Newton method for stochastic control problems

Emmanuel Gobet, Maxime Grangereau

Newton method for stochastic control problems *

Emmanuel GOBET † and Maxime GRANGEREAU ‡

Abstract

We develop a new iterative method based on Pontryagin principle to solve stochastic control problems. This method is nothing else than the Newton method extended to the framework of stochastic controls, where the state dynamics is given by an ODE with stochastic coefficients. Each iteration of the method is made of two ingredients: computing the Newton direction, and finding an adapted step length. The Newton direction is obtained by solving an affine-linear Forward-Backward Stochastic Differential Equation (FBSDE) with random coefficients. This is done in the setting of a general filtration. We prove that solving such an FBSDE reduces to solving a Riccati Backward Stochastic Differential Equation (BSDE) and an affine-linear BSDE, as expected in the framework of linear FBSDEs or Linear-Quadratic stochastic control problems. We then establish convergence results for this Newton method. In particular, sufficient regularity of the second-order derivative of the cost functional is required to obtain (local) quadratic convergence. A restriction to the space of essentially bounded stochastic processes is needed to obtain such regularity. To choose an appropriate step length while fitting our choice of space of processes, an adapted backtracking line-search method is developed. We then prove global convergence of the Newton method with the proposed line-search procedure, which occurs at a quadratic rate after finitely many iterations. An implementation with regression techniques to solve BSDEs arising in the computation of the Newton step is developed. We apply it to the control problem of a large number of batteries providing ancillary services to an electricity network.

1 Introduction

In this paper, we introduce a new method to solve stochastic control problems, which is a generalization of the Newton method to the particular (infinite-dimensional) setting of stochastic control problems. We consider problems in general filtrations with a linear dynamic. For the sake of simplicity, we restrict ourselves to the one-dimensional setting (meaning the state and control variables are real-valued stochastic processes). The general form of problems we consider is:

$$\left.\begin{array}{r}\mathcal{J}(u) := \mathbb{E}\left[\int_0^T l\left(t,\omega,u_{t,\omega},X^u_{t,\omega}\right)\mathrm{d}t + \Psi(\omega, X^u_{T,\omega})\right] \\ \text{s.t. } X^u_{t,\omega} = x_0 + \int_0^t (\alpha_{s,\omega}u_{s,\omega} + \beta_{s,\omega}X^u_{s,\omega})\mathrm{d}s + M_{t,\omega}.\end{array}\right\} \longrightarrow \min_u . \tag{1.1}$$

To properly introduce the Newton method for stochastic control problems, we give an overview of state-of-the-art numerical methods for this class of problems, then a brief introduction to the Newton method for the optimization of functions taking values in $\mathbb{R}^d$ with $d \in \mathbb{N}$.

State of the art of numerical methods for stochastic control problems Standard approaches to solve stochastic control problems are based either on Bellman dynamic programming principle, either on Pontryagin's optimality principle.

*This work has benefited from several supports: Siebel Energy Institute (Calls for Proposals #2, 2016), ANR project CAESARS (ANR-15-CE05-0024), Association Nationale de la Recherche Technique (ANRT) and Electricité De France (EDF), *Finance for Energy Market* (FiME) Lab (Institut Europlace de Finance), Chair "Stress Test, RISK Management and Financial Steering" of the Foundation Ecole Polytechnique.

†Email: emmanuel.gobet@polytechnique.edu. Centre de Mathématiques Appliquées (CMAP), CNRS, Ecole Polytechnique, Institut Polytechnique de Paris, Route de Saclay, 91128 Palaiseau Cedex, France. Corresponding author.

‡Email: maxime.grangereau@edf.fr. Electricité de France, Department OSIRIS, 7 Boulevard Gaspard Monge, 91120 Palaiseau cedex, France.

The dynamic programming principle gives rise to a non-linear Partial Differential Equation (PDE) called Bellman's equation, which is satisfied by the value function under reasonable conditions [Pha09; Kry08]. Finite-difference methods to solve this type of PDE have been studied in [GŠ09] and [BZ03] for instance, which allow to reduce the problem to a high dimensional non-linear system of equations. Among other methods based on Bellman's principle, the Howard's policy improvement algorithm is an iterative algorithm where a linearized version of the Bellman's equation is solved at each step. This method has been introduced by Howard in [How60], in the context of Markovian decision processes. A global convergence rate for Howard policy improvement algorithm (and a variant) for stochastic control problems have been recently established in [KŠS20b]. Deep-learning methods have also been applied to solve the non-linear Bellman PDE arising in this context [HJW18].

Another approach to solve stochastic optimal control is the Pontryagin principle, which gives rise to a Forward-Backward Stochastic Differential Equation, see [Zha17]. Methods to solve this type of equations include fixed-point methods such as Picard iterations or the method of Markovian iterations for coupled FBSDEs in [BZ+08]. They converge for small time horizon under the assumption of Lipschitz coefficients, but convergence can be proved for arbitrary time horizon under some monotony assumption [PT99], using norms with appropriate exponential weights. Among methods related to fixed-point iteration, the Method of Successive Approximations is an iterative method based on Pontryagin's principle, which was proposed in [CL82] for deterministic control problems. However, convergence is not guaranteed in general. This method is refined in the yet to be published work [KŠS20a], for stochastic control problems, using an a modification of the Hamiltonian, called augmented Hamiltonian, which allows to show global convergence, and even establish a convergence rate for particular structures of problems. Solvability of FBSDEs for arbitrary time horizon under monotony conditions can be proved using the continuation method in [HP95; Yon97; PW99]. In this method, the interval is divided in smaller sub-intervals, with boundary conditions ensuring consistency of the overall representation. Then a fixed point argument is done on each sub-interval. This method is well developed theoretically to prove the existence of solutions of a FBSDE constructively, but rarely used to design algorithms to solve the problem numerically. We mention however [Ang+19] which uses the continuation method to design numerical schemes to solve Mean-Field Games. Another method to solve FBSDEs is the Four step scheme introduced in [MPY94], which allows to compute a so-called decoupling field as solution of a quasi-linear PDE. This decoupling field allows to express the adjoint variable as a feedback of the state variable. Some Deep-learning based algorithms have been recently proposed to solve FBSDEs in [Ji+20] and [HL20].

The case of linear FBSDEs and of linear quadratic stochastic optimal control problems has been extensively studied see for instance [Bis76; Yon99; Yon06]. Our result builds on these works, as our algorithm is based on successive linearizations of non-linear FBSDEs obtained by a Taylor expansion.

Preliminary on the Newton method in $\mathbb{R}^d$. Consider a twice-differentiable convex function $f : x \in \mathbb{R}^d \mapsto \mathbb{R}$. We wish to solve the minimization problem $\min_{x\in\mathbb{R}^d} f(x)$ If f is strongly convex and its second-order derivative is Lipschitz-continuous [Kan48; NW06; BV04] or if f is strictly convex and self-concordant [NN94; BV04] (meaning that f is three times differentiable and $\frac{d}{d\alpha}\nabla^2 f(x+\alpha y)\big|_{\alpha=0} \leq 2\sqrt{y^T\nabla^2 f(x)\, y}\,\nabla^2 f(x)$, for all x, y in $\mathbb{R}^d$), then the Newton method gives a sequence of points which converges locally quadratically to the global minimizer of f. This means that if the initial point is sufficiently close to the optimum, the convergence of the sequence to this point is very fast. The pseudo-code for the Newton method is given in Algorithm 1.

Algorithm 1 Newton's method

$\varepsilon > 0$, $k = 0$, $x^{(0)}$ fixed
while $|\nabla f(x^{(k)})| > \varepsilon$ **do**
    Compute Newton direction Δ_x solution of the linear equation $\nabla^2 f(x^{(k)})(\Delta_x) = -\nabla f(x^{(k)})$
    Compute new iterate $x^{(k+1)} = x^{(k)} + \Delta_x$.
    $k \leftarrow k + 1$
end while
return $x^{(k)}$

To obtain global convergence of the Newton method (i.e., convergence to the optimum no matter the choice of the initial point), a common procedure is to use line search methods, allowing to choose a step size σ and to define new iterates by the formula $x^{(k+1)} = x^{(k)} + \sigma\Delta_x$ instead of the standard Newton method which considers the case $\sigma = 1$. Among them, given $\beta \in (0, 1)$, the backtracking line search procedure allows to find the largest value among $\{1, \beta, \beta^2, \beta^3, ...\}$ which satisfies a sufficient decrease condition. Its pseudo-code is given in Algorithm 2. The combination of Newton method with backtracking line search gives a globally convergent method [BV04] in $\mathbb{R}^d$.

Algorithm 2 Backtracking line search procedure

```
Inputs: Current point x ∈ ℝ, Current search direction Δ_x, β ∈ (0,1), γ ∈ (0,1).
σ = 1.
while f(x + σΔ_x) > f(x) + γσ∇f(x) · Δ_x  do
    σ ← βσ.
end while
return  x + σΔ_x.
```

Our contributions. To solve numerically the problem proposed in (1.1), we extend the Newton method to the infinite-dimensional setting of convex stochastic control problems in general filtrations, where the dynamics is an affine-linear Ordinary Differential Equation (ODE). This iterative method generates a sequence of points which are solution of successive Linear-Quadratic approximation of the stochastic control problem around a current estimate of the solution, see Proposition 3.2. Equivalently, it can be interpreted as successive linearizations of the Forward Backward Stochastic Differential Equation (FBSDE) arising from the stochastic Pontryagin's principle around a current estimate of the solution, see Proposition 3.3.

In section 3.2, a full methodology is proposed to solve affine-linear FBSDEs with random coefficients in general filtrations, which arise when computing the Newton step. The methodology is quite standard, though the framework is a bit unusual as we do not assume Brownian filtrations. In particular, we show that solving the solution of linear FBSDEs or the computation of the Newton step require solving a Riccati BSDE and an affine-linear BSDE, see Theorem 3.7 and Corollary 3.8.

The convergence of Newton's method typically requires sufficient regularity of the second-order derivative of the cost functional, see [Kan48; NW06] for the case of a Lipschitz second order derivative and [NN94] for the self-concordant case. Such regularity is not guaranteed in our case: a counter-example (Example 3.9) is given to show that even under strong assumptions (namely, the regularity of the running and terminal costs), the second-order derivative of the cost functional $\mathcal{J}$ may fail to be sufficiently regular in the infinite-dimensional space considered. To tackle this issue, we show that an appropriate restriction of the problem to essentially bounded processes allows to obtain the desired regularity of the second-order derivative of the cost function to minimize, see Theorem 3.11. Local quadratic convergence can thus be expected in this framework [Kan48]-[NW06, Theorem 3.5, p. 44]. However, as in the case of Newton's method in $\mathbb{R}^n$, global convergence may fail in our infinite-dimensional setting, even when the function to minimize is strongly convex with Lipschitz-continuous bounded second-order derivative. We give a counter-example (Example 3.12) to show that such pathological behaviors may occur in our setting. To obtain global convergence, a new line-search method tailored to our infinite-dimensional framework is proposed (see Algorithm 4) and global convergence results are derived for the Newton method combined with this line-search method (see Algorithm 5) under convexity assumptions, see Theorem 3.15.

We then apply our results to solve an energy management problem, which consists in a set of many weakly-interacting symmetric batteries controlled to minimize the total operational costs and power imbalance. A Markovian framework is assumed and regression techniques are used to compute efficiently all the conditional expectations required for the computation of the Newton direction. This allows to obtain a fully implementable version of the Newton method with Backtracking line search (see Algorithm 8).

Numerical results show the performance of the Newton method which of the proposed Backtracking line-search procedure. On the other hand, we show numerically that the natural extension of the standard Backtracking line search is not adapted to our infinite-dimensional setting: the algorithm takes ridiculously small steps and the gradient norm does not decrease after a few iterations, see Figures 6b and 6d. The numerical results are consistent with what we expect, with the asymmetric loss function allowing to penalize more heavily positive than negative power imbalance. We then discuss the choice of some hyper-parameters of the regression methods used to solve the BSDEs.

Organization of the paper. Section 2 introduces the general framework of stochastic control problems studied. Classical results under suitable assumptions are derived: well-posedness, existence and uniqueness of a minimizer, Gateaux and Fréchet differentiability, as well as necessary and sufficient optimality conditions 2.6. We then prove the second-order differentiability of the problem (Proposition 2.8) and show that the second-order differential is valued in the space of isomorphisms of the ambient process space (Corollary 3.8). Section 3 defines the Newton step and its two equivalent interpretations (Propositions 3.2 and3.3). We show that the computation of the Newton step amounts to solve an affine-linear FBSDE, for which we show existence and uniqueness of the solutions, and we prove the computation of the Newton step reduces to solving a Backward Riccati Stochatic Differential equation and an affine-linear BSDE (Theorem 3.7 and Corollary 3.8). We then show Lipschitz-continuity of the second-order derivative of the cost when considering bounded processes. An adapted line-search, called Gradient Backtracking Line Search, as well as Newton's method with line search are given. We prove global convergence for this method as well as quadratic convergence after finitely many iterations (Theorem 3.15). Section 4 provides a full implementation and the numerical results of the Newton method on the stochastic optimal control problem of a large number of batteries tracking power imbalance. Some proofs are postponed in Section 5 to ease the reading.

Notations. We list the most common notations used in all this work.

▷ *Numbers, vectors, matrices.* $\mathbb{R}$, $\mathbb{N}$, $\mathbb{N}^*$ denote respectively the set of real numbers, integers, positive integers. For $n \in \mathbb{N}^*$, $[n]$ denotes the set of integers $\{1, ..., n\}$, and for $m, p \in \mathbb{N}$ with $m \leq p$, $[m : p]$ denotes the set $\{m, ..., p\}$.The notation $|x|$ stands for the Euclidean norm of a vector x, without further reference to its dimension. For a given matrix $A \in \mathbb{R}^p \otimes \mathbb{R}^d$, $A^\top \in \mathbb{R}^d \otimes \mathbb{R}^p$ refers to its transpose. Its norm is that induced by the Euclidean norms in $\mathbb{R}^p$ and $\mathbb{R}^d$, i.e. $|A| := \sup_{x \in \mathbb{R}^d, |x|=1} |Ax|$. Recall that $|A^\top| = |A|$. For $p \in \mathbb{N}^*$, Id_p stands for the identity matrix of size $p \times p$.

▷ *Functions, derivatives.* When a function (or a process) ψ depends on time, we write indifferently $\psi_t(z)$ or $\psi(t, z)$ for the value of ψ at time t, where z represents all other arguments of ψ.

For a smooth function $g : \mathbb{R}^q \mapsto \mathbb{R}^p$, g_x represents the partial derivative of g with respect to x. However, a subscript x_t refers to the value of a process x at time t (and not to a partial derivative with respect to t).

▷ *Probability.* To model the random uncertainty on the time interval $[0, T]$ ($T > 0$ fixed), we consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. We assume that the filtration $\mathbb{F} := \{\mathcal{F}_t\}_{0 \leq t \leq T}$ is right-continuous, augmented with the $\mathbb{P}$-null sets. For a vector/matrix-valued random variable V, its conditional expectation with respect to the sigma-field $\mathcal{F}_t$ is denoted by $\mathbb{E}_t[Z] = \mathbb{E}[Z|\mathcal{F}_t]$. Denote by $\mathcal{P}$ the σ-field of predictable sets of $[0, T] \times \Omega$.

All the quantities impacted by the control u are upper-indexed by u, like Z^u for instance.

As usually, càdlàg processes stand for processes that are right continuous with left-hand limits. All the martingales are considered with their càdlàg modifications.

▷ *Spaces.* Let $k \in \mathbb{N}^*$. We define $\mathbb{L}^2([0, T], \mathbb{R}^k)$ (resp. $\mathbb{L}^\infty([0, T], \mathbb{R}^k)$) as the Banach space of square integrable (resp. bounded) deterministic functions f on $[0, T]$ with values in $\mathbb{R}^k$. Since the arrival space $\mathbb{R}^k$ will be unimportant, we will skip the reference to it in the notation and write the related norms as

$$\|f\|_{\mathbb{L}^2_T} := \left(\int_0^T |f(t)|^2 \mathrm{d}t \right)^{\frac{1}{2}}, \qquad \|f\|_{\mathbb{L}^\infty_T} := \sup_{t \in [0,T]} |f(t)|.$$

The Banach space of $\mathbb{R}^k$-valued square integrable random variables X is denoted by $\mathbb{L}^2(\Omega, \mathbb{R}^k)$, or simply $\mathbb{L}^2_\Omega$. We also define the Banach space of $\mathbb{R}^k$-valued essentially bounded random variables X, denoted by $\mathbb{L}^\infty(\Omega, \mathbb{R}^k)$, or simply $\mathbb{L}^\infty_\Omega$. The associated norms are

$$\|X\|_{\mathbb{L}^2_\Omega} := \mathbb{E}\left[|X|^2\right]^{\frac{1}{2}}; \qquad \|X\|_{\mathbb{L}^\infty_\Omega} := \mathrm{essup}|\mathrm{X}| = \inf_{\mathrm{M}} \{\mathrm{M} \mid \mathbb{P}(|\mathrm{X}| \leq \mathrm{M}) = 1\}.$$

The Banach space $\mathbb{H}^{2,2}([0, T] \times \Omega, \mathbb{R}^k)$ (resp. $\mathbb{H}^{2,2}_{\mathcal{P}}([0, T] \times \Omega, \mathbb{R}^k)$) is the set of all $\mathbb{F}$-adapted (resp. $\mathbb{F}$-predictable) processes $\psi : [0, T] \times \Omega \to \mathbb{R}^k$ such that $\mathbb{E}\left[\left(\int_0^T |\psi_t|^2 \mathrm{d}t\right)\right] < +\infty$. The Banach space $\mathbb{H}^{\infty,2}\left([0, T] \times \Omega, \mathbb{R}^k\right)$ stands for the elements of $\mathbb{H}^{2,2}\left([0,T]\times\Omega,\mathbb{R}^k\right)$ satisfying $\mathbb{E}\left[\sup_{t\in[0,T]}|\psi_t|^2\right] < +\infty$. The Banach space $\mathbb{H}^{\infty,\infty}\left([0,T]\times\Omega,\mathbb{R}^k\right)$ (resp. $\mathbb{H}^{\infty,\infty}_{\mathcal{P}}\left([0,T]\times\Omega,\mathbb{R}^k\right)$) stands for the space of essentially bounded processes in $\mathbb{H}^{2,2}\left([0,T]\times\Omega,\mathbb{R}^k\right)$ (resp. $\mathbb{H}^{2,2}_{\mathcal{P}}\left([0,T]\times\Omega,\mathbb{R}^k\right)$). Here again we will omit the reference to $\mathbb{R}^k$ and $[0,T]\times\Omega$, which will be clear from the context. The associated norms are:

$$||\psi||_{\mathbb{H}^{2,2}} := \mathbb{E}\left[\left(\int_0^T |\psi_t|^2 \mathrm{d}t\right)\right]^{\frac{1}{2}}; \qquad ||\psi||_{\mathbb{H}^{\infty,2}} := \mathbb{E}\left[\sup_{t\in[0,T]}|\psi_t|^2\right]^{\frac{1}{2}}; \qquad ||\psi||_{\mathbb{H}^{\infty,\infty}} := \text{essup}\sup_{t\in[0,T]}|\psi_t|.$$

The space of martingales in $\mathbb{H}^{\infty,2}$ is denoted $\mathcal{M}^2$ and the space of martingales vanishing at $t=0$ is denoted $\mathcal{M}^2_0$.

2 Control problem: setting, assumptions and preliminary results

2.1 Setting and assumptions

We consider a stochastic control problem where the state dynamic is given by an ordinary differential equation, which is relevant for applications such as control of energy storage/conversion systems. Problems with a more general state dynamics, given by a stochastic differential equation with uncontrolled diffusion or jump terms can also be embedded in this framework, see Remark 2.1. In the field of energy management, this can be used to model water dams for instance, where the level of stored water depends on decisions (pumping, ...) and exogenous random processes, like water inflows arising from the rain or the ice melting in the mountains. More generally it allows to model a problem of control of an energy storage system subject to an exogenous random environment, which makes sense in a context of high renewable penetration.

We assume that the coefficients of the control problem are random, without any Markovian assumption and we do not suppose that the filtration is Brownian. We do not consider control nor state constraints. For clarity of the presentation, the results are established in the one dimensional-case, i.e., both the state and control variables are real-valued processes. However, they could be established in a higher dimension setting.

$$\left.\begin{aligned} &\mathcal{J}(u) := \mathbb{E}\left[\int_0^T l\left(t,\omega,u_{t,\omega},X^u_{t,\omega}\right)\mathrm{d}t + \Psi(\omega, X^u_{T,\omega})\right] \\ &\text{s.t. } X^u_{t,\omega} = x_0 + \int_0^t \alpha_{s,\omega}u_{s,\omega}\mathrm{d}s. \end{aligned}\right\} \longrightarrow \min_{u\in\mathbb{H}^{2,2}_{\mathcal{P}}}. \tag{2.1}$$

We consider the following regularity assumptions on the problem:

(Reg-1) The function $l : (t,\omega,u,x) \in [0,T]\times\Omega\times\mathbb{R}\times\mathbb{R} \mapsto l(t,\omega,u,x) \in \mathbb{R}$ is $\mathcal{P}\otimes\mathcal{B}(\mathbb{R})\otimes\mathcal{B}(\mathbb{R})$-measurable. The function $\Psi : (\omega,x) \in \Omega\times\mathbb{R} \mapsto \psi(\omega,x) \in \mathbb{R}$ is $\mathcal{F}_T\otimes\mathcal{B}(\mathbb{R})$-measurable. Besides, l and Ψ satisfy the growth conditions:

$$|l(t,\omega,u,x)| \le C^{(l)}_{t,\omega} + C\left(|u|^2+|x|^2\right),$$
$$|\Psi(\omega,x)| \le C^{(\Psi)}_{\omega} + C|x|^2,$$

with $C^{(l)} \in \mathbb{H}^{1,1}$, $C^{(\Psi)} \in \mathbb{L}^\infty_T$ and $C > 0$ a deterministic constant. We assume besides $x_0 \in \mathbb{L}^2_\Omega$.

