left[\sup_{0\leq s\leq t}||x_s^{i,N}||^k\right] \leq Kt^{\frac{1}{2}} \ , \quad \forall i = 1,\dots,N. \tag{5.230}$$

Proof. We will prove the first claim (the proof of the second being essentially identical). By Itô's Lemma, we have

$$\begin{aligned}||x_t||^{2k} = ||x_0||^{2k} &+ \int_0^t 2k||x_s||^{2k-2}\langle x_s, B(\theta,x_s,\mu_s)\rangle\mathrm{d}s \\ &+ \int_0^t k||x_s||^{2k-2}\mathrm{Tr}[I_d + (k-2)[x_s^i x_s^j]_{i,j=1}^d ||x_s||^{-2}]\mathrm{d}s \\ &+ \int_0^t 2k||x_s||^{2k-2}\langle x_s, \mathrm{d}w_s\rangle\end{aligned} \tag{5.231}$$

It follows, taking the supremum and taking expectations, that

$$\mathbb{E}\left[\sup_{0\leq s\leq t}||x_t||^{2k}\right] \leq \mathbb{E}\left[||x_0||^{2k}\right] + \underbrace{2k\int_0^t \mathbb{E}\left[\left|||x_s||^{2k-2}\langle x_s, B(\theta, x_s, \mu_s)\rangle\right|\right]\mathrm{d}s}_{\Pi_t^1} \tag{5.232}$$
$$+ \underbrace{k\int_0^t \mathbb{E}\left[\left|||x_s||^{2k-2}\mathrm{Tr}[I_d + (k-2)[x_s^i x_s^j]_{i,j=1}^d ||x_s||^{-2}]\right|\right]\mathrm{d}s}_{\Pi_t^2}$$
$$+ \underbrace{2k\mathbb{E}\left[\sup_{0\leq s\leq t}\int_0^t ||x_s||^{2k-2}\langle x_s, \mathrm{d}w_s\rangle\right]}_{\Pi_t^3}$$

We begin by bounding the first term. First note that, by Conditions B.1 - B.2, there exists a positive constant C such that

$$\langle x_s, B(\theta, x_s, \mu_s)\rangle \leq -(\alpha - L_2)||x_s||^2 + C||x_s|| + L_1||x_s||\mathbb{E}\left[||x_s||\right]$$

It then follows that

$$\begin{aligned}
\Pi_t^1 &\leq K\int_0^t \mathbb{E}\left[||x_s||^{2k}\right] + \mathbb{E}\left[||x_s||^{2k-1}\right] + \mathbb{E}\left[||x_s||^{2k-1}\right]\mathbb{E}\left[||x_s||\right]\mathrm{d}s \\
&\leq K\int_0^t \mathbb{E}\left[||x_s||^{2k}\right] + \mathbb{E}\left[||x_s||^{2k}\right]^{\frac{2k-1}{2k}} + \mathbb{E}\left[||x_s||^{2k}\right]^{\frac{2k-1}{2k}}\mathbb{E}\left[||x_s||^2\right]^{\frac{1}{2}}\mathrm{d}s \\
&\leq Kt\left[1 + \int_{\mathbb{R}^d} x^{2k}\mu_0(\mathrm{d}x) + \left[\int_{\mathbb{R}^d} x^{2k}\mu_0(\mathrm{d}x)\right]^{\frac{2k-1}{2k}}\left[1 + \left(\int_{\mathbb{R}^d} x^2\mu_0(\mathrm{d}x)\right)^{\frac{1}{2}}\right]\right] \leq Kt
\end{aligned} \tag{5.233}$$

where in the penultimate line we have used Hölder's inequality, and in the final line we have used Proposition 5.2 (moment bounds for the McKean-Vlasov SDE) and Condition D.1. Similarly, for the second term in (5.232), we have

$$\Pi_t^2 \leq K\int_0^t \mathbb{E}\left[||x_s||^{2k}\right]\mathrm{d}s \leq Kt\left[1 + \int_{\mathbb{R}^d} x^{2k}\mu_0(\mathrm{d}x)\right] \leq Kt. \tag{5.234}$$

It remains to bound the final term in (5.232). For this term, we use the Burkholder-Davis-

Gundy inequality, Proposition 5.2, and Condition D.1 to obtain

$$\Pi_t^3 \leq K\mathbb{E}\left[\int_0^t ||x_s||^{4k-4} x_s^T x_s \mathrm{d}s\right]^{\frac{1}{2}} \leq K\mathbb{E}\left[\int_0^t ||x_s||^{4k-2} \mathrm{d}s\right]^{\frac{1}{2}} \leq Kt^{\frac{1}{2}}. \tag{5.235}$$

Combining equations (5.232), (5.233), (5.234), and (5.235), and using the Hölder inequality, we conclude that, for all $t > 0$, there exists a positive constant K such that

$$\mathbb{E}\left[\sup_{0\leq s\leq t} ||x_s||^k\right] \leq \mathbb{E}\left[\sup_{0\leq s\leq t} ||x_s||^{2k}\right]^{\frac{1}{2}} \leq Kt^{\frac{1}{2}}. \tag{5.236}$$

□

Lemma 5.5. *Assume that Conditions B.1 - B.2 and D.1 hold. Suppose that, for all $\theta \in \mathbb{R}^p$, $f(\theta, \cdot) : \mathbb{R}^d \to \mathbb{R}$ is locally Lipschitz, and satisfies a polynomial growth condition, viz*

$$||f(\theta, x) - f(\theta, y)|| \leq K||x - y|| \left[1 + ||x||^q + ||y||^q\right]. \tag{5.237}$$

Then, for all $\theta \in \mathbb{R}^p$, $x, y \in \mathbb{R}^d$, $t \geq 0$, there exist positive constants $q, K > 0$ such that

$$\left|\mathbb{E}_x\left[f(\theta, x_t)\right] - \int_{\mathbb{R}^d} f(\theta, z)\mu_\infty(\mathrm{d}z)\right| \leq K[1 + ||x||^q]e^{-\lambda t}. \tag{5.238}$$

$$\left|\mathbb{E}_x\left[f(\theta, x_t)\right] - \mathbb{E}_y\left[f(\theta, x_t)\right]\right| \leq K[1 + ||x||^q + ||y||^q]e^{-\lambda t}. \tag{5.239}$$

Alternatively, suppose that, for all $\theta \in \mathbb{R}^p$, $f(\theta, \cdot) : (\mathbb{R}^d)^N \to \mathbb{R}$ is locally Lipschitz and satisfies a polynomial growth condition in the sense that

$$\begin{aligned}\left|f(\theta, \hat{x}^N) - f(\theta, \hat{y}^N)\right| &\leq K\left[1 + ||x^{i,N}||^q + ||y^{i,N}||^q + \frac{1}{N}\sum_{j=1}^N ||x^{j,N}||^q + \frac{1}{N}\sum_{j=1}^N ||y^{j,N}||^q\right] \\ &\quad \cdot \left[||x^{i,N} - y^{i,N}|| + \left(\frac{1}{N}\sum_{j=1}^N ||x^{j,N} - y^{j,N}||^2\right)^{\frac{1}{2}}\right]\end{aligned} \tag{5.240}$$

where $\hat{x}^N = (x^{1,N}, \ldots, x^{N,N}) \in (\mathbb{R}^d)^N$. Then, for all $i = 1, \ldots, N$, and for all $\theta \in \mathbb{R}^p$, there exist positive constants $q, K > 0$ such that

$$\left|\mathbb{E}_{\hat{x}^N}[f(\theta, \hat{x}_t^N)] - \int_{(\mathbb{R}^d)^N} f(\theta, \hat{z}^N)\hat{\mu}_\infty^N(\mathrm{d}\hat{z}^N)\right| \leq K\left[1 + ||x^{i,N}||^q + \frac{1}{N}\sum_{j=1}^N ||x^{j,N}||^q\right]e^{-\lambda t} \tag{5.241}$$

$$\left|\mathbb{E}_{\hat{x}^N}[f(\theta,\hat{x}_t^N)] - \mathbb{E}_{\hat{y}^N}[f(\theta,\hat{x}_t^N)]\right| \leq K\left[1 + ||x^{i,N}||^q + ||y^{i,N}||^q + \frac{1}{N}\sum_{j=1}^N (||x^{j,N}||^q + ||y^{j,N}||^q)\right]e^{-\lambda t} \tag{5.242}$$

for all $\hat{x}^N, \hat{y}^N \in (\mathbb{R}^d)^N$, and for all $t \geq 0$.

Proof. We will focus on the first statement of the first part of the Lemma. Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$, and $\pi \in \Pi(\mu,\nu)$. Then, using the Hölder inequality and the local Lipschitz assumption, it follows that

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |f(\theta,y) - f(\theta,z)|\,\pi(\mathrm{d}y,\mathrm{d}z) \tag{5.243}$$

$$\leq K\bigg[\int_{\mathbb{R}^d\times\mathbb{R}^d} ||y-z||^2\pi(\mathrm{d}y,\mathrm{d}z)\bigg]^{\frac{1}{2}}\bigg[1 + \bigg[\int_{\mathbb{R}^d\times\mathbb{R}^d} ||y||^{2q}\pi(\mathrm{d}y,\mathrm{d}z)\bigg]^{\frac{1}{2}} + \bigg[\int_{\mathbb{R}^d\times\mathbb{R}^d} ||z||^{2q}\pi(\mathrm{d}y,\mathrm{d}z)\bigg]^{\frac{1}{2}}\bigg] \tag{5.244}$$

$$= K\bigg[\int_{\mathbb{R}^d\times\mathbb{R}^d} ||y-z||^2\pi(\mathrm{d}y,\mathrm{d}z)\bigg]^{\frac{1}{2}}\bigg[1 + \bigg[\int_{\mathbb{R}^d} ||y||^{2q}\mu(\mathrm{d}y)\bigg]^{\frac{1}{2}} + \bigg[\int_{\mathbb{R}^d} ||z||^{2q}\mu(\mathrm{d}z)\bigg]^{\frac{1}{2}}\bigg] \tag{5.245}$$

Let x_t be a solution of the McKean-Vlasov SDE starting from $x \in \mathbb{R}^d$. Let μ_t^x denote the law of x_t, and let μ_∞ denote the invariant measure of the McKean-Vlasov SDE. Moreover, let $\pi_t^{x,\infty}$ denote an arbitrary coupling of μ_t^x and μ_∞. It then follows straightforwardly from the previous inequality that

$$\left|\mathbb{E}_x\left[f(\theta,x_t)\right] - \int_{\mathbb{R}^d} f(\theta,z)\mu_\infty(\mathrm{d}z)\right| = \left|\int_{\mathbb{R}^d} f(\theta,y)\mu_t^x(\mathrm{d}y) - \int_{\mathbb{R}^d} f(\theta,z)\mu_\infty(\mathrm{d}z)\right| \tag{5.246}$$

$$\leq \int_{\mathbb{R}^d\times\mathbb{R}^d} |f(\theta,y) - f(\theta,z)|\pi_t^{x,\infty}(\mathrm{d}y,\mathrm{d}z) \tag{5.247}$$

$$\leq K\bigg[\int_{\mathbb{R}^d\times\mathbb{R}^d} ||y-z||^2\pi_t^{x,\infty}(\mathrm{d}y,\mathrm{d}z)\bigg]^{\frac{1}{2}} \tag{5.248}$$

$$\cdot\bigg[1 + \bigg[\int_{\mathbb{R}^d} ||y||^{2q}\mu_t^x(\mathrm{d}y)\bigg]^{\frac{1}{2}} + \bigg[\int_{\mathbb{R}^d} ||z||^{2q}\mu_\infty(\mathrm{d}z)\bigg]^{\frac{1}{2}}\bigg]$$

Finally, using the fact that the chosen coupling was arbitrary, and using Lemma 5.3 (the bounded moments of the invariant measure of the McKean-Vlasov SDE), Proposition 5.2 (the moment bounds for the McKean-Vlasov SDE), Proposition 5.3 (exponential contractivity of the McKean-Vlasov SDE), and Condition D.1, the previous inequality implies

$$\left|\mathbb{E}_x\left[f(\theta,x_t)\right]-\int_{\mathbb{R}^d} f(\theta,z)\mu_\infty(\mathrm{d}z)\right| \leq K\mathbb{W}_2(\mu_t^x,\mu_\infty)\left[1+\left[\int_{\mathbb{R}^d}||y||^{2q}\mu_t^x(\mathrm{d}y)\right]^{\frac{1}{2}}+K'\right] \tag{5.249}$$

$$\leq K\mathbb{W}_2(\mu_0^x,\mu_\infty)\big[1+||x||^q\big]e^{-\lambda t} \tag{5.250}$$

$$\leq K\big[1+||x||^q\big]e^{-\lambda t} \tag{5.251}$$

This completes the proof of the first statement of the first part of the Lemma. The proof of the second statement is essentially identical, this time considering an arbitrary coupling of μ_t^x and μ_t^y, and making use of the bound $\mathbb{W}_2(\mu_t^x,\mu_t^y)\leq e^{-\lambda t}\mathbb{W}_2(\mu_0^x,\mu_0^y)$. Finally, the proof of the second part of the Lemma follows closely the previous proof, now using the statements in Lemma 5.3, Proposition 5.2, and Proposition 5.4 that are relevant to the IPS. □

Lemma 5.6. *Assume that Condition C.1 holds. Then, for $k=0,1,2,3$, there exist constants $q,K<\infty$, such that $\nabla_\theta^k L(\theta,x,\mu)$, satisfy the following polynomial growth conditions:*

$$||\nabla_\theta^k L(\theta,x,\mu)||\leq K\left[1+||x||^q+\mu(||\cdot||^q)\right]. \tag{5.252}$$

Proof. We first observe that, by Condition C.1(ii), there exist constants $q_k,K_k<\infty$ such that $\nabla_\theta^k b(\theta,x)\leq K_k(1+||x||^{q_k})$ and $\nabla_\theta^k\phi(\theta,x,y)\leq K_k(1+||x||^{q_k}+||y||^{q_k})$. It follows from the definition of $B(\theta,x,\mu)$, c.f. (5.3), that

$$\nabla_\theta^k B(\theta,x,\mu)=\nabla_\theta^k b(\theta,x,\mu)+\int_{\mathbb{R}^d}\nabla_\theta^k\phi(\theta,x,y)\mu(\mathrm{d}y) \tag{5.253}$$

$$=K_k(1+||x||^{q_k})+K_k\int_{\mathbb{R}^d}(1+||x||^{q_k}+||y||^{q_k})\mu(\mathrm{d}y) \tag{5.254}$$

$$\leq K_k\left[1+||x||^{q_k}+\mu(||\cdot||^{q_k})\right] \tag{5.255}$$

where we allow the values of q_k,K_k to vary from line to line. Thus, from the definition of $G(\theta,x,\mu)$, c.f. (5.8), we have that

$$||\nabla_\theta^k G(\theta,x,\mu)||=||\nabla_\theta^k B(\theta,x,\mu)-\nabla_\theta^k B(\theta_0,x,\mu)||\leq K_k\left[1+||x||^{q_k}+\mu(||\cdot||^{q_k})\right]. \tag{5.256}$$

It now follows, recalling the definition of $L(\theta,x,\mu)$, c.f. (5.9), that

$$||L(\theta,x,\mu)||=\frac{1}{2}||G(\theta,x,\mu)||^2\leq K_0^2\left[1+||x||^{q_0}+\mu(||\cdot||^{q_0})\right]^2\leq K\left[1+||x||^q+\mu(||\cdot||^q)\right] \tag{5.257}$$

where in the final inequality we have set $K = 3K_0^2$ and $q = 2q_0$, after applying Hölder's inequality. The bounds for $\nabla_\theta^k L(\theta, x, \mu)$, for $k = 1, 2, 3$, are obtained in an almost identical fashion. $\square$

Remark. This result implies, substituting $x = x^{i,N}$ and $\mu = \mu^N$, and recalling that $\hat{L}^{i,N}(\theta, \hat{x}^N) := L(\theta, x^{i,N}, \hat{\mu}^N)$, that for all $i = 1, \ldots, N$, $N \in \mathbb{N}$, we have

$$||\nabla_\theta^k \hat{L}^{i,N}(\theta, \hat{\mu}^N)|| \leq K \left[1 + ||x^{i,N}||^q + \frac{1}{N}\sum_{j=1}^N ||x^{j,N}||^q\right] \tag{5.258}$$

Lemma 5.7. *Assume that Condition C.1 holds. Then, for $k = 0, 1, 2, 3$, there exist constants $q, K < \infty$ such that $\nabla_\theta^k L(\theta, x, \mu)$ satisfy*

$$||\nabla_\theta^k L(\theta, x, \mu) - \nabla_\theta^k L(\theta, x', \mu')|| \leq K \left[||x - x'|| + \mathbb{W}_2(\mu, \mu')\right] \tag{5.259}$$

$$\cdot \left[1 + ||x||^q + ||x'||^q + \mu(||\cdot||^q) + \mu'(||\cdot||^q)\right] \tag{5.260}$$

Proof. We begin by recalling that, from Condition C.1(ii), there exist constants $q, K < \infty$ such that $\nabla_\theta^k \phi(\theta, x, y) \leq K[||x - x'|| + ||y - y'||][1 + ||x||^q + ||x'||^q + ||y||^q + ||y'||^q]$. It follows, letting $\pi \in \Pi(\mu, \mu')$ and using the Hölder inequality, that

$$\left|\left|\int_{\mathbb{R}^d} \nabla_\theta^k \phi(\theta, x, y)\mu(\mathrm{d}y) - \int_{\mathbb{R}^d} \nabla_\theta^k \phi(\theta, x, y')\mu'(\mathrm{d}y')\right|\right| \tag{5.261}$$

$$\leq K\left[\int_{\mathbb{R}^d\times\mathbb{R}^d} ||y - y'|| \left[1 + ||y||^q + ||y'||^q\right] \pi(\mathrm{d}y, \mathrm{d}y')\right] \tag{5.262}$$

$$\leq K\left[\int_{\mathbb{R}^d\times\mathbb{R}^d} ||y - y'||^2 \pi(\mathrm{d}y, \mathrm{d}y')\right]^{\frac{1}{2}} \left[1 + \left[\int_{\mathbb{R}^d} ||y||^{2q}\mu(\mathrm{d}y)\right]^{\frac{1}{2}} + \left[\int_{\mathbb{R}^d} ||y'||^{2q}\mu(\mathrm{d}y')\right]^{\frac{1}{2}}\right] \tag{5.263}$$

$$\leq K\mathbb{W}_2(\mu, \mu') \left[1 + \mu(||\cdot||^{2q})^{\frac{1}{2}} + \mu'(||\cdot||^{2q})^{\frac{1}{2}}\right]. \tag{5.264}$$

We then have, via the triangle inequality, the bound (5.261), and another application of both Condition C.1(ii) and the Hölder inequality, that

$$\left|\left|\int_{\mathbb{R}^d} \nabla_\theta^k \phi(\theta, x, y)\mu(\mathrm{d}y) - \int_{\mathbb{R}^d} \nabla_\theta^k \phi(\theta, x', y')\mu'(\mathrm{d}y')\right|\right| \tag{5.265}$$

$$\leq \left|\left|\int_{\mathbb{R}^d} \nabla_\theta^k \phi(\theta, x, y)\mu(\mathrm{d}y) - \int_{\mathbb{R}^d} \nabla_\theta^k \phi(\theta, x, y')\mu'(\mathrm{d}y')\right|\right| + \int_{\mathbb{R}^d} \left|\left|\nabla_\theta^k \phi(\theta, x, y') - \nabla_\theta^k \phi(\theta, x', y')\right|\right| \mu'(\mathrm{d}y') \tag{5.266}$$

$$\leq K\mathbb{W}_2(\mu, \mu') \left[1 + \mu(||\cdot||^{2q})^{\frac{1}{2}} + \mu'(||\cdot||^{2q})^{\frac{1}{2}}\right] + K||x - x'|| \left[1 + \mu'(||\cdot||^{2q})^{\frac{1}{2}}\right] \tag{5.267}$$

$$\leq K\left[||x-x'||+\mathbb{W}_2(\mu,\mu')\right]\left[1+\mu(||\cdot||^{2q})^{\frac{1}{2}}+\mu'(||\cdot||^{2q})^{\frac{1}{2}}\right]. \tag{5.268}$$

Thus, recalling the definition of $B(\theta,x,\mu)$, c.f. (5.3), and once more making use Condition C.1(ii), we obtain

$$\begin{aligned} ||\nabla_\theta^k B(\theta,x,\mu)-\nabla_\theta^k B(\theta,x',\mu')|| &\leq \left|\left|\nabla_\theta^k b(\theta,x)-\nabla_\theta^k b(\theta,x')\right|\right| &(5.269)\\ &+\left|\left|\int_{\mathbb{R}^d}\nabla_\theta^k\phi(\theta,x,y)\mu(\mathrm{d}y)-\int_{\mathbb{R}^d}\nabla_\theta^k\phi(\theta,x',y')\mu'(\mathrm{d}y')\right|\right| &(5.270)\\ &\leq K\left[||x-x'||+\mathbb{W}_2(\mu,\mu')\right] &(5.271)\\ &\quad\cdot\left[1+||x||^q+||x'||^q+\mu(||\cdot||^{2q})^{\frac{1}{2}}+\mu'(||\cdot||^{2q})^{\frac{1}{2}}\right]. \end{aligned}$$

From the definition of $G(\theta,x,\mu)$, c.f. (5.8), we trivially then have

$$\begin{aligned} ||\nabla_\theta^k G(\theta,x,\mu)-\nabla_\theta^k G(\theta,x',\mu')|| &\leq K\left[||x-x'||+\mathbb{W}_2(\mu,\mu')\right] &(5.272)\\ &\quad\cdot\left[1+||x||^q+||x'||^q+\mu(||\cdot||^{2q})^{\frac{1}{2}}+\mu'(||\cdot||^{2q})^{\frac{1}{2}}\right]. \end{aligned}$$