(Reg-2) Assumption **(Reg-1)** holds and the function l is C^1 with respect to (u,x) and Ψ is C^1 with respect to x with derivatives satisfying:

$$|l'_v(t,\omega,u,x)| \le C^{(l')}_{t,\omega} + C'\left(|u|+|x|\right), \qquad v \in \{u,x\},$$
$$|\Psi'_x(\omega,x)| \le C^{(\Psi')}_{\omega} + C'|x|,$$

with $C^{(l')} \in \mathbb{H}^{2,2}$, $C^{(\Psi')} \in \mathbb{L}^2_T$ and $C' > 0$ a deterministic constant.

(Reg-3) Assumptions **(Reg-1)**-**(Reg-2)** hold and the functions l and Ψ are C^1 with respect to (u,x) with Lipschitz continuous derivatives.

(Reg-4) Assumptions **(Reg-1)**-**(Reg-2)**-**(Reg-3)** hold and the functions l and Ψ are C^2 with respect to (u, x) with bounded second derivatives (uniformly in (t, ω, u, x)).

(Reg-5) Assumptions **(Reg-1)**-**(Reg-2)**-**(Reg-3)**-**(Reg-4)** hold and the mappings l and Ψ have lipschitz-continuous second derivatives. Besides, the bounds $C^{(l)}$ and $C^{(l')}$ introduced earlier are in $\mathbb{H}^{\infty,\infty}$, and the constants $C^{(\Psi)}$ and $C^{(\Psi')}$ are in $\mathbb{L}^\infty_T$. We assume besides $x_0 \in \mathbb{L}^\infty_\Omega$.

We introduce the assumption of linearity of the dynamic:

(Lin-Dyn) The dynamic is affine-linear given by $\Phi : (t, \omega, u, x) \mapsto \alpha_{t,\omega} u_{t,\omega}$, with $\alpha \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$.

We consider the following convexity assumptions on the problem:

(Conv-1) The mapping l is convex in (u, x) and Ψ is convex in x.

(Conv-2) Assumption **(Conv-1)** holds and the mapping l is μ-strongly convex in u, with $\mu > 0$. In particular, under Assumption **(Reg-4)**, l''_{uu} is uniformly bounded from below by μ.

Remark 2.1. *The assumption* **(Lin-Dyn)** *is not restrictive and one could consider general affine-linear dynamics of the form:*

$$X^u_{t,\omega} = x_0 + \int_0^t (\alpha_{s,\omega} u_{s,\omega} + \beta_{s,\omega} X^u_{s,\omega}) \mathrm{d}s + M_{t,\omega},$$

with M an uncontrolled $(\mathcal{F}_t)$-adapted càdlàg uncontrolled process. Without loss of generality, we can assume $M = 0$, as we can reformulate the obtained problem in terms of $\tilde{X}^u = X^u - M$, up to minor modifications of l and Ψ. In the case $M = 0$, we can directly show for general $\beta \in \mathbb{H}^{\infty,\infty}$:

$$X^u_{t,\omega} = \exp\left(\int_0^t \beta_{s,\omega} \mathrm{d}s\right)\left(x_0 + \int_0^t \alpha_{s,\omega} u_{s,\omega} \exp\left(-\int_0^s \beta_{r,\omega} \mathrm{d}r\right) u_{s,\omega} \mathrm{d}s\right)$$

and thus the problem is equivalent to:

$$\left.\begin{array}{l}\mathcal{J}(u) := \mathbb{E}\left[\int_0^T \tilde{l}\left(t, \omega, u_{t,\omega}, \tilde{X}^u_{t,\omega}\right) \mathrm{d}t + \tilde{\Psi}(\tilde{X}^u_{T,\omega})\right] \\ \text{s.t. } \tilde{X}^u_{t,\omega} = x_0 + \int_0^t \tilde{\alpha}_{s,\omega} u_{s,\omega} \mathrm{d}s.\end{array}\right\} \longrightarrow \min_{u \in \mathbb{H}^{2,2}_{\mathcal{P}}}.$$

with:

$$\begin{cases}\tilde{l}(t, \omega, u, x) := l\left((t, \omega, u, \exp\left(\int_0^t \beta_{s,\omega} \mathrm{d}s\right) x\right), \\ \tilde{\alpha}_{t,\omega} := \alpha_{t,\omega} \exp\left(-\int_0^t \beta_{s,\omega} \mathrm{d}s\right), \\ \tilde{\Psi}(\omega, x) = \Psi\left(\omega, \exp\left(\int_0^T \beta_{s,\omega} \mathrm{d}s\right) x\right).\end{cases}$$

2.2 Well-posedness, existence and uniqueness of an optimal control

Proposition 2.2. *Under Assumption* **(Reg-1)** *and* **(Lin-Dyn)***, for any $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$, one can define $X^u \in \mathbb{H}^{\infty,2}$ by:*

$$X^u_t = x_0 + \int_0^t \alpha_s u_s \mathrm{d}s. \tag{2.2}$$

Besides, we have $\|X^u\|_{\mathbb{H}^{\infty,2}} \leq \sqrt{T}\|\alpha\|_{\mathbb{H}^{\infty,\infty}}\|u\|_{\mathbb{H}^{2,2}} + \|x_0\|_{\mathbb{L}^2_\Omega}$, $\|X^u - X^v\|_{\mathbb{H}^{\infty,2}} \leq \sqrt{T}\|\alpha\|_{\mathbb{H}^{\infty,\infty}}\|u - v\|_{\mathbb{H}^{2,2}}$ and $\mathcal{J}(u) < +\infty$.

Proposition 2.3. *Under Assumption* **(Reg-2)**-**(Lin-Dyn)**-**(Conv-2)***, $\mathcal{J}$ is continuous and strongly convex, coercive (i.e., $\lim_{\|u\|_{\mathbb{H}^{2,2}}} \mathcal{J}(u) = +\infty$) and hence $\mathcal{J}$ has a unique minimizer in $\mathbb{H}^{2,2}_{\mathcal{P}}$.*

Proof. The continuity of $u \in \mathbb{H}^{2,2}_{\mathcal{P}} \mapsto X^u \in \mathbb{H}^{\infty,2}$ holds thanks to **(Lin-Dyn)**, by Lebesgue's continuity theorem. The continuity of $\mathcal{J}$ stems from this fact, Lebesgue's continuity theorem and **(Reg-2)**. Under assumption **(Conv-2)**, $\mathcal{J}$ is μ-strongly convex and coercive. Besides, $\mathbb{H}^{2,2}_{\mathcal{P}}$ is reflexive, as it is a Hilbert space. Therefore, by [Bre10, Corollary 3.23, pp.71], $\mathcal{J}$ has a unique minimizer $u^* \in \mathbb{H}^{2,2}_{\mathcal{P}}$. □

2.3 First-order necessary and sufficient optimality conditions

We first prove first-order differentiability properties of the state variable and the cost function with respect to the control variable under suitable assumptions.

Lemma 2.4. *The application $\Phi_X : u \in \mathbb{H}^{2,2}_{\mathcal{P}} \mapsto X^u \in \mathbb{H}^{\infty,2}$ is Fréchet-differentiable. Besides, for any $(u,v) \in (\mathbb{H}^{2,2}_{\mathcal{P}})^2$, the derivative of Φ_X at point u in direction v is independent from u and given by:*

$$\dot{X}^v := \left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Phi_X(u+\varepsilon v)\right)_{|\varepsilon=0} = \int_0^t \alpha_s v_s \mathrm{d}s = X^v - x_0. \tag{2.3}$$

Besides, we have the following estimate:

$$\|\dot{X}^v\|_{\mathbb{H}^{\infty,2}} \leq C\|v\|_{\mathbb{H}^{2,2}}.$$

Proof. The application $u \in \mathbb{H}^{2,2}_{\mathcal{P}} \mapsto X^u \in \mathbb{H}^{\infty,2}$ is Gateaux-differentiable at $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ in direction $v \in \mathbb{H}^{2,2}_{\mathcal{P}}$ with derivative $\dot{X}^v := X^v - x_0 \in \mathbb{H}^{\infty,2}$. In particular, Φ_X is continuously differentiable and therefore Fréchet-differentiable. The bound $\|\dot{X}^v\|_{\mathbb{H}^{\infty,2}} \leq C\|v\|_{\mathbb{H}^{2,2}}$ arises from Cauchy-Schwarz inequality and the assumption $\alpha \in \mathbb{H}^{\infty,\infty}$. □

Proposition 2.5. *Suppose Assumptions* **(Reg-3)** *and* **(Lin-Dyn)** *hold. Then for any $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$, consider $X^u \in \mathbb{H}^{\infty,2}$ given in* (2.2) *and define Y^u by:*

$$Y^u_t = \mathbb{E}_t\left[\Psi'_x(X^u_T) + \int_t^T l'_x(s,u_s,X^u_s)\mathrm{d}s\right]. \tag{2.4}$$

Then, Y^u is well-defined and in $\mathbb{H}^{\infty,2}$. Besides, $\mathcal{J}$ is Fréchet-differentiable and admits a gradient at u denoted $\nabla\mathcal{J}(u) \in \mathbb{H}^{2,2}_{\mathcal{P}}$ given by:

$$\forall u \in \mathbb{H}^{2,2}_{\mathcal{P}}, \mathrm{d}\mathbb{P}\otimes\mathrm{d}t - a.e., \quad (\nabla\mathcal{J}(u))_t = l'_u(t,u_t,X^u_t) + \alpha_t Y^u_{t-}. \tag{2.5}$$

Besides, we have the following estimates for a deterministic constant C independent of u and v:

$$\forall u \in \mathbb{H}^{2,2}_{\mathcal{P}}, \qquad \|Y^u\|_{\mathbb{H}^{\infty,2}} + \|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{2,2}} \leq C(1 + \|u\|_{\mathbb{H}^{2,2}} + \|x_0\|_{\mathbb{L}^2_\Omega}),$$
$$\forall (u,v) \in (\mathbb{H}^{2,2}_{\mathcal{P}})^2, \qquad \|Y^u - Y^v\|_{\mathbb{H}^{\infty,2}} + \|\nabla\mathcal{J}(u) - \mathcal{J}(v)\|_{\mathbb{H}^{2,2}} \leq C\|u-v\|_{\mathbb{H}^{2,2}}.$$

Proof. The regularity assumptions combined with the estimation on $\|X^u\|_{\mathbb{H}^{\infty,2}}$ directly show that $Y^u \in \mathbb{H}^{\infty,2}$ with:

$$\|Y^u\|_{\mathbb{H}^{\infty,2}} \leq C(1 + \|u\|_{\mathbb{H}^{2,2}} + \|x_0\|_{\mathbb{L}^2_\Omega})$$
$$\|Y^u - Y^v\|_{\mathbb{H}^{\infty,2}} \leq C\|u-v\|_{\mathbb{H}^{2,2}}.$$

Admitting first the expression (2.5) for $\nabla\mathcal{J}(u)$, we can deduce from the regularity assumptions the bounds claimed.

Let us now prove that $\mathcal{J}$ is Fréchet-differentiable as well as the expression of $\nabla\mathcal{J}(u)$. By Lebesgue's differentiation theorem and Lemma 2.4, the application $\mathcal{J}$ is Gateaux-differentiable at u in direction v with derivative given by:

$$\dot{\mathcal{J}}(u,v) := \left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathcal{J}(u+\varepsilon v)\right)_{|\varepsilon=0} = \mathbb{E}\left[\Psi'_x(X^u_T)\dot{X}^v_T + \int_0^T \left(l'_x(s,u_s,X^u_s)\dot{X}^v_s + l'_u(s,u_s,X^u_s)v_s\right)\mathrm{d}s\right]. \tag{2.6}$$

Define the martingale $M^u \in \mathbb{H}^{\infty,2}$ by $M^u_t := \mathbb{E}_t\left[\psi'_x(X^u_T) + \int_0^T l'_x(s,u_s,X^u_s)\mathrm{d}s\right]$. Then $Y^u_t = M^u_t - \int_0^t l'_x(s,u_s,X^u_s)\mathrm{d}s$ so that (Y^u, M^u) satisfies the following BSDE in $(Y,M) \in \mathbb{H}^{2,2} \times \mathcal{M}^2_0$:

$$\begin{cases} -\mathrm{d}Y_t = l'_x(t,u_t,X^u_t)\mathrm{d}t - \mathrm{d}M_t, \\ Y_T = \Psi'_x(X^u_T). \end{cases}$$

Then, applying Integration by Parts Formula in [Pro03, Corollary 2, p. 68] to the product $Y^u \cdot \dot{X}^v$ between 0 and T yields, using $\dot{X}^v_0 = 0$, $Y^u_T = \Psi'_x(X^u_T)$, the fact that $\dot{X}^v$ is continuous with finite variations and the fact that $\int_{0^+}^t \dot{X}^v_s \mathrm{d}M^u_s$ is a càdlàg martingale in $\mathbb{H}^{\infty,2}$, see [Pro03, Theorem 20 p.63, Corollary 3 p.73, Theorem 29 p.75]:

$$\begin{aligned}\dot{\mathcal{J}}(u,v) &= \mathbb{E}\left[Y^u_T \dot{X}^v_T + \int_0^T \left(l'_x(s,u_s,X^u_s)\dot{X}^v_s + l'_u(s,u_s,X^u_s)v_s\right)\mathrm{d}s\right] \\ &= \mathbb{E}\left[\int_0^T \left(l'_x(s,u_s,X^u_s)\dot{X}^v_s + l'_u(s,u_s,X^u_s)v_s + Y^u_s\alpha_s v_s - \dot{X}^v_s l'_x(s,u_s,X^u_s)\right)\mathrm{d}s\right] \\ &= \mathbb{E}\left[\int_0^T \left(\alpha_s Y^u_s + l'_u(s,u_s,X^u_s)\right)v_s\mathrm{d}s\right] \\ &= \mathbb{E}\left[\int_0^T \left(\alpha_s Y^u_{s-} + l'_u(s,u_s,X^u_s)\right)v_s\mathrm{d}s\right].\end{aligned}$$

In the last inequality, we used the fact that Y^u has countably many jumps, and that the Lebesgue integral is left unchanged by changing the integrand on a countable set of points. This yields the expression of $\nabla\mathcal{J}(u)$. In particular, the previous estimates imply that $\mathcal{J}$ has a Lipschitz continuous gradient and is therefore Fréchet-differentiable.

□

Theorem 2.6 (First order necessary and sufficient optimality conditions)**.** *1. Suppose Assumptions* **(Reg-3)** *and* **(Lin-Dyn)** *hold. Assume $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ is a minimizer of $\mathcal{J}$. Define $X^u \in \mathbb{H}^{\infty,2}$ by* (2.2) *and $Y^u \in \mathbb{H}^{\infty,2}$ by* (2.4)*. Then, necessarily,*

$$l'_u(t,u_t,X^u_t) + \alpha_t Y^u_{t-} = 0, \quad \mathrm{d}\mathbb{P}\otimes\mathrm{d}t - a.e. \tag{2.7}$$

2. *Under Assumptions* **(Reg-3)**, **(Conv-1)** *and* **(Lin-Dyn)**, *if $(u, X^u, Y^u) \in \mathbb{H}^{2,2}_{\mathcal{P}} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ satisfies* (2.7) *with X^u given by* (2.2) *and Y^u given by* (2.4)*, then u is a solution of* (2.1)*, i.e., a minimizer of $\mathcal{J}$.*

Proof. 1. Under **(Reg-3)**, $\mathcal{J}$ is Gateaux-differentiable and an optimal control is necessary a critical point of $\mathcal{J}$, hence $\nabla\mathcal{J}(u) = 0$, which yields (2.7).

2. Under **(Conv-1)** and **(Lin-Dyn)**, $\mathcal{J}$ is convex and $\mathcal{J}$ is Gateaux-differentiable under **(Reg-3)**, so that (2.7) is a sufficient optimality condition.

□

2.4 Second-order differentiability

We now turn to second-order differentiability of the cost functional, necessary for the Newton method. We then prove a key result showing the invertibility of the second order derive of $\mathcal{J}$, and the form of the inverse. This shows the existence and provides a characterization of the Newton step.

Lemma 2.7. *Suppose Assumptions* **(Reg-4)** *and* **(Lin-Dyn)** *hold. Then the mapping*

$$\Phi_Y : \begin{cases}\mathbb{H}^{2,2}_{\mathcal{P}} \mapsto \mathbb{H}^{\infty,2} \\ u \mapsto Y^u\end{cases}$$

is Gateaux-differentiable. Furthermore, for all u, v in $\mathbb{H}^{2,2}_{\mathcal{P}}$, $D\Phi_Y(u)(v) = \dot{Y}^{u,v}$ is defined by the following affine-linear BSDE with Lipschitz coefficients:

$$\dot{Y}^{u,v}_t = \mathbb{E}_t\left[\Psi''_{xx}(X^u_T)\dot{X}^v_T + \int_t^T \left(l''_{xu}(s,u_s,X^u_s)v_s + l''_{xx}(s,u_s,X^u_s)\dot{X}^v_s\right)\mathrm{d}s\right]. \tag{2.8}$$

Besides, we have the following estimate:

$$\|\dot{Y}^{u,v}\|_{\mathbb{H}^{\infty,2}} \leq C\|v\|_{\mathbb{H}^{2,2}}.$$

Proof. By our assumptions, $\dot{Y}^{u,v}$ is well defined and by Lebesgue's differentiation theorem and Lemma 2.4, it is straightforward that Φ_Y is Gateau-differentiable and that: $D\Phi_Y(u)(v) = \dot{Y}^{u,v}$. Applying Lebesgue's theorem and using our estimation on $\|\dot{X}^v\|_{\mathbb{H}^{\infty,2}}$, one gets $\|\dot{Y}^{u,v}\|_{\mathbb{H}^{\infty,2}} \leq C\|v\|_{\mathbb{H}^{2,2}}$. ◻

Proposition 2.8 (Second-order differentiability)**.** *Suppose Assumptions* **(Reg-4)** *and* **(Lin-Dyn)** *hold. Then the mapping $\mathcal{J}$ is twice Gateaux differentiable and its second-order derivative $\nabla^2\mathcal{J} : \mathbb{H}^{2,2}_{\mathcal{P}} \mapsto \mathcal{L}(\mathbb{H}^{2,2}_{\mathcal{P}})$ is given by:*

$$\forall v \in \mathbb{H}^{2,2}, \mathrm{d}\mathbb{P}\otimes \mathrm{d}t - a.e., \quad \left(\nabla^2\mathcal{J}(u)(v)\right)_t = l''_{uu}(t, u_t, X^u_t)v_t + l''_{ux}(t, u_t, X^u_t)\dot{X}^v_t + \alpha_t \dot{Y}^{u,v}_{t-}. \tag{2.9}$$

Besides, we have the following estimate:

$$\|\nabla^2\mathcal{J}(u)(v)\|_{\mathbb{H}^{2,2}} \leq C\|v\|_{\mathbb{H}^{2,2}},$$

where C is a constant independent of u. In other words, for any $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$, $\nabla^2\mathcal{J}(u)$ is a continuous endomorphism of $\mathbb{H}^{2,2}_{\mathcal{P}}$.

Proof. Applying Lebesgue's differentiation theorem to $\nabla\mathcal{J}(u)$ given by (2.5) yields (2.9), using Lemmas 2.4 and 2.7. The continuity of $\nabla^2\mathcal{J}(u)$ for all $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ results from the previous estimates, our assumptions and Lebesgue's differentiation theorem. ◻

The computation of the Newton step Δu amounts to solve the equation $\nabla^2\mathcal{J}(u)(\Delta u) = -\nabla\mathcal{J}(u)$, as expected by an infinite-dimensional generalization of the Newton method in $\mathbb{R}^d$. The following theorem is the key result which guarantees that this infinite dimensional equation has a unique solution, as we show invertibility of the second order derivative of the cost function at any admissible point. This theorem also makes the connection between the computation of the Newton step and the solution of an auxiliary Linear-Quadratic stochastic control problem, or equivalently, with the solution of an affine-linear FBSDE with random coefficients.

Theorem 2.9. *Suppose Assumptions* **(Conv-2)**, **(Reg-4)** *and* **(Lin-Dyn)** *hold. Let $(u, w) \in \mathbb{H}^{2,2}_{\mathcal{P}} \times \mathbb{H}^{2,2}_{\mathcal{P}}$ and define $X^u \in \mathbb{H}^{\infty,2}$ by* (2.2)*. We introduce the following auxiliary (linear-quadratic) stochastic control problem:*

$$\begin{aligned} &\min_{v\in\mathbb{H}^{2,2}_{\mathcal{P}}} \tilde{\mathcal{J}}^{quad,u,w}(v) \\ &\text{s.t. } \tilde{X}_t = \int_0^t \alpha_s v_s \mathrm{d}s. \end{aligned} \tag{2.10}$$

where $\tilde{\mathcal{J}}^{quad,u,w}(v)$ is defined by:

$$\mathbb{E}\left[\int_0^T \left\{\frac{1}{2}l''_{uu}\left(t, u_t, X^u_t\right)v_t^2 + \frac{1}{2}l''_{xx}\left(t, u_t, X^u_t\right)\tilde{X}_t^2 + l''_{ux}\left(t, u_t, X^u_t\right)\tilde{X}_t v_t - w_t v_t\right\}\mathrm{d}t + \frac{1}{2}\Psi''_{xx}(X^u_T)\tilde{X}_T^2\right].$$

Then $\tilde{\mathcal{J}}^{quad,u,w}$ has a unique minimizer $\tilde{u}^{u,w} \in \mathbb{H}^{2,2}_{\mathcal{P}}$ defined by

$$l''_{uu}(t, u_t, X^u_t)\tilde{u}^{u,w}_t + l''_{ux}(t, u_t, X^u_t)\tilde{X}^{u,w}_t + \alpha_t \tilde{Y}^{u,w}_{t-} = w_t,$$

where $(\tilde{X}^{u,w}, \tilde{Y}^{u,w}) \in \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ are given by:

$$\begin{cases} \tilde{X}^{u,w}_t = \int_0^t \alpha_s \tilde{u}^{u,w}_s \mathrm{d}s, \\ \tilde{Y}^{u,w}_t = \mathbb{E}_t\left[\Psi''_{xx}(X^u_T)\tilde{X}^{u,w}_T + \int_t^T \left(l''_{xu}(s, u_s, X^u_s)\tilde{u}^{u,w}_s + l''_{xx}(s, u_s, X^u_s)\tilde{X}^{u,w}_s\right)\mathrm{d}s\right]. \end{cases}$$

In particular, $(\tilde{u}^{u,w}, \tilde{X}^{u,w}, \tilde{Y}^{u,w}) \in \mathbb{H}^{2,2}_{\mathcal{P}} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ is the unique solution of an affine-linear FBSDE with random coefficients (which depend on the stochastic process $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$). Besides, for any $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$, $\nabla^2\mathcal{J}(u) \in \mathcal{L}(\mathbb{H}^{2,2}_{\mathcal{P}})$ is invertible and for any $w \in \mathbb{H}^{2,2}_{\mathcal{P}}$, $\left(\nabla^2\mathcal{J}(u)\right)^{-1}(w) = \tilde{u}^{u,w}$.

Proof. Introduce the auxiliary running cost function $\tilde{l}^{u,w} : [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ defined by $\tilde{l}^{u,w}(t, \omega, \tilde{u}, \tilde{x}) = \frac{1}{2}l''_{uu}\left(t, u_t, X^u_t\right)\tilde{u}_t^2 + \frac{1}{2}l''_{xx}\left(t, u_t, X^u_t\right)\tilde{x}_t^2 + l''_{ux}\left(t, u_t, X^u_t\right)\tilde{u}_t\tilde{x}_t - w_t\tilde{u}_t$, where we dropped the reference to ω for simplicity. Introduce as well the auxiliary terminal cost function $\tilde{\Psi}^{u,w} : \Omega \times \mathbb{R} \mapsto \mathbb{R}$ defined by $\tilde{\Psi}^{u,w}(\omega, \tilde{x}) = \frac{1}{2}\Psi''_{xx}(X^u_T)\tilde{x}^2$. Then

Assumptions **(Reg-4)**, **(Lin-Dyn)** and **(Conv-2)** are verified for the auxiliary minimization problem of $\tilde{\mathcal{J}}^{quad,u,w}$, with l and Ψ respectively replaced by $\tilde{l}^{u,w}$ and $\tilde{\Psi}^{u,w}$. Applying Proposition 2.3 to the auxiliary problem shows the existence and uniqueness of a minimizer, denoted $\tilde{u}^{u,w}$. Applying Theorem 2.6 to the auxiliary problem, we have existence and uniqueness of $(\tilde{X}^{u,w}, \tilde{Y}^{u,w}) \in \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ such that $(\tilde{u}^{u,w}, \tilde{X}^{u,w}, \tilde{Y}^{u,w}) \in \mathbb{H}^{2,2}_{\mathcal{P}} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ is the (unique) solution of the FBSDE:

$$\begin{cases} \tilde{X}^{u,w}_t = \int_0^t \alpha_s \tilde{u}^{u,w}_s \mathrm{d}s, \\ \tilde{Y}^{u,w}_t = \mathbb{E}_t \left[\Psi''_{xx}(X^u_T)\tilde{X}^{u,w}_T + \int_t^T \left(l''_{xu}(s,u_s,X^u_s)\tilde{u}^{u,w}_s + l''_{xx}(s,u_s,X^u_s)\tilde{X}^{u,w}_s \right) \mathrm{d}s \right], \\ l''_{uu}(t,u_t,X^u_t)\tilde{u}^{u,w}_t + l''_{ux}(t,u_t,X^u_t)\tilde{X}^{u,w}_t + \alpha_t \tilde{Y}^{u,w}_{t-} = w_t. \end{cases}$$

In particular, one has $\tilde{X}^{u,w} = \dot{X}^{\tilde{u}^{u,w}}$ and $\tilde{Y}^{u,w} = \dot{Y}^{\tilde{u}^{u,w}}$. The last equation therefore writes $l''_{uu}(t,u_t,X^u_t)\tilde{u}^{u,w}_t + l''_{ux}(t,u_t,X^u_t)\dot{X}^{\tilde{u}^{u,w}}_t + \alpha_t \dot{Y}^{\tilde{u}^{u,w}}_{t-} = w_t$. Recognizing $\nabla^2\mathcal{J}(u)(\tilde{u}^{u,w})$ in the left-hand side (see (2.9)), we get in particular the existence and uniqueness of $\tilde{u}^{u,w} \in \mathbb{H}^{2,2}_{\mathcal{P}}$ solution of the equation $\nabla^2\mathcal{J}(u)(\tilde{u}^{u,w}) = w$. This holds for all $w \in \mathbb{H}^{2,2}_{\mathcal{P}}$. We thus obtain the invertibility of $\nabla^2\mathcal{J}(u)$ and the expression $(\nabla^2\mathcal{J}(u))^{-1}(w) = \tilde{u}^{u,w}$, for all $u, w \in \mathbb{H}^{2,2}_{\mathcal{P}}$. □

3 The Newton method for stochastic control problems

3.1 Definition and interpretation of the Newton step

Definition 3.1. *Suppose Assumptions* **(Conv-2)**, **(Reg-4)** *and* **(Lin-Dyn)** *hold. Let $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ and define $X^u \in \mathbb{H}^{\infty,2}$ by* (2.2). *The Newton step Δ_u of $\mathcal{J}$ at the point $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ is defined by $\Delta_u = -(\nabla^2\mathcal{J}(u))^{-1}(\nabla\mathcal{J}(u)) \in \mathbb{H}^{2,2}_{\mathcal{P}}$.*

The following Proposition shows that computation of the Newton step at point u, $-(\nabla^2\mathcal{J}(u))^{-1}(\nabla\mathcal{J}(u))$ amounts to solve a Linear-Quadratic approximation of the original problem around the current point u, based on a point-wise second order expansion of the cost and a first order expansion of the dynamic.

Proposition 3.2. *Suppose Assumptions* **(Conv-2)**, **(Reg-4)** *and* **(Lin-Dyn)** *hold. Let $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ and define $X^u \in \mathbb{H}^{\infty,2}$ by* (2.2). *Denote by $\theta^u_t := (t, u_t, X^u_t)$. The Newton step $\Delta_u = -(\nabla^2\mathcal{J}(u))^{-1}(\nabla\mathcal{J}(u)) \in \mathbb{H}^{2,2}_{\mathcal{P}}$ of $\mathcal{J}$ at the point u is the unique minimizer in $\mathbb{H}^{2,2}_{\mathcal{P}}$ of the Linear-Quadratic approximation $\mathcal{J}^{LQ,u}$ of $\mathcal{J}$ around u, defined by:*

$$\begin{aligned} \forall v \in \mathbb{H}^{2,2}_{\mathcal{P}}, \quad \mathcal{J}^{LQ,u}(v) := &\mathbb{E}\left[\int_0^T \left\{ \frac{1}{2} l''_{uu}(\theta^u_t) v_t^2 + \frac{1}{2} l''_{xx}(\theta^u_t)(\dot{X}^v_t)^2 + l''_{ux}(\theta^u_t)\dot{X}^v_t v_t + l'_u(\theta^u_t) v_t + l'_x(\theta^u_t)\dot{X}^v_t + l(\theta^u_t) \right\} \mathrm{d}t \right] \\ &+ \mathbb{E}\left[\frac{1}{2}\Psi''_{xx}(X^u_T)(\dot{X}^v_T)^2 + \Psi'_x(X^u_T)\dot{X}^v_T + \Psi(X^u_T) \right], \end{aligned}$$

where $\dot{X}^v_t = \int_0^t \alpha_s v_s \mathrm{d}s$.

Proof. Introduce the auxiliary running cost function $l^{LQ,u} : [0,T]\times\Omega\times\mathbb{R}\times\mathbb{R} \mapsto \mathbb{R}$ defined by $l^{LQ,u}(t,\tilde{u},\tilde{x}) = \frac{1}{2} l''_{uu}\left(\theta^u_t\right)\tilde{u}^2_t + \frac{1}{2} l''_{xx}\left(\theta^u_t\right)\tilde{x}^2_t + l''_{ux}\left(\theta^u_t\right)\tilde{u}\tilde{x} + l'_u\left(\theta^u_t\right)\tilde{u} + l'_x\left(\theta^u_t\right)\tilde{x} + l\left(\theta^u_t\right)$, where we dropped the reference to ω for simplicity. Introduce as well the auxiliary terminal cost function $\Psi^{LQ,u} : \Omega \times \mathbb{R} \mapsto \mathbb{R}$ defined by $\Psi^{LQ,u}(\omega, \tilde{x}) = \frac{1}{2}\Psi''_{xx}(X^u_T)\tilde{x}^2 + \Psi'_x(X^u_T)\tilde{x} + \Psi(X^u_T)$. Then Assumptions **(Reg-4)**, **(Lin-Dyn)** and **(Conv-2)** are verified for the auxiliary minimization problem of $\mathcal{J}^{LQ,u}$, with l and Ψ respectively replaced by $l^{LQ,u}$ and $\Psi^{LQ,u}$. Applying Proposition 2.3 to the auxiliary problem shows the existence and uniqueness of a minimizer of $\mathcal{J}^{LQ,u}$, denoted by $\hat{u}$. We can then apply Theorem 2.6 to the auxiliary problem and get existence and uniqueness of $(\hat{X}, \hat{Y}) \in \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ such that $(\hat{u}, \hat{X}, \hat{Y}) \in \mathbb{H}^{2,2}_{\mathcal{P}} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ is the (unique) triple satisfying:

$$\begin{cases} \hat{X}_t = \int_0^t \alpha_s \hat{u}_s \mathrm{d}s, \\ \hat{Y}_t = \mathbb{E}_t \left[\Psi''_{xx}(X^u_T)\hat{X}_T + \Psi'_x(X^u_T) + \int_t^T \left(l''_{xu}(s,u_s,X^u_s)\hat{u}_s + l''_{xx}(s,u_s,X^u_s)\hat{X}_s + l'_x(s,u_s,X^u_s) \right) \mathrm{d}s \right], \\ l''_{uu}(t,u_t,X^u_t)\hat{u}_t + l''_{ux}(t,u_t,X^u_t)\hat{X}_t + l'_u(t,u_t,X^u_t) + \alpha_t \hat{Y}_{t-} = 0. \end{cases}$$

In particular, we have $\hat{X} = \dot{X}^{\hat{u}}$ by (2.3), $\hat{Y} = Y^u + \dot{Y}^{u,\hat{u}}$ by (2.4) and(2.8). Besides, the last equation is equivalent to $\nabla\mathcal{J}(u) + \nabla^2\mathcal{J}(u)(\hat{u}) = 0$ by (2.5) and (2.9). In particular, the minimizer $\hat{u}$ of $\mathcal{J}^{LQ,u}$ is nothing else than the Newton step of $\mathcal{J}$ at point u. □

Without surprises, solving such a linear-quadratic stochastic control problem is equivalent to solving an affine-linear FBSDE. This gives a second interpretation of the Newton step as solution of the linearized first order optimality conditions of the control problem, summed up in the following Proposition.

Proposition 3.3. *Suppose Assumptions* **(Conv-2)**, **(Reg-4)** *and* **(Lin-Dyn)** *hold. Let* $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ *and define* $X^u \in \mathbb{H}^{\infty,2}$ *by* (2.2). *Let* $\Delta_u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ *be the Newton step of* $\mathcal{J}$ *at the point* $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$. *Define* $(X^u, Y^u, \dot{X}^{\Delta_u}, \dot{Y}^{\Delta_u}) \in \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ *by:*

$$\begin{cases} X^u_t = x_0 + \int_0^t \alpha_s u_s \mathrm{d}s, \\ Y^u_t = \mathbb{E}_t\left[\Psi'_x(X^u_T) + \int_t^T l'_x(s, u_s, X^u_s)\mathrm{d}s\right], \\ \dot{X}^{\Delta_u}_t = \int_0^t \alpha_s(\Delta_u)_s \mathrm{d}s, \\ \dot{Y}^{\Delta_u}_t = \mathbb{E}_t\left[\Psi''_x(X^u_T)\dot{X}^{\Delta_u}_T + \int_t^T \left(l''_{xx}(s, u_s, X^u_s)\dot{X}^{\Delta_u}_s + l''_{xu}(s, u_s, X^u_s)(\Delta_u)_s\right)\mathrm{d}s\right]. \end{cases}$$

Then $X^u + \dot{X}^{\Delta_u}$ *is the first-order approximation of* $v \mapsto X^{u+v}$ *evaluated at* $v = \Delta_u$, $Y^u + \dot{Y}^{\Delta_u}$ *is the first-order approximation of* $v \mapsto Y^{u+v}$ *evaluated at* $v = \Delta_u$ *and* Δ_u *is the zero of the linearized gradient of* $\mathcal{J}$ *around* u, *i.e.,* $\nabla\mathcal{J}(u) + \nabla^2\mathcal{J}(u)(\Delta_u) = 0$.

Proof. This is a direct consequence of Lemmas 2.4 and 2.7, as well as the definition of the Newton step 3.1. □

Remark 3.4. *In Brownian filtrations, one can derive similar results with more general dynamics of the state variable. Consider the following dynamic (controlled diffusion):*

$$X_t = x_0 + \int_0^t (\alpha_s u_s + \beta_s X_s + \gamma_s)\mathrm{d}s + \int_0^t (\bar{\alpha}_s u_s + \bar{\beta}_s X_s + \bar{\gamma}_s)\mathrm{d}W_s.$$

We make the same assumptions on the cost functional. One can apply Pontryagin principle to show necessary and sufficient optimality condition of order 1. The computation of the second-order derivative can be adapted to this particular case as well. Then, one can show that the Newton step can be defined in this setting, and that its computation amounts to solve an affine-linear FBSDE with stochastic coefficients. In this case, as the filtration is Brownian, the affine-linear FBSDE is computable using directly the results in [Yon06].

3.2 Solution of affine-linear FBSDEs and computation of the Newton step

Computing the Newton step is equivalent to compute the inverse of $(\nabla^2\mathcal{J}(u))^{-1}(w)$ with $w = -\nabla\mathcal{J}(u)$. By Theorem 2.9, computing this quantity is equivalent to solving the following affine-linear FBSDE:

$$\begin{cases} \tilde{X}^{u,w}_t = \int_0^t \left\{-\frac{\alpha_s l''_{ux}(s, u_s, X^u_s)}{l''_{uu}(s, u_s, X^u_s)}\tilde{X}^{u,w}_s - \frac{\alpha_s^2}{l''_{uu}(s, u_s, X^u_s)}\tilde{Y}^{u,w}_s + \frac{\alpha_s w_s}{l''_{uu}(s, u_s, X^u_s)}\right\}\mathrm{d}s, \\ \tilde{Y}^{u,w}_t = \mathbb{E}_t\left[\Psi''_{xx}(X^u_T)\tilde{X}^{u,w}_T + \int_t^T \left(\frac{l''_{uu}(s, u_s, X^u_s)l''_{xx}(s, u_s, X^u_s) - (l''_{ux}(s, u_s, X^u_s))^2}{l''_{uu}(s, u_s, X^u_s)}\tilde{X}^{u,w}_s - \frac{l''_{xu}(s, u_s, X^u_s)\alpha_s}{l''_{uu}(s, u_s, X^u_s)}\tilde{Y}^{u,w}_s + \frac{l''_{xu}(s, u_s, X^u_s)w_s}{l''_{uu}(s, u_s, X^u_s)}\right)\mathrm{d}s\right], \end{cases} \tag{3.1}$$

and then $(\nabla^2\mathcal{J}(u))^{-1}(w)$ is given by:

$$\left((\nabla^2\mathcal{J}(u))^{-1}(w)\right)_t = -\frac{1}{l''_{uu}(t, u_t, X^u_t)}\left(l''_{ux}(t, u_t, X^u_t)\tilde{X}^{u,w}_t + \alpha_t \tilde{Y}^{u,w}_{t-} - w_t\right). \tag{3.2}$$

Note that $(\tilde{X}^{u,w}, \tilde{Y}^{u,w})$ are solution of an affine-linear FBSDE with stochastic coefficients which has the following structure:

$$\begin{cases} X_t = x + \int_0^t (A_s X_s + B_s Y_s + a_s)\,\mathrm{d}s, \\ Y_t = \mathbb{E}_t\left[\Gamma X_T + \eta + \int_t^T (C_s X_s + A_s Y_s + b_s)\,\mathrm{d}s\right] \end{cases} \tag{3.3}$$

with $C \geq 0$ (convexity of l with respect to (u, x) and strong convexity of l with respect to u), $\Gamma \geq 0$, $B \leq 0$. A, B and C are in $\mathbb{H}^{\infty,\infty}$, a and b are in $\mathbb{H}^{2,2}$, $\Gamma \in \mathbb{L}^{\infty}(\mathcal{F}_T)$, $\eta \in \mathbb{L}^2(\mathcal{F}_T)$ and $x \in \mathbb{L}^2(\mathcal{F}_0)$.

Affine-linear FBSDEs have been studied in the literature, and a solution method is based on a so-called decoupling field, assumed affine-linear, by the structure of the equation. The assumption is that the solution verifies $Y = PX + \Pi$ for some P and Π to determine. Standard results in the literature [Yon06] show that P solves a matrix Riccati BSDE and Π solves an affine-linear BSDE. However, we are outside of the scope of [Yon06], which assumes a Brownian filtration. The following Lemma gives some results on solutions of Riccati BSDEs in general filtrations.

Lemma 3.5 (One-dimensional Riccati-BSDE under general filtrations)**.** *Let A, B, C be processes in $\mathbb{H}^{\infty,\infty}$ and $\Gamma \in \mathbb{L}^{\infty}(\mathcal{F}_T)$. Suppose additionally that $\Gamma \geq 0$ $\mathrm{d}\mathbb{P}$-a.e., $B_t \leq 0$ $\mathrm{d}\mathbb{P} \otimes \mathrm{d}t$-a.e., $C_t \geq 0$ $\mathrm{d}\mathbb{P} \otimes \mathrm{d}t$-a.e. Then the following Riccati BSDE with unknown P and stochastic coefficients:*

$$P_t = \mathbb{E}_t\left[\Gamma + \int_t^T \left(2A_sP_s + B_sP_s^2 + C_s\right)\mathrm{d}s\right] \tag{3.4}$$

has a unique solution in $\mathbb{H}^{\infty,\infty}([0, T])$ and we have the estimation $0 \leq P_t \leq \overline{P}_t$ $\mathrm{d}\mathbb{P} \otimes \mathrm{d}t$-a.e. with:

$$\overline{P}_t = \mathbb{E}_t\left[\Gamma \exp\left(2\|A\|_{\mathbb{H}^{\infty,\infty}}(T - t)\right) + \int_t^T C_s \exp\left(2\|A\|_{\mathbb{H}^{\infty,\infty}}(s - t)\right)\mathrm{d}s\right]. \tag{3.5}$$

A general result of Bismut for existence and uniqueness of the solution of Riccati BSDE can be found in [Bis76, Theorem 6.1]. In section 5.1, we provide our own proof in the one-dimensional case. It is based on the comparison principle for BSDEs, and allows to prove that the solution of the Riccati BSDE coincides with the solution of a BSDE with a truncated drift, globally Lipschitz continuous. As a limitation, the comparison principle applies only for one-dimensional BSDEs. Therefore, our proof cannot be expected to be generalized to higher dimension, except if an analogous comparison principle for BSDEs with square symmetric matrix unknown is developed, using the order defined by the cone of positive semi-definite matrices.

We now give a result on the second ingredient allowing to solve coupled linear FBSDEs.

Lemma 3.6. *[1-dimensional affine-linear BSDE in general filtrations] Let A, B, C and P be as before. Suppose $a, b \in \mathbb{H}^{2,2}$ and $\eta \in \mathbb{L}^2(\mathcal{F}_T)$. Define $\Pi \in \mathbb{H}^{\infty,2}$ by:*

$$\Pi_t = \mathbb{E}_t\left[\eta \exp\left(\int_t^T (P_sB_s + A_s)\mathrm{d}s\right) + \int_t^T (a_sP_s + b_s)\exp\left(\int_t^s (P_rB_r + A_r)\mathrm{d}r\right)\mathrm{d}s\right]. \tag{3.6}$$

Then Π is the unique solution in $\mathbb{H}^{\infty,2}$ of the BSDE:

$$\Pi_t = \mathbb{E}_t\left[\eta + \int_t^T \left((P_sB_s + A_s)\Pi_s + a_sP_s + b_s\right)\mathrm{d}s\right]. \tag{3.7}$$

Additionally, we have the estimation:

$$\|\Pi\|_{\mathbb{H}^{\infty,2}} \leq \left(\|\eta\|_{\mathbb{L}^2} + \sqrt{T}\|a\|_{\mathbb{H}^{2,2}}\|P\|_{\mathbb{H}^{\infty,\infty}} + \sqrt{T}\|b\|_{\mathbb{H}^{2,2}}\right)e^{\|PB+A\|_{\mathbb{H}^{\infty,\infty}}T}.$$

The proof is given in Section 5.2. We are now in position of deriving a verification Theorem for the solution of affine-linear FBSDEs, based on the two previous Lemmas.

Theorem 3.7. *[Scalar affine-linear FBSDEs with exogenous noise under general filtrations] Let A, B, C be processes in $\mathbb{H}^{\infty,\infty}$ and $\Gamma \in \mathbb{L}^{\infty}(\mathcal{F}_T)$. Let $a, b \in \mathbb{H}^{2,2}$, $x \in \mathbb{L}^2(\mathcal{F}_0)$ and $\eta \in \mathbb{L}^2(\mathcal{F}_T)$. Suppose additionally that $\Gamma \geq 0$ $\mathrm{d}\mathbb{P}$-a.e., $B_t \leq 0$ $\mathrm{d}\mathbb{P} \otimes \mathrm{d}t$-a.e., $C_t \geq 0$ $\mathrm{d}\mathbb{P} \otimes \mathrm{d}t$-a.e. Then:*

1. *The following FBSDE has a unique solution $(X, Y) \in (\mathbb{H}^{\infty,2})^2$:*

$$\begin{cases} X_t = x + \int_0^t (A_sX_s + B_sY_s + a_s)\,\mathrm{d}s, \\ Y_t = \mathbb{E}_t\left[\Gamma X_T + \eta + \int_t^T (C_sX_s + A_sY_s + b_s)\,\mathrm{d}s\right]. \end{cases}$$

2. *Define $P \in \mathbb{H}^{\infty,\infty}$ by* (3.4) *and $\Pi \in \mathbb{H}^{\infty,2}$ by* (3.6)*. Define as well X by:*

$$X_t = x + \int_0^t ((A_s + B_s P_s)X_s + B_s \Pi_s + a_s)\,\mathrm{d}s. \tag{3.8}$$

and

$$Y = PX + \Pi. \tag{3.9}$$

The processes X and Y are well-defined and in $\mathbb{H}^{\infty,2}$ with the estimates:

$$\|X\|_{\mathbb{H}^{\infty,2}} \le \left(\|x\|_{\mathbb{L}^2_\Omega} + T\|B\|_{\mathbb{H}^{\infty,\infty}}\|\Pi\|_{\mathbb{H}^{\infty,2}} + \sqrt{T}\|a\|_{\mathbb{H}^{2,2}}\right) e^{\|A+PB\|_{\mathbb{H}^{\infty,\infty}}T},$$

and

$$\|Y\|_{\mathbb{H}^{\infty,2}} \le \|P\|_{\mathbb{H}^{\infty,\infty}}\|X\|_{\mathbb{H}^{\infty,2}} + \|\Pi\|_{\mathbb{H}^{\infty,2}}.$$

Besides, $(X, Y) \in (\mathbb{H}^{\infty,2})^2$ is the unique solution of the FBSDE:

$$\begin{cases} X_t = x + \int_0^t (A_s X_s + B_s Y_s + a_s)\,\mathrm{d}s, \\ Y_t = \mathbb{E}_t\left[\Gamma X_T + \eta + \int_t^T (C_s X_s + A_s Y_s + b_s)\,\mathrm{d}s\right]. \end{cases}$$

The proof of this Theorem is postponed to Section 5.3. Applying the above result to (3.1) and using (3.2), we get the following Corollary.