Finally, recalling the definition of $L(\theta,x,\mu)$, c.f. (5.9), and combining (5.256) and (5.272), we obtain

$$\begin{aligned} ||L(\theta,x,\mu)-L(\theta,x',\mu')|| &= \frac{1}{2}\left|\left|G^T(\theta,x,\mu)G(\theta,x,\mu)-G^T(\theta,x',\mu')G(\theta,x',\mu')\right|\right| &(5.273)\\ &\leq \frac{1}{2}||G(\theta,x,\mu)-G(\theta,x',\mu')||\,||G(\theta,x,\mu)+G(\theta,x',\mu')|| &(5.274)\\ &\leq K\left[||x-x'||+\mathbb{W}_2(\mu,\mu')\right] &(5.275)\\ &\quad\cdot\left[1+||x||^q+||x'||^q+\mu(||\cdot||^{2q})^{\frac{1}{2}}+\mu'(||\cdot||^{2q})^{\frac{1}{2}}\right]^2\\ &\leq K\left[||x-x'||+\mathbb{W}_2(\mu,\mu')\right] &(5.276)\\ &\quad\cdot\left[1+||x||^q+||x'||^q+\mu(||\cdot||^{q})+\mu'(||\cdot||^{q})\right]. \end{aligned}$$

where in the final line we have replaced the unimportant constant $q\to 2q$, after applying the Hölder inequality. The bounds for $\nabla_\theta^k L(\theta,x,\mu)$, $k=0,1,2$, follow analogously. □

Remark. In the case that substituting $x=x^{i,N}$, $x'=y^{i,N}$, $\mu=\mu_x^N$ and $\mu'=\mu_y^N$, and recalling that $\hat{L}^{i,N}(\theta,\hat{x}^N):=L(\theta,x^{i,N},\hat{\mu}^N)$, that for all $i=1,\dots,N$, $N\in\mathbb{N}$, it holds that

$$\left|\nabla_\theta^k L(\theta,\hat{x}^N)-\nabla_\theta^k L(\theta,\hat{y}^N)\right|\leq K\left[||y^{i,N}-z^{i,N}||+\left(\frac{1}{N}\sum_{j=1}^N||y^{j,N}-z^{j,N}||^2\right)^{\frac{1}{2}}\right] \tag{5.277}$$

$$\cdot \Big[1 + ||x^{i,N}||^q + ||y^{i,N}||^q + \frac{1}{N}\sum_{j=1}^N ||x^{j,N}||^q + \frac{1}{N}\sum_{j=1}^N ||y^{j,N}||^q\Big] \tag{5.278}$$

$$\tag{5.279}$$

Lemma 5.8. *Assume that Conditions B.1 - B.2, C.1 and D.1 hold. Then, for $k = 0, 1, 2$, there exist $K, K' > 0$ such that, for all $\theta \in \mathbb{R}^p$, $||\nabla_\theta^k \tilde{\mathcal{L}}(\theta)|| \leq K$ and $||\nabla_\theta^k \tilde{\mathcal{L}}^{i,N}(\theta)|| \leq K'$.*

Proof. Using the definition of $\nabla_\theta^k \tilde{\mathcal{L}}(\theta)$ (Lemma 5.4.A), the polynomial growth property of $\nabla_\theta^k L(\theta, x, \mu)$ (Lemma 5.6), and the finite moments of the invariant measure of the McKean-Vlasov SDE (Lemma 5.3), we have that

$$||\nabla_\theta^k \tilde{\mathcal{L}}(\theta)|| \leq \int_{\mathbb{R}^d} ||\nabla_\theta^k L(\theta, x, \mu_\infty)|| \mu_\infty(\mathrm{d}x) \tag{5.280}$$

$$\leq K \int_{\mathbb{R}^d} \big[1 + ||x||^q + [\int_{\mathbb{R}^d} ||y||^q \mu_\infty(\mathrm{d}y)\big] \mu_\infty(\mathrm{d}y) \tag{5.281}$$

$$\leq K \int_{\mathbb{R}^d} (1 + ||x||^q)\mu_\infty(\mathrm{d}x) \leq K. \tag{5.282}$$

The bound for $\nabla_\theta^k \tilde{\mathcal{L}}^{i,N}(\theta)$ follows identically, this time using the defintion of $\nabla_\theta^k \tilde{\mathcal{L}}^{i,N}(\theta)$ (Lemma 5.4.B), and the finite moments of the invariant measure of the IPS (Lemma 5.3). □

5.C.2 Additional Lemmas for Lemma 5.4.C

Lemma 5.9. *Assume that Conditions B.1 - B.2 and D.1 hold. For all Lipschitz functions φ, there exists $K > 0$ such that, for all $t \geq 0$, for all $N \in \mathbb{N}$,*

$$\mathbb{E}\left[\Big|\Big|\int_{\mathbb{R}^d} \varphi(y)\mu_t(\mathrm{d}y) - \frac{1}{N}\sum_{i=1}^N \varphi(x_t^{i,N})\Big|\Big|^2\right] \leq \frac{K}{N} \tag{5.283}$$

Proof. Let x_t^i, $i = 1, \ldots, N$ denote independent solutions of the McKean-Vlasov SDE (5.1) - (5.2). We then have, using the elementary inequality $||a + b||^2 \leq 2||a||^2 + 2||b||^2$, that

$$\mathbb{E}\Big[||\int_{\mathbb{R}^d} \varphi(y)\mu_t(\mathrm{d}y) - \frac{1}{N}\sum_{i=1}^N \varphi(x_t^{i,N})||^2\Big] \leq 2\mathbb{E}\Big[||\int_{\mathbb{R}^d} \varphi(y)\mu_t(\mathrm{d}y) - \frac{1}{N}\sum_{i=1}^N \varphi(x_t^i)||^2\Big] \tag{5.284}$$

$$+ 2\mathbb{E}\Big[||\frac{1}{N}\sum_{i=1}^N \Big(\varphi(x_t^i) - \varphi(x_t^{i,N})\Big)||^2\Big]$$

For the first term, we observe, using the independence of the variables x_t^i, $i = 1, \ldots, N$, that

$$\mathbb{E}\Big[\big|\big|\int_{\mathbb{R}^d} \varphi(y)\mu_t(\mathrm{d}y) - \frac{1}{N}\sum_{i=1}^N \varphi(x_t^i)\big|\big|^2\Big] \leq \frac{1}{N}\mathbb{E}\left[\big|\big|\varphi(x_t^1) - \mathbb{E}[\varphi(x_t^1)]\big|\big|^2\right] \tag{5.285}$$

We also have that $\mathbb{E}[(\varphi(x_t^1) - \mathbb{E}[\varphi(x_t^1)])^2] = \mathrm{Var}[\varphi(x_t^1)] = \mathrm{Var}[\varphi(x_t^1) - \varphi(\mathbb{E}[x_t^1])] \leq \mathbb{E}[(\varphi(x_t^1) - \mathbb{E}[\varphi(x_t^1)])^2]$. It follows, using also the fact that φ is Lipschitz, and Proposition 5.2 (the bounded moments of the McKean-Vlasov SDE), that

$$\mathbb{E}\Big[\big|\big|\int_{\mathbb{R}^d} \varphi(y)\mu_t(\mathrm{d}y) - \frac{1}{N}\sum_{i=1}^N \varphi(x_t^i)\big|\big|^2\Big] \leq \frac{1}{N}\mathbb{E}\left[\big|\big|\varphi(x_t^1) - \varphi(\mathbb{E}[x_t^1])\big|\big|^2\right] \leq \frac{K}{N}. \tag{5.286}$$

where, as previously, the value of the constant K is allowed to vary from line to line. For the second term, using the Cauchy-Schwarz inequality, the fact that φ is Lipschitz, and Proposition 5.5 (uniform-in-time propagation of chaos), we obtain

$$\mathbb{E}[\big|\big|\frac{1}{N}\sum_{i=1}^N \left(\varphi(x_t^i) - \varphi(x_t^{i,N})\right)\big|\big|^2] \leq \frac{K}{N}\sum_{i=1}^N \mathbb{E}[\big|\big|x_t^i - x_t^{i,N}\big|\big|^2] \leq \frac{K}{N}. \tag{5.287}$$

The result follows immediately. $\square$

Lemma 5.10. *Assume that Conditions B.1 - B.2 and D.1 hold. Let x_t^i denote a solution of the McKean-Vlasov SDE, driven by $w^i = (w_t^i)_{t\geq 0}$. Then, for all Lipschitz functions φ, there exists $K > 0$ such that, for all $t \geq 0$, for all $N \in \mathbb{N}$,*

$$\mathbb{E}\left[\left|\left|\int_{\mathbb{R}^d} \varphi(x_t^i, y)\mu_t(\mathrm{d}y) - \frac{1}{N}\sum_{i=1}^N \varphi(x_t^{i,N}, x_t^{j,N})\right|\right|^2\right] \leq \frac{K}{N} \tag{5.288}$$

Proof. The is an immediate corollary of Lemma 5.9. Indeed, using the Hölder inequality, and that φ is Lipschitz, we have

$$\big|\big|\int_{\mathbb{R}^d} \varphi(x_t^i, y)\mu_t(\mathrm{d}y) - \frac{1}{N}\sum_{j=1}^N \varphi(x_t^{i,N}, x_t^{j,N})\big|\big|^2 \tag{5.289}$$

$$\leq 2\big|\big|\varphi(x_t^i, y)\mu_t(\mathrm{d}y) - \frac{1}{N}\sum_{j=1}^N \varphi(x_t^i, x_t^{j,N})\big|\big|^2 + 2\big|\big|\frac{1}{N}\sum_{j=1}^N [\varphi(x_t^i, x_t^{j,N}) - \varphi(x_t^{i,N}, x_t^{j,N})]\big|\big|^2 \tag{5.290}$$

$$\leq 2\big|\big|\varphi(x_t^i, y)\mu_t(\mathrm{d}y) - \frac{1}{N}\sum_{j=1}^N \varphi(x_t^i, x_t^{j,N})\big|\big|^2 + \frac{2K}{N}\sum_{j=1}^N \big|\big|x_t^i - x_t^{i,N}\big|\big|^2 \tag{5.291}$$

It follows immediately, as required, that

$$\mathbb{E}\left[\left|\left|\int_{\mathbb{R}^d}\varphi(x_t^i,y)\mu_t(\mathrm{d}y)-\frac{1}{N}\sum_{i=1}^N\varphi(x_t^{i,N},x_t^{j,N})\right|\right|^2\right] \tag{5.292}$$

$$\leq K\,\underbrace{\mathbb{E}\left[||\varphi(x_t^i,y)\mu_t(\mathrm{d}y)-\frac{1}{N}\sum_{j=1}^N\varphi(x_t^i,x_t^{j,N})||^2\right]}_{\leq\frac{K}{N}\text{ by Lemma 5.9}}+\underbrace{\mathbb{E}\left[\frac{1}{N}\sum_{j=1}^N||x_t^i-x_t^{i,N}||^2\right]}_{\leq\frac{K}{N}\text{ by Proposition 5.5}}\leq\frac{K}{N}. \tag{5.293}$$

□

5.C.3 Additional Lemmas for Lemma 5.4.D

5.C.3.1 Main Lemmas

The lemmas in this section are variations of Lemmas 3.1 - 3.5 in [420]. For convenience, and since modified versions of these lemmas are also required for the proofs of Theorems 5.3* and 5.3† (see [413]), we provide the proofs of these results in full, appropriately adapted to the current setting.

Lemma 5.11. *Assume that Conditions B.1 - B.2, C.1, D.1, and F.1 hold. Define, with $\hat{x}^N=(x^{1,N},\dots,x^{N,N})$, the function*

$$\Gamma_{k,\eta}=\int_{\tau_k}^{\sigma_{k,\eta}}\gamma_s\left(\nabla_\theta\hat{L}^{i,N}(\theta_s^{i,N},\hat{x}_s^N)-\nabla_\theta\tilde{\mathcal{L}}^{i,N}(\theta_s^{i,N})\right)\mathrm{d}s. \tag{5.294}$$

Then, a.s. , $||\Gamma_{k,\eta}||\to 0$ as $k\to\infty$.

Proof. Let $\hat{x}^N=(x^{1,N},\dots,x^{N,N})\in(\mathbb{R}^d)^N$. Consider the function

$$S^{i,N}(\theta,\hat{x}^N)=\nabla_\theta\hat{L}^{i,N}(\theta,\hat{x}^N)-\nabla_\theta\tilde{\mathcal{L}}^{i,N}(\theta). \tag{5.295}$$

We begin by noting that this function is 'centred' with respect to the invariant measure $\hat{\mu}_\infty(\cdot)$, using the definition of $\nabla_\theta\tilde{\mathcal{L}}^{i,N}(\cdot)$ from Lemma 5.4.B. In addition, observe that, by Lemma 5.17 (see Appendix 5.C.3.3), the function $S^{i,N}(\theta,\hat{x})\in\mathcal{C}^{2,\alpha}(\mathbb{R}^p,(\mathbb{R}^d)^N)$, and there exist positive constants $q,K>0$ such that, for $j=0,1,2$,

$$|\partial_\theta^j S^{i,N}(\theta,\hat{x}^N)|\leq K(1+||x_i||^q+\frac{1}{N}\sum_{j=1}^N||x_j||^q), \tag{5.296}$$

Thus, the function $S^{i,N} : \mathbb{R}^p \times (\mathbb{R}^d)^N \to \mathbb{R}^p$ satisfies the conditions of Lemma 5.16. It follows that, for all $i = 1, \dots, N$, the Poisson equation

$$\mathcal{A}_{\hat{x}} v^{i,N}(\theta, \hat{x}^N) = S^{i,N}(\theta, \hat{x}^N) \quad , \quad \int\limits_{(\mathbb{R}^d)^N} v^{i,N}(\theta, \hat{x}^N) \hat{\mu}_\infty^N(\mathrm{d}\hat{x}^N) = 0 \tag{5.297}$$

has a unique twice differentiable solution which satisfies

$$\sum_{j=0}^{2} \left| \frac{\partial^j v^{i,N}}{\partial \theta^i}(\theta, \hat{x}^N) \right| + \left| \frac{\partial^2 v^{i,N}}{\partial \theta \partial x}(\theta, \hat{x}^N) \right| \leq K \bigg(1 + ||x^{i,N}||^q + \frac{1}{N} \sum_{j=1}^{N} ||x^{j,N}||^q \bigg). \tag{5.298}$$

Let $u^{i,N}(t, \theta, \hat{x}^N) = \gamma_t v^{i,N}(\theta, \hat{x}^N)$. Applying Ito's formula to each component of this vector-valued function, we obtain, for $l = 1, \dots, p$,

$$\tag{5.299}$$

$$u_l^{i,N}(t_2, \theta_{t_2}^{i,N}, \hat{x}_{t_2}^N) - u_l^{i,N}(t_1, \theta_{t_1}^{i,N}, \hat{x}_{t_1}^N) = \int\limits_{t_1}^{t_2} \partial_s u_l^{i,N}(s, \theta_s^{i,N}, \hat{x}_s^N) \mathrm{d}s \tag{5.300}$$

$$+ \int\limits_{t_1}^{t_2} \mathcal{A}_{\hat{x}} u_l^{i,N}(s, \theta_s^{i,N}, \hat{x}_s^N) \mathrm{d}s + \int\limits_{t_1}^{t_2} \mathcal{A}_\theta u_l^{i,N}(s, \theta_s^{i,N}, \hat{x}_s^N) \mathrm{d}s \tag{5.301}$$

$$+ \int\limits_{t_1}^{t_2} \gamma_s \mathrm{Tr} \Big[\nabla_\theta \hat{B}^{i,N}(\theta_s^{i,N}, \hat{x}_s^N) \partial_\theta \partial_{\hat{x}} u_l^{i,N}(s, \theta_s^{i,N}, \hat{x}_s^N) \Big] \mathrm{d}s \tag{5.302}$$

$$+ \int\limits_{t_1}^{t_2} \partial_{\hat{x}} u_l^{i,N}(s, \theta_s^{i,N}, \hat{x}_s^N) \cdot \mathrm{d}\hat{w}_s^N \tag{5.303}$$

$$+ \int\limits_{t_1}^{t_2} \gamma_s \partial_\theta u_l^{i,N}(s, \theta_s^{i,N}, \hat{x}_s^N) \cdot \nabla_\theta \hat{B}^{i,N}(\theta_s^{i,N}, \hat{x}_s^N) \mathrm{d}w_s^i \tag{5.304}$$

where $\mathcal{A}_{\hat{x}}$ and $\mathcal{A}_\theta$ are the infinitesimal generators of $\hat{x}^N$ and θ, respectively, and we recall from (5.96) that $\hat{w}_t^N = (w_t^1, \dots, w_t^N)^T$. Rearranging this identity, and also recalling that $v^{i,N}(\theta, \hat{x}^N)$ is the solution of the Poisson equation, we obtain

$$\Gamma_{k,\eta} = \int\limits_{\tau_k}^{\sigma_{k,\eta}} \gamma_s \mathcal{A}_{\hat{x}} v^{i,N}(\theta_s^{i,N}, \hat{x}_s^N) \mathrm{d}s \tag{5.305}$$

$$
\begin{aligned}
&= \gamma_{\sigma_{k,\eta}} v^{i,N}(\theta^{i,N}_{\sigma_{k,\eta}}, \hat{x}^N_{\sigma_{k,\eta}}) - \gamma_{\tau_k} v^{i,N}(\theta^{i,N}_{\tau_k}, \hat{x}^N_{\tau_k}) - \int_{\tau_k}^{\sigma_{k,\eta}} \dot{\gamma}_s v^{i,N}(\theta^{i,N}_s, \hat{x}^N_s)\mathrm{d}s \\
&- \int_{\tau_k}^{\sigma_{k,\eta}} \gamma_s \mathcal{A}_\theta v^{i,N}(\theta^{i,N}_s, \hat{x}^N_s)\mathrm{d}s - \int_{\tau_k}^{\sigma_{k,\eta}} \gamma_s^2 \mathrm{Tr}\big[\nabla_\theta \hat{B}^{i,N}(\theta^{i,N}_s, \hat{s}^N_x)\partial_\theta \partial_{\hat{x}} v^{i,N}(\theta^{i,N}_s, \hat{x}^N_s)\big]\mathrm{d}s \\
&- \int_{\tau_k}^{\sigma_{k,\eta}} \gamma_s \partial_{\hat{x}} v^{i,N}(\theta^{i,N}_s, \hat{x}^N_s) \cdot \mathrm{d}\hat{w}^N_s - \int_{\tau_k}^{\sigma_{k,\eta}} \gamma_s^2 \partial_\theta v^{i,N}(\theta^{i,N}_s, \hat{x}^N_s) \cdot \nabla_\theta \hat{B}^{i,N}(\theta^{i,N}_s, \hat{x}^N_s)\mathrm{d}w^i_s
\end{aligned}
$$

We now prove the convergence of each term on the right hand side of this equation. As previously, we allow the value of K to change from line to line. First define

$$
J^{(1)}_t = \gamma_t ||v^{i,N}(\theta^{i,N}_t, \hat{x}^N_t)|| \tag{5.306}
$$

We have, make use of the polynomial growth of $v^{i,N}(\theta, \hat{x}^N)$, and Proposition 5.2 (the bounded moments of the IPS), that

$$
\mathbb{E}[|J^{(1)}_t|^2] \leq K\gamma_t^2 \bigg(1 + \mathbb{E}[||x^{i,N}_t||^q] + \frac{1}{N}\sum_{j=1}^N \mathbb{E}[||x^{j,N}_t||^q]\bigg) \leq K\gamma_t^2. \tag{5.307}
$$

Applying the Borel-Cantelli argument as in [422, Appendix B], it follows that $J^{(1)}_t$ converges to zero with probability one. We next consider the term

$$
\begin{aligned}
J^{(2)}_{0,t} = \int_0^t \partial_s \dot{\gamma}_s v^{i,N}(\theta^{i,N}_s, \hat{x}^N_s)\mathrm{d}s + \int_0^t \gamma_s \mathcal{A}_\theta v^{i,N}(\theta^{i,N}_s, \hat{x}^N_s)\mathrm{d}s \\
+ \int_0^t \gamma_s^2 \mathrm{Tr}\big[\nabla_\theta \hat{B}^{i,N}(\theta^{i,N}_s, \hat{x}^N_s)\partial_\theta \partial_x v^{i,N}(\theta^{i,N}_s, \hat{x}^N_s)\big]\mathrm{d}s
\end{aligned} \tag{5.308}
$$

This term obeys the bound

$$
\begin{aligned}
\sup_{t>0} \mathbb{E}|J^{(2)}_{0,t}| &\leq K \int_0^\infty (|\dot{\gamma}_s| + \gamma_s^2)(1 + \mathbb{E}[||x^{i,N}_s||^q] + \frac{1}{N}\sum_{j=1}^N \mathbb{E}[||x^{j,N}_s||^q])\mathrm{d}s && (5.309) \\
&\leq K \int_0^\infty (|\dot{\gamma}_s| + \gamma_s^2)\mathrm{d}s < \infty. && (5.310)
\end{aligned}
$$

Here, the first inequality follows from the growth properties of the $v^{i,N}(\theta, \hat{x}^N)$ in (5.298), the second inequality from Proposition 5.2 (the bounded moments of the IPS), and the final inequality from Condition F.1 (the properties of the learning rate). It follows that there exists a finite random variable $J_{0,\infty}^{(2)}$ such that, with probability one,

$$J_{0,t}^{(2)} \to J_{0,\infty}^{(2)} , \quad \text{as } t \to \infty. \tag{5.311}$$

The last term to consider is the stochastic integral

$$J_{0,t}^{(3)} = \int_0^t \gamma_s \partial_{\hat{x}} v^{i,N}(\theta_s^{i,N}, \hat{x}_s^N) \cdot \mathrm{d}\hat{w}_s^N + \int_{\tau_k}^{\sigma_{k,\eta}} \gamma_s^2 \partial_\theta v^{i,N}(\theta_s^{i,N}, \hat{x}_s^N) \cdot \nabla_\theta \hat{B}^{i,N}(\theta_s^{i,N}, \hat{x}_s^N) \mathrm{d}w_s^i \tag{5.312}$$

In this case, using the BDG inequality, and the same bounds as above, we have

$$\mathbb{E}\left[|J_{0,t}^{(3)}|^2\right] \leq K \int_0^\infty (\gamma_s^2 + \gamma_s^4)\left[1 + \mathbb{E}[||x_s^{i,N}||^q] + \frac{1}{N}\sum_{j=1}^N \mathbb{E}[||x_s^{j,N}||^q]\right] \mathrm{d}s \leq K \int_0^\infty \gamma_s^2 \mathrm{d}s < \infty. \tag{5.313}$$

Thus, by Doob's martingale convergence theorem, there exists a square integrable random variable $J_{0,\infty}^{(3)}$ such that, both a.s. and in $\mathbb{L}^2$,

$$J_{0,t}^{(3)} \to J_{0,\infty}^{(3)} , \quad \text{as } t \to \infty. \tag{5.314}$$