Corollary 3.8 (Explicit computation of the inverse of $\nabla^2\mathcal{J}(u)$)**.** *Suppose assumptions* **(Reg-4)**, **(Conv-2)** *and* **(Lin-Dyn)** *hold. Let $u, w \in \mathbb{H}^{2,2}_{\mathcal{P}}$ and define X^u and Y^u in $\mathbb{H}^{\infty,2}$ as in* (2.2) *and* (2.4)*. Define as well:*

$$\begin{cases} A^u_t = -\dfrac{\alpha_t l''_{ux}(t, u_t, X^u_t)}{l''_{uu}(t, u_t, X^u_t)}, \\ B^u_t = -\dfrac{\alpha_t^2}{l''_{uu}(t, u_t, X^u_t)}, \\ C^u_t = \dfrac{l''_{uu}(t, u_t, X^u_t) l''_{xx}(t, u_t, X^u_t) - (l''_{ux}(t, u_t, X^u_t))^2}{l''_{uu}(t, u_t, X^u_t)}, \\ \Gamma^u = \Psi''_{xx}(X^u_T), \\ a^{u,w}_t = \dfrac{\alpha_t w_t}{l''_{uu}(t, u_t, X^u_t)}, \\ b^{u,w}_t = \dfrac{l''_{ux}(t, u_t, X^u_T) w_t}{l''_{uu}(t, u_t, X^u_t)}. \end{cases} \tag{3.10}$$

Then the following Riccati BSDE with unknown P:

$$P_t = \mathbb{E}_t\left[\Gamma^u + \int_t^T \left(2A^u_s P_s + B^u_s P_s^2 + C^u_s\right)\mathrm{d}s\right] \tag{3.11}$$

has a unique solution in $\mathbb{H}^{\infty,\infty}$, denoted P^u and the following affine-linear BSDE with unknown Π:

$$\Pi_t = \mathbb{E}_t\left[\int_t^T ((P^u_s B^u_s + A^u_s)\Pi_s + a^{u,w}_s P^u_s + b^{u,w}_s)\,\mathrm{d}s\right]. \tag{3.12}$$

has a unique solution in $\mathbb{H}^{\infty,2}$, which is denoted $\Pi^{u,w}$. Define $\tilde{X}^{u,w} \in \mathbb{H}^{\infty,2}$ as the unique solution of the following ordinary differential equation:

$$\tilde{X}_t = \int_0^t \left((A^u_s + B^u_s P^u_s)\tilde{X}_s + B^u_s \Pi^{u,w}_s + a^{u,w}_s\right)\mathrm{d}s. \tag{3.13}$$

Then:

$$\left((\nabla^2\mathcal{J}(u))^{-1}(w)\right)_t = -\frac{1}{l''_{uu}(t,u_t,X^u_t)}\left(\{l''_{ux}(t,u_t,X^u_t) + \alpha_t P^u_{t-}\}\,\tilde{X}^{u,w}_t + \alpha_t \Pi^{u,w}_{t-} - w_t\right). \tag{3.14}$$

Besides, we have:

$$\begin{aligned} &\|P^u\|_{\mathbb{H}^{\infty,\infty}} \le C,\\ &\|\Pi^{u,w}\|_{\mathbb{H}^{\infty,2}} \le C\|w\|_{\mathbb{H}^{2,2}},\\ &\|(\nabla^2\mathcal{J}(u))^{-1}(w)\|_{\mathbb{H}^{2,2}} \le C\|w\|_{\mathbb{H}^{2,2}}, \end{aligned}$$

for some constant C independent of u and w. In particular, $\nabla^2\mathcal{J}(u)$ is a bi-continuous isomorphism of $\mathbb{H}^{2,2}_{\mathcal{P}}$ for any $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$.

The proof of this Corollary is postponed to Section 5.4.

3.3 Global convergence of Newton's method with an adapted line-search method

For linear-quadratic problems with random coefficient, the Newton direction is equal to the optimal solution of the problem minus the current point, so that the Newton method converges in one iteration.

In the finite-dimensional case, the local quadratic convergence of Newton's method typically requires the function to minimize to have a Lipschitz continuous second-order derivative, see [Kan48]-[NW06, Theorem 3.5, p. 44], or to be self-concordant [NN94]. As self-concordance is not a notion well-defined in our setting, we focus on the first assumption.

We provide next a -example of $\mathcal{J} : \mathbb{H}^{2,2}_{\mathcal{P}} \mapsto \mathbb{R}$ to show that, even under assumption **(Reg-5)**, $\mathcal{J}$ may have a second-order derivative $\nabla^2\mathcal{J} : \mathbb{H}^{2,2} \mapsto \mathcal{L}(\mathbb{H}^{2,2})$ which is not Lipschitz-continuous. This shows that the (local) convergence of the Newton method should be established in a different space.

Example 3.9. *Let us assume $T = 1$ and consider $\mathcal{J}$ given by:*

$$\forall u \in \mathbb{H}^{2,2}_{\mathcal{P}}, \quad \mathcal{J}(u) := \mathbb{E}\left[\int_0^1 l(u_t)\mathrm{d}t\right], \tag{3.15}$$

$$s.t. \quad X^u_t = 0, \forall t \in [0,1]. \tag{3.16}$$

where l is deterministic twice continuously differentiable with derivative given by the oscillating function represented in Figure 1, which is Lipschitz-continuous, bounded and non-negative. In particular, Assumptions **(Reg-5)**, **(Lin-Dyn)** *and* **(Conv-2)**, *therefore, $\mathcal{J}$ is 1-strongly-convex, twice continuously differentiable, with second order-derivative $\nabla^2\mathcal{J} : \mathbb{H}^{2,2}_{\mathcal{P}} \mapsto \mathcal{L}(\mathbb{H}^{2,2}_{\mathcal{P}})$ given by:*

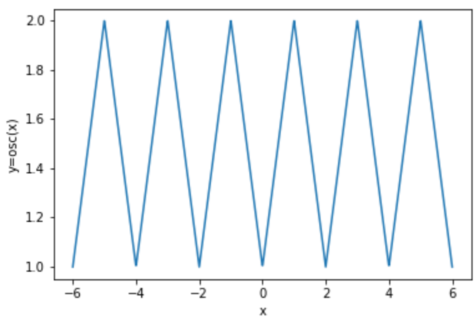


Figure 1: Oscillating function l''

$$\forall (u,v) \in \mathbb{H}^{2,2}_{\mathcal{P}}, \forall t \in [0,1], \quad (\nabla^2\mathcal{J}(u)(v))_t = l''(u_t)v_t.$$

This second-order derivative is bounded, and for all $u \in \mathbb{H}^{2,2}$*,* $\nabla^2\mathcal{J}(u) : \mathbb{H}^{2,2}_{\mathcal{P}} \mapsto \mathbb{H}^{2,2}_{\mathcal{P}}$ *is a bi-continuous endomorphism of* $\mathbb{H}^{2,2}_{\mathcal{P}}$*, i.e.,* $\nabla^2\mathcal{J}(u)$ *is a continuous invertible endomorphism of* $\mathbb{H}^{2,2}_{\mathcal{P}}$*, and its inverse is bounded as well. We have:*

$$\forall (u,w) \in (\mathbb{H}^{2,2}_{\mathcal{P}})^2, \forall t \in [0,1], \quad ((\nabla^2\mathcal{J}(u))^{-1}(w))_t = \frac{w_t}{l''(u_t)}.$$

In particular, Assumptions **(Reg-4)***,* **(Lin-Dyn)** *and* **(Conv-2)** *are verified. However, for* $n \in \mathbb{N}_*$*, let us define* $u^{(n)} \in \mathbb{H}^{2,2}_{\mathcal{P}}$ *the constant process with value* $U^{(n)}$ *given by a Bernoulli random variable with parameter* $p^{(n)} = 1/n$*, i.e.,* $U^{(n)} = 1$ *with probability* $1/n$ *and* 0 *else. Let* $v = 0 \in \mathbb{H}^{2,2}_{\mathcal{P}}$*. Then we have, using* $l''(1) = 1$ *and* $l''(0) = 0$*,* $\mathbb{E}\left[(U^{(n)})^2\right] = \mathbb{E}\left[(U^{(n)})^4\right] = \frac{1}{n}$*:*

$$\begin{aligned}
\frac{\|\nabla^2\mathcal{J}(u^{(n)}) - \nabla^2\mathcal{J}(v)\|_{\mathcal{L}(\mathbb{H}^{2,2})}}{\|u^{(n)} - v\|_{\mathbb{H}^{2,2}}} &\geq \frac{\|\nabla^2\mathcal{J}(u^{(n)})(u^{(n)}) - \nabla^2\mathcal{J}(v)(u^{(n)})\|_{\mathbb{H}^{2,2}}}{\|u^{(n)}\|_{\mathbb{H}^{2,2}}\|u^{(n)} - v\|_{\mathbb{H}^{2,2}}} \\
&= \frac{\mathbb{E}\left[(l''(U^{(n)}) - l''(0))^2 (U^{(n)})^2\right]^{1/2}}{\mathbb{E}\left[(U^{(n)})^2\right]} \\
&= \frac{\mathbb{E}\left[(U^{(n)})^4\right]^{1/2}}{\mathbb{E}\left[(U^{(n)})^2\right]} \\
&= \sqrt{n} \underset{n \longrightarrow +\infty}{\longrightarrow} +\infty.
\end{aligned}$$

In particular $\nabla^2\mathcal{J}$ *is not Lipschitz-continuous in* $\mathbb{H}^{2,2}_{\mathcal{P}}$ *endowed with* $\|.\|_{\mathbb{H}^{2,2}}$*.*

Under additional uniform boundedness and regularity assumptions, one can actually show that $\mathcal{J}$ actually defines an operator from the space of uniformly bounded process in $\mathbb{H}^{\infty,\infty}$ to reals, that $\nabla^2\mathcal{J}$ send $\mathbb{H}^{\infty,\infty}$ on $\mathcal{L}(\mathbb{H}^{\infty,\infty})$ and is Lipschitz-continuous.

Lemma 3.10. *Let the conditions of Theorem 3.7 hold. Define* $P \in \mathbb{H}^{\infty,\infty}$ *as in* (3.4)*,* $\Pi \in \mathbb{H}^{\infty,2}$ *by* (3.6)*,* $X \in \mathbb{H}^{\infty,2}$ *as in* (3.8)*,* $Y = PX + \Pi \in \mathbb{H}^{\infty,2}$*. Suppose additionally that* $a, b \in \mathbb{H}^{\infty,\infty}$ *and* $x, \eta \in \mathbb{L}^\infty_\Omega$*. Then* Π*,* X *and* Y *are in* $\mathbb{H}^{\infty,\infty}$ *and we have the estimates:*

$$\begin{aligned}
\|\Pi\|_{\mathbb{H}^{\infty,\infty}} &\leq \left(\|\eta\|_{\mathbb{L}^\infty_\Omega} + T\|a\|_{\mathbb{H}^{\infty,\infty}}\|P\|_{\mathbb{H}^{\infty,\infty}} + T\|b\|_{\mathbb{H}^{\infty,\infty}}\right) e^{\|PB+A\|_{\mathbb{H}^{\infty,\infty}} T}, \\
\|X\|_{\mathbb{H}^{\infty,\infty}} &\leq \left(\|x\|_{\mathbb{L}^\infty_\Omega} + T\|B\|_{\mathbb{H}^{\infty,\infty}}\|\Pi\|_{\mathbb{H}^{\infty,\infty}} + T\|a\|_{\mathbb{H}^{\infty,\infty}}\right) e^{\|A+PB\|_{\mathbb{H}^{\infty,\infty}} T}, \\
\|Y\|_{\mathbb{H}^{\infty,\infty}} &\leq \|P\|_{\mathbb{H}^{\infty,\infty}}\|X\|_{\mathbb{H}^{\infty,\infty}} + \|\Pi\|_{\mathbb{H}^{\infty,\infty}}.
\end{aligned}$$

Proof. The fact that $\Pi \in \mathbb{H}^{\infty,\infty}$ and the estimate on Π are immediate using formula (3.6). The fact that $X \in \mathbb{H}^{\infty,\infty}$ and the estimate on X are a consequence of this latter fact, from definition (3.8) and from Gronwall's lemma. The fact that $Y \in \mathbb{H}^{\infty,\infty}$ and the estimate on Y directly come from the fact that P, X and Π are in $\mathbb{H}^{\infty,\infty}$. □

Theorem 3.11 (Stability of $\mathbb{H}^{\infty,\infty}$)**.** *Suppose assumptions* **(Reg-5)***,* **(Lin-Dyn)** *and* **(Conv-2)** *hold. Then, for all* $(u,v) \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$*,* $X^u, Y^u, \nabla\mathcal{J}(u), \dot{X}^v, \dot{Y}^{u,v}, \nabla^2\mathcal{J}(u)(v)$ *and* $(\nabla^2\mathcal{J}(u))^{-1}(v)$ *are all in* $\mathbb{H}^{\infty,\infty}$ *and besides:*

$$\begin{aligned}
&\forall (u,v,w) \in (\mathbb{H}^{\infty,\infty}_{\mathcal{P}})^3, \\
&\|P^u\|_{\mathbb{H}^{\infty,\infty}} \leq C, \\
&\|\Pi^{u,w}\|_{\mathbb{H}^{\infty,\infty}} \leq C\|w\|_{\mathbb{H}^{\infty,\infty}}, \\
&\|X^u\|_{\mathbb{H}^{\infty,\infty}} + \|Y^u\|_{\mathbb{H}^{\infty,\infty}} + \|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}} \leq C(1 + \|u\|_{\mathbb{H}^{\infty,\infty}}), \\
&\|X^u - X^v\|_{\mathbb{H}^{\infty,\infty}} + \|Y^u - Y^v\|_{\mathbb{H}^{\infty,\infty}} + \|\nabla\mathcal{J}(u) - \nabla\mathcal{J}(v)\|_{\mathbb{H}^{\infty,\infty}} \leq C\|u - v\|_{\mathbb{H}^{\infty,\infty}}, \\
&\|\dot{X}^v\|_{\mathbb{H}^{\infty,\infty}} + \|\dot{Y}^{u,v}\|_{\mathbb{H}^{\infty,\infty}} + \|\nabla^2\mathcal{J}(u)(v)\|_{\mathbb{H}^{\infty,\infty}} \leq C\|v\|_{\mathbb{H}^{\infty,\infty}}, \\
&\|\dot{Y}^{u,w} - \dot{Y}^{v,w}\|_{\mathbb{H}^{\infty,\infty}} + \|\nabla^2\mathcal{J}(u)(w) - \nabla^2\mathcal{J}(v)(w)\|_{\mathbb{H}^{\infty,\infty}} \leq C\|u - v\|_{\mathbb{H}^{\infty,\infty}}\|w\|_{\mathbb{H}^{\infty,\infty}}, \\
&\|(\nabla^2\mathcal{J}(u))^{-1}(w)\|_{\mathbb{H}^{\infty,\infty}} \leq C\|w\|_{\mathbb{H}^{\infty,\infty}},
\end{aligned}$$

where the (generic) constant $C > 0$ is independent of u, v, w. Note that this implies that $\nabla^2\mathcal{J}$ defines a Lipschitz-continuous operator from $\mathbb{H}^{\infty,\infty}_{\mathcal{P}}$ to the space of continuous endomorphisms of $\mathbb{H}^{\infty,\infty}_{\mathcal{P}}$, and that $\nabla^2\mathcal{J}(u)$ and $(\nabla^2\mathcal{J}(u))^{-1}$ are bounded linear operators, uniformly in u. This also implies that for any $u \in \mathbb{H}^{\infty,\infty}$, the Newton direction $-\left(\nabla^2\mathcal{J}(u)\right)^{-1}(\nabla\mathcal{J}(u))$ is also in $\mathbb{H}^{\infty,\infty}_{\mathcal{P}}$.

Proof. Note that for any $(X, Y) \in (\mathbb{H}^{\infty,\infty})^2$, $XY \in \mathbb{H}^{\infty,\infty}$ and:

$$\|XY\|_{\mathbb{H}^{\infty,\infty}} \leq \|X\|_{\mathbb{H}^{\infty,\infty}}\|Y\|_{\mathbb{H}^{\infty,\infty}}.$$

Throughout the proof, C denotes a generic deterministic constant depending only on T and the bounds on the data of the problem and their derivatives in $\mathbb{H}^{\infty,\infty}$. We have immediately for $(u, v) \in (\mathbb{H}^{\infty,\infty})^2$:

$$\begin{aligned}
&\|X^u\|_{\mathbb{H}^{\infty,\infty}} \leq C(1 + \|u\|_{\mathbb{H}^{\infty,\infty}}),\\
&\|X^u - X^v\|_{\mathbb{H}^{\infty,\infty}} \leq C\|u - v\|_{\mathbb{H}^{\infty,\infty}},\\
&\|\dot{X}^v\|_{\mathbb{H}^{\infty,\infty}} \leq C\|v\|_{\mathbb{H}^{\infty,\infty}}.
\end{aligned}$$

We also have:

$$\begin{aligned}
&\|Y^u\|_{\mathbb{H}^{\infty,\infty}} \leq C(1 + \|u\|_{\mathbb{H}^{\infty,\infty}} + \|X^u\|_{\mathbb{H}^{\infty,\infty}}),\\
&\|Y^u - Y^v\|_{\mathbb{H}^{\infty,\infty}} \leq C(\|u - v\|_{\mathbb{H}^{\infty,\infty}} + \|X^u - X^v\|_{\mathbb{H}^{\infty,\infty}}),\\
&\|\dot{Y}^{u,v}\|_{\mathbb{H}^{\infty,\infty}} \leq C(\|v\|_{\mathbb{H}^{\infty,\infty}} + \|\dot{X}^v\|_{\mathbb{H}^{\infty,\infty}}).
\end{aligned}$$

Combining the first two upper bounds with the estimates on X^u yields the inequality on Y^u. Combining the third estimate with the bound on $\dot{X}^v$ yields:

$$\|\dot{Y}^{u,v}\|_{\mathbb{H}^{\infty,\infty}} \leq C\|v\|_{\mathbb{H}^{\infty,\infty}}.$$

Using the Lipschitz-continuity of the second-order derivatives of l and Ψ, one gets:

$$\|\dot{Y}^{u,w} - \dot{Y}^{v,w}\|_{\mathbb{H}^{\infty,\infty}} \leq C\left(\|u - v\|_{\mathbb{H}^{\infty,\infty}} + \|X^u - X^v\|_{\mathbb{H}^{\infty,\infty}}\right)\left(\|w\|_{\mathbb{H}^{\infty,\infty}} + \|\dot{X}^w\|_{\mathbb{H}^{\infty,\infty}}\right).$$

We can use all our previous estimates to get the claimed bound on $\|\dot{Y}^{u,w} - \dot{Y}^{v,w}\|_{\mathbb{H}^{\infty,\infty}}$.

We have:

$$\begin{aligned}
&\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}} \leq C(1 + \|u\|_{\mathbb{H}^{\infty,\infty}} + \|X^u\|_{\mathbb{H}^{\infty,\infty}} + \|Y^u\|_{\mathbb{H}^{\infty,\infty}}),\\
&\|\nabla\mathcal{J}(u) - \nabla\mathcal{J}(v)\|_{\mathbb{H}^{\infty,\infty}} \leq C(\|u - v\|_{\mathbb{H}^{\infty,\infty}} + \|X^u - X^v\|_{\mathbb{H}^{\infty,\infty}} + \|Y^u - Y^v\|_{\mathbb{H}^{\infty,\infty}}).
\end{aligned}$$

Combining this with the estimates on X^u and Y^u yields the estimates on $\nabla\mathcal{J}(u)$. Using the expression of $\nabla^2\mathcal{J}(u)(v)$ derived earlier, we easily get the estimate:

$$\begin{aligned}
&\|\nabla^2\mathcal{J}(u)(v)\|_{\mathbb{H}^{\infty,\infty}} \leq C(\|v\|_{\mathbb{H}^{\infty,\infty}} + \|\dot{X}^v\|_{\mathbb{H}^{\infty,\infty}} + \|\dot{Y}^{u,v}\|_{\mathbb{H}^{\infty,\infty}}),\\
&\|\nabla^2\mathcal{J}(u)(w) - \nabla^2\mathcal{J}(v)(w)\|_{\mathbb{H}^{\infty,\infty}} \leq C(\|u - v\|_{\mathbb{H}^{\infty,\infty}} + \|X^u - X^v\|_{\mathbb{H}^{\infty,\infty}})(\|w\|_{\mathbb{H}^{\infty,\infty}} + \|\dot{X}^w\|_{\mathbb{H}^{\infty,\infty}}) + C\|\dot{Y}^{u,w} - \dot{Y}^{v,w}\|_{\mathbb{H}^{\infty,\infty}}.
\end{aligned}$$

We get the claimed estimates on $\|\nabla^2\mathcal{J}(u)(v)\|_{\mathbb{H}^{\infty,\infty}}$ and $\|\nabla^2\mathcal{J}(u)(w) - \mathcal{J}(v)(w)\|_{\mathbb{H}^{\infty,\infty}}$ using the previous bounds.