It remains only to observe, combining (5.311) and (5.314), we have

$$||\Gamma_{k,\eta}|| \leq J_{\sigma_{k,\eta}}^{(1)} + J_{\tau_k}^{(1)} + J_{\tau_k,\sigma_{k,\eta}}^{(2)} + J_{\tau_k,\sigma_{k,\eta}}^{(3)} \overset{k\to\infty}{\to} 0. \tag{5.315}$$

□

Lemma 5.12. *Assume that Conditions B.1 - B.2, C.1, D.1 and F.1 hold. Let $\rho > 0$ be such that, for a given $\kappa > 0$, it is true that $3\rho + \frac{\rho}{4\kappa} = \frac{1}{2L}$, where L denotes the Lipschitz constant of $\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta)$. For k large enough, and for $\eta > 0$ small enough (potentially random, and depending on k), one has*

$$\int_{\tau_k}^{\sigma_{k,\eta}} \gamma_s \mathrm{d}s > \rho \quad \textit{and, a.s.,} \quad \frac{\rho}{2} \leq \int_{\tau_k}^{\sigma_k} \gamma_s \mathrm{d}s \leq \rho. \tag{5.316}$$

Proof. We proceed by contradiction. Let us assume that $\int_{\tau_k}^{\sigma_{k,\eta}} \gamma_s \mathrm{d}s \leq \rho$. Choose arbitrary $\varepsilon > 0$ such that $\varepsilon \leq \frac{\rho}{8}$. We begin with the observation that, via the Itô isometry, we have that

$$\sup_{t\geq 0} \mathbb{E}\Big|\Big| \int_0^t \gamma_s \frac{\kappa}{||\nabla \tilde{\mathcal{L}}^{i,N}(\theta_{\tau_k}^{i,N})||} \nabla_\theta \hat{B}^{i,N}(\theta_s^{i,N}, \hat{x}_s^N) \mathrm{d}w_s^i \Big|\Big|^2 \leq \int_0^t K\gamma_s^2 \left(1 + \mathbb{E}\left[||\hat{x}_s^N||^q\right]\right) \mathrm{d}s < \infty \tag{5.317}$$

where, we have used the polynomial growth of $\nabla_\theta \hat{B}^{i,N}(\theta,\hat{x})$ (see the proof of Lemma 5.17), Proposition 5.2 (the bounded moments of the IPS), and Condition F.1 (the properties of the learning rate). Thus, by the Doob's martingale convergence theorem, there exists a finite random variable M such that, both a.s. and in $\mathbb{L}^2$,

$$\int_0^t \gamma_s \frac{\kappa}{||\nabla\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k})||}\nabla_\theta \hat{B}^{i,N}(\theta^{i,N}_s,\hat{x}^N_s)\mathrm{d}w^i_s \to M \tag{5.318}$$

It follows that, for the chosen $\varepsilon>0$, there exists k such that

$$\int_{\tau_k}^{\sigma_{k,\eta}} \gamma_s \frac{\kappa}{||\nabla\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k})||}\nabla_\theta \hat{B}^{i,N}(\theta^{i,N}_s,\hat{x}^N_s)\mathrm{d}w^i_s < \varepsilon. \tag{5.319}$$

Let us now also assume that, for the given k, η is small enough such that for all $s\in[\tau_k,\sigma_{k,\eta}]$, we have $||\nabla_\theta\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_s)|| \leq 3||\nabla_\theta\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k})||$. We can then compute

$$\tag{5.320}$$

$$||\theta^{i,N}_{\sigma_{k,\eta}} - \theta^{i,N}_{\tau_k}|| = \Big|\Big| \int_{\tau_k}^{\sigma_{k,\eta}} \gamma_s \nabla_\theta \hat{L}^{i,N}(\theta^{i,N}_s,\hat{x}^N_s)\mathrm{d}s + \int_{\tau_k}^{\sigma_{k,\eta}} \gamma_s\langle \nabla_\theta \hat{B}^{i,N}(\theta^{i,N}_s,\hat{x}^N_s),\mathrm{d}w^i_s\rangle\Big|\Big| \tag{5.321}$$

$$\begin{aligned}
&\leq 3||\nabla_\theta\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k})||\int_{\tau_k}^{\sigma_{k,\eta}}\gamma_s\mathrm{d}s + \Big|\Big|\int_{\tau_k}^{\sigma_{k,\eta}}\gamma_s[\nabla_\theta \hat{L}^{i,N}(\theta^{i,N}_s,\hat{x}^N_s) - \nabla_\theta\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_s)]\mathrm{d}s\Big|\Big| \\
&+ \frac{||\nabla\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k})||}{\kappa}\Big|\Big|\int_{\tau_k}^{\sigma_{k,\eta}}\gamma_s\frac{\kappa}{||\nabla\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k})||}\langle\nabla_\theta B(\theta^{i,N}_s,\hat{x}^N_s),\mathrm{d}w^i_s\rangle\Big|\Big| && (5.322)\\
&\leq 3||\nabla_\theta\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k})||\rho + \varepsilon + \frac{||\nabla\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k})||}{\kappa}\varepsilon && (5.323)\\
&\leq ||\nabla_\theta\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k})||\left[3\rho + \frac{\rho}{4\kappa}\right] && (5.324)
\end{aligned}$$

where in the penultimate line we have used Lemma 5.11 and (5.319), and in the final line we have used the fact that our choice of ε satisfies $\varepsilon \leq \frac{\rho}{8}$. We thus obtain

$$||\theta^{i,N}_{\sigma_{k,\eta}} - \theta^{i,N}_{\tau_k}|| \leq ||\nabla_\theta\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k})||\left[3\rho + \frac{\rho}{4\kappa}\right] \leq ||\nabla_\theta\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k})||\frac{1}{2L}. \tag{5.325}$$

Thus, using also the definition of the Lipschitz constant L, we obtain

$$||\nabla_\theta\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\sigma_{k,\eta}}) - \nabla_\theta\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k})|| \leq L||\theta^{i,N}_{\sigma_{k,\eta}} - \theta^{i,N}_{\tau_k}|| \leq \frac{1}{2}||\nabla_\theta\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k})|| \tag{5.326}$$

which then yields

$$\frac{1}{2}||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k})|| \leq ||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\sigma_{k,\eta}})|| \leq 2||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k})||. \tag{5.327}$$

But this implies that $\sigma_{k,\eta} \in [\tau_k, \sigma_k]$, which is a contradiction. Thus we do indeed have $\int_{\tau_k}^{\sigma_{k,\eta}} \gamma_s \mathrm{d}s > \rho$. We now turn our attention to the second part of the Lemma. By definition, we have that $\int_{\tau_k}^{\sigma_k} \gamma_s \mathrm{d}s \leq \rho$. Thus, it remains only to show that $\frac{\rho}{2} \leq \int_{\tau_k}^{\sigma_k} \gamma_s \mathrm{d}s$. From the first part of the Lemma, we have that $\int_{\tau_k}^{\sigma_{k,\eta}} \gamma_s \mathrm{d}s > \rho$. Moreover, for k sufficiently large and η sufficiently small, we must have $\int_{\sigma_k}^{\sigma_{k,\eta}} \gamma_s \mathrm{d}s \leq \frac{\rho}{2}$. We thus obtain

$$\int_{\tau_k}^{\sigma_k} \gamma_s \mathrm{d}s \geq \rho - \int_{\sigma_k}^{\sigma_{k,\eta}} \gamma_s \mathrm{d}s \geq \rho - \frac{\rho}{2} = \frac{\rho}{2}. \tag{5.328}$$

□

Lemma 5.13. *Assume that Conditions B.1 - B.2, C.1, D.1 and F.1 hold. Suppose that there are an infinite number of intervals $[\tau_k, \sigma_k)$. Then there exists a fixed constant $\beta = \beta(\kappa) > 0$ such that, for k large enough, a.s.,*

$$\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\sigma_k}) - \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k}) \geq \beta. \tag{5.329}$$

Proof. By Itô's formula, we have that

$$\begin{aligned}
&\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\sigma_k}) - \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k}) \qquad (5.330)\\
&= \underbrace{\int_{\tau_k}^{\sigma_k} \gamma_s ||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_s)||^2 \mathrm{d}s}_{A^{i,N}_{1,k}} \\
&+ \underbrace{\int_{\tau_k}^{\sigma_k} \gamma_s \langle \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_s), \nabla_\theta \hat{L}^{i,N}(\theta^{i,N}_s, \hat{x}^N_s) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_s)\rangle \mathrm{d}s}_{A^{i,N}_{2,k}} \\
&+ \underbrace{\int_{\tau_k}^{\sigma_k} \gamma_s \langle \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_s), \nabla_\theta \hat{B}^{i,N}(\theta^{i,N}_s, \hat{x}^N_s) \mathrm{d}w^i_s\rangle}_{A^{i,N}_{3,k}} \\
&+ \underbrace{\int_{\tau_k}^{\sigma_k} \frac{1}{2}\gamma_s^2 \mathrm{Tr}\left[\nabla_\theta \hat{B}^{i,N}(\theta^{i,N}_s, \hat{x}^N_s) \nabla_\theta \hat{B}^{i,N}(\theta^{i,N}_s, \hat{x}^N_s)^T \nabla^2_\theta \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_s) \mathrm{d}s\right]}_{A^{i,N}_{4,k}}
\end{aligned}$$

We will deal with each of these terms individually. First consider $A^{i,N}_{1,k}$. For this term, we have that

$$A_{1,k}^{i,N} = \int_{\tau_k}^{\sigma_k} \gamma_s ||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_s^{i,N})||^2 \mathrm{d}s \geq \frac{||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_{\tau_k}^{i,N})||^2}{4} \int_{\tau_k}^{\sigma_k} \gamma_s \mathrm{d}s \geq \frac{||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_{\tau_k}^{i,N})||^2}{8}\rho \tag{5.331}$$

where, in the first inequality, we have used the definition of the $\{\tau_k\}_{k\geq 0}$, namely, that $||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_s^{i,N})|| \geq \frac{1}{2}||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_{\tau_k}^{i,N})||$ for all $s \in [\tau_k, \sigma_k]$, and in the second inequality we have used Lemma 5.12.

We now turn our attention to $A_{2,k}^{i,N}$. We will handle this term using a very similar to approach to that used in the proof of Lemma 5.11. Let us consider the function

$$T^{i,N}(\theta, \hat{x}^N) = \langle \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta), \nabla_\theta \hat{L}^{i,N}(\theta, \hat{x}^N) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta)\rangle. \tag{5.332}$$

By Lemma 5.18, we have that $T^{i,N}(\theta, \hat{x}^N) \in \mathcal{C}^{2,\alpha}(\mathbb{R}^p, \mathbb{R}^d)$, and that $||\partial_\theta^j T^{i,N}(\theta, \hat{x}^N)| \leq K(1+||x^{i,N}||^q + \frac{1}{N}\sum_{j=1}^N ||x_j||^q)$, for $j = 0, 1, 2$. Moreover, it is straightforward to show that this function satisfies $\int_{(\mathbb{R}^d)^N} T^{i,N}(\theta, \hat{x}^N)\hat{\mu}_\infty(\mathrm{d}\hat{x}^N) = 0$. Thus, Lemma 5.16, the Poisson equation

$$\mathcal{A}_{\hat{x}} v^{i,N}(\theta, \hat{x}^N) = T^{i,N}(\theta, \hat{x}^N) \quad , \quad \int_{(\mathbb{R}^d)^N} v^{i,N}(\theta, \hat{x}^N)\mu_\infty(\mathrm{d}\hat{x}^N) = 0 \tag{5.333}$$

has a unique twice differentiable solution which satisfies

$$\sum_{j=0}^{2} \left|\frac{\partial^j v^{i,N}}{\partial \theta^i}(\theta, \hat{x}^N)\right| + \left|\frac{\partial^2 v^{i,N}}{\partial\theta\partial x}(\theta, \hat{x}^N)\right| \leq K\left[1 + ||x^{i,N}||^q + \frac{1}{N}\sum_{j=1}^{N}||x^{j,N}||^q\right]. \tag{5.334}$$

and, using the same steps as in the proof of Lemma 5.11, we can prove that, a.s.,

$$\left|\left|\int_{\tau_k}^{\sigma_k} \gamma_s \langle \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_s^{i,N}), \nabla_\theta \hat{L}^{i,N}(\theta_s^{i,N}, \hat{x}_s^N) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_s^{i,N})\rangle \mathrm{d}s\right|\right| \overset{k\to\infty}{\to} 0. \tag{5.335}$$

We next consider $A_{3,k}^{i,N}$. Using Itô's isometry, Lemma 5.8, the polynomial growth of the function $\nabla_\theta \hat{B}^{i,N}(\theta, \hat{x})$ (see the proof of Lemma 5.17), Proposition 5.2 (the moment bounds for solutions of the IPS) and Condition F.1 (the square summability of the learning rate), we have that

$$\sup_{t\geq 0}\mathbb{E}\left[\left|\int_0^t \gamma_s \langle \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_s^{i,N}), \nabla_\theta \hat{B}^{i,N}(\theta_s^{i,N}, \hat{x}_s^N)\mathrm{d}w_s^i\rangle\right|^2\right] \tag{5.336}$$

$$\leq K\mathbb{E}\int_0^\infty \gamma_s^2 ||\nabla_\theta \hat{B}^{i,N}(\theta_s^{i,N}, \hat{x}_s^N)||^2 \mathrm{d}s \tag{5.337}$$

$$\leq K\int_0^\infty \gamma_s^2 (1 + \mathbb{E}\left[||x_s^{i,N}||^q\right] + \frac{1}{N}\sum_{j=1}^N \mathbb{E}\left[||x_s^{j,N}||^q\right] \mathrm{d}s < \infty. \tag{5.338}$$

Thus, by Doob's martingale convergence theorem, there exists a finite random variable $A_{3,\infty}^{i,N}$ such that, both a.s. and in $\mathbb{L}^2$,

$$\int_0^t \gamma_s \langle \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_s^{i,N}), \nabla_\theta \hat{B}^{i,N}(\theta_s^{i,N}, \hat{x}_s^N)\mathrm{d}w_s^i\rangle \to A_{3,\infty}^{i,N}. \tag{5.339}$$

as $t \to \infty$. It follows that $A_{3,k}^{i,N} \to 0$ a.s. as $k \to \infty$. Finally, we turn our attention to $A_{4,k}^{i,N}$. For this term, we observe that

$$\sup_{t\geq 0} \mathbb{E}\left|\left|\int_0^t \frac{1}{2}\gamma_s^2 \mathrm{Tr}\left[\nabla_\theta \hat{B}^{i,N}(\theta_s^{i,N}, \hat{x}_s^N)\nabla_\theta \hat{B}^{i,N}(\theta_s^{i,N}, \hat{x}_s^N)^T \nabla_\theta^2 \tilde{\mathcal{L}}^{i,N}(\theta_s^{i,N})\right]\mathrm{d}s\right|\right| \tag{5.340}$$

$$\leq K\int_0^\infty \gamma_s^2 (1 + \mathbb{E}\left[||x_s^{i,N}||^q\right] + \frac{1}{N}\sum_{j=1}^N \mathbb{E}\left[||\hat{x}_s^{j,N}||^q\right])\mathrm{d}s < \infty, \tag{5.341}$$

where, as above, we have used Lemma 5.8, the polynomial growth of $\nabla_\theta \hat{B}^{i,N}(\theta, \hat{x})$, Proposition 5.2, and Condition F.1. It follows that the random variable

$$\int_0^\infty \frac{1}{2}\gamma_s^2 \mathrm{Tr}\left[\nabla_\theta \hat{B}^{i,N}(\theta_s^{i,N}, \hat{x}_s^N)\nabla_\theta \hat{B}^{i,N}(\theta_s^{i,N}, \hat{x}_s^N)^T \nabla_\theta^2 \tilde{\mathcal{L}}^{i,N}(\theta_s^{i,N})\right]\mathrm{d}s \tag{5.342}$$

is finite a.s., which in turn implies that there exists a finite random variable $A_{4,\infty}^{i,N}$ such that

$$\int_0^t \frac{1}{2}\gamma_s^2 \mathrm{Tr}\left[\nabla_\theta \hat{B}^{i,N}(\theta_s^{i,N}, \hat{x}_s^N)\nabla_\theta \hat{B}^{i,N}(\theta_s^{i,N}, \hat{x}_s^N)^T \nabla_\theta^2 \tilde{\mathcal{L}}^{i,N}(\theta_s^{i,N})\right]\mathrm{d}s \to A_4^\infty \quad \text{a.s.} \tag{5.343}$$

It follows, in particular, that $A_{4,k}^{i,N} \to 0$ a.s. as $k \to \infty$. Summarising, we thus have that, for all $\varepsilon > 0$, there exists k such that

$$\tilde{\mathcal{L}}^{i,N}(\theta_{\sigma_k}^{i,N}) - \tilde{\mathcal{L}}^{i,N}(\theta_{\tau_k}^{i,N}) = A_{1,k}^{i,N} + A_{2,k}^{i,N} + A_{3,k}^{i,N} + A_{4,k}^{i,N} \tag{5.344}$$

$$\geq A_{1,k}^{i,N} - ||A_{2,k}^{i,N}|| - ||A_{3,k}^{i,N}|| - ||A_{4,k}^{i,N}|| \tag{5.345}$$

$$= \frac{||\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta_{\tau_k}^{i,N})||^2}{8}\rho - 3\varepsilon \tag{5.346}$$

The claim follows by setting $\varepsilon = \frac{\rho(\kappa)\kappa^2}{32}$ and $\beta = \frac{\rho(\kappa)\kappa^2}{32}$. $\square$

Lemma 5.14. *Assume that Conditions B.1 - B.2, C.1, D.1 and F.1 hold. Suppose that there are an infinite number of intervals $[\tau_k, \sigma_k)$. Then there exists a fixed constant $0 < \beta_1 < \beta$ such that, for k large enough,*

$$\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k}) - \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\sigma_{k-1}}) \geq -\beta_1. \tag{5.347}$$

Proof. Using Itô's formula, we have that

$$\begin{aligned}\tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\tau_k}) - \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_{\sigma_{k-1}}) \geq & \underbrace{\int_{\sigma_{k-1}}^{\tau_k} \gamma_s \langle \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_s), \nabla_\theta \hat{L}^{i,N}(\theta^{i,N}_s, \hat{x}^N_s) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_s)\rangle \mathrm{d}s}_{B^{i,N}_{1,k}} \\ & + \underbrace{\int_{\sigma_{k-1}}^{\tau_k} \gamma_s \langle \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_s), \nabla_\theta \hat{B}^{i,N}(\theta^{i,N}_s, \hat{x}^N_s)\mathrm{d}w^i_s\rangle}_{B^{i,N}_{2,k}} \\ & + \underbrace{\int_{\sigma_{k-1}}^{\tau_k} \frac{1}{2}\gamma_s^2 \mathrm{Tr}\left[\nabla_\theta \hat{B}^{i,N}(\theta^{i,N}_s, \hat{x}^N_s)\nabla_\theta \hat{B}^{i,N}(\theta^{i,N}_s, \hat{x}^N_s)^T \nabla^2_\theta \tilde{\mathcal{L}}^{i,N}(\theta^{i,N}_s)\mathrm{d}s\right]}_{B^{i,N}_{3,k}}.\end{aligned} \tag{5.348}$$

Arguing as in the proof of Lemma 5.13, the magnitude of each of the terms converges to zero a.s. as $k \to \infty$. This is sufficient for the conclusion. $\square$

5.C.3.2 Technical Lemmas: On A Related Poisson Equation

Lemma 5.15. *Assume that Conditions B.1 - B.2 and D.1 hold. Suppose that, for all $\theta \in \mathbb{R}^p$, $f(\theta, \cdot) : (\mathbb{R}^d)^N \to \mathbb{R}$ satisfies a polynomial growth condition of the form*

$$||f(\theta, \hat{x}^N)|| \leq K\left(1 + ||x^{i,N}||^q + \frac{1}{N}\sum_{j=1}^N ||x^{j,N}||^q\right) \tag{5.349}$$

Moreover, suppose that $f(\theta, \cdot)$ is centred, in the sense that $\int_{(\mathbb{R}^d)^N} f(\theta, \hat{x}^N)\mathrm{d}\hat{\mu}^N_\infty(\mathrm{d}\hat{x}^N) = 0$. Then, for all $N \in \mathbb{N}$, the function

$$F(\theta, \hat{x}^N) = \int_0^\infty \mathbb{E}_{\hat{x}^N, \theta_0}\left[f(\theta, \hat{x}^N_t)\right]\mathrm{d}t \tag{5.350}$$

is a well defined, continuous function of Sobolev class $\cap_{p\geq 1} W^2_{p,\mathrm{loc}}$, which satisfies the Poisson equation

$$\mathcal{A}_{\hat{x}^N,\theta^*} F(\theta,\hat{x}^N) = -f(\theta,\hat{x}^N). \tag{5.351}$$

Moreover, F is centred, in the sense that $\int_{(\mathbb{R}^d)^N} F(\theta,\hat{x}^N)\hat{\mu}^N_\infty(\mathrm{d}\hat{x}^N) = 0$, and there exist constants $q, K > 0$ such that

$$|F(\theta,\hat{x}^N)| \leq K\left[1 + ||x^{i,N}||^q + \frac{1}{N}\sum_{j=1}^N ||x^{j,N}||^q\right] \tag{5.352}$$

$$|\nabla_{\hat{x}^N} F(\theta,\hat{x})^N| \leq K\left[1 + ||x^{i,N}||^q + \frac{1}{N}\sum_{j=1}^N ||x^{j,N}||^q\right] \tag{5.353}$$

Remark. This is essentially a statement of [371, Theorem 1], adapted appropriately to the current statement. In our case, however, since we are interested in the solution of the Poisson equation associated with the generator of the IPS $\hat{x}^N = (x^{1,N},\dots,x^{N,N}) \in (\mathbb{R}^d)^N$ for any $N \in \mathbb{N}$, a little care is needed in places to ensure that arguments in the proof of [371, Theorem 1], in particular those used to establish that the solution is well defined, and that it satisfies the bounds in (5.352) - (5.353), are independent of N. Indeed, we are interested in the solution of this Poisson equation for arbitrarily large N, since we will later take the limit as $N \to \infty$. As an example, if we were to use [371, Theorem 1] directly, we would only have, in place of (5.352), the bound $|F(\theta,\hat{x})| \leq K(1 + ||\hat{x}^N||^q)$, which, due to the $||\hat{x}^N||^q$ term, is unbounded in the limit as $N \to \infty$.

Proof. We begin by showing that the function $F(\theta,\hat{x}^N)$ is well defined, and that it satisfies (5.352). Let $\hat{x}^N_t$ denote a solution of the IPS starting from $\hat{x}^N \in (\mathbb{R}^d)^N$. Let $\hat{\mu}^N_t$ denote the law of $\hat{x}^N_t$. Using the bounds in Lemma 5.5, and that f is centred, we have

$$\left|\mathbb{E}_{\hat{x}^N}\left[f(\theta,\hat{x}^N_t)\right]\right| = \left|\mathbb{E}_{\hat{x}^N}\left[f(\theta,\hat{x}^N_t)\right] - \int_{(\mathbb{R}^d)^N} f(\theta,\hat{z}^N)\hat{\mu}^N_\infty(\mathrm{d}\hat{z}^N)\right| \tag{5.354}$$

$$\leq K\left[1 + ||x^{i,N}||^q + \frac{1}{N}\sum_{j=1}^N ||x^{j,N}||^q\right]e^{-\lambda t} \tag{5.355}$$

We remark that, crucially, the constants $q, K, \lambda > 0$ are independent of N. Thus, for all $N \in \mathbb{N}$, the function F, as defined in (5.350), is absolutely integrable, and thus well defined. Moreover, via the triangle inequality, we immediately obtain the bound in (5.352).

The remaining statements in Lemma 5.15 now follow directly from [371, Theorem 1]. In particular, the arguments in the proof of [371, Theorem 1(b), 1(c), 1(d), 1(f)] show that (5.350) defines a continuous, centred solution, unique in the class of solutions belonging to $\cap_{p\geq 1} W^2_{p,\mathrm{loc}}$, of the Poisson equation (5.351).

Finally, we can obtain the bound in (5.353) using the argument in the proof of [371, Theorem 1(e)], replacing the intermediate bound on $||F(\theta,\cdot,\cdot)||$ by (5.352), and the intermediate bound on $||f(\theta,\cdot,\cdot)||$ by our condition on the polynomial growth of $f(\cdot)$.[9]. This completes the proof. $\square$

Lemma 5.16. *Assume that Conditions B.1 - B.2 and D.1 hold. Suppose that the function $f(\theta,\hat{x}^N)\in\mathcal{C}^{\alpha,2}(\mathbb{R}^p,(\mathbb{R}^d)^N)$, for some $\alpha>0$, is centred in the same sense as Lemma 5.15, and satisfies*

$$|f(\theta,\hat{x}^N)|+|\partial_\theta f(\theta,\hat{x}^N)|+|\partial_\theta^2 f(\theta,\hat{x}^N)|\leq K\left[1+||x^{i,N}||^q+\frac{1}{N}\sum_{j=1}^N||x^{j,N}||^q\right] \tag{5.356}$$

where $\hat{x}=(x^{1,N},\dots,x^{N,N})$. Then the solution (5.350) of the Poisson equation (5.351) satisfies $F(\cdot,\hat{x}^N)\in\mathcal{C}^2$ for all $\hat{x}^N\in(\mathbb{R}^d)^N$. Moreover, there exist $q',K'>0$ such that

$$\sum_{k=0}^2\left|\frac{\partial^k F}{\partial\theta^k}\right|+\left|\frac{\partial^2 F}{\partial\hat{x}\partial\theta}\right|\leq K\left[1+||x^{i,N}||^q+\frac{1}{N}\sum_{j=1}^N||x^{j,N}||^q\right] \tag{5.357}$$

Proof. The first statement of the Theorem follows directly from [372, Theorem 3]. Now, observe that, since $\partial_\theta^k f$, $k=0,1,2$, satisfies a polynomial growth condition in the required sense, $\partial_\theta^k f^{i,N}$ can be shown to satisfy bounds of the form given in Lemma 5.5. It follows, arguing as in (5.354) - (5.355), that

$$\left|\mathbb{E}_{\hat{x}^N}\left[\frac{\partial^k f}{\partial\theta^k}(\theta,\hat{x}_t^N)\right]\right|\leq K\left[1+||x^{i,N}||^q+\frac{1}{N}\sum_{j=1}^N||x^{j,N}||^q\right]e^{-\lambda t} \tag{5.358}$$

We thus have that, allowing the value of the constant K to change from line to line, that

$$\left|\frac{\partial^k F}{\partial\theta^k}(\theta,\hat{x}^N)\right|\leq\int_0^\infty\left|\mathbb{E}_{\hat{x}^N}\left[\frac{\partial^k f}{\partial\theta^k}(\theta,\hat{x}^N)\right]\right|\mathrm{d}t\leq K\int_0^\infty\left[1+||x^{i,N}||^q+\frac{1}{N}\sum_{j=1}^N||x^{j,N}||^q\right]e^{-\lambda t}\mathrm{d}t \tag{5.359}$$

$$\leq K\left[1+||x^{i,N}||^q+\frac{1}{N}\sum_{j=1}^N||x^{j,N}||^q\right]. \tag{5.360}$$

Finally, the bound on the mixed derivative follows from (5.353) in Lemma 5.15. $\square$

5.C.3.3 Technical Lemmas: Miscellaneous

Lemma 5.17. *Assume that Conditions B.1 - B.2, C.1 and D.1 hold. Then, for all $i=1,\dots,N$, $N\in\mathbb{N}$, the function $S^{i,N}(\theta,\hat{x}^N)=\nabla_\theta\hat{L}^{i,N}(\theta,\hat{x}^N)-\nabla_\theta\tilde{\mathcal{L}}^{i,N}(\theta)$ is in*

[9] In the original notation, these are the bounds on $||u||$ and $||Lu||$, respectively. See [371, pg. 1070]

$\mathcal{C}^{2,\alpha}(\mathbb{R}^p, (\mathbb{R}^d)^N)$. *Moreover, for $k = 0, 1, 2$, there exists q and K such that*

$$||\nabla_\theta^k S^{i,N}(\theta, \hat{x}^N)|| \leq K\bigg[1 + ||x^{i,N}||^q + \frac{1}{N}\sum_{j=1}^N ||x^{j,N}||^q\bigg]. \tag{5.361}$$

Proof. By definition, we have that, for $k = 0, 1, 2$,

$$\nabla_\theta^k S^{i,N}(\theta, \hat{x}^N) = \nabla_\theta^{k+1}\hat{L}^{i,N}(\theta, \hat{x}^N) - \nabla_\theta^{k+1}\tilde{\mathcal{L}}^{i,N}(\theta). \tag{5.362}$$

By Condition C.1(i), $\nabla_\theta b(\theta, x) \in \mathcal{C}^{2,\alpha}(\mathbb{R}^p, \mathbb{R}^d)$, and $\nabla_\theta \phi(\theta, x, y) \in \mathcal{C}^{2,\alpha,\alpha}(\mathbb{R}^p, \mathbb{R}^d, \mathbb{R}^d)$. It follows from the definitions, c.f. (5.97), (5.98) and (5.99), that $\nabla_\theta \hat{B}^{i,N}(\theta, \hat{x}^N)$, $\nabla_\theta \hat{G}^{i,N}(\theta, \hat{x}^N)$, and $\nabla_\theta \hat{L}^{i,N}(\theta, \hat{x}^N)$ are in $\mathcal{C}^{2,\alpha}(\mathbb{R}^p, (\mathbb{R}^d)^N)$. It also follows from the definition (Lemma 5.4.B) that $\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta)$ is in $\mathcal{C}^2(\mathbb{R}^p)$. Thus, as claimed, $S^{i,N}(\theta, \hat{x})$ is in $\mathcal{C}^{2,\alpha}(\mathbb{R}^p, (\mathbb{R}^d)^N)$. It remains to note that the bound (5.361) follows immediately from Lemma 5.6 and Lemma 5.8 □

Lemma 5.18. *Assume that Conditions B.1 - B.2, C.1 and D.1 hold. Then, for all $i = 1, \dots, N$, $N \in \mathbb{N}$, the function $T^{i,N}(\theta, \hat{x}^N) = \langle \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta), \nabla_\theta \hat{L}^{i,N}(\theta, \hat{x}^N) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta)\rangle$ is in $\mathcal{C}^{2,\alpha}(\mathbb{R}^p, (\mathbb{R}^d)^N)$. Moreover, for $k = 0, 1, 2$, there exists q, K such that*

$$||\nabla_\theta^k T^{i,N}(\theta, \hat{x}^N)|| \leq K\bigg[1 + ||x^{i,N}||^q + \frac{1}{N}\sum_{j=1}^N ||x^{j,N}||^q\bigg]. \tag{5.363}$$

Proof. This lemma follows almost immediately from Lemma 5.17. First note that, by definition, we can write $T^{i,N}(\theta, \hat{x}^N) = \langle \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta), S^{i,N}(\theta, \hat{x}^N)\rangle$. By Lemma 5.17, $S^{i,N}(\theta, \hat{x}^N)$ is in $\mathcal{C}^{2,\alpha}(\mathbb{R}^p, (\mathbb{R}^d)^N)$ and $\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta)$ is in $\mathcal{C}^2(\mathbb{R}^p)$. It follows immediately that also $T^{i,N}(\theta, \hat{x}^N) \in \mathcal{C}^{2,\alpha}(\mathbb{R}^p, (\mathbb{R}^d)^N)$. Finally, the bound (5.363) follows from Lemma 5.8 and Lemma 5.17, via an application of Holdër's inequality. □

5.D Proof of Lemma for Theorem 5.4

Lemma 5.19. *Assume that Conditions A.1, B.1 - B.2, C.1, and D.1 hold. Let $i = 1, \dots, N$, and $N \in \mathbb{N}$. Then, for all $\theta \in \mathbb{R}^p$, there exists $K < \infty$ such that*

$$||\nabla_\theta \tilde{\mathcal{L}}(\theta) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta)|| \leq \frac{K}{N^{\frac{1}{2}}}\ , \quad a.s. \tag{5.364}$$

Proof. Let us define $g : \mathbb{R}^p \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ as the function which satisfies $G(\theta, x, \mu) = \int_{\mathbb{R}^d} g(\theta, x, y)\mu(\mathrm{d}y)$, where $G(\theta, x, \mu)$ is defined in (5.8). Thus, in particular,

$$g(\theta, x, y) = [b(\theta, x) + \phi(\theta, x, y)] - [b(\theta_0, x) + \phi(\theta_0, x, y)] \tag{5.365}$$

From the definition of $L(\theta, x, \mu)$, c.f. (5.9), we have that

$$\nabla_\theta L(\theta, x, \mu) = -\nabla_\theta^T G(\theta, x, \mu) G(\theta, x, \mu) = -\int_{\mathbb{R}^d \times \mathbb{R}^d} \nabla_\theta^T g(\theta, x, y) g(\theta, x, z) \mu(\mathrm{d}y) \mu(\mathrm{d}z) \tag{5.366}$$

We can thus define $l : \mathbb{R}^p \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^p$ as the function which satisfies $\nabla_\theta L(\theta, x, \mu) = \int_{\mathbb{R}^d \times \mathbb{R}^d} l(\theta, x, y, z) \mu(\mathrm{d}y) \mu(\mathrm{d}z)$. In particular, we identify

$$l(\theta, x, y, z) = -\nabla_\theta^T g(\theta, x, y) g(\theta, x, z). \tag{5.367}$$

We note that, via Condition C.1(ii), $l(\theta, \cdot, \cdot, \cdot)$ is locally Lipschitz with polynomial growth. That is, for all $x, x', y, y', z, z' \in \mathbb{R}^d$, we have

$$\begin{aligned} ||l(\theta, x, y, z) - l(\theta, x', y', z')|| \leq K & \left[[1 + ||x||^q + ||x'||^q + ||y||^q + ||y'||^q + ||z||^q + ||z'||^q \right] \\ & \cdot \left[||x - x'|| + ||y - y'|| + ||z - z'|| \right] \end{aligned} \tag{5.368}$$

In terms of this function, we can now write

$$\nabla_\theta L(\theta, x^i, \mu_\infty) = \int_{\mathbb{R}^d \times \mathbb{R}^d} l(\theta, x^i, x^j, x^k) \mu_\infty(\mathrm{d}x^j) \mu_\infty(\mathrm{d}x^k) \tag{5.369}$$

$$= \frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N \int_{\mathbb{R}^d \times \mathbb{R}^d} l(\theta, x^i, x^j, x^k) \mu_\infty(\mathrm{d}x^j) \mu_\infty(\mathrm{d}x^k) \tag{5.370}$$

$$\nabla_\theta L(\theta, x^{i,N}, \mu^N) = \frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N l(\theta, x^{i,N}, x^{j,N}, x^{k,N}). \tag{5.371}$$

where, in the second line, we have simply summed over the dummy variables x^j and x^k. and thus, from the definitions (see Lemmas 5.4.A - 5.4.B),

$$\nabla_\theta \tilde{\mathcal{L}}(\theta) = \frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N \int_{\mathbb{R}^d} \left[\int_{\mathbb{R}^d \times \mathbb{R}^d} l(\theta, x^i, x^j, x^k) \mu_\infty(\mathrm{d}x^j) \mu_\infty(\mathrm{d}x^k) \right] \mu_\infty(\mathrm{d}x^i) \tag{5.372}$$

$$= \frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N \int_{(\mathbb{R}^d)^N} l(\theta, x^i, x^j, x^k) \mu_\infty(\mathrm{d}x^1) \cdots \mu_\infty(\mathrm{d}x^N) \tag{5.373}$$

$$\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta) = \frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N \int_{(\mathbb{R}^d)^N} l(\theta, x^{i,N}, x^{j,N}, x^{k,N}) \hat{\mu}_\infty^N(\mathrm{d}\hat{x}). \tag{5.374}$$

where in the second line we have simply integrated with respect to the invariant probability measure μ_∞ over additional dummy variables, which does not change the value of the integral. Let $\hat{x}_t^N$ denote a solution of the IPS starting from $\hat{x}_0^N = (x_0^{1,N}, \ldots, x_0^{N,N})$, and let $x_t^{[N]}$ denote N independent solutions of the McKean-Vlasov SDE starting from $x_0^{[N]} = (x_0^1, \dots, x_0^N)$. Then, using the definition of an invariant measure, we can write

$$\nabla_\theta \tilde{\mathcal{L}}(\theta) = \frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N \int_{(\mathbb{R}^d)^3} \mathbb{E}_{(x_0^i, x_0^j, x_0^k)} \left[l(\theta, x^i, x^j, x^k) \right] \mu_\infty(\mathrm{d}x^1) \cdots \mu_\infty(\mathrm{d}x^N) \tag{5.375}$$

$$\nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta) = \frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N \int_{(\mathbb{R}^d)^N} \mathbb{E}_{(x_0^{i,N}, x_0^{j,N}, x_0^{k,N})} \left[l(\theta, x^{i,N}, x^{j,N}, x^{k,N}) \right] \hat{\mu}_\infty^N(\mathrm{d}\hat{x}). \tag{5.376}$$

Let $\pi^\infty \in \Pi(\hat{\mu}_\infty^N, \mu_\infty^{\otimes N})$ denote an arbitrary coupling of $\hat{\mu}_\infty^N$ and $\mu_\infty^{\otimes N}$. Then, using (5.375) - (5.376), it follows straightforwardly that

$$||\nabla_\theta \tilde{\mathcal{L}}(\theta) - \nabla_\theta \tilde{\mathcal{L}}^{i,N}(\theta)|| \tag{5.377}$$

$$\leq \frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N \int_{(\mathbb{R}^d)^N \times (\mathbb{R}^d)^N} \mathbb{E}_{(x_0^i, x_0^j, x_0^k, x_0^{i,N}, x_0^{j,N}, x_0^{k,N})} \tag{5.378}$$

$$\left[||l(\theta, x_s^i, x_s^j, x_s^k) - l(\theta, x_s^{i,N}, x_s^{j,N}, x_s^{k,N})|| \right] \pi^\infty(\mathrm{d}\hat{x}^N, \mathrm{d}x^{[N]})$$

Now, using the growth property (5.368) and Hölder's inequality, we obtain (now suppressing dependence of the expectation on the initial conditions)

$$\mathbb{E}\left[||l(\theta, x_s^i, x_s^j, x_s^k) - l(\theta, x_s^{i,N}, x_s^{j,N}, x_s^{k,N})|| \right] \tag{5.379}$$

$$\leq \left[1 + \mathbb{E}\left[||x_s^i||^{2q} \right]^{\frac{1}{2}} + \cdots + \left[\mathbb{E}||x_s^{k,N}||^{2q} \right]^{\frac{1}{2}} \right] \tag{5.380}$$

$$\cdot \left[\mathbb{E}\left[||x_s^i - x_s^{i,N}||^2 \right]^{\frac{1}{2}} + \mathbb{E}\left[||x_s^j - x_s^{j,N}||^2 \right]^{\frac{1}{2}} + \mathbb{E}\left[||x_s^k - x_s^{k,N}||^2 \right]^{\frac{1}{2}} \right]$$

$$\leq \frac{K}{N^{\frac{1}{2}}}, \tag{5.381}$$

where in the final line we have used Proposition 5.2 (the bounded moments of the McKean-Vlasov SDE and the IPS) and Proposition 5.5 (uniform in time propagation of chaos). Finally, substituting (5.379) - (5.381) into (5.378), the result follows. □

6 Conclusions

6.1 Final Remarks

Stochastic gradient descent is undoubtedly one of the most popular and widely applicable methods used in stochastic optimisation. In this thesis, we have made several contributions to both the theory and applications of stochastic gradient descent in continuous time. In both theory and application, our primary focus has been on two-timescale algorithms, which arise naturally in bilevel optimisation problems. We have also considered an important optimisation problem - recursive parameter estimation - arising in McKean-Vlasov SDEs. Let us briefly recap the main contributions of this thesis.

In Chapter 2, we analysed the asymptotic properties of two-timescale stochastic gradient descent in continuous time under general noise and stability conditions. We considered algorithms with both additive, state-dependent noise, and those with non-additive, state dependent noise. The obtained results cover a broad class of non-linear, two-timescale stochastic gradient descent algorithms in continuous time.

Chapter 3 considered the problem of joint online parameter estimation and optimal sensor placement for a partially observed diffusion process. We proposed a continuous-time, two-timescale stochastic gradient descent algorithm for this problem, which is both highly principled and computationally efficient. In particular, our method seeks to simultaneously maximise the asymptotic log-likelihood of the observations, and minimise the uncertainty in the state estimate. This approach is entirely novel: until now, the problems of recursive estimation and optimal sensor placement have been treated separately.

In Chapter 4, we demonstrated how to apply the joint online parameter estimation and optimal sensor placement algorithm to a dynamic spatio-temporal model governed by the stochastic advection-diffusion partial differential equation. Our results in this chapter not only demonstrate that our approach is highly effective in several scenarios of practical interest, but also illustrate the significant advantages of tackling the problems of parameter estimation and optimal sensor placement together.

Finally, in Chapter 5 we turned our attention to the problem of online parameter estimation for a stochastic McKean-Vlasov equation, and the associated system of interacting particles. We proposed a principled solution to this problem in the form of a continuous-time stochastic gradient descent algorithm, and provide a rigorous analysis of its asymptotic properties. For completeness, we also obtained asymptotic results for the offline maximum likelihood estimator. Our theoretical results were illustrated via several numerical examples, including a model commonly used in the study of opinion dynamics.

6.2 Future Work

There are several interesting avenues for future research based on the work presented in this thesis. Regarding the two-timescale stochastic gradient descent algorithm studied in Chapter 2, an important open problem is to obtain sufficient conditions for the assumption that the algorithm iterates remain a.s. bounded. While this assumption is necessary in order to prove almost sure convergence, it is generally far from automatic, and not very straightforward to establish. While methods for verifying stability are now relatively well understood in the single-timescale case (e.g., [44, 59, 265]), to our knowledge, there is only one existing result [281] along these lines in the two-timescale case. Currently, the most promising approaches to this task appear to be extensions of the randomly varying truncations method in [108], the stopping-times approach in [46, 420, 430], or the recent results in [346], to the two-timescale setting.

In terms of the joint online parameter estimation and optimal sensor placement algorithm considered in Chapter 3, the main open problem is to obtain conditions on the generative model (i.e., the partially observed diffusion process) which are easy to verify, sufficient for convergence, and not overly restrictive (see the discussion in Section 3.2.4.6). This problem is particularly challenging when the filter is approximate. In this case, even if the latent signal is ergodic, there is no guarantee that the filter is ergodic, let alone the tangent filter.

There are several important extensions to the work presented in Chapter 4. From a theoretical perspective, the main open problem is to obtain rigorous convergence results for the parameter estimates and optimal sensor placements generated by the finite-dimensional approximation of the joint online parameter estimation and optimal sensor placement algorithm to the stationary points of the true, infinite-dimensional asymptotic log-likelihood and asymptotic sensor placement objective function. This requires a careful extension of the results in Chapter 3 to include the case in which the latent signal process is infinite-dimensional. There are, in fact, several existing results along these lines for the optimal sensor placement (e.g., [80, 488]). The main obstacle, therefore, is to obtain corresponding results for the parameter estimates.

From a computational perspective, a natural extension of the numerical results in Chapter 4 is to consider the case in which the advection-diffusion operator is no longer spatially or temporally invariant, with the drift, diffusion, and damping parameters allowed to vary in space (see, e.g., [306]) or in time. It is also of interest to consider alternative, more complex spatial domains and boundary conditions, which are typical of environmental monitoring applications in urban settings. One may also be interested in considering other sensor configurations, perhaps allowing explicitly for the possibility of mobile sensors whose motion is governed by some controlled ODEs, or for differing levels of communication between sensors (e.g., [80, 156]). Finally, we would like to extend the results in this chapter to partially observed diffusion processes governed by non-linear dissipative SPDEs, such as the stochastic Navier-Stokes or Kuramoto-Sivashinski equations [68, 232]. In such models, of course, it will no longer be possible to compute the filter or tangent filter analytically, and it will be necessary to replace these quantities by suitable approximations (e.g., [47, 238]).