We have the bounds on parameters appearing in(3.10):

$$\begin{aligned}
&\|A^u\|_{\mathbb{H}^{\infty,\infty}} + \|B^u\|_{\mathbb{H}^{\infty,\infty}} + \|C^u\|_{\mathbb{H}^{\infty,\infty}} + \|\Gamma^u\|_{\mathbb{L}^\infty_\Omega} \leq C,\\
&\|a^{u,w}\|_{\mathbb{H}^{\infty,\infty}} + \|b^{u,w}\|_{\mathbb{H}^{\infty,\infty}} \leq C\|w\|_{\mathbb{H}^{\infty,\infty}}.
\end{aligned}$$

Therefore, we are in the framework of application of Lemma 3.10 for $\Pi = \Pi^{u,w}$, $P = P^u$ and $X = \tilde{X}^{u,w}$, which are defined in (3.11), (3.12) and (3.13) (with $\eta = 0$ and $x = 0$). This yields:

$$\begin{aligned}
&\|P^u\|_{\mathbb{H}^{\infty,\infty}} \leq C,\\
&\|\Pi^{u,w}\|_{\mathbb{H}^{\infty,\infty}} \leq C(1 + \|P^u\|_{\mathbb{H}^{\infty,\infty}})\|w\|_{\mathbb{H}^{\infty,\infty}},
\end{aligned}$$

$$\|\tilde{X}^{u,w}\|_{\mathbb{H}^{\infty,\infty}} \leq C(\|w\|_{\mathbb{H}^{\infty,\infty}} + \|\Pi^{u,w}\|_{\mathbb{H}^{\infty,\infty}}).$$

We have by (3.14):

$$\|(\nabla^2 \mathcal{J}(u))^{-1}(w)\|_{\mathbb{H}^{\infty,\infty}} \leq C(\|w\|_{\mathbb{H}^{\infty,\infty}} + \|\Pi^{u,w}\|_{\mathbb{H}^{\infty,\infty}} + (1 + \|P^u\|_{\mathbb{H}^{\infty,\infty}})\|\tilde{X}^{u,w}\|_{\mathbb{H}^{\infty,\infty}}),$$

which gives the estimate on $\|(\nabla^2 \mathcal{J}(u))^{-1}(w)\|_{\mathbb{H}^{\infty,\infty}}$. □

It is well known that, for strongly convex functions, with Lipschitz-continuous second order derivative, local convergence can be shown for the Newton method, see for instance [Kan48] - [NW06, Theorem 3.5, p. 44]. However, even under such demanding assumptions, it is well-known in a finite-dimensional setting that Newton method is not guaranteed to converge globally. We provide next a counter-example to the global convergence of Newton method in our infinite-dimensional framework.

Example 3.12 (Counter-example to global convergence of Newton method)**.** *We consider $\mathcal{J} : \mathbb{H}^{\infty,\infty} \mapsto \mathbb{R}$ given by:*

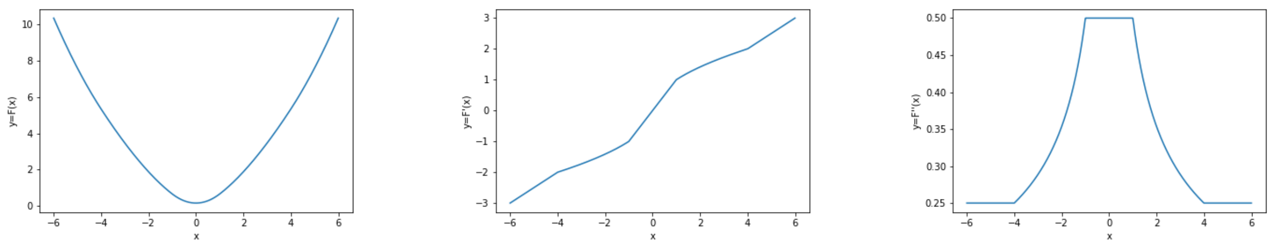
$$\forall u \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}, \quad \mathcal{J}(u) := \mathbb{E}\left[\int_0^1 F(u_t)\mathrm{d}t\right], \tag{3.17}$$

$$s.t. \quad X^u_t = 0, \forall t \in [0,1], \tag{3.18}$$

where $F : x \mapsto \mathbb{R}$ is defined by:

$$F(x) := \begin{cases} \frac{x^2}{4} + \frac{4}{3}, & \text{if } |x| > 4, \\ \frac{2|x|^{\frac{3}{2}}}{3}, & \text{if } 1 \leq |x| \leq 4, \\ \frac{x^2}{2} + \frac{1}{6}, & \text{if } |x| < 1. \end{cases}$$

Then F is strongly convex, twice continuously differentiable with bounded second order derivative, see Figure 2.



(a) Graph of function F (b) Graph of the first derivative F' (c) Graph of the second derivative F''

Figure 2: Graphs of F, its first and second derivatives

Assumptions **(Reg-5)***,* **(Lin-Dyn)** *and* **(Conv-2)** *hold. Let $u^{(0)} \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$ be any stochastic process such that $1 < |u^{(0)}_{t,\omega}| < 4$, $\mathrm{d}\mathbb{P} \otimes \mathrm{d}t$-a.e. Due to the particular structure of our problem, Newton's method reduces to the Newton method applied to $F : \mathbb{R} \mapsto \mathbb{R}$ applied ω by ω and t by t. For $1 < |x_0| < 4$, Newton's method (in $\mathbb{R}$) applied to F with initial guess x_0 produces the sequence $(x_k)_{k\in\mathbb{N}}$ with general term $x_k = (-1)^k x_0$. Indeed, for $1 \leq |x| \leq 4$, we have $F'(x) = sign(x)\sqrt{|x|}$ and $F''(x) = \frac{1}{2\sqrt{|x|}}$, so that the Newton iteration is given by:*

$$x_{k+1} = x_k - \frac{F'(x_k)}{F''(x_k)} = -x_k.$$

Therefore, Newton's method (in $\mathbb{H}^{\infty,\infty}$) applied to $\mathcal{J}$ with initial guess $u^{(0)}$ produces the sequence $(u^{(k)})_{k\in\mathbb{N}}$ with general term $u^{(k)} = (-1)^k u^{(0)}$, which does not converge.

The previous counter-example motivates globalization procedures. We consider Backtracking-line search methods, which are iterative procedures which allow to select appropriate step lengths such that the Goldstein conditions, presented in [NW06, p. 36], hold.

The standard Backtracking line-search method is given in Algorithm 3. It is directly adapted from Backtracking line-search method in $\mathbb{R}^n$ [NW06, Algorithm 3.1, p.37] or [BV04, Algorithm 9.2, p. 464]. Under the assumption of a strongly convex function with bounded and Lipschitz second order derivative, it can be shown in the finite dimensional case that the Newton method with the Standard Backtracking line search converges globally, see [BV04, Section 9.5.3, pp. 488-491].

Algorithm 3 Standard Backtracking line search (compact generic version)

Inputs: Current point $u \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$, Current search direction $\Delta u \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$, $\beta \in (0,1)$, $\gamma \in (0,1)$.
$\sigma = 1$.
while $\mathcal{J}(u + \sigma\Delta u) > \mathcal{J}(u) + \gamma\sigma\langle\nabla\mathcal{J}(u), \Delta u\rangle_{\mathbb{H}^{2,2}}$ **do**
    $\sigma \leftarrow \beta\sigma$.
end while
return $u + \sigma\Delta u$.

However, since we do not work in a finite-dimensional setting, the global convergence of the method is not guaranteed to our knowledge. This is due to the fact that $\mathcal{J}$ is a criteria in expectation, whereas we are working with the norm in $\mathbb{H}^{\infty,\infty}$. We shall see in our numerical applications that the Standard Backtracking line-search may prevent the Newton method to converge, see Figures 6b and 6d.

To alleviate this issue, we propose a new Backtracking line-search rule, based on the sup-essential norm of the gradient, described in Algorithm 4. Hence, we are only working with the norm in $\mathbb{H}^{\infty,\infty}$, which makes more sense as we are working with variables in $\mathbb{H}^{\infty,\infty}$.

Algorithm 4 Gradient Backtracking line search (compact generic version)

Inputs: Current point $u \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$, Current search direction $\Delta u \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$, $\beta \in (0,1)$, $\gamma \in (0,1)$.
$\sigma = 1$.
while $\|\nabla\mathcal{J}(u + \sigma\Delta u)\|_{\mathbb{H}^{\infty,\infty}} > (1 - \gamma\sigma)\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}}$ **do**
    $\sigma \leftarrow \beta\sigma$.
end while
return $u + \sigma\Delta u$.

We show that the Newton method combined with the Gradient Backtracking line search algorithm 4 ensures global convergence of the method, and that after a finite number of iterations, the algorithm takes full Newton steps, which ensures quadratic convergence.

Lemma 3.13. *Suppose assumptions* **(Reg-5)**, **(Lin-Dyn)** *and* **(Conv-2)** *hold. For any $u \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$, the Gradient Backtracking line search terminates in finitely many iterations for the Newton step $\Delta u := -(\nabla^2\mathcal{J}(u))^{-1}(\nabla\mathcal{J}(u)) \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$.*

Besides, if the algorithm returns $\sigma = 1$, then the new point $u + \Delta u$ satisfies:

$$\|\nabla\mathcal{J}(u + \Delta u)\|_{\mathbb{H}^{\infty,\infty}} \leq \min(1 - \gamma, C\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}})\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}},$$

where $C = \frac{L_{\nabla^2\mathcal{J}} C^2_{(\nabla^2\mathcal{J})^{-1}}}{2}$ with $L_{\nabla^2\mathcal{J}}$ the Lipschitz constant of $\nabla^2\mathcal{J} : \mathbb{H}^{\infty,\infty} \mapsto \mathcal{L}(\mathbb{H}^{\infty,\infty})$, and

$$C_{(\nabla^2\mathcal{J})^{-1}} := \sup_{u \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}} \|(\nabla^2\mathcal{J}(u))^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})},$$

which is finite by Theorem 3.11). Conversely, if $C\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}} \leq (1 - \gamma)$ then the algorithm returns $\sigma = 1$. On the other hand, if the algorithm returns $\sigma < 1$, then the new iterate $u + \sigma\Delta u$ satisfies:

$$\|\nabla\mathcal{J}(u + \sigma\Delta u)\|_{\mathbb{H}^{\infty,\infty}} \leq \|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}} - \frac{\beta\gamma(1-\gamma)}{C}.$$

Proof. By Taylor-Lagrange formula, using $\Delta u = -(\nabla^2\mathcal{J}(u))^{-1}(\nabla\mathcal{J}(u))$, we have

$$\nabla\mathcal{J}(u + \sigma\Delta u) = \nabla\mathcal{J}(u) + \int_0^1 \nabla^2\mathcal{J}(u + s\sigma\Delta u)(\sigma\Delta u)\mathrm{d}s,$$

$$= (1-\sigma)\nabla\mathcal{J}(u) + \sigma\int_0^1 \left(\nabla^2\mathcal{J}(u+s\sigma\Delta u) - \nabla^2\mathcal{J}(u)\right)(\Delta u)\mathrm{d}s,$$

which yields by Lipschitz continuity of $\nabla^2\mathcal{J}$ in $\mathcal{L}(\mathbb{H}^{\infty,\infty})$:

$$\begin{aligned}
\|\nabla\mathcal{J}(u+\sigma\Delta u)\|_{\mathbb{H}^{\infty,\infty}} &\leq (1-\sigma)\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}} + \sigma\int_0^1 \|\nabla^2\mathcal{J}(u) - \nabla^2\mathcal{J}(u+s\sigma\Delta u)\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}\mathrm{d}s\|\Delta u\|_{\mathbb{H}^{\infty,\infty}} \\
&\leq (1-\sigma)\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}} + \frac{L_{\nabla^2\mathcal{J}}}{2}\sigma^2\|\Delta u\|^2_{\mathbb{H}^{\infty,\infty}} \\
&\leq \left(1-\sigma+\frac{L_{\nabla^2\mathcal{J}}C^2_{(\nabla^2\mathcal{J})^{-1}}}{2}\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}}\sigma^2\right)\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}}.
\end{aligned}$$

Notice that, since $0 < \gamma < 1$, for $\sigma > 0$ small enough, we have:

$$1-\sigma+C\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}}\sigma^2 \leq 1-\gamma\sigma. \tag{3.19}$$

In particular, the Gradient Backtracking line search terminates after finitely many iterations. Besides, if $\frac{C\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}}}{1-\gamma} \leq 1$ then the algorithm stops and returns $\sigma = 1$.

Suppose that the algorithm stops with $\sigma = 1$, then:

$$\|\nabla\mathcal{J}(u+\Delta u)\|_{\mathbb{H}^{\infty,\infty}} \leq \min(1-\gamma, C\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}})\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}},$$

by the termination criterion and the previous estimate.

Suppose that the algorithm returns $\sigma < 1$. Then by the termination criteria of the algorithm, (3.19) implies in particular that:

$$-(\sigma/\beta) + C\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}}(\sigma/\beta)^2 > -\gamma(\sigma/\beta).$$

This yields:

$$\frac{\beta(1-\gamma)}{C\|\nabla\mathcal{J}(u).\|_{\mathbb{H}^{\infty,\infty}}} < \sigma < 1.$$

This yields, using the termination criterion of the algorithm, the fact that σ, γ and β are in $(0,1)$ and the previous inequality:

$$\begin{aligned}
\|\nabla\mathcal{J}(u+\sigma\Delta u)\|_{\mathbb{H}^{\infty,\infty}} &\leq (1-\gamma\sigma)\,\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}} \\
&\leq \left(1-\frac{\beta\gamma(1-\gamma)}{C\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}}}\right)\|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}} \\
&\leq \|\nabla\mathcal{J}(u)\|_{\mathbb{H}^{\infty,\infty}} - \frac{\beta\gamma(1-\gamma)}{C}.
\end{aligned}$$

□

Remark 3.14. *Note that the properties of the backtracking line-search algorithm can be enforced for values of parameters which are independent on the problem, and in particular, independent from a priori unknown constants of the problem (bounds on derivatives, constant of strong convexity...). The properties of the problem (regularity, convexity) only impact the number of iterations of the algorithm.*

Theorem 3.15. *Suppose assumptions* **(Reg-5)**, **(Lin-Dyn)** *and* **(Conv-2)** *hold. Then $\mathcal{J} : \mathbb{H}^{\infty,\infty} \mapsto \mathbb{R}$ is twice continuously differentiable, with Lipschitz continuous first and second derivatives. Besides $\nabla^2\mathcal{J}$ is an invertible, bi-continuous endomorphism of $\mathbb{H}^{\infty,\infty}$. Let $u^{(0)} \in \mathbb{H}^{\infty,\infty}$. Let γ and β be parameters in $(0,1)$. Define $C = \frac{L_{\nabla^2\mathcal{J}}C^2_{(\nabla^2\mathcal{J})^{-1}}}{2}$ with $L_{\nabla^2\mathcal{J}}$ the Lipschitz constant of $\nabla^2\mathcal{J} : \mathbb{H}^{\infty,\infty}_{\mathcal{P}} \mapsto \mathcal{L}(\mathbb{H}^{\infty,\infty}_{\mathcal{P}})$ and*

$$C_{(\nabla^2\mathcal{J})^{-1}} := \sup_{u\in\mathbb{H}^{\infty,\infty}_{\mathcal{P}}} \|(\nabla^2\mathcal{J}(u))^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}.$$

Algorithm 5 Newton's method with Backtracking line search (compact generic version)

Inputs: $u^{(0)} \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$, $\gamma \in (0,1)$, $\beta \in (0,1)$, $\varepsilon > 0$. $k = 0$
while $\|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} > \varepsilon$ **do**
  Compute Newton direction $-(\nabla^2\mathcal{J}(u^{(k)}))^{-1}(\nabla\mathcal{J}(u^{(k)})) \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$
  Compute new iterate using Backtracking line-search rule $u^{(k+1)} = u^{(k)} - \sigma(\nabla^2\mathcal{J}(u^{(k)}))^{-1}(\nabla\mathcal{J}(u^{(k)})) \in \mathbb{H}^{\infty,\infty}_{\mathcal{P}}$
  $k \leftarrow k+1$
end while
return $u^{(k)}$

Define as well:

$$\eta = \frac{\beta\gamma(1-\gamma)}{C},$$
$$k_1 = \inf\left\{k \in \mathbb{N} \quad | \quad \|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} \leq \frac{1-\gamma}{C}\right\}.$$

Then k_1 is finite. Besides, after k_1 iterations, the step length is always 1 and Newton method with Gradient Backtracking line search converges quadratically, i.e.,

$$\forall k \geq k_1, \quad C\|\nabla\mathcal{J}(u^{(k+1)})\|_{\mathbb{H}^{\infty,\infty}} \leq \left(C\|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}}\right)^2,$$
$$\forall k \geq k_1, \quad \|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} \leq \frac{(1-\gamma)^{2^{k-k_1}}}{C}.$$

Besides, $u^{(k)}$ converges to $u^* \in \mathbb{H}^{\infty,\infty}$ which is the minimizer of $\mathcal{J}$, and the asymptotic convergence is quadratic, i.e.,

$$\forall k \geq k_1, \quad C\|u^{(k+1)} - u^*\|_{\mathbb{H}^{\infty,\infty}} \leq C^2\|u^{(k)} - u^*\|^2_{\mathbb{H}^{\infty,\infty}},$$
$$\forall k \geq k_1+1, \quad \|u^{(k)} - u^*\|_{\mathbb{H}^{\infty,\infty}} \leq \frac{\|(\nabla^2\mathcal{J})^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}}{\gamma C}(1-\gamma)^{2^{k-k_1}}.$$

Proof. By the previous lemma, the sequence $\left(\|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}}\right)_{k\in\mathbb{N}}$ is monotone decreasing.

We first prove that $k_1 < +\infty$. Let $k \leq k_1$. Then $\|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} > \frac{1-\gamma}{C}$. If at iteration k, the gradient backtracking line search returns a unit step length $\sigma = 1$, then by Lemma 3.13:

$$\begin{aligned}\|\nabla\mathcal{J}(u^{(k+1)})\|_{\mathbb{H}^{\infty,\infty}} &\leq (1-\gamma)\|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} \\ &\leq \|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} - \eta,\end{aligned}$$

where we used the assumption $\|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} > \frac{1-\gamma}{C}$ and $\beta \in (0,1)$.

Else, at iteration k, the gradient backtracking line search returns a step length $\sigma < 1$ and still by Lemma 3.13,

$$\|\nabla\mathcal{J}(u^{(k+1)})\|_{\mathbb{H}^{\infty,\infty}} \leq \|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} - \eta.$$

This yields:

$$\forall k \leq k_1, \quad \|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} \leq \|\nabla\mathcal{J}(u^{(0)})\|_{\mathbb{H}^{\infty,\infty}} - k\eta.$$

Since $\eta > 0$, this yields existence and finiteness of k_1, which is bounded from above by $\frac{\|\nabla\mathcal{J}(u^{(0)})\|_{\mathbb{H}^{\infty,\infty}}}{\eta} + 1$.

Besides, for all $k \geq k_1$, we have $\|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} \leq \frac{1-\gamma}{C}$ since the sequence $\left(\|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}}\right)_{k\in\mathbb{N}}$ is monotone decreasing. In that case, Lemma 3.13 shows that the algorithm takes unit step length and:

$$C\|\nabla\mathcal{J}(u^{(k+1)})\|_{\mathbb{H}^{\infty,\infty}} \leq C^2\|\nabla\mathcal{J}(u^{(k+1)})\|^2_{\mathbb{H}^{\infty,\infty}}.$$

This combined with $C\|\nabla\mathcal{J}(u^{(k_1)})\| \leq 1-\gamma$ yields:

$$\forall k \geq k_1, \quad C\|\nabla\mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} \leq (1-\gamma)^{2^{k-k_1}}.$$

From that and since the algorithm takes unit step length after iteration k_1, we deduce:

$$\begin{aligned}\forall k \geq k_1, \quad \|u^{(k+1)} - u^{(k)}\|_{\mathbb{H}^{\infty,\infty}} &= \|(\nabla^2 \mathcal{J}(u^{(k)}))^{-1}(\nabla \mathcal{J}(u^{(k)}))\|_{\mathbb{H}^{\infty,\infty}} \\ &\leq \|(\nabla^2 \mathcal{J})^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})} \|\nabla \mathcal{J}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} \\ &\leq \|(\nabla^2 \mathcal{J})^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})} \frac{(1-\gamma)^{2^{k-k_1}}}{C}.\end{aligned}$$

In particular, this yields the absolute convergence of the series $S_n = \sum_{k=0}^{n} v_n$ of general term $v_0 = u^{(0)}$ and $v_k = u^{(k)} - u^{(k-1)}$ for $k > 1$. Hence $(S_n)_{n\in\mathbb{N}}$ converges in $\mathbb{H}^{\infty,\infty}$ and so does $(u^{(k)})_{k\in\mathbb{N}}$. Denote $u^* \in \mathbb{H}^{\infty,\infty}$ the limit point of $(u^{(k)})_{k\in\mathbb{N}}$. By continuity of $\nabla\mathcal{J} : \mathbb{H}^{\infty,\infty} \mapsto \mathbb{H}^{\infty,\infty}$, u^* is a critical point of $\mathcal{J}$ and hence, by strong convexity of $\mathcal{J}$, u^* is the unique minimizer of $\mathcal{J}$. Besides, as the algorithm takes unit step length $\sigma = 1$ after k_1 iterations,

$$\begin{aligned}\forall k \geq k_1, \|u^{(k+1)} - u^*\|_{\mathbb{H}^{\infty,\infty}} &= \|u^{(k)} - (\nabla^2 \mathcal{J}(u^{(k)}))^{-1}(\nabla \mathcal{J}(u^{(k)})) - u^*\|_{\mathbb{H}^{\infty,\infty}} \\ &= \|u^{(k)} - u^* - (\nabla^2 \mathcal{J}(u^{(k)}))^{-1}(\nabla \mathcal{J}(u^{(k)}) - \nabla \mathcal{J}(u^*))\|_{\mathbb{H}^{\infty,\infty}} \\ &\leq \|(\nabla^2 \mathcal{J}(u^{(k)}))^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})} \|\nabla \mathcal{J}(u^{(k)}) - \nabla \mathcal{J}(u^*) - \nabla^2 \mathcal{J}(u^{(k)})(u^{(k)} - u^*)\|_{\mathbb{H}^{\infty,\infty}} \\ &\leq \frac{\|(\nabla^2 \mathcal{J}(u^{(k)}))^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})} L_{\nabla^2 \mathcal{J}}}{2} \|u^{(k)} - u^*\|^2_{\mathbb{H}^{\infty,\infty}},\end{aligned}$$

by Taylor-Lagrange's formula and Lipschitz continuity of $\nabla^2\mathcal{J} : \mathbb{H}^{\infty,\infty} \mapsto \mathcal{L}(\mathbb{H}^{\infty,\infty})$.