Regarding the results in Chapter 5, in the offline case a natural extension is to establish a non-asymptotic $\mathbb{L}^p$ convergence rate for the MLE in both the mean-field (large N) and long time (large T) regimes, extending the recent results in [110] to a more general class of IPSs. In the online case, it is of interest to obtain a central limit theorem for the recursive estimator, extending the results in [422] to non-linear McKean-Vlasov diffusions. One could also aim to extend our results to the case in which the diffusion coefficient is unknown, and must be estimated online (see [420] for online estimation of the diffusion coefficient in the linear case, and [221] for offline estimation of the diffusion coefficient in IPSs). Finally, there has been significant recent interest in weakly interacting diffusions on random graphs (e.g., [118, 145, 325]), for which the parameter estimation problem remains almost entirely unexplored, even in the offline case.

Bibliography

[1] J. A. Acebrón, L. L. Bonilla, C. J. Pérez Vicente, F. Ritort, and R. Spigler. The Kuramoto model: a simple paradigm for synchronization phenomena. *Reviews of Modern Physics*, 77(1):137–185, 2005.

[2] S. E. Aidarous, M. R. Gevers, and M. J. Installé. Optimal sensors' allocation strategies for a class of stochastic distributed systems. *International Journal of Control*, 22(2):197–213, 1975.

[3] S. E. Aidarous, M. R. Gevers, and M. J. Installé. On the asymptotic behaviour of sensors' allocation algorithm in stochastic distributed systems. In A. Ruberti, editor, *Distributed Parameter Systems: Modelling and Identification. Lecture Notes in Control and Information Science, Volume 1.* Springer-Verlag, Berlin, Heidelberg, 1978.

[4] S. I. Aihara. Regularized maximum likelihood estimate for an infinite-dimensional parameter in stochastic parabolic systems. *SIAM Journal on Control and Optimization*, 30(4):745–764, 1992.

[5] S. I. Aihara and Y. Sunahara. Identification of an infinite-dimensional parameter for stochastic diffusion equations. *SIAM Journal on Control and Optimization*, 26(5):1062–1075, 1988.

[6] S. I. Aihara. Maximum likelihood estimate for discontinuous parameter in stochastic hyperbolic systems. *Acta Applicandae Mathematica*, 35(1):131–151, 1994.

[7] S. I. Aihara and A. Bagchi. Parameter identification for hyperbolic stochastic systems. *Journal of Mathematical Analysis and Applications*, 160(2):485–499, 1991.

[8] A. E. Albert and L. A. Gardner. *Stochastic Approximation and Non Linear Regression.* MIT Press, Cambridge, MA, 1967.

[9] M. Amouroux, J. P. Babary, and C. Malandrakis. Optimal location of sensors for linear stochastic distributed parameter systems. In A. Ruberti, editor, *Distributed Parameter Systems: Modelling and Identification. Lecture Notes in Control and Information Science, Volume 1*, pages 92–113. Springer-Verlag, Berlin, Heidelberg, 1978.

[10] C. Andrieu, A. Doucet, and V. Z. B. Tadic. Online parameter estimation in general state-space models. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 332–337, New York, 2005. IEEE.

[11] M. Athans. On the determination of optimal costly measurement strategies for linear stochastic systems. *Automatica*, 8:397–412, 1972.

[12] A. C. Atkinson. Optimum experimental design. In M. Lovric, editor, *International Encyclopedia of Statistical Science*, pages 1037–1039. Springer-Verlag, Berlin, Heidelberg, 2011.

[13] F. Bach and E. Moulines. Non-strongly-convex smooth stochastic approximation with convergence rate O(1/n). In *Proceedings of the 26th International Conference on Neural Information Processing Systems (NIPS 2013)*, pages 773–781, Lake Tahoe, NV, 2013.

[14] A. Bagchi. Consistent estimates of parameters in continuous time systems. In O. L. R. Jacobs, M. H. A. Davis, M. A. H. Dempster, C. J. Harris, and P. C. Parks, editors, *Analysis and Optimization of Stochastic Systems*, pages 437–450. Academic Press, New York, 1980.

[15] A. Bagchi and V. Borkar. Parameter identification in infinite dimensional linear systems. In *20th IEEE Conference on Decision and Control including the Symposium on Adaptive Processes*, pages 62–66, 1981.

[16] A. Bagchi. Continuous time systems identification with unknown noise covariance. *IFAC Proceedings Volumes*, 8(1):311–318, 1975.

[17] A. Bagchi. A note on asymptotically efficient estimates of parameters in continuous-time dynamical systems. *SIAM Journal on Control and Optimization*, 19(2):169–176, 1981.

[18] A. Bagchi and V. S. Borkar. Parameter identification in infinte dimensional linear systems. *Stochastics*, 12(3-4):201–213, 1984.

[19] A. Bain and D. Crisan. *Fundamentals of Stochastic Filtering*. Springer-Verlag, New York, 2009.

[20] L. C. Baird. Reinforcement learning in continuous time: advantage updating. In *Proceedings of the 1994 IEEE International Conference on Neural Networks (ICNN 1994)*, volume 4, pages 2448–2453, Orlando, FL, 1994.

[21] J. Baladron, D. Fasoli, O. Faugeras, and J. Touboul. Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons. *The Journal of Mathematical Neuroscience*, 2(1):10, 2012.

[22] A. V. Balakrishnan. *Stochastic Differential Systems I*, volume 84 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, Berlin, Heidelberg, 1973.

[23] A. V. Balakrishnan. Identification and stochastic control of a class of distributed systems with boundary noise. In A. Bensoussan and J. Lions, editors, *Control Theory, Numerical Methods and Computer Systems Modelling*, pages 163–178. Springer-Verlag, New York, 1975.

[24] A. V. Balakrishnan. Identification inverse problems for partial differential equations: a stochastic formulation. In G. I. Marchuk, editor, *Optimization Techniques IFIP Technical Conference. Lecture Notes in Computer Science.*, pages 1–12. Springer-Verlag, Berlin, Heidelberg, 1975.

[25] F. Ball and D. Sirl. Stochastic SIR in structured populations. In T. Britton and E. Pardoux, editors, *Stochastic Epidemic Models with Inference, Part II*, pages 123–240. Springer-Verlag, Cham, Switzerland, 2020.

[26] M. Ballesio and A. Jasra. Unbiased estimation of the gradient of the log-likelihood for a class of continuous-time state-space models. *Monte Carlo Methods and Applications*, 28(1):61–83, 2022.

[27] H. T. Banks and K. Kunisch. The linear regulator problem for parabolic systems. *SIAM Journal on Control and Optimization*, 22(5):684–698, 1984.

[28] H. T. Banks and K. Kunisch. *Estimation Techniques for Distributed Parameter Systems*. Springer-Verlag, 2012.

[29] J. S. Baras and V. S. Borkar. A learning algorithm for Markov decision processes with adaptive state aggregation. In *Proceedings of the 39th IEEE Conference on Decision and Control*, volume 4, pages 3351–3356, 2000.

[30] J. S. Baras and A. Bensoussan. Optimal sensor scheduling in nonlinear filtering of diffusion processes. *SIAM Journal on Control and Optimization*, 27(4):786–813, 1989.

[31] P. Barrera and B. Spagnolo. A stochastic model for the advection-diffusion equation of air pollution in the atmosphere. *WIT Transactions on Ecology and the Environment*, 21, 1997.

[32] M. Bauer, T. Meyer-Brandis, and F. Proske. Strong solutions of mean-field stochastic differential equations with irregular drift. *Electronic Journal of Probability*, 23:1–35, 2018.

[33] J. Baumeister, W. Scondo, M. A. Demetriou, and I. G. Rosen. Online parameter estimation for infinite-dimensional dynamical systems. *SIAM Journal on Control and Optimization*, 35(2):678–713, 1997.

[34] S. Benachour, B. Roynette, D. Talay, and P. Vallois. Nonlinear self-stabilizing processes - I Existence, invariant probability, propagation of chaos. *Stochastic Processes and their Applications*, 75(2):173–201, 1998.

[35] M. Benaïm. A dynamical system approach to stochastic approximations. *SIAM Journal on Control and Optimization*, 34(2):437–472, 1996.

[36] M. Benaïm. Dynamics of stochastic approximation algorithms. In J. Azéma, M. Émery, M. Ledoux, and M. Yor, editors, *Séminaire de Probabilités XXXIII*, pages 1–68. Springer-Verlag, Berlin, Heidelberg, 1999.

[37] M. Benaïm, J. Hofbauer, and S. Sorin. Stochastic approximations and differential inclusions. *SIAM Journal on Control and Optimization*, 44(1):328–348, 2005.

[38] D. Benedetto, E. Caglioti, and M. Pulvirenti. A kinetic equation for granular media. *Mathematical Modelling and Numerical Analysis*, 31(5):615–641, 1997.

[39] V. E. Beneš. Exact finite-dimensional filters for certain diffusions with nonlinear drift. *Stochastics*, 5(1-2):65–92, 1981.

[40] V. E. Beneš. New exact nonlinear filters with large Lie algebras. *Systems & Control Letters*, 5(4):217–221, 1985.

[41] A. Bensoussan. Optimisation of sensors' location in a distributed filtering problem. In *Stability of Stochastic Dynamical Sysems. Lecture Notes in Mathematics*, volume 294, pages 62–84. Springer-Verlag, Berlin, Germany, 1972.

[42] A. Bensoussan, G. Da Prato, M. Delfour, and S. Mitter. *Representation and Control of Infinite Dimensional Systems*. Birkhäuser Basel, 2nd edition, 2007.

[43] A. Bensoussan. *Filtrage optimal des systèmes linéaires*. Dunod, Paris, 1971.

[44] A. Benveniste, M. Metivier, and P. Priouret. *Adaptive Algorithms and Stochastic Approximations*. Springer-Verlag, Berlin, Heidelberg, 1990.

[45] L. Bertini, G. Giacomin, and K. Pakdaman. Dynamical aspects of mean field plane rotators and the Kuramoto model. *Journal of Statistical Physics*, 138(1):270–290, 2010.

[46] D. P. Bertsekas and J. N. Tsitsiklis. Gradient convergence in gradient methods with errors. *SIAM Journal on Optimization*, 10(3):627–642, 2000.

[47] A. Beskos, D. Crisan, A. Jasra, N. Kantas, and H. Ruzayqat. Score-based parameter estimation for a class of continuous-time state space models. *SIAM Journal on Scientific Computing*, 43(4):2555–2580, 2021.

[48] S. Bhatnagar, M. C. Fu, S. I. Marcus, and P. J. Fard. Optimal structured feedback policies for ABR flow control using two-timescale SPSA. *IEEE/ACM Transactions on Networking*, 9(4):479–491, 2001.

[49] S. Bhatnagar, M. C. Fu, S. I. Marcus, and S. Bhatnagar. Two-timescale algorithms for simulation optimization of hidden Markov models. *IIE Transactions*, 33(3):245–258, 2001.

[50] T. Bhudisaksang and Á. Cartea. Online drift estimation for jump-diffusion processes. *Bernoulli*, 27(4):2494–2518, 2021.

[51] N. Bingham, C. Goldie, and J.L.Teugles. *Regular Variation*. Cambridge University Press, Cambridge, UK, 1987.

[52] J. P. N. Bishwal. *Parameter Estimation in Stochastic Differential Equations*. Springer-Verlag, Berlin, Heidelberg, 2008.

[53] J. P. N. Bishwal. Estimation in interacting diffusions: continuous and discrete sampling. *Applied Mathematics*, 2(9):1154–1158, 2011.

[54] V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis. On Krause's multi-agent consensus model with state-dependent connectivity. *IEEE Transactions on Automatic Control*, 54(11):2586–2597, 2009.

[55] J. R. Blum. Approximation methods which converge with probability one. *The Annals of Mathematical Statistics*, 25(2):382–386, 1954.

[56] J. R. Blum. Multidimensional stochastic approximation methods. *The Annals of Mathematical Statistics*, 25(4):737–744, 1954.

[57] F. Bolley, I. Gentil, and A. Guillin. Uniform convergence to equilibrium for granular media. *Archive for Rational Mechanics and Analysis*, 208(2):429–445, 2013.

[58] M. Bongini, M. Fornasier, M. Hansen, and M. Maggioni. Inferring interaction rules from observations of evolutive systems I: The variational approach. *Mathematical Models and Methods in Applied Sciences*, 27(05):909–951, 2017.

[59] V. S. Borkar and S. P. Meyn. The O.D.E. method for convergence of stochastic approximation and reinforcement learning. *SIAM Journal on Control and Optimization*, 38(2):447–469, 2000.

[60] V. S. Borkar and S. Pattathil. Concentration bounds for two time scale stochastic approximation. In *56th Annual Allerton Conference on Communication, Control, and Computing*, pages 504–511, 2018.

[61] V. S. Borkar. Stochastic approximation with two time scales. *Systems & Control Letters*, 29(5):291–294, 1997.

[62] V. S. Borkar. *Stochastic Approximation: A Dynamical Systems Viewpoint*. Cambridge University Press, Cambridge, UK, 2008.

[63] V. S. Borkar and A. Bagchi. Parameter estimation in continuous-time stochastic processes. *Stochastics*, 8(3):193–212, 1982.

[64] A. Borovykh, N. Kantas, P. Parpas, and G. Pavliotis. Optimizing interacting Langevin dynamics using spectral gaps. In *Proceedings of the 38th International Conference on Machine Learning (ICML 2021)*, 2021.

[65] L. Bottou. Large-scale machine learning with stochastic gradient descent. In *COMPSTAT 2010*, pages 177–186, Paris, France, 2010.

[66] L. Bottou, F. E. Curtis, and J. Nocedal. Optimization methods for large-scale machine learning. *SIAM Review*, 60(2):223–311, 2018.

[67] O. Brandière. The dynamic system method and the traps. *Advances in Applied Probability*, 30(1):137–151, 1998.

[68] D. Breit. An introduction to stochastic Navier-Stokes equations. In M. Bulíček, E. Feireisl, and M. Pokorný, editors, *New Trends and Results in Mathematical Description of Fluid Flows*, pages 1–51. Springer-Verlag, Cham, Switzerland, 2018.

[69] D. Brigo, B. Hanzon, and F. LeGland. A differential geometric approach to nonlinear filtering: the projection filter. *IEEE Transactions on Automatic Control*, 43(2):247–252, 1998.

[70] D. Brigo, B. Hanzon, and F. Le Gland. Approximate nonlinear filtering by projection of exponential manifolds of densities. *Bernoulli*, 5(3):495–534, 1999.

[71] P. E. Brown, P. J. Diggle, M. E. Lord, and P. C. Young. Spacetime calibration of radar rainfall data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 50(2):221–241, 2001.

[72] P. E. Brown, K. F. Karesen, G. O. Roberts, and S. Tonellato. Blur-generated nonseparable space-time models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(4):847–860, 2000.

[73] C. Brugna and G. Toscani. Kinetic models of opinion formation in the presence of personal conviction. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 92(5):52818, 2015.

[74] R. Buckdahn, J. Li, and J. Ma. A mean-field stochastic control problem with partial observations. *The Annals of Applied Probability*, 27(5):3201–3245, 2017.

[75] A. Budhiraja. Asymptotic stability, ergodicity and other asymptotic properties of the nonlinear filter. *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, 39(6):919–941, 2003.

[76] M. Burger, V. Capasso, and D. Morale. On an aggregation model with long and short range interactions. *Nonlinear Analysis: Real World Applications*, 8(3):939–958, 2007.

[77] J. A. Burns, E. M. Cliff, and C. Rautenberg. A distributed parameter control approach to optimal filtering and smoothing with mobile sensor networks. In *Proceedings of the 17th Mediterranean Conference on Control and Automation*, pages 181–186, Thessaloniki, Greece, 2009.

[78] J. A. Burns and C. Rautenberg. Solutions and approximations to the Riccati integral equation with values in a space of compact operators. *SIAM Journal on Control and Optimization*, 53(5):2846–2877, 2015.

[79] J. A. Burns and C. N. Rautenberg. Bochner integrable solutions to Riccati partial differential equations and optimal sensor placement. In *Proceedings of the 2011 American Control Conference*, pages 2368–2373, San Francisco, CA, 2011.

[80] J. A. Burns and C. N. Rautenberg. The infinite-dimensional optimal filtering problem with mobile and stationary sensor networks. *Numerical Functional Analysis and Optimization*, 36(2):181–224, 2015.

[81] A. G. Bustos. *Approximations des distributions d'equilibre de certains systems stochastiques avec interactions McKean-Vlasov.* Phd thesis, Universite de Nice Sophia Antipolis, 2008.

[82] M. Cameletti, F. Lindgren, D. Simpson, and H. Rue. Spatio-temporal modeling of particulate matter concentration through the SPDE approach. *AStA Advances in Statistical Analysis*, 97(2):109–131, 2013.

[83] F. Campillo and F. Le Gland. MLE for partially observed diffusions: direct maximization vs. the em algorithm. *Stochastic Processes and their Applications*, 33(2):245–274, 1989.

[84] J. R. Cannon and R. E. Klein. Optimal selection of measurement locations in a conductor for approximate determination of temperature distributions. *Journal of Dynamic Systems, Measurement, and Control*, 93(3):193–199, 1971.

[85] C. Canuto, F. Fagnani, and P. Tilli. An Eulerian approach to the analysis of Krause's consensus models. *SIAM Journal on Control and Optimization*, 50(1):243–265, 2012.

[86] O. Cappé. Online sequential Monte Carlo EM algorithm. In *2009 IEEE/SP 15th Workshop on Statistical Signal Processing*, pages 37–40, 2009.

[87] O. Cappé and E. Moulines. On-line expectation-maximization algorithm for latent data models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(3):593–613, 2009.

[88] O. Cappé, E. Moulines, and T. Ryden. *Inference in Hidden Markov Models*. Springer-Verlag, New York, 2005.

[89] P. Caravani, G. Di Pillo, and L. Grippo. Optimal location of a measurement point in a diffusion process. *IFAC Proceedings Volumes*, 8(1):62–68, 1975.

[90] A. Carè, B. C. Csáji, B. Gerencsér, L. Gerencsér, and M. Rásonyi. On the Poisson equation of parameter-dependent Markov chains. *arXiv preprint*, 2019.

[91] R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications I*. Springer-Verlag, Cham, Switzerland, 2018.

[92] J. A. Carrillo, R. S. Gvalani, G. A. Pavliotis, and A. Schlichting. Long-time behaviour and phase transitions for the McKean-Vlasov equation on the torus. *Archive for Rational Mechanics and Analysis*, 235(1):635–690, 2020.

[93] J. A. Carrillo, R. J. McCann, and C. D. Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Archive for Rational Mechanics and Analysis*, 179(2):217–263, 2006.

[94] J. A. Carrillo, R. McCann, and C. Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Revista Matematica Iberoamericana*, 19(3):971–1018, 2003.

[95] P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non-uniformly convex case. *Probability Theory and Related Fields*, 140(1):19–40, 2008.

[96] C. D. Charalambous, R. J. Elliott, and V. Krishnamurthy. Conditional moment generating functions for integrals and stochastic integrals. *SIAM Journal on Control and Optimization*, 42(5):1578–1603, 2003.

[97] C. D. Charalambous and A. Logothetis. Maximum likelihood parameter estimation from incomplete data via the sensitivity equations: the continuous-time case. *IEEE Transactions on Automatic Control*, 45(5):928–934, 2000.

[98] C. D. Charalambous and R. J. Elliott. Classes of nonlinear partially observable stochastic optimal control problems with explicit optimal control laws. *SIAM Journal on Control and Optimization*, 36(2):542–578, 1998.

[99] C. D. Charalambous and R. J. Elliott. New explicit filters and smoothers for diffusions with nonlinear drift and measurements. *Systems & Control Letters*, 33(2):89–103, 1998.

[100] P. E. Chaudru de Raynal. Strong well posedness of McKeanVlasov stochastic differential equations with Hölder drift. *Stochastic Processes and their Applications*, 130(1):79–107, 2020.

[101] B. Chazelle, Q. Jiu, Q. Li, and C. Wang. Well-posedness of the limiting equation of a noisy consensus model in opinion dynamics. *Journal of Differential Equations*, 263(1):365–397, 2017.

[102] H. F. Chen. Least squares identification for continuous-time systems. In A. Bensoussan and J. L. Lions, editors, *Analysis and Optimization of Systems*, pages 264–277. Springer-Verlag, Berlin, Heidelberg, 1980.

[103] H. F. Chen. On continuous-time stochastic approximation. In A. Bensoussan and J. L. Lions, editors, *Analysis and Optimization of Systems*, pages 203–214. Springer-Verlag, Berlin, Heidelberg, 1982.

[104] H. F. Chen. Continuous time stochastic approximation with measurement error being process of dependent increment. *Acta Mathematica Scientia*, 3(1):57–70, 1983.

[105] H. F. Chen. Asymptotic normality of the continuous-time stochastic approximation algorithm. *Acta Mathematicae Applicatae Sinica*, 1(1):31–43, 1984.

[106] H. F. Chen. Continuous-time stochastic approximation: Convergence and asymptotic efficiency. *Stochastics and Stochastic Reports*, 51(1-2):111–132, 1994.

[107] H. F. Chen. *Stochastic Approximation and Its Applications*. Springer-Verlag, Boston, MA, 2002.

[108] H. F. Chen, C. Wu, and Y. Zhu. Continuous-time stochastic approximation procedure with randomly varying truncations. *Acta Mathematica Scientia*, 7(1):43–55, 1987.

[109] W. Chen and J. H. Seinfeld. Optimal location of process measurements. *International Journal of Control*, 21(6):1003–1014, 1975.

[110] X. Chen. Maximum likelihood estimation of potential energy in interacting particle systems from single-trajectory data. *Electronic Communications in Probability*, 26:1–13, 2021.

[111] T.-S. Chiang, C.-R. Hwang, and S. J. Sheu. Diffusion for global optimization in $\mathbb{R}^n$. *SIAM Journal on Control and Optimization*, 25(3):737–753, 1987.

[112] K. L. Chung. On a stochastic approximation method. *The Annals of Mathematical Statistics*, 25(3):463–483, 1954.

[113] I. Cialenco. Statistical inference for SPDEs: an overview. *Statistical Inference for Stochastic Processes*, 21(2):309–329, 2018.

[114] G. Colantuoni and L. Padmanabhan. Optimal sensor selection in sequential estimation problems. *International Journal of Control*, 28(6):821–845, 1978.

[115] I. B. Collings, V. Krishnamurthy, and J. B. Moore. On-line identification of hidden Markov models via recursive prediction error techniques. *IEEE Transactions on Signal Processing*, 42(12):3535–3539, 1994.

[116] I. B. Collings and T. Ryden. A new maximum likelihood gradient algorithm for on-line hidden Markov model identification. In *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 4, pages 2261–2264, 1998.

[117] F. Coppini. *Weakly interacting diffusions on graphs.* Phd thesis, Universite de Paris, 2020.

[118] F. Coppini, H. Dietert, and G. Giacomin. A law of large numbers and large deviations for interacting diffusions on Erdös-Rényi graphs. *Stochastics and Dynamics*, 20(2), 2019.

[119] S. L. Corff, G. Fort, and E. Moulines. Online expectation maximization algorithm to solve the SLAM problem. In *2011 IEEE Statistical Signal Processing Workshop (SSP)*, pages 225–228, 2011.

[120] I. D. Couzin, J. Krause, N. R. Franks, and S. A. Levin. Effective leadership and decision-making in animal groups on the move. *Nature*, 433(7025):513–516, 2005.

[121] D. Cox and N. Reid. *The Theory of the Design of Experiments.* Chapman and Hall/CRC, London, UK, 2000.

[122] N. Cressie and H. Huang. Classes of nonseparable, spatio-temporal stationary covariance functions. *Journal of the American Statistical Association*, 94(1):1330–1340, 1999.

[123] N. Cressie and C. K. Wikle. Space-time Kalman filter. In A. E. Shaarawi and W. Piegorsch, editors, *Encyclopedia of Environmetrics*, pages 2045–2049. Wiley & Sons, Chichester, UK, 2002.

[124] N. Cressie. *Statistics for Spatial Data.* John Wiley & Sons, Hoboken, NJ, 1993.

[125] N. Cressie and C. K. Wikle. *Statistics for Spatio-Temporal Data.* John Wiley & Sons, Hoboken, NJ, 2011.

[126] D. Crisan and J. Xiong. Approximate McKean-Vlasov representations for a class of SPDEs. *Stochastics*, 82(1):53–68, 2010.

[127] R. F. Curtain. A survey of infinite-dimensional filtering. *SIAM Review*, 17(3):395–411, 1975.

[128] R. F. Curtain, A. Ichikawa, and E. Ryan. Optimal location of sensors for filtering for distributed systems. In A. Ruberti, editor, *Distributed Parameter Systems: Modelling and Identification*, pages 236–255. Springer-Verlag, Berlin, Heidelberg, 1978.

[129] R. F. Curtain and A. J. Pritchard. *Infinite Dimensional Linear Systems Theory.* Springer-Verlag, Berlin, Heidelberg, 1978.

[130] R. F. Curtain and H. Zwart. *An Introduction to Infinite-Dimensional Linear Systems Theory.* Springer-Verlag, New York, 1995.

[131] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions.* Cambridge University Press, Cambridge, UK, 2nd edition, 2014.

[132] T. E. Dabbous and N. U. Ahmed. Parameter identification for partially observed diffusions. *Journal of Optimization Theory and Applications*, 75(1):33–50, 1992.

[133] G. Dalal, B. Szörényi, and G. Thoppe. A tale of two-timescale reinforcement learning with the tightest finite-time bound. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(4):3701–3708, 2020.

[134] G. Dalal, G. Thoppe, B. Szörényi, and S. Mannor. Finite sample analysis of two-timescale stochastic approximation with applications to reinforcement learning. In *Proceedings of the 31st Conference On Learning Theory (COLT)*, volume 75, pages 1199–1233, 2018.

[135] N. Darivandi, K. Morris, and A. Khajepour. An algorithm for LQ optimal actuator location. *Smart Materials and Structures*, 22(3):35001, 2013.

[136] F. Daum. Exact finite-dimensional nonlinear filters. *IEEE Transactions on Automatic Control*, 31(7):616–622, 1986.

[137] F. Daum. Solution of the Zakai equation by separation of variables. *IEEE Transactions on Automatic Control*, 32(10):941–943, 1987.

[138] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 2: Functional and Variational Methods.* Springer-Verlag, Berlin, Heidelberg, 2000.

[139] D. A. Dawson. Critical dynamics and fluctuations for a mean-field model of cooperative behavior. *Journal of Statistical Physics*, 31(1):29–85, 1983.

[140] A. De Santis, A. Germani, and L. Jetto. Approximation of the algebraic Riccati equation in the Hilbert space of Hilbert-Schmidt operators. *SIAM Journal on Control and Optimization*, 31(4):847–874, 1993.

[141] J. de Wiljes, S. Reich, and W. Stannat. Long-time stability and accuracy of the ensemble Kalman-Bucy filter for fully observed processes and small measurement noise. *SIAM Journal on Applied Dynamical Systems*, 17(2):1152–1181, 2018.

[142] P. Del Moral and L. Miclo. Branching and interacting particle systems approximations of feynman-kac formulae with applications to non-linear filtering. In J. Azéma, M. Ledoux, M. Émery, and M. Yor, editors, *Séminaire de Probabilités XXXIV*, pages 1–145. Springer-Verlag, Berlin, Heidelberg, 2000.

[143] P. Del Moral and J. Tugaut. On the stability and the uniform propagation of chaos properties of Ensemble Kalman-Bucy filters. *The Annals of Applied Probability*, 28(2):790–850, 2018.

[144] P. Del Moral, A. Kurtzmann, and J. Tugaut. On the stability and the concentration of extended Kalman-Bucy filters. *Electronic Journal of Probability*, 23:1–30, 2018.

[145] S. Delattre, G. Giacomin, and E. Luçon. A note on dynamical models on random graphs and Fokker-Planck equations. *Journal of Statistical Physics*, 165(4):785–798, 2016.

[146] L. Della Maestra and M. Hoffmann. Nonparametric estimation for interacting particle systems: McKean-Vlasov models. *Probability Theory and Related Fields*, 182(1):551–613, 2022.

[147] B. Delyon. General results on the convergence of stochastic algorithms. *IEEE Transactions on Automatic Control*, 41(9):1245 – 1255, 1996.

[148] A. Dembo and O. Zeitouni. Parameter estimation of partially observed continuous time stochastic processes via the EM algorithm. *Stochastic Processes and their Applications*, 23(1):91–113, 1986.

[149] M. A. Demetriou. Integrated actuator-sensor placement and hybrid controller design of flexible structures under worst case spatiotemporal disturbance variations. *Journal of Intelligent Material Systems and Structures*, 15(12):901–921, 2004.

[150] M. A. Demetriou. Power management of sensor networks for detection of a moving source in 2-D spatial domains. In *Proceedings of the 2006 American Control Conference*, pages 1144–1149, Minneapolis, MN, 2006.

[151] M. A. Demetriou. Guidance of mobile actuator-plus-sensor networks for improved control and estimation of distributed parameter systems. *IEEE Transactions on Automatic Control*, 55(7):1570–1584, 2010.

[152] M. A. Demetriou. Adaptive control of 2-D PDEs using mobile collocated actuator / sensor pairs with augmented vehicle dynamics. *IEEE Transactions on Automatic Control*, 57(12):2979–2993, 2012.

[153] M. A. Demetriou. Enforcing consensus on adaptive parameter estimation of structurally perturbed infinite dimensional systems. *IEEE Transactions on Automatic Control*, 57(12):3147–3152, 2012.

[154] M. A. Demetriou. Gradient-ascent schemes for the guidance of mobile sensors used in the estimation of spatially distributed systems. In *Proceedings of the 2015 American Control Conference*, pages 553–558, Chicago, IL, 2015.

[155] M. A. Demetriou. Closed-loop guidance of mobile sensors for the estimation of spatially distributed processes. In *Proceedings of the 2018 American Control Conference*, pages 931–936, Milwaukee, VI, 2018.

[156] M. A. Demetriou and I. Hussein. Estimation of spatially distributed processes using mobile spatially distributed sensor network. *SIAM Journal on Control and Optimization*, 48(1):266–291, 2009.

[157] D. G. DeWolf and D. M. Wiberg. An ordinary differential equation technique for continuous-time parameter estimation. *IEEE Transactions on Automatic Control*, 38(4):514–528, 1993.

[158] T. T. Doan and J. Romberg. Linear two-timescale stochastic approximation: a finite-time analysis. In *57th Annual Allerton Conference on Communication, Control, and Computing*, pages 399–406, Monticello, IL, 2019.

[159] T. T. Doan. Finite-time analysis and restarting scheme for linear two-timescale stochastic approximation. *SIAM Journal on Control and Optimization*, 59(4):2798–2819, 2021.

[160] T. T. Doan. Finite-time convergence rates of nonlinear two timescale stochastic approximation under Markovian noise. *arXiv preprint*, 2021.

[161] T. T. Doan. Nonlinear two-timescale stochastic approximation: convergence and finite-time performance. In *Proceedings of the 3rd Conference on Learning for Dynamics and Control*, volume 144, 2021.

[162] T. T. Doan, C. L. Beck, and R. Srikant. On the convergence rate of distributed gradient methods for finite-sum optimization under communication delays. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 1(2), 2017.

[163] C. Doléans-Dade. On the existence and unicity of solutions of stochastic integral equations. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 36(2):93–101, 1976.

[164] J. L. Doob. *Stochastic Processes.* Wiley, New York, 1953.

[165] A. Doucet and V. Z. B. Tadić. Parameter estimation in general state-space models using particle methods. *Annals of the Institute of Statistical Mathematics*, 55(2):409–422, 2003.

[166] K. Doya. Reinforcement Learning in Continuous Time and Space. *Neural Computation*, 12(1):219–245, 2000.

[167] M. Driml and J. Nedoma. Stochastic approximations for continuous random functions. In *Transactions of the Second Prague Conference on Information Theory*, pages 145–158, New York, 1961. Academic Press.

[168] J. Duchi, E. Hazan, and Y. Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12(61):2121–2159, 2011.

[169] M. Duflo. *Algorithmes stochastiques.* Springer-Verlag, Berlin, Heidelberg, 1996.

[170] A. Durmus, A. Eberle, A. Guillin, and R. Zimmer. An elementary approach to uniform in time propagation of chaos. *Proceedings of the American Mathematical Society*, 148:5387–5398, 2020.

[171] A. Eberle, A. Guillin, and R. Zimmer. Quantitative Harris-type theorems for diffusions and McKean-Vlasov processes. *Transactions of the American Mathematical Society*, 371:7135–7173, 2019.

[172] R. J. Elliott, J. J. Ford, and J. B. Moore. On-line consistent estimation of hidden Markov models. Technical report, Department of Systems Engineering, Australian National University, Canberra, Autralia, 2000.

[173] R. J. Elliott and V. Krishnamurthy. Exact finite-dimensional filters for maximum likelihood parameter estimation of continuous-time linear Gaussian systems. *SIAM Journal on Control and Optimization*, 35(6):1908–1923, 1997.

[174] M. Emery. Stabilité des solutions des équations différentielles stochastiques application aux intégrales multiplicatives stochastiques. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 41(3):241–262, 1978.

[175] V. Fabian. On asymptotic normality in stochastic approximation. *The Annals of Mathematical Statistics*, 39(4):1327–1332, 1968.

[176] P. L. Falb. Infinite-dimensional filtering: The Kalman-Bucy filter in Hilbert space. *Information and Control*, 11(1):102–137, 1967.

[177] Y. Feng, T. Gao, L. Li, J.-G. Liu, and Y. Lu. Uniform-in-time weak error analysis for stochastic gradient descent algorithms via diffusion approximation. *Communications in Mathematical Sciences*, 18(1):163–188, 2020.

[178] Z. G. Feng, K. L. Teo, and V. Rehbock. Optimal sensor scheduling in continuous time. *Dynamic Systems and Applications*, 17:331–350, 2008.

[179] T. Flury and N. Shephard. Bayesian inference based only on simulated likelihood: particle filter analysis of dynamic economic models. *Econometric Theory*, 27(5):933–956, 2011.

[180] J. J. Ford. *Adaptive hidden Markov model estimation and applications*. Phd thesis, Australian National University, Canberra, 1998.

[181] J. J. Ford and J. B. Moore. On adaptive HMM state estimation. *IEEE Transactions on Signal Processing*, 46(2):475–486, 1998.

[182] L. Franceschi, P. Frasconi, S. Salzo, R. Grazzi, and M. Pontil. Bilevel programming for hyperparameter optimization and meta-learning. In *Proceedings of the 35th International Conference on Machine Learning (ICML 2018)*, pages 1563–1572, Stockholm, Sweden, 2018.

[183] T. Frank. *Nonlinear Fokker-Planck Equations*. Springer-Verlag, Berlin, Heidelberg, 2005.

[184] H. Frydman and P. Lakner. Maximum likelihood estimation of hidden Markov processes. *The Annals of Applied Probability*, 13(4):1296–1312, 2003.

[185] M. C. Fu. *Handbook of Simulation Optimization*. Springer-Verlag, New York, 2015.

[186] M. Fuentes and A. E. Raftery. Model evaluation and spatial interpolation by Bayesian combination of observations with outputs from numerical models. *Biometrics*, 61(1):36–45, 2005.

[187] T. Funaki. A certain class of diffusion processes associated with nonlinear parabolic equations. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 67(3):331–348, 1984.

[188] L. I. Gal'chuk. Existence and uniqueness of a solution for stochastic equations with respect to semimartingales. *Theory of Probability & Its Applications*, 23(4):751–763, 1979.

[189] R. Ge, F. Huang, C. Jin, and Y. Yuan. Escaping from saddle points - online stochastic gradient for tensor decomposition. In *Proceedings of the 28th Annual Conference on Learning Theory (COLT)*, 2015.

[190] A. E. Gelfand, S. Banerjee, and D. Gamerman. Spatial process modelling for univariate and multivariate dynamic spatial data. *Environmetrics*, 16(5):465–479, 2005.

[191] S. B. Gelfand and S. K. Mitter. Simulated annealing type algorithms for multivariate optimization. *Algorithmica*, 6(1):419–436, 1991.

[192] S. Geman and C.-R. Hwang. Diffusions for global optimization. *SIAM Journal on Control and Optimization*, 24(5):1031–1043, 1986.

[193] V. Genon-Catalot and J. Jacod. Estimation of the diffusion coefficient for diffusion processes: random sampling. *Scandinavian Journal of Statistics*, 21(3):193–221, 1994.

[194] V. Genon-Catalot and C. Larédo. Parametric inference for small variance and long time horizon McKean-Vlasov diffusion models. *Electronic Journal of Statistics*, 15(2):5811–5854, 2021.

[195] V. Genon-Catalot and C. Larédo. Probabilistic properties and parametric inference of small variance nonlinear self-stabilizing stochastic differential equations. *Stochastic Processes and their Applications*, 142:513–548, 2021.

[196] L. Gerencsér, I. Gyöngy, and G. Michaletzky. Continuous-time recursive maximum likelihood method: a new approach to Ljung's scheme. *IFAC Proceedings Volumes*, 17(2):683–686, 1984.

[197] L. Gerencsér and V. Prokaj. Recursive identification of continuous-time linear stochastic systems - convergence w.p.1 and in Lq. In *Proceedings of the 2009 European Control Conference (ECC)*, pages 1209–1214, 2009.

[198] A. Germani, L. Jetto, and M. Piccioni. Galerkin approximation for optimal linear filtering of infinite-dimensional linear systems. *SIAM Journal on Control and Optimization*, 26(6):1287–1305, 1988.

[199] S. Ghadimi and G. Lan. Stochastic approximation methods and their finite-time convergence properties. In *Handbook of Simulation Optimization.* Springer-Verlag, New York, 2015.

[200] S. Ghadimi and M. Wang. Approximation methods for bilevel programming. *arXiv preprint*, 2018.

[201] K. Giesecke, G. Schwenkler, and J. A. Sirignano. Inference for large financial systems. *Mathematical Finance*, 30(1):3–46, 2020.

[202] T. Gneiting. Nonseparable, stationary covariance functions for space-time data. *Journal of the American Statistical Association*, 97(458):590–600, 2002.

[203] T. Gneiting, M. G. Genton, and P. Guttorp. Geostatistical space-time models, stationarity, separability, and full symmetry. In B. Finkenstadt, L. Held, and V. Isham, editors, *Statistical Methods for Spatio-Temporal Systems*, pages 151–175. Chapman and Hall/CRC, Boca Raton, FL, 2006.

[204] S. N. Gomes and G. A. Pavliotis. Mean field limit for interacting diffusions in a two-scale potential. *Journal of Nonlinear Science*, 28(3):905–941, 2018.

[205] S. N. Gomes, A. M. Stuart, and M.-T. Wolfram. Parameter estimation for macroscopic pedestrian dynamics models from microscopic data. *SIAM Journal on Applied Mathematics*, 79(4):1475–1500, 2019.

[206] H. Gupta, R. Srikant, and L. Ying. Finite-time performance bounds and adaptive learning rate selection for two time-scale reinforcement learning. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems (NeurIPS 2019)*, pages 4706–4715, Vancouver, Canada, 2019.

[207] M. S. Handcock and M. L. Stein. A Bayesian analysis of kriging. *Technometrics*, 35(4):403–410, 1993.

[208] U. G. Haussmann and E. Pardoux. A conditionally almost linear filtering problem with non-gaussian initial condition. *Stochastics*, 23(2):241–275, 1988.

[209] R. Hegselmann and U. Krause. Opinion dynamics and bounded confidence models, analysis and simulation. *Journal of Artificial Societies and Social Simulation*, 5(3), 2002.

[210] J. Heng, J. Houssineau, and A. Jasra. On unbiased score estimation for partially observed diffusions. *arXiv preprint*, 2021.

[211] K. Herring and J. Melsa. Optimum measurements for estimation. *IEEE Transactions on Automatic Control1*, 19(3):264–266, 1974.

[212] S. Herrmann and J. Tugaut. Non-uniqueness of stationary measures for self-stabilizing processes. *Stochastic Processes and their Applications*, 120(7):1215–1246, 2010.

[213] S. Herrmann and J. Tugaut. Self-stabilizing processes: uniqueness problem for stationary measures and convergence rate in the small-noise limit. *ESAIM: Probability and Statistics*, 16:277–305, 2012.

[214] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter. GANs trained by a two timescale update rule converge to a local nash equilibrium. In *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS 2017)*, pages 6629–6640, Red Hook, NY, 2017.

[215] C. C. Heyde. A quasi-likelihood approach to estimating parameters in diffusion-type processes. *Journal of Applied Probability*, 31:283–290, 1994.

[216] M. Hintermüller, C. Rautenberg, M. Mohammadi, and M. Kanitsar. Optimal sensor placement: a robust approach. *SIAM Journal on Control and Optimization*, 55(6):3609–3639, 2017.

[217] M. W. Hirsch. *Differential Topology.* Springer-Verlag, New York, 1976.

[218] M. Hong, H.-T. Wai, Z. Wang, and Z. Yang. A two-timescale framework for bilevel optimization: complexity analysis and application to actor-critic. *arXiv preprint*, 2020.

[219] M. B. Hooten, M. J. Garlick, and J. A. Powell. Computationally efficient statistical differential equation modeling using homogenization. *Journal of Agricultural, Biological, and Environmental Statistics*, 18(3):405–428, 2013.

[220] M. B. Hooten and C. K. Wikle. A hierarchical Bayesian non-linear spatio-temporal model for the spread of invasive species with application to the Eurasian Collared-Dove. *Environmental and Ecological Statistics*, 15(1):59–70, 2008.

[221] X. Huang and F.-Y. Wang. Distribution dependent SDEs with singular coefficients. *Stochastic Processes and their Applications*, 129(11):4747–4770, 2019.

[222] M. Huebner. A characterization of asymptotic behaviour of maximum likelihood estimators for stochastic PDEs. *Mathematical Methods of Statistics*, 6(4):395–415, 1997.

[223] M. Huebner. Asymptotic properties of the maximum likelihood estimator for stochastic PDEs distributed by small noise. *Statistical Inference for Stochastic Processes*, 2(1):57–68, 1999.

[224] M. Huebner, R. Khasminskii, and B. L. Rozovskii. Two examples of parameter estimation for stochastic partial differential equations. In S. Cambanis, J. K. Ghosh, R. L. Karandikar, and P. K. Sen, editors, *Stochastic Processes: A Festschrift in Honour of Gopinath Kallianpur*, pages 149–160. Springer-Verlag, New York, 1993.

[225] M. Huebner, S. Lototsky, and B. L. Rozovskii. Asymptotic properties of an approximate maximum likelihood estimator for stochastic PDEs. In Y. M. Kabanov, B. L. Rozovskii, and A. N. Shiryaev, editors, *Statistics and Control of Stochastic Processes*, pages 139–155. World Scientific, Singapore, 1997.

[226] M. Huebner and B. L. Rozovskii. On asymptotic properties of maximum likelihood estimators for parabolic stochastic PDE's. *Probability Theory and Related Fields*, 103(2):143–163, 1995.

[227] G. Huerta, B. Sansó, and J. R. Stroud. A spatiotemporal model for Mexico City ozone levels. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 53(2):231–248, 2004.

[228] N. Ikeda and S. Watanabe. A comparison theorem for solutions of stochastic differential equations and its applications. *Osaka Journal of Mathematics*, 14(3):619–633, 1977.

[229] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*. North-Holland Publishing Company, Amsterdam, The Netherlands, 2nd edition, 1989.

[230] K. Ito and K. Xiong. Gaussian filters for nonlinear filtering problems. *IEEE Transactions on Automatic Control*, 45(5):910–927, 2000.

[231] M. R. James and F. Le Gland. Consistent parameter estimation for partially observed diffusions with small noise. *Applied Mathematics and Optimization*, 32(1):47–72, 1995.

[232] M. Jardak, I. M. Navon, and M. Zupanski. Comparison of sequential data assimilation methods for the Kuramoto Sivashinsky equation. *International Journal for Numerical Methods in Fluids*, 62(4):374–402, 2010.

[233] S. Joshi and S. Boyd. Sensor selection via convex optimization. *IEEE Transactions on Signal Processing*, 57(2):451–462, 2009.

[234] B. Jourdain, S. Meleard, and W. A. Woyczynski. Nonlinear SDEs driven by Levy processes and related PDEs. *Alea*, 4:1–29, 2008.

[235] M. Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 3: Contributions to Astronomy and Physics*, pages 171–197, Berkeley, CA, 1956.

[236] M. Kaledin, E. Moulines, A. Naumov, V. Z. B. Tadic, and H.-T. Wai. Finite time analysis of linear two-timescale stochastic approximation with Markovian noise. In *Proceedings of 33rd Conference on Learning Theory (COLT '20)*, pages 2144–2203, 2020.

[237] R. E. Kalman and R. S. Bucy. New results in linear filtering and prediction theory. *Journal of Basic Engineering*, 83(1):95–103, 1961.