Besides, for all $k > k_1$,

$$\begin{aligned}\|u^{(k)} - u^*\|_{\mathbb{H}^{\infty,\infty}} &\leq \sum_{j=k}^{+\infty} \|u^{(j+1)} - u^{(j)}\|_{\mathbb{H}^{\infty,\infty}} \\ &\leq \frac{\|(\nabla^2 \mathcal{J})^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}}{C} \left(\sum_{j=k}^{+\infty} (1-\gamma)^{2^{j-k_1}} \right) \\ &= \frac{\|(\nabla^2 \mathcal{J})^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}}{C} \left(\sum_{j=0}^{+\infty} (1-\gamma)^{2^{j+k-k_1}} \right) \\ &\leq \frac{\|(\nabla^2 \mathcal{J})^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}}{C} (1-\gamma)^{2^{k-k_1}} \left(1 + \sum_{j=0}^{+\infty} (1-\gamma)^{2^j} \right) \\ &\leq \frac{\|(\nabla^2 \mathcal{J})^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}}{C} (1-\gamma)^{2^{k-k_1}} \left(\sum_{j=0}^{+\infty} (1-\gamma)^{j} \right) \\ &\leq \frac{\|(\nabla^2 \mathcal{J})^{-1}\|_{\mathcal{L}(\mathbb{H}^{\infty,\infty})}}{C\gamma} (1-\gamma)^{2^{k-k_1}}\end{aligned}$$

where we used $k > k_1$ and the fact that $2^{j+k-k_1} \geq 2^j + 2^{k-k_1}$ for all $j \geq 1$. □

4 Application: energy storage system control for power balancing

4.1 Problem setting

We consider N identical batteries with energy capacity $\mathcal{E}_{\max}$ operated in order to balance production and consumption on an electricity network. Other types of energy storage systems could be considered. For instance, one could replace batteries in this application by a large population of Thermostatically Controlled Loads (TCLs), which include water heaters, Air Conditioners, Heat pumps,... provided a first order affine-linear model of their temperature dynamic is used. The global consumption on the network is given by a deterministic function $N\mathrm{P}^{\mathrm{cons}}$, where $\mathrm{P}^{\mathrm{cons}}$ is the total consumption divided by the number of batteries. The assumption of a deterministic consumption profile can be justified by the fact that it is the aggregation of a large number of small independent consumption profiles, which allows to use the law of large numbers. We assume additionally a total solar power production $N\mathrm{P}^{\mathrm{sun}}$, i.e. $\mathrm{P}^{\mathrm{sun}}$ is the total solar production divided by the number of batteries on the network (we do not account for wind power, although this could easily be included in the model). We follow [Bad+18] by setting $\mathrm{P}^{\mathrm{sun}} = \mathrm{P}^{\mathrm{sun,max}}\mathrm{X}^{\mathrm{sun}}$ where $\mathrm{P}^{\mathrm{sun,max}} : [0,T] \mapsto \mathbb{R}$ is a deterministic function (the clear sky model) represented in 3a and $\mathrm{X}^{\mathrm{sun}}$ solves a Fisher-Wright type SDE which dynamics is

$$\mathrm{d}\mathrm{X}_t^{\mathrm{sun}} = -\rho^{\mathrm{sun}}(\mathrm{X}_t^{\mathrm{sun}} - \mathrm{x}_t^{\mathrm{sun,ref}})\mathrm{d}t + \sigma^{\mathrm{sun}}(\mathrm{X}_t^{\mathrm{sun}})^{k_1}(1-\mathrm{X}_t^{\mathrm{sun}})^{k_2}\mathrm{d}\tilde{W}_t, \tag{4.1}$$

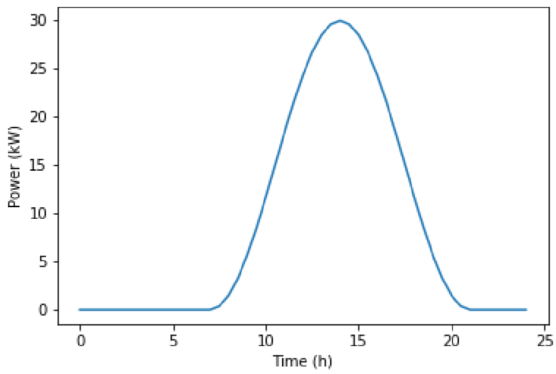
with $k_1, k_2 \geq 1/2$. As proved in [Bad+18], there is a strong solution to the above SDE and the solution $\mathrm{X}^{\mathrm{sun}}$ takes values in $[0,1]$. Since the drifts are affine-linear, the conditional expectation of the solution of (4.1) is known in closed forms (this property is intensively used in [BSS05]):

$$\mathbb{E}_t\left[\mathrm{P}_s^{\mathrm{sun}}\right] = \left(\frac{\mathrm{P}_t^{\mathrm{sun}}}{\mathrm{P}_t^{\mathrm{sun,max}}}\exp(-\rho^{\mathrm{sun}}(s-t)) + \int_t^s \rho^{\mathrm{sun}}\mathrm{x}_\tau^{\mathrm{sun,ref}}\exp(-\rho^{\mathrm{sun}}(s-\tau))\mathrm{d}\tau\right)\mathrm{P}_s^{\mathrm{sun,max}}, \tag{4.2}$$

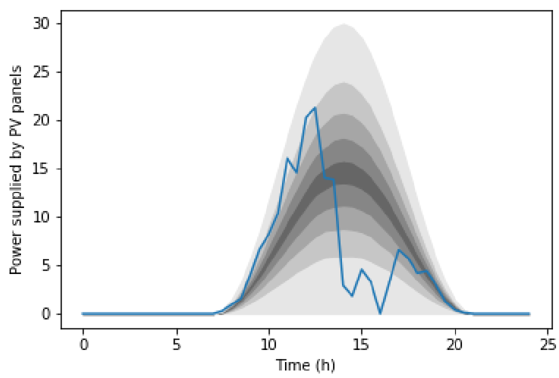
for $s \geq t$. This will allow us to speed up computations of the conditional expectations $\mathbb{E}_t\left[\mathrm{P}_s^{\mathrm{sun}}\right]$ and $\mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]$ as required when deriving the optimal control. The value of the parameters used are given in the following table. Empirical quantile plot (obtained by simulation of 10000 i.i.d. trajectories) as well as one example trajectory of $\mathrm{P}^{\mathrm{sun}}$ are given in Figure 3b.

Table 1: Parameter values for the simulation of PV power production

ρ^{sun}	$\mathrm{x}^{\mathrm{sun,ref}}$	σ^{sun}	k_1	k_2
0.75 h^{-1}	0.5	0.8	0.8	0.7



(a) Time evolution of $\mathrm{P}^{\mathrm{sun,max}}$



(b) Empirical quantiles of $\mathrm{P}^{\mathrm{sun}}$, obtained with $M = 10000$ samples, and one realization of $\mathrm{P}^{\mathrm{sun}}$

Figure 3: Graphical statistics of the evolution of $\mathrm{P}^{\mathrm{sun}}$

Our goal is to minimize global cost for the control of N batteries, which are composed of operational costs for battery management and a penalization for power balance imbalance, represented in Figure 4[1].

We assume a production profile per battery $\mathrm{P}_t^{\mathrm{prod}} = \mathrm{P}_t^{\mathrm{cons}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]$, which can be easily computed using the model on $\mathrm{P}^{\mathrm{sun}}$.

Denoting by $u^{(n)}$ the power supplied by battery $n \in [N]$ and by $X^{(n)}$ its normalized state of charge, we wish to solve the following stochastic control problem:

$$\min_{u\in\mathbb{H}^{2,2}_{\mathcal{P}}} \mathbb{E}\left[\int_0^T \left\{\frac{1}{N}\sum_{n=1}^N \frac{\mu}{2}(u_t^{(n)})^2 + \frac{1}{N}\sum_{n=1}^N \frac{\nu}{2}\left(X_t^{(n)} - \frac{1}{2}\right)^2 + \mathcal{L}\left(\frac{1}{N}\sum_{n=1}^N u_t^{(n)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]\right)\right\}\mathrm{d}t + \frac{1}{N}\sum_{n=1}^N \frac{\rho}{2}\left(X_T^{(n)} - \frac{1}{2}\right)^2\right], \tag{4.3}$$

[1] Icons made by Freepik and Smashicons from www.flaticon.com

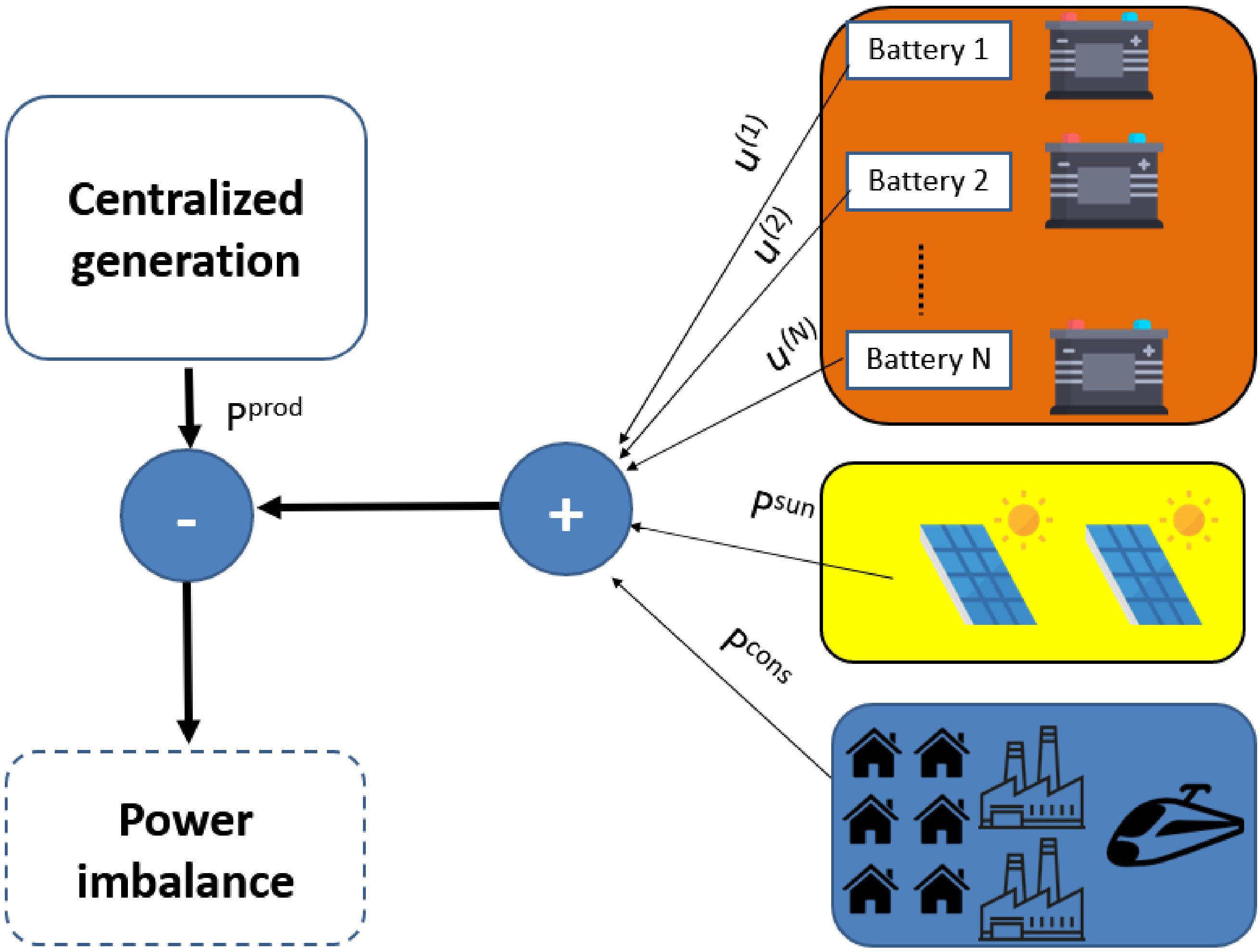


Figure 4: Power imbalance on the network

$$s.t. \quad X_t^{(n)} = x_0^{(n)} - \int_0^t \frac{u_s^{(n)}}{\mathcal{E}_{\max}} \mathrm{d}s. \tag{4.4}$$

The first two terms as well as the last term in the cost functional represent the sum of the operational costs for individual batteries. We penalize quadratically power supplied or absorbed by the batteries (first term) and penalize deviations of the normalized states of charge of the batteries from the reference value $1/2$ (second and last term). The third term represents a penalization term for the power imbalance $\frac{1}{N}\sum_{n=1}^N u_t^{(n)} + \mathrm{P}_t^{\mathrm{sun}} - \mathrm{P}_t^{\mathrm{cons}} + \mathrm{P}_t^{\mathrm{prod}} = \frac{1}{N}\sum_{n=1}^N u_t^{(n)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]$, using $\mathrm{P}_t^{\mathrm{prod}} = \mathrm{P}_t^{\mathrm{cons}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]$. The state variable $X^{(n)}$ represents the normalized state of charge of battery n, i.e., the energy stored divided by the maximal capacity $\mathcal{E}_{\max} = 150$ kWh. In particular, the total installed storage capacity corresponds to 5 hours of the PV panels production at full capacity, which is 30 kW, which corresponds to about 300 squared meters of photo-voltaic panels, with the current technology. Equivalently, assuming a availability rate of 12% for solar (accounting for seasonality, intermittency and unavailability at night), about 40 hours of the average solar production, where the availability rate is defined as the average power production divided by the maximal power capacity. We consider simple ideal batteries with charging and discharging efficiencies equal to 1. We do not enforce the state constraints $X^{(n)} \in [0, 1]$. We will consider a non-quadratic loss function $\mathcal{L}$ given by:

$$\mathcal{L}(x) = \begin{cases} \frac{1}{2}\lambda x^2 - \frac{\delta}{6\varepsilon}x^3 & \text{if } -\varepsilon \le x \le \varepsilon, \\ \frac{1}{2}(\lambda - \delta)x^2 + \frac{\delta\varepsilon}{2}x - \frac{\delta}{6}\varepsilon^2 & \text{if } x \ge \varepsilon, \\ \frac{1}{2}(\lambda + \delta)x^2 + \frac{\delta\varepsilon}{2}x + \frac{\delta}{6}\varepsilon^2 & \text{if } x \le -\varepsilon. \end{cases}$$

In this case, $\mathcal{L}$ is C^2 with $\frac{\delta}{\varepsilon}$-Lipschitz-continuous second order derivatives and is represented in Figure 5 for $\lambda = 2, \delta = 1, \varepsilon = 0.1$.

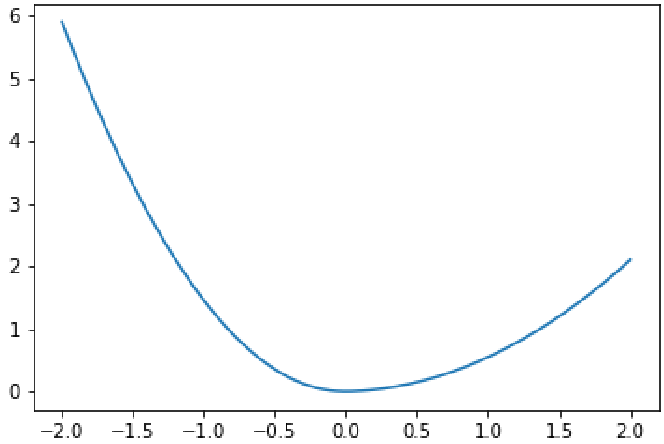


Figure 5: Loss function $\mathcal{L}$ with $\lambda = 2, \delta = 1, \varepsilon = 0.1$

The function $\mathcal{L}$ penalizes more energy production deficit (as compared to its expected value). Indeed, such situation possibly requires the use of extra production units with high carbon footprint, which is clearly to discard as often as possible. We will use the following parameter values for the cost functional.

Table 2: Parameter values for the cost functional

μ	ν	ρ	λ	δ	ε
1 kW^{-2} h^{-1}	5 h^{-1}	10000	10 kW^{-2} h^{-1}	5 kW^{-2} h^{-1}	1 kW

4.2 Solving the stochastic control problem

Applying a similar methodology as in Theorem 2.6, one can show that it is equivalent to solve the stochastic control problem (4.3) and the following high-dimension coupled FBSDE:

$$\forall n \in [N], \quad \begin{cases} X_t^{(n)} = x_0^{(n)} - \int_0^t \frac{u_s^{(n)}}{\mathcal{E}_{\max}} \mathrm{d}s, \\ Y_t^{(n)} = \mathbb{E}_t\left[\rho(X_T^{(n)} - 1/2) + \int_t^T \nu(X_s^{(n)} - 1/2)\mathrm{d}s\right], \\ \mu u_t^{(n)} + \mathcal{L}'\left(\frac{1}{N}\sum_{j=1}^N u_t^{(j)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]\right) - \frac{Y_{t-}^{(n)}}{\mathcal{E}_{\max}} = 0. \end{cases} \tag{4.5}$$

There exists a unique solution $(u^{(n)}, X^{(n)}, Y^{(n)})_{n\in[N]} \in (\mathbb{H}_{\mathcal{P}}^{2,2} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2})^N$ of the above FBSDE, and $(u^{(n)})_{n\in[N]}$ is the unique solution of the stochastic control problem (4.3). The above FBSDE (4.5) is a high-dimensional fully coupled FBSDE. To solve it, introduce

$$(\bar{u}, \bar{X}, \bar{Y}) := \left(\frac{1}{N}\sum_{j=1}^N u^{(j)}, \frac{1}{N}\sum_{j=1}^N X^{(j)}, \frac{1}{N}\sum_{j=1}^N Y^{(j)}\right).$$

Introduce as well $\left(u^{(n),\Delta}, X^{(n),\Delta}, Y^{(n),\Delta}\right)_{n\in[N]} := \left(u^{(n)} - \bar{u}, X^{(n)} - \bar{X}, Y^{(n)} - \bar{Y}\right)$. Then for all $n \in [N]$, $\left(u^{(n),\Delta}, X^{(n),\Delta}, Y^{(n),\Delta}\right)$ is a solution of the linear FBSDE:

$$\begin{cases} X_t^{(n),\Delta} = x_0^{(n)} - \bar{x}_0 - \int_0^t \frac{u_s^{(n),\Delta}}{\mathcal{E}_{\max}} \mathrm{d}s, \\ Y_t^{(n),\Delta} = \mathbb{E}_t \left[\rho X_T^{(n),\Delta} + \int_t^T \nu X_s^{(n),\Delta} \mathrm{d}s \right], \\ \mu u_t^{(n),\Delta} - \frac{Y_{t-}^{(n),\Delta}}{\mathcal{E}_{\max}} = 0. \end{cases} \tag{4.6}$$

To solve efficiently this linear FBSDE, we use Theorem **??**. Let us turn to the computation of $(\bar{u}, \bar{X}, \bar{Y})$. Notice that $(\bar{u}, \bar{X}, \bar{Y})$ is solution of the following FBSDE, where we denoted $\bar{x}_0 = \frac{1}{N} \sum_{j=1}^N x_0^{(j)}$:

$$\begin{cases} \bar{X}_t = \bar{x}_0 - \int_0^t \frac{\bar{u}_s}{\mathcal{E}_{\max}} \mathrm{d}s, \\ \bar{Y}_t = \mathbb{E}_t \left[\rho(\bar{X}_T - 1/2) + \int_t^T \nu(\bar{X}_s - 1/2) \mathrm{d}s \right], \\ \mu \bar{u}_t + \mathcal{L}' \left(\bar{u}_t + \mathrm{P}_t^{\mathsf{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathsf{sun}}\right] \right) - \frac{\bar{Y}_t}{\mathcal{E}_{\max}} = 0. \end{cases} \tag{4.7}$$

We will assume $\bar{x}_0 = 0.5$ in our case study. This FBSDE fully characterizes the solution of the following stochastic control problem, called coordination problem:

$$\begin{aligned} \min_{\bar{u} \in \mathbb{H}^{2,2}} \quad & \bar{\mathcal{J}}(\bar{u}) := \mathbb{E}\left[\int_0^T \left\{ \frac{\mu}{2} \bar{u}_t^2 + \frac{\nu}{2} \left(\bar{X}_t - \frac{1}{2} \right)^2 + \mathcal{L}(\bar{u}_t + \mathrm{P}_t^{\mathsf{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathsf{sun}}\right]) \right\} \mathrm{d}t + \frac{\rho}{2} \left(\bar{X}_T - \frac{1}{2} \right)^2 \right] \\ s.t. \quad & \bar{X}_t = \bar{x}_0 - \int_0^t \frac{\bar{u}_s}{\mathcal{E}_{\max}} \mathrm{d}s. \end{aligned}$$

Proposition 2.3 shows that there exists a unique solution of the coordination problem. By applying Theorem 2.6, we deduce the existence and uniqueness of a solution $(\bar{u}, \bar{X}, \bar{Y}) \in \mathbb{H}_{\mathcal{P}}^{2,2} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$. To solve the non-linear FBSDE (4.7), we use Newton's method globalized with (Gradient) Backtracking Line Search, noting that the random parameters of the problem are uniformly bounded.