[238] N. Kantas, A. Doucet, S. S. Singh, J. Maciejowski, and N. Chopin. On particle methods for parameter estimation in state-space models. *Statistical Science*, 30(3):328–351, 2015.

[239] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer-Verlag, New York, 2nd edition, 1998.

[240] B. Karimi, B. Miasojedow, E. Moulines, and H.-T. Wai. Non-asymptotic analysis of biased stochastic approximation scheme. In *Proceedings of the 32nd Conference on Learning Theory (COLT '19)*, pages 1944–1974, Phoenix, AZ, 2019.

[241] P. Karmakar and S. Bhatnagar. Two time-scale stochastic approximation with controlled Markov noise and off-policy temporal-difference learning. *Mathematics of Operations Research*, 43(1):130–151, 2018.

[242] R. A. Kasonga. Maximum likelihood theory for large interacting systems. *SIAM Journal on Applied Mathematics*, 50(3):865–875, 1990.

[243] R. Khasminskii. *Stochastic Stability of Differential Equations.* Springer-Verlag, Berlin, Heidelberg, 2 edition, 2012.

[244] W. Khreich, E. Granger, A. Miri, and R. Sabourin. A survey of techniques for incremental learning of HMM parameters. *Information Sciences*, 197:105–130, 2012.

[245] J. Kiefer and J. Wolfowitz. Stochastic estimation of the maximum of a regression function. *The Annals of Mathematical Statistics*, 23(3):462–466, 1952.

[246] J. Kim, J. Shin, and I. Yang. Hamilton-Jacobi deep Q-learning for deterministic continuous-time systems with Lipschitz continuous controls. *Journal of Machine Learning Research*, 22(206):1–34, 2021.

[247] Y. Kim. Parameter estimation for an infinite dimensional stochastic differential equation. *Journal of the Korean Statistical Society*, 25(2):161–173, 1996.

[248] D. Kingma and J. Ba. Adam: a method for stochastic optimisation. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR '15)*, pages 1–13, San Diego, CA, 2015.

[249] V. R. Konda and J. N. Tsitsiklis. Linear stochastic approximation driven by slowly varying Markov chains. *Systems & Control Letters*, 50(2):95–102, 2003.

[250] V. R. Konda and J. N. Tsitsiklis. On actor-critic algorithms. *SIAM Journal on Control and Optimization*, 42(4):1143–1166, 2003.

[251] V. R. Konda and J. N. Tsitsiklis. Convergence rate of linear two-time-scale stochastic approximation. *The Annals of Applied Probability*, 14(2):796–819, 2004.

[252] V. R. Konda and V. S. Borkar. Actor-critic type learning algorithms for Markov decision processes. *SIAM Journal on Control and Optimization*, 38(1):94–123, 1999.

[253] F. Konecny. Maximum likelihood estimation of a drift parameter from a partially observed diffusion in the case of small measurement noise. *Statistics & Risk Modeling*, 8(2):115–130, 1990.

[254] J. Korbicz and D. Uciński. Sensor allocation for state and parameter estimation of distributed systems. In W. Gutkowski and J. Bauer, editors, *Discrete Structural Optimization*, pages 178–189. Springer-Verlag, Berlin, Heidelberg, 1994.

[255] T. Koski and W. Loges. On minimum-contrast estimation for hilbert space-valued stochastic differential equations. *Stochastics*, 16(3-4):217–225, 1986.

[256] L. C. Kramer. *On stochastic control and optimal measurement strategies*. Phd thesis, Massachusetts Institute of Technology, 1971.

[257] U. Krause. A discrete nonlinear and non-autonomous model of consensus formation. In *Communications in Difference Equations*, pages 227–236. Gordon and Breach, Amsterdam, The Netherlands, 2000.

[258] V. Krishnamurthy and G. G. Yin. Recursive algorithms for estimation of hidden Markov models and autoregressive models with Markov regime. *IEEE Transactions on Information Theory*, 48(2):458–476, 2002.

[259] C. S. Kubrusly. Distributed parameter systems identification: a survey. *International Journal of Control*, 26(4):509–535, 1977.

[260] S. Kulkarni and C. Horn. An alternative proof for convergence of stochastic approximation algorithms. *IEEE Transactions on Automatic Control*, 41(3):419 – 424, 1996.

[261] P. Kumar, T. E. Unny, and K. Ponnambalam. Stochastic partial differential equations in groundwater hydrology. *Stochastic Hydrology and Hydraulics*, 5(3):239–251, 1991.

[262] S. Kumar and J. Seinfeld. Optimal location of measurements for distributed parameter estimation. *IEEE Transactions on Automatic Control*, 23(4):690–698, 1978.

[263] H. Kunita. Ergodic properties of nonlinear filtering processes. In K. S. Alexander and J. C. Watkins, editors, *Spatial Stochastic Processes*, pages 233–256. Birkhäuser Boston, Boston, MA, 1991.

[264] Y. Kuramoto. Rhythms and turbulence in populations of chemical oscillators. *Physica A: Statistical Mechanics and its Applications*, 106(1):128–143, 1981.

[265] H. J. Kushner and G. G. Yin. *Stochastic approximation algorithms and applications.* Springer-Verlag, New York, 1997.

[266] H. J. Kushner. On the differential equations satisfied by conditional probability densities of Markov processes, with applications. *Journal of the Society for Industrial and Applied Mathematics: Series A (Control)*, 2(1):106–119, 1964.

[267] H. J. Kushner and H. Huang. Rates of convergence for stochastic approximation type algorithms. *SIAM Journal on Control and Optimization*, 17(5):607–617, 1979.

[268] H. J. Kushner and A. Shwartz. An invariant measure approach to the convergence of stochastic approximations with state dependent noise. *SIAM Journal on Control and Optimization*, 22(1):13–27, 1984.

[269] H. J. Kushner and F. J. Vázquez-Abad. Stochastic approximation methods for systems over an infinite horizon. *SIAM Journal on Control and Optimization*, 34(2):712–756, 1996.

[270] H. J. Kushner and J. Yang. Stochastic approximation with averaging of the iterates: optimal asymptotic rate of convergence for general processes. *SIAM Journal on Control and Optimization*, 31(4):1045–1062, 1993.

[271] H. J. Kushner and G. G. Yin. *Stochastic approximation and recursive algorithms and applications.* Springer-Verlag, New York, 2003.

[272] H. J. Kushner and D. S. Clark. *Stochastic approximation methods for constrained and unconstrained systems.* Springer-Verlag, New York, 1978.

[273] Y. A. Kutoyants. *Parameter estimation for stochastic processes.* Heldermann, 1984.

[274] Y. A. Kutoyants. *Identification of Dynamical Systems with Small Noise.* Springer-Verlag, Dordrecht, 1994.

[275] Y. A. Kutoyants. Parameter estimation for continuous time hidden Markov processes. *Automation and Remote Control*, 81(3):445–468, 2020.

[276] Y. A. Kutoyants. *Statistical Inference for Ergodic Diffusion Processes.* Springer-Verlag, London, 2004.

[277] Y. A. Kutoyants. On parameter estimation of the hidden OrnsteinUhlenbeck process. *Journal of Multivariate Analysis*, 169:248–263, 2019.

[278] Y. A. Kutoyants and L. Zhou. On parameter estimation of the hidden Gaussian process in perturbed SDE. *Electronic Journal of Statistics*, 15(1):211–234, 2021.

[279] N. Kuznetsov, R. Liptser, and A. Serebrovskii. Optimal control and data processing in continuous time (linear system and quadratic functional). *Automation and Remote Control*, 41:1369–1374, 1980.

[280] P. C. Kyriakidis and A. G. Journel. Geostatistical space-time models: a review. *Mathematical Geology*, 31(6):651–684, 1999.

[281] C. Lakshminarayanan and S. Bhatnagar. A stability criterion for two timescale stochastic approximation schemes. *Automatica*, 79(C):108–114, 2017.

[282] J. Latz. Analysis of stochastic gradient descent in continuous time. *Statistics and Computing*, 31(4):39, 2021.

[283] N. Lazrieva, T. Sharia, and T. Toronjadze. The Robbins-Monro type stochastic differential equations. I. Convergence of solutions. *Stochastics and Stochastic Reports*, 61(1-2):67–87, 1997.

[284] N. Lazrieva, T. Sharia, and T. Toronjadze. Semimartingale stochastic approximation procedure and recursive estimation. *Journal of Mathematical Sciences*, 153(3):211–261, 2008.

[285] N. Lazrieva, T. Sharia, and T. Toronjadze. The Robbins-Monro type stochastic differential equations. II. Asymptotic behaviour of solutions. *Stochastics and Stochastic Reports*, 75(3):153–180, 2003.

[286] N. Lazrieva and T. Toronjadze. The Robbins-Monro type stochastic differential equations. III. Polyak's averaging. *Stochastics*, 82(2):165–188, 2010.

[287] N. Lazrieva and T. Toronjadze. Recursive parameter estimation in the trend coefficient of a diffusion process. *Georgian Mathematical Journal*, 17(4):683–704, 2010.

[288] N. Lazrieva and T. Toronjadze. Recursive estimation procedures for one-dimensional parameter of statistical models associated with semimartingales. *Transactions of A. Razmadze Mathematical Institute*, 171(1):57–75, 2017.

[289] S. Le Corff and G. Fort. Online expectation maximization based algorithms for inference in hidden Markov models. *Electronic Journal of Statistics*, 7:763–792, 2013.

[290] H. W. J. Lee and K. L. Teo. Control parametrization enhancing technique for solving a special ODE class with state dependent switch. *Journal of Optimization Theory and Applications*, 118(1):55–66, 2003.

[291] H. W. J. Lee, K. L. Teo, and A. E. B. Lim. Sensor scheduling in continuous time. *Automatica*, 37(12):2017–2023, 2001.

[292] H. W. J. Lee, K. L. Teo, V. Rehbock, and L. S. Jennings. Control parametrization enhancing technique for optimal discrete-valued control problems. *Automatica*, 35:1401–1407, 1999.

[293] F. LeGland. Nonlinear filtering and problem of parametric estimation. In M. Hazewinkel and J. Willems, editors, *Stochastic Systems: The Mathematics of Filtering and Identification and Applications*, pages 613–620. D. Reidel Publishing Co., Boston, MA, 1980.

[294] F. LeGland and L. Mevel. Recursive identification of HMMs with observations in a finite set. In *Proceedings of the 34th IEEE Conference on Decision and Control*, pages 216–221, New Orleans, LA, 1995.

[295] F. LeGland and L. Mevel. Recursive identification in hidden Markov models. In *Proceedings of 36th IEEE Conference on Decision and Control*, pages 3468–3473, San Diego, CA, 1997.

[296] D. Levanony, A. Shwartz, and O. Zeitouni. Recursive identification in continuous-time stochastic processes. *Stochastic Processes and their Applications*, 49(2):245–275, 1994.

[297] J. Lévine. Finite dimensional realizations of stochastic p.d.e.'s and application to filtering. *Stochastics and Stochastic Reports*, 37(1-2):75–103, 1991.

[298] Q. Li, C. Tai, and W. E. Stochastic modified equations and adaptive stochastic gradient algorithms. In *Proceedings of the 34th International Conference on Machine Learning (ICML 2017)*, pages 2101–2110, Sydney, Australia, 2017.

[299] Q. Li, C. Tai, and W. E. Stochastic modified equations and dynamics of stochastic gradient algorithms I: mathematical foundations. *Journal of Machine Learning Research*, 20(40):1–47, 2019.

[300] R. Li, K. L. Teo, K. H. Wong, and G. R. Duan. Control parameterization enhancing transform for optimal control of switched systems. *Mathematical and Computer Modelling*, 43(11):1393–1403, 2006.

[301] Z. Li, F. Lu, M. Maggioni, S. Tang, and C. Zhang. On the identifiability of interaction functions in systems of interacting particles. *Stochastic Processes and their Applications*, 132:135–163, 2021.

[302] F. Lindgren, H. Rue, and J. Lindstrom. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498, 2011.

[303] J. L. Lions. *Optimal Control of Systems Governed by Partial Differential Equations.* Springer-Verlag, Berlin, Heidelberg, 1971.

[304] R. S. Liptser and A. N. Shiryaev. *Statistics of Random Processes.* Springer-Verlag, Berlin, Heidelberg, 2nd edition, 2001.

[305] J. Liu. *Monte Carlo Strategies in Scientific Computing.* Springer-Verlag, Berlin, 2001.

[306] M. Liu and H. Qiao. Parameter estimation of path-dependent McKean-Vlasov stochastic differential equations. *arXiv preprint*, 2020.

[307] Q. Liu. Stein variational gradient descent as gradient flow. In *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS 2017)*, pages 3118–3126, Red Hook, NY, 2017.

[308] W. Liu, Y. Zhang, Z. Deng, J. Zhao, and L. Tong. A hybrid quantum-classical conditional generative adversarial network algorithm for human-centered paradigm in cloud. *EURASIP Journal on Wireless Communications and Networking*, 2021(1):37, 2021.

[309] X. Liu, K. Yeo, Y. Hwang, J. Singh, and J. Kalagnanam. A statistical modeling approach for air quality data based on physical dispersion processes and its application to ozone modeling. *The Annals of Applied Statistics*, 10:756–785, 2016.

[310] L. Ljung. Analysis of recursive stochastic algorithms. *IEEE Transactions on Automatic Control*, 22(4):551–575, 1977.

[311] L. Ljung, S. K. Mitter, and J. M. F. Moura. Optimal recursive maximum likelihood estimation. *IFAC Proceedings Volumes*, 20(5):241–242, 1987.

[312] L. Ljung. On the consistency of prediction error identification methods. In R. K. Mehra, D. G. B. T. M. i. S. Lainiotis, and Engineering, editors, *System Identification Advances and Case Studies*, volume 126, pages 121–164. Elsevier, 1976.

[313] L. Ljung. *System Identification: Theory for the User.* Prentice-Hall, Englewood Cliffs, NJ, 2nd edition, 1999.

[314] L. Ljung, G. C. Pflug, and H. Walk. *Stochastic Approximation and Optimization of Random Systems.* Birkhäuser, Basel, 1992.

[315] L. Ljung and T. Soderstrom. *Theory and Practice of Recursive Identification.* MIT Press, Cambridge, MA, 1983.

[316] F. P. Llopis, N. Kantas, A. Beskos, and A. Jasra. Particle filtering for stochastic Navier-Stokes signal observed with linear additive noise. *SIAM Journal on Scientific Computing*, 40(3):1544–1565, 2018.

[317] C. Long, Q. Zhang, B. Li, H. Yang, and X. Guan. Non-cooperative power control for wireless ad hoc networks with repeated games. *IEEE Journal on Selected Areas in Communications*, 25(6):1101–1112, 2007.

[318] J. A. N. Lorenz. Continuous opinion dynamics under bounded confidence: a survey. *International Journal of Modern Physics C*, 18(12):1819–1838, 2007.

[319] S. V. Lototsky. Statistical inference for stochastic parabolic equations: a spectral approach. *Publicacions Matemàtiques*, 53(1):3–45, 2009.

[320] S. V. Lototsky and B. L. Rosovskii. Spectral asymptotics of some functionals arising in statistical inference for SPDEs. *Stochastic Processes and their Applications*, 79(1):69–94, 1999.

[321] S. V. Lototsky and B. L. Rozovsky. *Stochastic Partial Differential Equations*. Springer-Verlag, Cham, Switzerland, 2017.

[322] F. Lu, M. Maggioni, and S. Tang. Learning interaction kernels in heterogeneous systems of agents from multiple trajectories. *Journal of Machine Learning Research*, 22(32):1–67, 2021.

[323] F. Lu, M. Maggioni, and S. Tang. Learning interaction kernels in stochastic systems of interacting particles from multiple trajectories. *Foundations of Computational Mathematics*, 21(4), 2021.

[324] F. Lu, M. Zhong, S. Tang, and M. Maggioni. Nonparametric inference of interaction laws in systems of agents from trajectory data. *Proceedings of the National Academy of Sciences*, 116(29):14424–14433, 2019.

[325] E. Luçon. Quenched asymptotics for interacting diffusions on inhomogeneous random graphs. *Stochastic Processes and their Applications*, 130(11):6783–6842, 2020.

[326] C. Ma. Families of spatio-temporal stationary covariance models. *Journal of Statistical Planning and Inference*, 116(2):489–501, 2003.

[327] F. Malrieu. Logarithmic Sobolev inequalities for some nonlinear PDE's. *Stochastic Processes and their Applications*, 95(1):109–132, 2001.

[328] F. Malrieu. Convergence to equilibrium for granular media equations and their Euler schemes. *The Annals of Applied Probability*, 13(2):540–560, 2003.

[329] C. Man and C. W. Tsai. Stochastic partial differential equation-based model for suspended sediment transport in surface water flows. *Journal of Engineering Mechanics*, 133(4):422–430, 2007.

[330] S. Mandt, M. D. Hoffman, and D. M. Blei. A Variational Analysis of Stochastic Gradient Algorithms. In *Proceedings of the 33rd International Conference on Machine Learning (ICML 2016)*, pages 354–363, New York, 2016.

[331] X. Mao. Exponential stability for nonlinear stochastic differential equations with respect to semimartingales. *Stochastics and Stochastic Reports*, 28(4):343–355, 1989.

[332] M. C. Maurel and D. Michel. Des resultats de non existence de filtre de dimension finie. *Stochastics*, 13(1-2):83–102, 1984.

[333] I. W. McKeague. Estimation for diffusion processes under misspecified models. *Journal of Applied Probability*, 21(3):511–520, 1984.

[334] H. P. McKean. A class of Markov processes associated with nonlinear parabolic equations. *Proceedings of the National Academy of Sciences of the United States of America*, 56(6):1907–1911, 1966.

[335] H. P. McKean. Propagation of chaos for a class of non-linear parabolic equations. In *Lecture Series in Differential Equations*, volume 7, pages 41–57. Catholic University, 1967.

[336] H. P. McKean. *Stochastic Integrals*. Academic Press, New York, 1969.

[337] L. Meier, J. Peschon, and R. Dressler. Optimal control of measurement subsystems. *IEEE Transactions on Automatic Control*, 12(5):528–536, 1967.

[338] S. Méléard. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. *Probabilistic Models for Nonlinear Partial Differential Equations*, 1627:42–95, 1996.

[339] A. V. Mel'nikov. On the theory of stochastic equations in components of semimartingales. *Mathematics of the USSR-Sbornik*, 38(3):381–394, 1981.

[340] A. V. Mel'nikov. On properties of strong solutions of stochastic equations with respect to semimartingales. *Stochastics*, 8(2):103–119, 1982.

[341] A. V. Mel'nikov. Procedures of stochastic approximation in statistics of semimartingale type processes. *Uspekhi Matematicheskikh Nauk*, 434:215–216, 1988.

[342] A. V. Mel'nikov. Stochastic approximation procedures for semimartingales. In A. A. Novikov and A. N. Shiryaev, editors, *Statistics and Control of Stochastic Processes*, pages 147–156. Nauka, Moscow, Russia, 1989.

[343] A. V. Mel'nikov. Stochastic differential equations: singularity of coefficients, regression models, and stochastic approximation. *Russian Mathematical Surveys*, 51(5):819–909, 1996.

[344] A. V. Mel'nikov, A. E. Rodkina, and E. Valkeila. On a general class of stochastic approximation algorithms. In *Proceedings of the Third Finnish Soviet Symposium on Probability Theory and Mathematical Statistics*, pages 183–196, Turku, Finland, 1992.

[345] J. Mercer and A. R. Forsyth. XVI. Functions of positive and negative type, and their connection the theory of integral equations. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 209(441-458):415–446, 1909.

[346] P. Mertikopoulos, N. Hallak, A. Kavis, and V. Cevher. On the almost sure convergence of stochastic gradient descent in non-convex problems. *arXiv preprint*, 2020.

[347] P. Mertikopoulos and M. Staudigl. On the convergence of gradient-like flows with noisy gradient input. *SIAM Journal on Optimization*, 28(1):163–197, 2018.

[348] M. Metivier and P. Priouret. Applications of a Kushner and Clark lemma to general classes of stochastic algorithms. *IEEE Transactions on Information Theory*, 30(2):140–151, 1984.

[349] B. M. Miller. Generalized optimisation in the problems of observation control. *Automation and Remote Control*, 52:83–92, 1991.

[350] B. M. Miller and W. J. Runggaldier. Optimization of observations: a stochastic control approach. *SIAM Journal on Control and Optimization*, 35(3):1030–1052, 1997.

[351] Y. S. Mishura and A. Y. Veretennikov. Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations. *Theory of Probability and Mathematical Statistics*, 103:59–101, 2020.

[352] J. Mohapl. Maximum likelihood estimation in linear infinite dimensional models. *Communications in Statistics. Stochastic Models*, 10(4):781–794, 1994.

[353] J. Mohapl. A stochastic advection-diffusion model for the rocky flats soil plutonium data. *Annals of the Institute of Statistical Mathematics*, 52(1):84–107, 2000.

[354] A. Mokkadem and M. Pelletier. Convergence rate and averaging of nonlinear two-time-scale stochastic approximation algorithms. *The Annals of Applied Probability*, 16(3):1671–1702, 2006.

[355] J. B. Moore. Convergence of continuous time stochastic ELS parameter estimation. *Stochastic Processes and their Applications*, 27:195–215, 1988.

[356] K. Morris. Linear-quadratic optimal actuator location. *IEEE Transactions on Automatic Control*, 56(1):113–124, 2011.

[357] S. Motsch and E. Tadmor. Heterophilious dynamics enhances consensus. *SIAM Review*, 56(4):577–621, 2014.

[358] J. M. F. Moura and S. K. Mitter. Identification and filtering: optimal recursive maximum likelihood approach. Technical report, Massachusetts Institute of Technology and Laboratory for Information and Decision Systems, 1986.

[359] R. Munos and P. Bourgine. Reinforcement learning for continuous stochastic control problems. In *Proceedings of the 10th International Conference in Neural Information Processing Systems (NIPS 1997)*, Denver, CO, 1997.

[360] Y. Nakamori, S. Miyamoto, S. Ikeda, and Y. Sawaragi. Measurement optimization with sensitivity criteria for distributed parameter systems. *IEEE Transactions on Automatic Control*, 25(5):889–901, 1980.

[361] M. Nevel'son and R. Has'minskii. Stochastic Approximation and Recursive Estimation. *Translations of Mathematical Monographs*, 47, 1976.

[362] J. L. Ny, E. Feron, and M. A. Dahleh. Scheduling Kalman filters in continuous time. In *Proceedings of the 2009 American Control Conference*, pages 3799–3805, St. Louis, MO, 2009.