Let $\bar{u}^{(k)}$ be the (candidate) control variable at iteration k. We define the associated state variable $\bar{X}^{(k)}$ at iteration k:

$$\bar{X}_t^{(k)} = \bar{x}_0 - \int_0^t \frac{\bar{u}_s^{(k)}}{\mathcal{E}_{\max}} \mathrm{d}s, \tag{4.8}$$

the adjoint variable $\bar{Y}^{(k)}$ at iteration k:

$$\bar{Y}_t^{(k)} = \mathbb{E}_t \left[\rho(\bar{X}_T^{(k)} - 1/2) + \int_t^T \nu(\bar{X}_s^{(k)} - 1/2) \mathrm{d}s \right]. \tag{4.9}$$

Applying Proposition 2.5 to $\bar{\mathcal{J}}$, the gradient of the cost at $\bar{u}^{(k)}$ is given by:

$$(\nabla \bar{\mathcal{J}}(\bar{u}^{(k)}))_t = \mu \bar{u}_t^{(k)} + \mathcal{L}' \left(\bar{u}_t^{(k)} + \mathrm{P}_t^{\mathsf{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathsf{sun}}\right] \right) - \frac{\bar{Y}_{t-}^{(k)}}{\mathcal{E}_{\max}}. \tag{4.10}$$

The Newton direction $\dot{u}^{(k)} = -(\nabla^2 \bar{\mathcal{J}}(\bar{u}^{(k)}))^{-1} (\nabla \bar{\mathcal{J}}(\bar{u}^{(k)}))$ at the point $\bar{u}^{(k)}$ is given by:

$$\dot{u}_t^{(k)} = \frac{\dot{Y}_{t-}^{(k)} + \bar{Y}_{t-}^{(k)} - \left\{ \mu \bar{u}_t^{(k)} + \mathcal{L}' \left(\bar{u}_t^{(k)} + \mathrm{P}_t^{\mathsf{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathsf{sun}}\right] \right) \right\} \mathcal{E}_{\max}}{\left\{ \mu + \mathcal{L}'' \left(\bar{u}_t^{(k)} + \mathrm{P}_t^{\mathsf{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathsf{sun}}\right] \right) \right\} \mathcal{E}_{\max}},$$

where $(\dot{X}^{(k)}, \dot{Y}^{(k)})$ satisfy:

$$\begin{cases} \dot{X}_t^{(k)} = -\int_0^t \frac{\dot{u}_s^{(k)}}{\mathcal{E}_{\max}} \mathrm{d}s, \\ \dot{Y}_t^{(k)} = \mathbb{E}_t \left[\rho \dot{X}_T^{(k)} + \int_t^T \nu \dot{X}_s^{(k)} \mathrm{d}s \right]. \end{cases}$$

This comes from the application of Theorem 2.9 which gives the expression of the inverse of the second order derivative at $\bar{u}^{(k)}$, applied to $-\nabla\bar{\mathcal{J}}(\bar{u}^{(k)})$. Eliminating $\dot{u}^{(k)}$, we obtain the following affine-linear FBSDE for $(\dot{X}^{(k)}, \dot{Y}^{(k)})$:

$$\begin{cases} \dot{X}_t^{(k)} = \displaystyle\int_0^t \frac{-\dot{Y}_s^{(k)} - \bar{Y}_s^{(k)} + \left\{\mu\bar{u}_s^{(k)} + \mathcal{L}'\left(\bar{u}_s^{(k)} + \mathrm{P}_s^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}}{\left\{\mu + \mathcal{L}''\left(\bar{u}_s^{(k)} + \mathrm{P}_s^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}^2}\mathrm{d}s, \\ \dot{Y}_t^{(k)} = \mathbb{E}_t\left[\rho\dot{X}_T^{(k)} + \int_t^T \nu\dot{X}_s^{(k)}\mathrm{d}s\right]. \end{cases}$$

To solve this FBSDE arising at iteration k, we apply Theorem 3.7 in our particular framework, with the following values for the parameters:

$$\begin{cases} A_t = 0, \\ B_t = \dfrac{-1}{\left\{\mu + \mathcal{L}''\left(\bar{u}_t^{(k)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}^2}, \\ C_t = \nu, \\ \Gamma = \rho, \\ a_t = \dfrac{-\bar{Y}_t^{(k)} + \mu\bar{u}_t^{(k)}\mathcal{E}_{\max} + \mathcal{L}'\left(\bar{u}_t^{(k)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]\right)\mathcal{E}_{\max}}{\left\{\mu + \mathcal{L}''\left(\bar{u}_t^{(k)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}^2}, \\ b_t = 0, \\ \eta = 0, \\ x = 0. \end{cases}$$

Introduce the Riccati BSDE with stochastic coefficients:

$$P_t^{(k)} = \mathbb{E}_t\left[\rho + \int_t^T \left(\nu - \frac{1}{\left\{\mu + \mathcal{L}''\left(\bar{u}_s^{(k)} + \mathrm{P}_s^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}^2}(P_s^{(k)})^2\right)\mathrm{d}s\right], \tag{4.11}$$

and the linear BSDE:

$$\Pi_t^{(k)} = \mathbb{E}_t\left[\int_t^T \left(\frac{-P_s^{(k)}}{\mathcal{E}_{\max}^2\left(\mu + \mathcal{L}''\left(\bar{u}_s^{(k)} + \mathrm{P}_s^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]\right)\right)}\Pi_s^{(k)} + \frac{\left(\mu\bar{u}_s^{(k)} + \mathcal{L}'\left(\bar{u}_s^{(k)} + \mathrm{P}_s^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]\right)\right)\mathcal{E}_{\max} - \bar{Y}_s^{(k)}}{\mathcal{E}_{\max}^2\left(\mu + \mathcal{L}''\left(\bar{u}_s^{(k)} + \mathrm{P}_s^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]\right)\right)}P_s^{(k)}\right)\mathrm{d}s\right]. \tag{4.12}$$

Then, $\dot{u}^{(k)}$ is given by the following feedback expression:

$$\dot{u}_t^{(k)} = \frac{P_{t-}^{(k)}\dot{X}_t^{(k)} + \Pi_{t-}^{(k)} + \bar{Y}_{t-}^{(k)} - \left\{\mu\bar{u}_t^{(k)} + \mathcal{L}'\left(\bar{u}_t^{(k)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}}{\left\{\mu + \mathcal{L}''\left(\bar{u}_t^{(k)} + \mathrm{P}_t^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_t^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}}. \tag{4.13}$$

The process $\dot{X}^{(k)}$ satisfies:

$$\dot{X}_t^{(k)} = \int_0^t \frac{-P_s^{(k)}\dot{X}_s^{(k)} - \Pi_s^{(k)} - \bar{Y}_s^{(k)} + \left\{\mu\bar{u}_s^{(k)} + \mathcal{L}'\left(\bar{u}_s^{(k)} + \mathrm{P}_s^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}}{\left\{\mu + \mathcal{L}''\left(\bar{u}_s^{(k)} + \mathrm{P}_s^{\mathrm{sun}} - \mathbb{E}\left[\mathrm{P}_s^{\mathrm{sun}}\right]\right)\right\}\mathcal{E}_{\max}^2}\mathrm{d}s,$$

and $\dot{Y}^{(k)} = P^{(k)}\dot{X}^{(k)} + \Pi^{(k)}$, according to Theorem 3.7.

To be able to practically implement the Newton method with (Gradient) Backtracking line search, the conditional expectations in the equations or expressions of $\bar{Y}^{(k)}$, $P^{(k)}$, $\Pi^{(k)}$ in (4.9), (4.11) and (4.12) need to be estimated. We focus on these aspects in the next section.

4.3 Practical implementation

The simulations have been performed on Python 3.7, with an Intel-Core i7 PC at 2.1 GHz with 16 Go memory. The process $\mathrm{X}^{\mathrm{sun}}$ is simulated using an Euler scheme with time step $h = \frac{T}{N_T} = 0.5$ h, with $T = 24$ h and $N_T = 48$. The number of Monte-Carlo simulations is $M = 10000$.

4.3.1 Linear Least-Squares Regression

To provide an implementation to compute the conditional expectation in the expression of $\bar{Y}^{(k)}$, $P^{(k)}$, $\Pi^{(k)}$ in (4.9), (4.11) and (4.12), we take advantage on the Markovian framework and use Linear-Least Square regression [GT16a]. We use this method to obtain closed-loop feedback expressions of the solutions of each BSDE, with respect to the Markovian underlying extended state process $(\mathtt{X}^{\mathrm{sun}}, \bar{X}^{(k)})$.

Notations To simplify the notations, we write $(u^{(k)}, X^{(k)}, Y^{(k)}) = (u^{(k)}_\tau, X^{(k)}_\tau, Y^{(k)}_\tau)_{\tau\in[N_T]}$ the discretized process associated $(\bar{u}^{(k)}, \bar{X}^{(k)}, \bar{Y}^{(k)})$ in the following on the time grid $(\tau h)_{\tau\in[N_T]}$(not be be confused with $(u^{(n)}, X^{(n)}, Y^{(n)})$ which is the optimal control, state and adjoint variable of battery n). We also use the notation $\mathtt{X}^{\mathrm{sun},h}_{\tau,m}$ for values of the m^{th} simulated (discretized) approximation of $\mathtt{X}^{\mathrm{sun}}$ at time τh.

Definition 4.1 (Linear Least-Squares Regression (LLSR) [GT16a])**.** *For $l \geq 1$ and for probability spaces $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\mathbb{R}^l, \mathcal{B}(\mathbb{R}^l), \nu)$, let S be a $\mathcal{F}\otimes\mathcal{B}(\mathbb{R}^l)$-measurable $\mathbb{R}$-valued function such that $S(\omega,\cdot) \in \mathbb{L}^2(\mathcal{B}(\mathbb{R}^l), \nu)$ for $\mathbb{P}$-a.e. $\omega \in \Omega$. Let $\mathcal{K} := span(\phi_f)_{f=1,\dots,N_f}$ be the vector space spanned by N_f deterministic functions $(\phi_f)_{f=1,\dots,N_f}$. The Least-Squares approximation of S in the space $\mathcal{K}$ with respect to ν is the $\mathrm{d}\mathbb{P}\otimes\mathrm{d}\nu$-a.e. unique $\mathcal{F}\otimes\mathcal{B}(\mathbb{R}^l)$-measurable function $S^\star$ given by:*

$$S^\star(\omega,\cdot) = \arg\inf_{\phi\in\mathcal{K}} \int |\phi(x) - S(\omega,x)|^2 \nu(\mathrm{d}x).$$

*We say that $S^\star$ solves **OLS**$(S, \mathcal{K}, \nu)$.*

In particular, if $\nu_M = \frac{1}{M}\sum_{m=1}^M \delta_{\chi^{(m)}}$ is a discrete probability measure on $(\mathbb{R}^l, \mathcal{B}(\mathbb{R}^l))$ where $\chi^{(1)}, \chi^{(2)}, \dots, \chi^{(M)} : \Omega \to \mathbb{R}^l$ are i.i.d. random variables with distribution ν. For an $\mathcal{F}\otimes\mathcal{B}(\mathbb{R}^l)$-measurable $\mathbb{R}$-valued function S such that $|S(\omega, \chi^{(m)}(\omega))| < \infty$ for any m and $\mathbb{P}$-a.e. $\omega \in \Omega$, the Least-Squares approximation of S in the space $\mathcal{K}$ with respect to ν_M is the $\mathbb{P} - a.e.$ unique $\mathcal{F}\otimes\mathcal{B}(\mathbb{R}^l)$-measurable function $S^\star$ given by:

$$S^\star(\omega,\cdot) = \arg\inf_{\phi\in\mathcal{K}} \frac{1}{M}\sum_{m=1}^M |\phi(\chi^{(m)}(\omega)) - S(\omega, \chi^{(m)}(\omega))|^2.$$

Informally, relying on the Markovian framework, we wish to use LLSR to obtain approximations of the solutions of the BSDE at time step τ in the form of closed-loop feedback with respect to the current value of the extended state variable $(\mathtt{X}^{\mathrm{sun},h}_\tau, X^{(k)}_\tau)$, i.e., we wish to determine $\Phi^{(k)}_{Y,\tau}$, $\Phi^{(k)}_{P,\tau}$, $\Phi^{(k)}_{\Pi,\tau}$ to obtain estimates of the form:

$$Y^{(k)}_\tau \simeq \Phi_{Y,\tau}(\mathtt{X}^{\mathrm{sun},h}_\tau, X^{(k)}_\tau); \quad P^{(k)}_t \simeq \Phi_{P,\tau}(\mathtt{X}^{\mathrm{sun},h}_\tau, X^{(k)}_t); \quad \Pi^{(k)}_\tau \simeq \Phi_{\Pi,\tau}(\mathtt{X}^{\mathrm{sun},h}_\tau, X^{(k)}_\tau).$$

We introduce the notation $\nu^{(k)}_{\tau,[M]} = \frac{1}{M}\sum_{m=1}^M \delta_{\mathtt{X}^{\mathrm{sun},h}_{[\tau:N_T],m}, u^{(k)}_{[\tau:N_T],m}, X^{(k)}_{[\tau:N_T],m}}$ for the empirical measure. We define $\mathcal{K}_\tau$ as the 3-dimensional vector space of functions spanned by $(\phi_{\tau,1}, \phi_{\tau,2}, \phi_{\tau,3})$ taking as arguments $(z_{[\tau:N_T]}, v_{[\tau:N_T]}, x_{[\tau:N_T]}) \in \mathbb{R}^{3(N_T-\tau+1)}$ and returning respectively 1, z_τ and x_τ. Hence ϕ_1 spans the vector space of constant functions, while ϕ_2 and ϕ_3 span the vector space of linear functions depending only on the two state variables at time τ.

We could consider more features in the function space $\mathcal{K}_\tau$ to allow more accurate functional representation. This is left for further investigation.

4.3.2 Ensuring the respect of a priori bounds for the solution of Riccati BSDE

In addition to this, we truncate $P^{(k)}$ at each iteration, using a priori upper and lower bounds. This helps to stabilize the LLSR algorithm [GT16b]. Let us introduce P^{UB} and P^{LB} the unique solutions of the Riccati ordinary differential equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}P^{UB}_t = \frac{(P^{UB}_t)^2}{(\lambda+\mu+\delta)\mathcal{E}^2_{\max}} - \nu, \qquad P^{UB}_T = \rho$$
$$\frac{\mathrm{d}}{\mathrm{d}t}P^{LB}_t = \frac{(P^{LB}_t)^2}{(\lambda+\mu-\delta)\mathcal{E}^2_{\max}} - \nu, \qquad P^{LB}_T = \rho.$$

Note that both equations can be easily solved analytically and numerically using similar arguments as when computing P^{Δ}. Using the (uniform) bounds:

$$\forall x \in \mathbb{R}, \quad \lambda - \delta \le \mathcal{L}''(x) \le \lambda + \delta,$$

it can be show by comparison principle for BSDEs in general filtrations [ØZ12, Theorem 3.4, p. 710] that

$$P_t^{LB} \le P_t^{(k)} \le P_t^{UB}, \tag{4.14}$$

for all $t \in [0, T]$, $d\mathbb{P}$-a.e. We use this to truncate the numerical approximation of $P^{(k)}$.

4.3.3 Other parameters and implementation details

We choose the parameter values $\gamma = \beta = 0.1$. The parameter γ is linked to the acceptance rate of the (possibly reduced) Newton step, and 0.1 is a good trade-off between the need of sufficient reduction and a high acceptance rate. The choice $\beta = 0.1$ ensures that when sufficient reduction is not achieved, the step is sufficiently reduced to provide an acceptable step length with high probability.

4.3.4 The algorithms considered

We implement and compare Newton's method combined with two backtracking line-search: the standard backtracking line-search Algorithm 3 and the Gradient Backtracking line-search 4 designed in this paper. To estimate the mean value of a random variable given an i.i.d. samples of this random variable, we use the empirical mean, which is an unbiased estimator. To estimate $\|X\|_{\mathbb{H}^{\infty,\infty}}$ given i.i.d. sample trajectories $(X_{\tau,m})_{\tau \in [N_T], m \in [M]}$, we use the estimator $\sup_{\tau \in [N_T], m \in [M]} X_{\tau,m}$, which is lower biased (neglecting the impact of time discretization on the bias). More accurate estimators based on extreme-value theory could be used, see for instance [ANR17]. Practical implementations of Algorithms 3, 4 and 5 are respectively given by Algorithms 6, 7 and 8. We use the initial guess $u^{(0)} = 0$ for Newton method. However, using the easily computable solution of the linear quadratic problem obtained by replacing the non-quadratic loss $\mathcal{L}$ by the quadratic loss function $\mathcal{L}^{\text{quad}} : x \mapsto \frac{\lambda x^2}{2}$ could allow to find a better initial guess (warm start). Though we do not show the results of such a procedure, numerical experiments show that this amounts to reduce by 1 the number of Newton iterations required to obtain a given accuracy.

4.3.5 On the stopping criteria of the Newton method

Ideally, the stopping criteria of the Newton method with Gradient Backtracking line-search should be $\|\nabla \bar{\mathcal{J}}(u^{(k)})\|_{\mathbb{H}^{\infty,\infty}} \le \varepsilon$. However, the norm of the gradient as estimated is erroneous, due to discretization and regression errors, and shold be estimated on a test set, distinct from the training set used in the algorithm. Hence, finding a relevant stopping criteria is a difficult task and left for further investigation. In practice, we shall replace the "while" loop by a "for" loop with a fixed number of iterations, and monitor the estimated norm of the gradient along iterations.

Algorithm 6 Standard Backtracking line search with Linear Least Square Regression

Inputs: Current control: $(u^{(k)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$, current state $(X^{(k)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$, $(\dot{u}^{(k)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$ Newton direction, $(\hat{Y}^{(k)}_{\tau})_{\tau\in[N_T]}$ regression functions for adjoint variable $Y^{(k)}$, $(\beta,\gamma)\in(0,1)^2$, M trajectories of solar irradiance $(\mathtt{X}^{\mathtt{sun},h}_{\tau,m})_{\tau\in[N_T],m\in[M]}$.
$\sigma = 1$.
repeat
  $u^{(k+1)} = u^{(k)} + \sigma \dot{u}^{(k)}$.
  Compute $(X^{(k+1)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$ by an Euler scheme. See (4.8).
  $\sigma \leftarrow \beta\sigma$.
  {*Computation of discretized gradient* (4.10).}
  $\nabla\mathcal{J}^{(k)}_{\tau,m} = \mu u^{(k)}_{\tau,m} + \mathcal{L}'\left(u^{(k)}_{\tau,m} + \mathtt{X}^{\mathtt{sun},h}_{\tau,m}\mathtt{P}^{\mathtt{sun,max}}_{\tau h} - \mathbb{E}\left[\mathtt{P}^{\mathtt{sun}}_{\tau h}\right]\right) - \frac{1}{\mathcal{E}_{\max}}\hat{Y}^{(k)}_{\tau}(\mathtt{X}^{\mathtt{sun},h}_{\tau,m}, u^{(k)}_{\tau,m}, X^{(k)}_{\tau,m})$.
  {*Computation of cost function.*}
  $\mathcal{J}^{(k+1)}_{m} = \sum_{\tau=1}^{N_T}\left(\frac{\mu}{2}(u^{(k+1)}_{\tau,m})^2 + \frac{\nu}{2}\left(X^{(k+1)}_{\tau,m} - \frac{1}{2}\right)^2 + \mathcal{L}\left(u^{(k+1)}_{\tau,m} + \mathtt{X}^{\mathtt{sun},h}_{\tau,m}\mathtt{P}^{\mathtt{sun,max}}_{\tau h} - \mathbb{E}\left[\mathtt{P}^{\mathtt{sun}}_{\tau h}\right]\right)\right)h + \frac{\rho}{2}\left(X^{(k+1)}_{N_T,m} - \frac{1}{2}\right)^2$.
  $\mathcal{J}^{(k)}_{m} = \sum_{\tau=1}^{N_T}\left(\frac{\mu}{2}(u^{(k)}_{\tau,m})^2 + \frac{\nu}{2}\left(X^{(k)}_{\tau,m} - \frac{1}{2}\right)^2 + \mathcal{L}\left(u^{(k)}_{\tau,m} + \mathtt{X}^{\mathtt{sun},h}_{\tau,m}\mathtt{P}^{\mathtt{sun,max}}_{\tau h} - \mathbb{E}\left[\mathtt{P}^{\mathtt{sun}}_{\tau h}\right]\right)\right)h + \frac{\rho}{2}\left(X^{(k)}_{N_T,m} - \frac{1}{2}\right)^2$.
until $\frac{1}{M}\sum_{m=1}^{M}\mathcal{J}^{(k+1)}_{m} \le \frac{1}{M}\sum_{m=1}^{M}\left(\mathcal{J}^{(k)}_{m} + \gamma\sigma\sum_{\tau=1}^{N_T}\nabla\mathcal{J}^{(k)}_{\tau,m}\dot{u}^{(k)}_{\tau,m}h\right)$ {*Sufficient decrease of cost*}
$\sigma \leftarrow \sigma/\beta$ {*Correction of* σ *which has been reduced one too many times.*}
for $\tau = N_T, .., 1$ **do**
  Define the empirical measure $\nu^{(k+1)}_{\tau,[M]} := \frac{1}{M}\sum_{m=1}^{M}\delta_{\mathtt{X}^{\mathtt{sun},h}_{[\tau:N_T],m}, u^{(k+1)}_{[\tau:N_T],m}, X^{(k+1)}_{[\tau:N_T],m}}$.
  {*Regression of adjoint variable. See* (4.9) *and Definition 4.1.*}
  Compute $\hat{Y}^{(k+1)}_{\tau}$ solution of **OLS**(S_{Y_τ}, $\mathcal{K}_\tau$, $\nu^{(k+1)}_{\tau,[M]}$) with $S_{Y_\tau}(z_{[\tau:N_T]}, v_{[\tau:N_T]}, x_{[\tau:N_T]}) = \rho(x_{N_T} - 1/2) + \sum_{j=\tau+1}^{N_T}\nu(x_j - 1/2)h$.
end for
return $u^{(k+1)}$, $X^{(k+1)}$, $\hat{Y}^{(k+1)}$.