[363] J. L. Ny, E. Feron, and M. A. Dahleh. Scheduling continuous-time Kalman filters. *IEEE Transactions on Automatic Control*, 56(6):1381–1394, 2011.

[364] D. L. Ocone, J. S. Baras, and S. I. Marcus. Explicit filters for diffusions with certain nonlinear drifts. *Stochastics*, 8(1):1–16, 1982.

[365] D. L. Ocone. Stochastic calculus of variations for stochastic partial differential equations. *Journal of Functional Analysis*, 79(2):288–331, 1988.

[366] D. L. Ocone and E. Pardoux. A Lie algebraic criterion for non-existence of finite dimensionally computable filters. In G. Da Prato and L. Tubaro, editors, *Stochastic Partial Differential Equations and Applications II*, pages 197–204. Springer-Verlag, Berlin, Heidelberg, 1989.

[367] K. Oelschlager. A martingale approach to the law of large numbers for weakly interacting stochastic processes. *The Annals of Probability*, 12(2):458–479, 1984.

[368] S. Omatu, S. Koide, and T. Soeda. Optimal sensor location problem for a linear distributed parameter system. *IEEE Transactions on Automatic Control*, 23(4):665–673, 1978.

[369] S. Omatu and J. Seinfeld. *Distributed Parameter Systems. Theory and Applications.* Oxford University Press, Oxford, UK, 1989.

[370] A. Orvieto and A. Lucchi. Continuous-time models for stochastic optimization algorithms. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems (NeurIPS 2019)*, Vancouver, Canada, 2019.

[371] E. Pardoux and A. Y. Veretennikov. On the Poisson equation and diffusion approximation 1. *The Annals of Probability*, 29(3):1061–1085, 2001.

[372] E. Pardoux and A. Y. Veretennikov. On the Poisson equation and diffusion approximation 2. *The Annals of Probability*, 31(3):1166–1192, 2003.

[373] M. Patan. *Optimal Sensor Networks Scheduling in Identification of Distributed Parameter Systems*. Springer-Verlag, Berlin, Heidelberg, 2012.

[374] G. A. Pavliotis. *Stochastic Processes and Applications*. Springer-Verlag, New York, 2014.

[375] G. A. Pavliotis and A. M. Stuart. *Multiscale Methods*. Springer-Verlag, New York, 2008.

[376] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer-Verlag, New York, 1983.

[377] R. Pemantle. Nonconvergence to unstable points in urn models and stochastic approximations. *The Annals of Probability*, 18(2):698–712, 1990.

[378] H. Pham. *Continuous-time Stochastic Control and Optimization with Financial Applications*. Springer-Verlag, Berlin, Heidelberg, 2009.

[379] M. Polis. The distributed system parameter identification problem. In *Proceedings of the 3rd IFAC Sympsoium on Control of Distributed Parameter Systems*, Toulouse, France, 1982.

[380] B. T. Polyak and A. B. Juditsky. Acceleration of stochastic approximation By averaging. *SIAM Journal on Control and Optimization*, 30(4):838–855, 1992.

[381] B. T. Polyak. New stochastic approximation type procedures. *Avtomatica i Telemekhanika*, 7:98–107, 1990.

[382] G. Poyiadjis, A. Doucet, and S. S. Singh. Particle approximations of the score and observed information matrix in state space models with application to parameter estimation. *Biometrika*, 98(1):65–80, 2011.

[383] P. E. Protter. On the Existence, Uniqueness, Convergence and Explosions of Solutions of Systems of Stochastic Integral Equations. *The Annals of Probability*, 5(2):243–261, 1977.

[384] N. Qian. On the momentum term in gradient descent learning algorithms. *Neural Networks*, 12(1):145–151, 1999.

[385] E. Rafajlowicz. Design of experiments for eigenvalue identification in distributed-parameter systems. *International Journal of Control*, 34(6):1079–1094, 1981.

[386] E. Rafajlowicz. Optimisation of measurements for state estimation in parabolic distributed systems. *Kybernetika*, 20(5):413–422, 1984.

[387] E. Rafajlowicz. Optimum choice of moving sensor trajectories for distributed-parameter system identification. *International Journal of Control*, 43(5):1441–1451, 1986.

[388] M. Raginsky and J. Bouvrie. Continuous-time stochastic mirror descent on a network: variance reduction, consensus, convergence. In *Proceedings of the 51st IEEE Conference on Decision and Control*, pages 6793–6800, Maui, HI, 2012.

[389] A. Rajeswaran, C. Finn, S. M. Kakade, and S. Levine. Meta-learning with implicit gradients. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems (NeurIPS 2019)*, Vancouver, Canada, 2019.

[390] B. L. S. P. Rao and H. Rubin. Asymptotic theory of estimation in nonlinear stochastic differential equations. *Sankhy: The Indian Journal of Statistics, Series A*, 43(2):170–189, 1981.

[391] C. N. Rautenberg. *A distributed parameter approach to optimal filtering and estimation with mobile sensor networks*. Phd thesis, Virginia Polytechnic Institute and State University, 2010.

[392] R. Rebolledo. Central limit theorems for local martingales. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 51(3):269–286, 1980.

[393] S. J. Reddi, S. Kale, and S. Kumar. On the convergence of Adam and beyond. In *Proceedings of the 6th International Conference on Learning Representations (ICLR '18)*, Vancouver, Canada, 2018.

[394] B. J. Reich, J. Eidsvik, M. Guindani, A. J. Nail, and A. M. Schmidt. A class of covariate-dependent spatiotemporal covariance functions. *The Annals of Applied Statistics*, 5(4):2265–2687, 2011.

[395] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer-Verlag, Berlin, Heidelberg, 3rd edition, 1999.

[396] H. Robbins and S. Monro. A stochastic approximation method. *The Annals of Mathematical Statistics*, 22(3):400–407, 1951.

[397] M. Röckner, X. Sun, and Y. Xie. Strong convergence order for slow-fast McKean-Vlasov stochastic differential equations. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 57(1):547–576, 2021.

[398] I. G. Rosen. Convergence of Galerkin approximations for operator Riccati equations - a nonlinear evolution equation approach. *Journal of Mathematical Analysis and Applications*, 155(1):226–248, 1991.

[399] A. Roth, J. Ullman, and Z. S. Wu. Watch and learn: optimizing from revealed preferences feedback. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing (STOC '16)*, pages 949–962, New York, 2016.

[400] S. Rouhani and H. Wackernagel. Multivariate geostatistical approach to space-time data analysis. *Water Resources Research*, 26(4):585–591, 1990.

[401] D. Ruppert. Efficient estimations from a slowly converging robbins-monro process. Technical report, Cornell University Operations Research and Industrial Engineering, 1988.

[402] D. Ruppert. Stochastic approximation. In B. K. Ghosh and P. Sen, editors, *Handbook in Sequential Analysis*, pages 503–529. Marcel Dekker, New York, 1991.

[403] J. Sacks. Asymptotic distribution of stochastic approximation procedures. *The Annals of Mathematical Statistics*, 29(2):373–405, 1958.

[404] H. Sakaguchi, S. Shinomoto, and Y. Kuramoto. Phase transitions and their bifurcation analysis in a large population of active rotators with mean-field coupling. *Progress of Theoretical Physics*, 79(3):600–607, 1988.

[405] M. Salgado, R. Middleton, and G. C. Goodwin. Connection between continuous and discrete Riccati equations with applications to Kalman filtering. *IEE Proceedings D - Control Theory and Applications*, 135(1):28–34, 1988.

[406] S. Särkkä. On unscented Kalman filtering for state estimation of continuous-time nonlinear systems. *IEEE Transactions on Automatic Control*, 52(9):1631–1641, 2007.

[407] S. Särkkä and A. Solin. *Applied Stochastic Differential Equations.* Cambridge University Press, Cambridge, UK, 2019.

[408] Y. Sawaragi, T. Soeda, and S. Omatu. *Modeling, Estimation, and Their Applications for Distributed Parameter Systems.* Springer-Verlag, Berlin, Heidelberg, 1978.

[409] B. Schwartz, S. Gannot, E. A. P. Habets, and Y. Noam. Recursive maximum likelihood algorithm for dependent observations. *IEEE Transactions on Signal Processing*, 67(5):1366–1381, 2019.

[410] G. R. Sell and Y. You. *Dynamics of Evolutionary Equations*. Springer-Verlag, New York, 2002.

[411] P. Sen and K. B. Athreya. On stochastic approximation procedures in continuous time. *Journal of the Indian Institute of Science*, 60(3):135–147, 1978.

[412] L. Sharrock and N. Kantas. Joint online parameter estimation for the partially observed stochastic advection diffusion equation. *SIAM /ASA Journal on Uncertainty Quantification*, 10(1):55–95, 2022.

[413] L. Sharrock and N. Kantas. Two-timescale stochastic gradient descent in continuous time with applications to joint online parameter estimation and optimal sensor placement. *To appear in Bernoulli*, 28(3), 2022.

[414] L. Sharrock, N. Kantas, P. Panos, and G. A. Pavliotis. Parameter estimation for the McKean-Vlasov stochastic differential equation. *arXiv preprint*, 2021.

[415] M. Shiino. Dynamical behavior of stochastic systems of infinitely many coupled nonlinear oscillators exhibiting phase transitions of mean-field type: H theorem on asymptotic approach to equilibrium and critical slowing down of order-parameter fluctuations. *Physical Review A*, 36(5):2393–2412, 1987.

[416] M. Shub. *Global Stability of Dynamical Systems*. Springer-Verlag, New York, 1987.

[417] F. Sigrist, H. R. Kunsch, and W. A. Stahel. A dynamic nonstationary spatio-temporal model for short term prediction of precipitation. *The Annals of Applied Statistics*, 6(4):1452–1477, 2012.

[418] F. Sigrist, H. R. Kunsch, and W. A. Stahel. Stochastic partial differential equation based modelling of large space-time data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77(1):3–33, 2015.

[419] S. Singh, N. Kantas, B. Vo, A. Doucet, and R. Evans. On the convergence of a two timescale stochastic optimisation algorithm for optimal observer trajectory planning. Technical report, Cambridge University Engineering Department, Cambridge, UK., 2005.

[420] J. Sirignano and K. Spiliopoulos. Stochastic gradient descent in continuous time. *SIAM Journal on Financial Mathematics*, 8(1):933–961, 2017.

[421] J. Sirignano and K. Spiliopoulos. Mean field analysis of neural networks: a law of large numbers. *SIAM Journal on Applied Mathematics*, 80(2):725–752, 2020.

[422] J. Sirignano and K. Spiliopoulos. Stochastic gradient descent in continuous time: a central limit theorem. *Stochastic Systems*, 10(2):124–151, 2020.

[423] M. W. A. Smith and A. P. Roberts. An exact equivalence between the discrete- and continuous-time formulations of the Kalman filter. *Mathematics and Computers in Simulation*, 20(2):102–109, 1978.

[424] V. Solo and X. Kong. *Adaptive Signal Processing Algorithms: Stability and Performance*. Prentice Hall, Hoboken, NJ, 1994.

[425] J. C. Spall. *Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control*. John Wiley & Sons, Inc., Hoboken, NJ, 2003.

[426] M. L. Stein. *Interpolation of Spatial Data: Some Theory of Kriging*. Springer-Verlag, New York, 1999.

[427] M. L. Stein. Space time covariance functions. *Journal of the American Statistical Association*, 100(469):310–321, 2005.

[428] G. Storvik, A. Frigessi, and D. Hirst. Stationary space-time Gaussian fields and their time autoregressive representation. *Statistical Modelling*, 2(2):139–161, 2002.

[429] J. R. Stroud, P. Muller, and B. Sanso. Dynamic models for spatiotemporal data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(4):673–689, 2001.

[430] S. C. Surace and J. Pfister. Online maximum-likelihood estimation of the parameters of partially observed diffusion processes. *IEEE Transactions on Automatic Control*, 64(7):2814–2829, 2019.

[431] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 2nd edition, 2018.

[432] R. S. Sutton, H. R. Maei, and C. Szepesvári. A convergent O(n) temporal-difference algorithm for off-policy learning with linear function approximation. In *Proceedings of the 23rd International Conference on Neural Information Processing Systems (NIPS 2009)*, pages 1609–1616, Vancouver, Canada, 2009.

[433] R. S. Sutton, H. R. Maei, D. Precup, S. Bhatnagar, D. Silver, C. Szepesvári, and E. Wiewiora. Fast gradient-descent methods for temporal-difference learning with linear function approximation. In *Proceedings of the 26th International Conference on Machine Learning (ICML 2009)*, pages 993–1000, New York, 2009.

[434] A.-S. Sznitman. Topics in propagation of chaos, Ecole d'Eté de Probabilités de Saint-Flour XIX - 1989. *Lecture Notes in Mathematics*, 1464:165–251, 1991.

[435] V. Z. B. Tadić and A. Doucet. Asymptotic properties of recursive particle maximum likelihood estimation. *IEEE Transactions on Information Theory*, 67(3):1825–1848, 2021.

[436] V. Z. B. Tadić. Convergence of stochastic approximation under general noise and stability conditions. In *Proceedings of the 36th IEEE Conference on Decision and Control*, volume 3, pages 2281–2286, San Diego, CA, 1997.

[437] V. Z. B. Tadić. Almost sure convergence of two time-scale stochastic approximation algorithms. In *Proceedings of the 2004 American Control Conference*, volume 4, pages 3802–3807, Boston, MA, 2004.

[438] V. Z. B. Tadić. Analyticity, convergence, and convergence rate of recrusive maximum-likelihood estimation in hidden Markov models. *IEEE Transactions on Information Theory*, 56(12):6406–6432, 2010.

[439] V. Z. B. Tadić. Convergence and convergence rate of stochastic gradient search in the case of multiple and non-isolated extrema. *Stochastic Processes and their Applications*, 125(5):1715–1755, 2015.

[440] V. Z. B. Tadić and A. Doucet. Asymptotic bias of stochastic gradient search. *The Annals of Applied Probability*, 27(6):3255–3304, 2017.

[441] V. Z. B. Tadić and A. Doucet. Asymptotic properties of recursive maximum likelihood estimation in nonlinear state-space models. *arXiv preprint*, 2021.

[442] V. Z. B. Tadić and S. P. Meyn. Asymptotic properties of two time-scale stochastic approximation algorithms with constant step sizes. In *Proceedings of the 2003 American Control Conference*, volume 5, pages 4426–4431, Denver, CO, 2003.

[443] Y. Tamura. On asymptotic behaviors of the solution of a nonlinear diffusion equation. *Journal of the Faculty of Science, University of Tokyo*, 31:195–221, 1984.

[444] H. Tanabe. On the equations of evolution in a Banach space. *Osaka Journal of Mathematics*, 12(2):363–376, 1960.

[445] S. Tang and K. A. Morris. Optimal sensor design for infinite-time Kalman filters. In *Proceedings of the 2017 IEEE 56th Annual Conference on Decision and Control*, pages 64–69, Melbourne, Australia, 2017.

[446] K. Teo, W. Lee, L. Jennings, S. Wang, and Y. Liu. Numerical solution of an optimal control problem with variable time points in the objective function. *ANZIAM Journal*, 43:463–478, 2002.

[447] T. Toronjadze. Stochastic equations in the problems of semimartingale parameter estimation. *Journal of Mathematical Sciences*, 121(6):2709–2840, 2004.

[448] Y. Tsypkin. *Foundations of the Theory of Learning Systems*. Academic Press, Cambridge, MA, 1973.

[449] J. Tugaut. Convergence to the equilibria for self-stabilizing processes in double-well landscape. *The Annals of Probability*, 41(3A):1427–1460, 2013.

[450] J. Tugaut. Phase transitions of McKean-Vlasov processes in double-wells landscape. *Stochastics*, 86(2):257–284, 2014.

[451] J. Tugnait. Identification and model approximation for continuous-time systems on finite parameter sets. *IEEE Transactions on Automatic Control*, 25(6):1202–1206, 1980.

[452] J. Tugnait. Continuous-time system identification on compact parameter sets. *IEEE Transactions on Information Theory*, 31(5):652–659, 1985.

[453] D. Uciński. *Optimal Measurement Methods for Distributed Parameter System Identification*. CRC Press, Boca Raton, FL, 2005.

[454] D. Uciński and M. Patan. Sensor network design for the estimation of spatially distributed processes. *International Journal of Applied Mathematics and Computer Science*, 20(3):459–481, 2010.

[455] T. E. Unny. Stochastic partial differential equations in groundwater hydrology. *Stochastic Hydrology and Hydraulics*, 3(2):135–153, 1989.

[456] E. Valkeila and A. V. Mel'nikov. Martingale models of stochastic approximation and their convergence. *Theory of Probability & Its Applications*, 44(2):333–360, 2000.

[457] A. Ventzel. On equations of the theory of conditional Markov processes. *Probability Theory and Its Applications*, 10:357–361, 1965.

[458] A. Y. Veretennikov. On ergodic measures for McKean-Vlasov stochastic equations. In H. Niederreiter and D. Talay, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2004*, pages 471–486. Springer-Verlag, Berlin, Heidelberg, 2006.

[459] A. A. Vlasov. The vibrational properties of an electron gas. *Soviet Physics Uspekhi*, 10(6):721–733, 1968.

[460] I.-J. Wang, E. K. P. Chong, and S. R. Kulkarni. Equivalent necessary and sufficient conditions on noise sequences for stochastic approximation algorithms. *Advances in Applied Probability*, 28(3):784–801, 1996.

[461] J.-G. Wang. A law of the iterated logarithm for stochastic integrals. *Stochastic Processes and their Applications*, 47:215–228, 1993.

[462] M. Wang, E. X. Fang, and H. Liu. Stochastic compositional gradient descent: algorithms for minimizing compositions of expected-value functions. *Mathematical Programming*, 161(1):419–449, 2017.

[463] M. T. Wasan. *Stochastic Approximation.* Cambridge University Press, Cambridge, UK, 1969.

[464] A. J. Weir. *Lebesgue Integration and Measure.* Cambridge University Press, Cambridge, UK, 1973.

[465] J. Wen, X. Wang, S. Mao, and X. Xiao. Maximum likelihood estimation of McKean-Vlasov stochastic differential equation and its application. *Applied Mathematics and Computation*, 274:237–246, 2016.

[466] P. Whittle. On stationary processes in the plane. *Biometrika*, 41(3-4):434–449, 1954.

[467] C. K. Wikle, R. F. Milliff, D. Nychka, and L. M. Berliner. Spatiotemporal hierarchical bayesian modeling: tropical ocean surface winds. *Journal of the American Statistical Association*, 96(454):382–397, 2001.

[468] C. K. Wikle. A kernel-based spectral model for non-Gaussian spatio-temporal processes. *Statistical Modelling*, 2(4):299–314, 2002.

[469] C. K. Wikle and N. Cressie. A dimension-reduced Kalman filtering approach to space-time. *Biometrika*, 86(4):815–829, 1999.

[470] S. Wojtowytsch. Stochastic gradient descent with noise of machine learning type. Part II: continuous time analysis. *arXiv preprint*, 2021.

[471] J. Wolfowitz. On the stochastic approximation method of Robbins and Monro. *The Annals of Mathematical Statistics*, 23(3):457–461, 1952.

[472] X. Wu, B. Jacob, and H. Elbern. Optimal control and observation locations for time-varying systems on a finite-time horizon. *SIAM Journal on Control and Optimization*, 54(1):291–316, 2016.

[473] X. Wu, K. Zhang, and C. Sun. Optimal scheduling of multiple sensors in continuous time. *ISA transactions*, 53(3):793–801, 2014.

[474] K. Xu, C. K. Wikle, and N. I. Fox. A kernel-based spatio-temporal dynamical model for nowcasting weather radar reflectivities. *Journal of the American Statistical Association*, 100(472):1133–1144, 2005.

[475] P. Xu, T. Wang, and Q. Gu. Accelerated stochastic mirror descent: from continuous-time dynamics to discrete-time algorithms. In *Proceedings of the 21st International Conference on Artificial Intelligence and Statistics (AISTATS 2018)*, pages 1087–1096, Lanzarote, Canary Islands, 2018.

[476] P. Xu, T. Wang, and Q. Gu. Continuous and discrete-time accelerated stochastic mirror descent for strongly convex functions. In *Proceedings of the 35th International Conference on Machine Learning (ICML 2018)*, pages 5492–5501, Stockholm, Sweden, 2018.

[477] T. Xu, S. Zou, and Y. Liang. Two time-scale off-policy TD learning: non-asymptotic analysis over Markovian samples. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems (NeurIPS 2019)*, pages 10634–10644, Vancouver, Canada, 2019.

[478] T. Yamada. On a comparison theorem for solutions of stochastic differential equations and its applications. *Journal of Mathematics of Kyoto University*, 13(3):497–512, 1973.

[479] J. Yang, W. Hu, and C. J. Li. On the fast convergence of random perturbations of the gradient flow. *Asymptotic Analysis*, 122(3-4):371–393, 2021.

[480] S. Yang, M. Wang, and E. X. Fang. Multilevel stochastic gradient methods for nested composition optimization. *SIAM Journal on Optimization*, 29(1):616–659, 2019.

[481] C. Yildiz, M. Heinonen, and H. Lähdesmäki. Continuous-time model-based reinforcement learning. In *Proceedings of the 38th International Conference on Machine Learning (ICML 2021)*, pages 12009–12018, Virtual, 2021.

[482] G. Yin. On extensions of Polyak's averaging approach to stochastic approximation. *Stochastics and Stochastic Reports*, 36(3-4):245–264, 1991.

[483] G. Yin and I. Gupta. On a continuous time stochastic approximation problem. *Acta Applicandae Mathematica*, 33(1):3–20, 1993.

[484] P. Young. *Recursive Estimation and Time-Series Analysis*. Springer-Verlag, Berlin, Heidelberg, 1984.

[485] T. K. Yu and J. H. Seinfeld. Observability and optimal measurement location in linear distributed parameter systems. *International Journal of Control*, 18(4):785–799, 1973.

[486] M. D. Zeiler. ADADELTA: an adaptive learning rate method. *arXiv preprint*, 2012.

[487] O. Zeitouni and B. Z. Bobrovsky. On the joint nonlinear filtering-smoothing of diffusion processes. *Systems & Control Letters*, 7(4):317–321, 1986.

[488] M. Zhang and K. Morris. Sensor choice for minimum error variance estimation. *IEEE Transactions on Automatic Control*, 63(2):315–330, 2018.