Algorithm 7 Gradient Backtracking line search with Linear Least Square Regression

Inputs: $(u^{(k)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$, $(X^{(k)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$, $(\dot{u}^{(k)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$, $(\hat{Y}^{(k)}_{\tau})_{\tau\in[N_T]}$, $(\beta,\gamma)\in(0,1)^2$, M trajectories of solar irradiance $(\mathtt{X}^{\mathtt{sun},h}_{\tau,m})_{\tau\in[N_T],m\in[M]}$.
$\sigma = 1$.
repeat
  $u^{(k+1)} = u^{(k)} + \sigma \dot{u}^{(k)}$.
  Compute $(X^{(k+1)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$ by an Euler scheme. See (4.8).
  for $\tau = N_T, .., 1$ **do**
    Define the empirical measure $\nu^{(k+1)}_{\tau,[M]} := \frac{1}{M}\sum_{m=1}^{M}\delta_{\mathtt{X}^{\mathtt{sun},h}_{[\tau:N_T],m}, u^{(k+1)}_{[\tau:N_T],m}, X^{(k+1)}_{[\tau:N_T],m}}$.
    {*Adjoint variable regression. See* (4.9) *and Definition 4.1.*}
    Compute $\hat{Y}^{(k+1)}_{\tau}$ solution of **OLS**(S_{Y_τ}, $\mathcal{K}_\tau$, $\nu^{(k+1)}_{\tau,[M]}$) with $S_{Y_\tau}(z_{[\tau:N_T]}, v_{[\tau:N_T]}, x_{[\tau:N_T]}) = \rho(x_{N_T} - 1/2) + \sum_{j=\tau+1}^{N_T}\nu(x_j - 1/2)h$.
  end for
  $\sigma \leftarrow \beta\sigma$
  {*Computation of discretized gradient* (4.10).}
  $\nabla\mathcal{J}^{(k+1)}_{\tau,m} = \mu u^{(k+1)}_{\tau,m} + \mathcal{L}'\left(u^{(k+1)}_{\tau,m} + \mathtt{X}^{\mathtt{sun},h}_{\tau,m}\mathtt{P}^{\mathtt{sun,max}}_{\tau h} - \mathbb{E}\left[\mathtt{P}^{\mathtt{sun}}_{\tau h}\right]\right) - \frac{1}{\mathcal{E}_{\max}}\hat{Y}^{(k+1)}_{\tau}(\mathtt{X}^{\mathtt{sun},h}_{\tau,m}, u^{(k+1)}_{\tau,m}, X^{(k+1)}_{\tau,m})$.
  $\nabla\mathcal{J}^{(k)}_{\tau,m} = \mu u^{(k)}_{\tau,m} + \mathcal{L}'\left(u^{(k)}_{\tau,m} + \mathtt{X}^{\mathtt{sun},h}_{\tau,m}\mathtt{P}^{\mathtt{sun,max}}_{\tau h} - \mathbb{E}\left[\mathtt{P}^{\mathtt{sun}}_{\tau h}\right]\right) - \frac{1}{\mathcal{E}_{\max}}\hat{Y}^{(k)}_{\tau}(\mathtt{X}^{\mathtt{sun},h}_{\tau,m}, u^{(k)}_{\tau,m}, X^{(k)}_{\tau,m})$.
until $\max_{\tau\in[N_T],m\in[M]}\left|\nabla\mathcal{J}^{(k+1)}_{\tau,m}\right| \le (1-\gamma\sigma)\left|\max_{\tau\in[N_T],m\in[M]}\left|\nabla\mathcal{J}^{(k)}_{\tau,m}\right|\right.$ {*Sufficient decrease condition*}
$\sigma \leftarrow \sigma/\beta$ {*Correction of* σ *which has been reduced one too many times.*}
return $u^{(k+1)}$, $X^{(k+1)}$, $\hat{Y}^{(k+1)}$.

Algorithm 8 Newton method with Least-Square Regression and backtracking line search

Initialization: M trajectories of solar irradiance $(\mathtt{X}^{\mathtt{sun},h}_{\tau,m})_{\tau\in[N_T],m\in[M]}$.

Set $k=0$.

Set $(u^{(0)}_{\tau,m})_{\tau\in[N_T],m\in[M]} = 0$.

Compute $(X^{(0)}_{\tau,m})_{\tau\in[N_T],m\in[M]}$ by an Euler scheme. See (4.8).

for $\tau = N_T, .., 1$ **do**

{*Regression function of adjoint state variable, see* (4.9) *and Definition 4.1.*}

Define the empirical measure $\nu^{(0)}_{\tau,[M]} := \frac{1}{M}\sum_{m=1}^{M}\delta_{\mathtt{X}^{\mathtt{sun},h}_{[\tau:N_T],m},\,u^{(0)}_{[\tau:N_T],m},\,X^{(0)}_{[\tau:N_T],m}}$.

$\hat{Y}^{(0)}_\tau$ solution of **OLS**$(S_{Y_\tau}, \mathcal{K}_\tau, \nu^{(0)}_{\tau,[M]})$ with $S_{Y_\tau}(z_{[\tau:N_T]}, v_{[\tau:N_T]}, x_{[\tau:N_T]}) = \rho(x_{N_T} - 1/2) + \sum_{j=\tau+1}^{N_T}\nu(x_j - 1/2)h$.

end for

while Stopping criteria not met **do**

for $\tau = N_T, ..., 1$ **do**

{*Regression functions for* $P^{(k)}$ *and* $\Pi^{(k)}$ *solutions of* (4.11) *and* (4.12). *See Definition 4.1. Truncation of the estimator of* $P^{(k)}$ *to verify the a priori bounds* (4.14).}

Define the empirical measure $\nu^{(k)}_{\tau,[M]} := \frac{1}{M}\sum_{m=1}^{M}\delta_{\mathtt{X}^{\mathtt{sun},h}_{[\tau:N_T],m},\,u^{(k)}_{[\tau:N_T],m},\,X^{(k)}_{[\tau:N_T],m}}$.

$\hat{P}^{(k)}_\tau$ as the projection on the convex set $[P^{LB}_\tau, P^{UB}_\tau]$ of the solution of **OLS**$(S_{P^{(k)}_\tau}, \mathcal{K}_\tau, \nu^{(k)}_{\tau,[M]})$ with

$$S_{P^{(k)}_\tau}(z_{[\tau:N_T]}, v_{[\tau:N_T]}, x_{[\tau:N_T]}) = \rho + \sum_{j=\tau+1}^{N_T}\left(-\frac{1}{\mathcal{E}^2_{\max}\left\{\mu + l''\left(z_j, v_j\right)\right\}}\left(\hat{P}^{(k)}_j(z_j, x_j)\right)^2 + \nu\right)h.$$

$\hat{\Pi}^{(k)}_\tau$ solution of **OLS**$(S_{\Pi^{(k)}_\tau}, \mathcal{K}_\tau, \nu^{(k)}_{\tau,[M]})$ with $S_{\Pi^{(k)}_\tau}(z_{[\tau:N_T]}, v_{[\tau:N_T]}, x_{[\tau:N_T]})$ given by

$$\sum_{j=\tau+1}^{N_T}\left(-\frac{\hat{P}^{(k)}_j(z_j, x_j)}{\mathcal{E}^2_{\max}\left\{\mu + l''_j(z_j, v_j)\right\}}\hat{\Pi}^{(k)}_j(z_j, x_j) + \frac{\left(\mu v_j + l'_j(z_j, v_j)\right)\mathcal{E}_{\max} - \hat{Y}^{(k)}_j(z_j, x_j)}{\mathcal{E}^2_{\max}\left\{\mu + l''_j(z_j, v_j)\right\}}\hat{P}^{(k)}_j(z_j, x_j)\right)h,$$

where we used the notations $l''_j(z, v) := \mathcal{L}''\left(v + z\mathtt{P}^{\mathtt{sun,max}}_{jh} - \mathbb{E}\left[\mathtt{P}^{\mathtt{sun}}_{jh}\right]\right)$ and $l'_j(z, v) := \mathcal{L}'\left(v + z\mathtt{P}^{\mathtt{sun,max}}_{jh} - \mathbb{E}\left[\mathtt{P}^{\mathtt{sun}}_{jh}\right]\right)$

end for

{*Computation of Newton step by feedback expression* (4.13).}

for $m = 1 \in [M]$ **do**

$\dot{X}^{(k)}_{1,m} = 0$.

for $\tau = 1, ..N_T$ **do**

Denote $\hat{P}^{(k)}_{\tau,m} = \hat{P}^{(k)}_\tau(\mathtt{X}^{\mathtt{sun},h}_{\tau,m}, X^{(k)}_{\tau,m})$, $\hat{\Pi}^{(k)}_{\tau,m} = \hat{\Pi}^{(k)}_\tau(\mathtt{X}^{\mathtt{sun},h}_{\tau,m}, X^{(k)}_{\tau,m})$ and $\hat{Y}^{(k)}_{\tau,m} = \hat{Y}^{(k)}_\tau(\mathtt{X}^{\mathtt{sun},h}_{\tau,m}, X^{(k)}_{\tau,m})$.

$$\dot{u}^{(k)}_{\tau,m} = \frac{\hat{P}^{(k)}_{\tau,m}\dot{X}^{(k)}_{\tau,m} + \hat{\Pi}^{(k)}_{\tau,m} + \hat{Y}^{(k)}_{\tau,m}}{\mathcal{E}_{\max}\left\{\mu + \mathcal{L}''\left(u^{(k)}_{\tau,m} + \mathtt{X}^{\mathtt{sun},h}_{\tau,m}\mathtt{P}^{\mathtt{sun,max}}_{\tau h} - \mathbb{E}\left[\mathtt{P}^{\mathtt{sun}}_{\tau h}\right]\right)\right\}} - \frac{\mu u^{(k)}_{\tau,m} + \mathcal{L}'\left(u^{(k)}_{\tau,m} + \mathtt{X}^{\mathtt{sun},h}_{\tau,m}\mathtt{P}^{\mathtt{sun,max}}_{\tau h} - \mathbb{E}\left[\mathtt{P}^{\mathtt{sun}}_{\tau h}\right]\right)}{\mu + \mathcal{L}''\left(u^{(k)}_{\tau,m} + \mathtt{X}^{\mathtt{sun},h}_{\tau,m}\mathtt{P}^{\mathtt{sun,max}}_{\tau h} - \mathbb{E}\left[\mathtt{P}^{\mathtt{sun}}_{\tau h}\right]\right)}$$

$\dot{X}^{(k)}_{\tau+1,m} = \dot{X}^{(k)}_{\tau,m} - \frac{1}{\mathcal{E}_{\max}}\dot{u}^{(k)}_{\tau,m}h$.

end for

end for

Backtracking line search to get $u^{(k+1)}$, $X^{(k+1)}$ and $\hat{Y}^{(k+1)}$.

$k \leftarrow k+1$.

end while

return $(u^\star, X^\star, \hat{Y}^\star) := (u^{(k)}, X^{(k)}, \hat{Y}^{(k)})$.

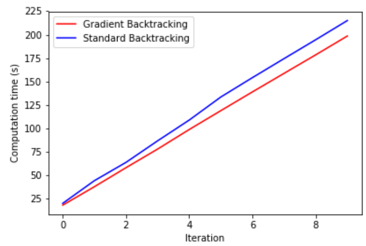
4.4 Analysis of the numerical performance

Figure 6b shows that initially both backtracking line-search methods return full Newton steps, which suggests that our initial guess $u^{(0)}$ is located in the quadratic convergence area for the Newton method. However, from iteration 3, the standard Backtracking line-search takes ridiculously small step lengths, as $\sigma = \beta^{13} = 10^{-13}$. Hence, the method fails to converge, as is suggest by Figures 6c and 6d which show that the $\mathbb{H}^{\infty,\infty}$ and $\mathbb{H}^{2,2}$ norm of the gradient is stationary from iteration 3. This shows that the Standard Backtracking line-search is not adapted to our setting.

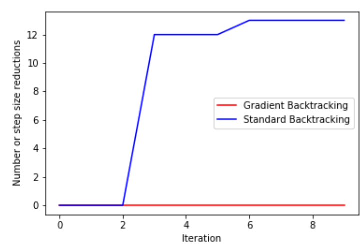
On the other hand, Figures 6c and 6d suggest that the Newton method with Gradient Backtracking line-search converges, as the norm of the gradient (as considered in $\mathbb{H}^{2,2}$ and $\mathbb{H}^{\infty,\infty}$) decreases along iterations. Hence, this shows that the Gradient Backtracking line-search procedure is better suited for our application than the (naive) standard backtracking line-search method.

Moreover, we would expect theoretically that $\left(||\nabla\bar{\mathcal{J}}(\bar{u}^{(k)})||_{\mathbb{H}^{\infty,\infty}}\right)_{k\in\mathbb{N}}$ decreases quadratically fast. However, this is not the case, see Figure 6c: after the third iteration, the convergence is not quadratic anymore, although the algorithm takes full steps $\sigma = 1$ at all iterations, see Figure 6b. We believe this comes from the regression steps, which introduce some residual errors in the computations of $\bar{Y}^{(k)}$, $P^{(k)}$ and $\Pi^{(k)}$.

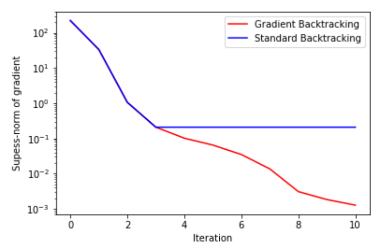
The cost decreases quickly for the first iterations for the Newton method combined with both Backtracking line-search procedures, see Figure 6e. From iteration 3, the Newton method with standard backtracking line-search does not make any progress (the step size is ridiculously small), while it is no longer decreasing for the Gradient Backtracking line-search, see Figure 6f. This is not surprising as our result states the convergence of the norm of the gradient (in $\mathbb{H}^{\infty,\infty}$) to 0, and not that the cost is decreasing along iterations. Besides, the number of samples considered $M = 10000$ may explain this non-monotonic behavior of the cost along iterations of the Newton method gradient backtracking line-search.



(a) Running computation time (in seconds)

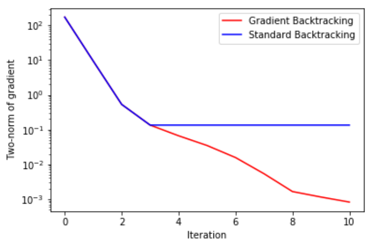


(b) Number of step size reductions by the backtracking line-search

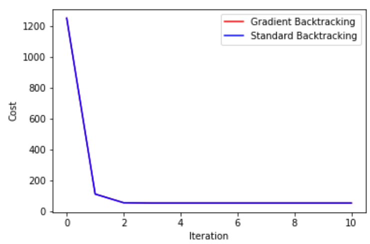


(c) Numerical approximation of $||\nabla\mathcal{J}(u^{(k)})||_{\mathbb{H}^{\infty,\infty}}$ along iterations

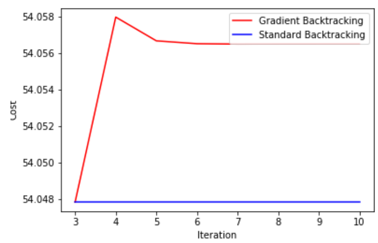
//



(d) Numerical approximation of $||\nabla\mathcal{J}(u^{(k)})||_{\mathbb{H}^{2,2}}$ along iterations



(e) Cost along iterations



(f) Cost along iterations from iteration 3

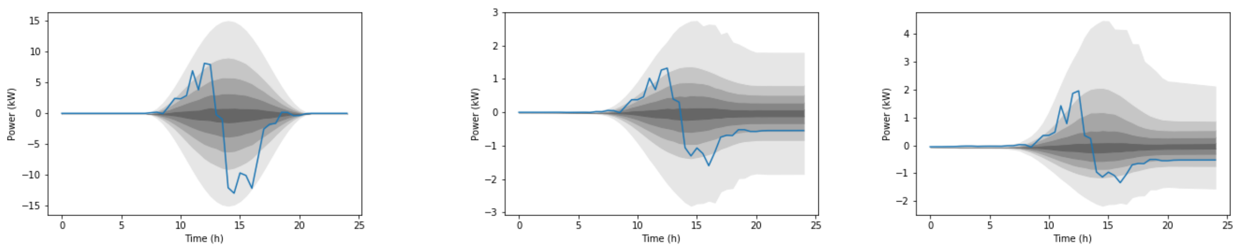
Figure 6: Comparison of performances of Newton method with the two Backtracking line search methods

4.5 Over-fitting, under-fitting, and automatic tuning of regression parameters by cross-validation

Over-fitting or under-fitting may occur in the regression steps and could be dealt with by introducing validation steps (by splitting the sample into a training set and a validation set) allowing automatic parameter tuning at each step. As examples, one could consider introducing regularization or use other functional spaces. Fine-tuning of the regression parameters by cross-validation would allow lower generalization errors, but at the expenses of a higher computational cost. Moreover, there are many heterogeneous regression steps, each of them requiring a particular treatment: good regression parameters have no reason to be the same for $Y^{(k)}$, $P^{(k)}$ and $\Pi^{(k)}$ and may change according to the time step considered and along the iterations. To give one example, one could consider regressing against the state variable $\bar{X}^{(k)}$ instead of $(\bar{X}^{(k)}, \mathtt{X}^{\mathrm{sun}})$ for the regression steps of processes at time steps after the sun set, because the solar irradiance does not play any role in the problem as it is canceled out by $\mathtt{P}^{\mathrm{sun,max}}$ which is null after sunset. We do not focus on these over-fitting or under-fitting issues in this paper, and simply consider a general regression procedure with a functional space spanned by affine-linear functions of the extended state $(\bar{X}^{(k)}, \mathtt{X}^{\mathrm{sun}})$. Incorporating cross-validation steps for automatic parameter tuning (functional space, regularization) in our algorithm is an interesting perspective of our work.

4.6 Analysis of the results from the application point of view

For completeness, we give some brief comments on the numerical results, from the point of view of the application. Figure 7a represents the evolution of the power imbalance without the control mechanism. Figure 7b represents the evolution of the power imbalance with a quadratic (symmetric) loss function $\mathcal{L}^{\mathrm{quad}} : x \mapsto \lambda x^2$, which gives a linear quadratic structure to the coordination sub-problem, which makes it particularly easy to solve. Figure 7c represents the evolution of the power imbalance with the asymmetric loss function $\mathcal{L}$. By comparing the uncontrolled case with the two controlled case, we can see that the power imbalance range has been significantly reduced (noticing the change of scale of the graphs). This shows efficiency of the proposed control mechanism to reduce the power imbalance. The asymmetric loss function $\mathcal{L}$ tends to penalize more heavily negative imbalance than positive imbalance, which creates an asymmetry in the probability distribution of the power imbalance, see Figure 7c, to be compared with the symmetry of the probability distribution of the power imbalance in the case of a symmetric loss function $\mathcal{L}^{\mathrm{quad}}$, see Figure 7b. For all plots, we add the realization of the power imbalance for one scenario of solar irradiance (the same scenario as the one plotted in Figure 3b). The power imbalance is null at night in the uncontrolled case, as there is no solar production. There is no power imbalance in the controlled case before sunrise in the controlled case. However, for some scenarios, there is a non-zero power imbalance in the controlled case, which arises from the fact that the state of charge of the batteries at sunset might be far away from its target terminal value $1/2$. Hence, the batteries are used after sunset in this case in order to take into account the target terminal value of the state of charge.



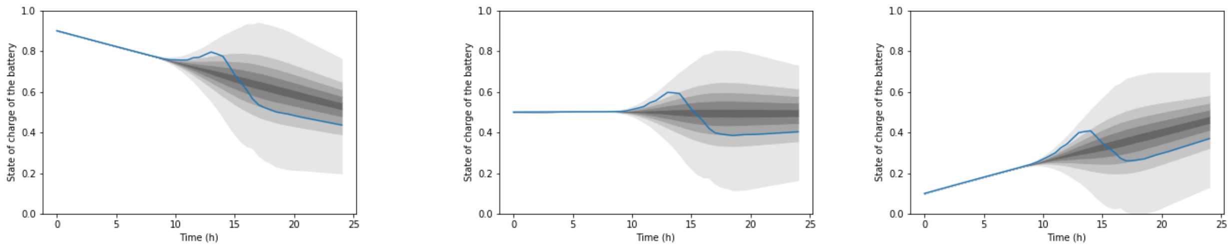
(a) Uncontrolled power imbalance $\mathtt{P}^{\mathrm{sun}} - \mathbb{E}[\mathtt{P}^{\mathrm{sun}}]$

(b) Controlled power imbalance $\bar{u} + \mathtt{P}^{\mathrm{sun}} - \mathbb{E}[\mathtt{P}^{\mathrm{sun}}]$ - symmetric loss $\mathcal{L}^{\mathrm{quad}}$

(c) Controlled power imbalance $\bar{u} + \mathtt{P}^{\mathrm{sun}} - \mathbb{E}[\mathtt{P}^{\mathrm{sun}}]$ - asymmetric loss $\mathcal{L}$

Figure 7: Power imbalance

One may wonder if the control mechanism proposed respects the constraint on the states of charge of the batteries, which must lie in $[0, 1]$, even for an initial state of charge close to 0 or 1. The quantiles of the state of charge of one of the batteries participating to the control mechanism is plotted in Figure 8, depending on the initial value of the state of charge. One can in particular see that, even for initial value of the state of charge close to 0 or 1, the state of charge of the battery remains between these two values with high probability. For all plots, we add the realization of the state of charge of the battery considered for one scenario of solar irradiance (the same scenario as the one plotted in Figure 3b).



(a) $x_0 = 0.9$ (b) $x_0 = 0.5$ (c) $x_0 = 0.1$

Figure 8: Percentiles of the state of charge of a battery participating to the control mechanism - Sensitivity to initial condition

5 Proofs

5.1 Proof of Lemma 3.5

Proof. For $M > 0$ we introduce:

$$\begin{cases} \overline{f}(t,p) := 2\|A\|_{\mathbb{H}^{\infty,\infty}}|p| + C_t, \\ f^{(M)}(t,p) := C_M\left(2A_t p + B_t p^2 + C_t\right), \\ \underline{f}(t,p) := -2\|A\|_{\mathbb{H}^{\infty,\infty}}|p| - \|B\|_{\mathbb{H}^{\infty,\infty}}p^2, \end{cases}$$

where the clipping operator C_M is defined by $C_M : x \in \mathbb{R} \mapsto \min(\max(x,-M),M) = \max(\min(x,M),-M) \in [-M,M]$. Notice that for any $p \in \mathbb{R}$ and any $M > 0$, by our assumptions, we have $\mathrm{d}\mathbb{P}\otimes\mathrm{d}t$-a.e. on $[0,T]$:

$$\begin{aligned} \underline{f}(t,p) &= \min(\underline{f}(t,p), M) && (M > 0, \underline{f}(t,p) \le 0) \\ &\le \min(2A_t p + B_t p^2 + C_t, M) && \text{(Monotony of min)} \\ &\le C_M\left(2A_t p + B_t p^2 + C_t\right) && (\forall x \in \mathbb{R}, \min(x,M) \le C_M(x)) \\ &= f^{(M)}(t,p) \\ &\le \max\left(-M, 2A_t p + B_t p^2 + C_t\right) && (\forall x \in \mathbb{R}, C_M(x) \le \max(-M,x)) \\ &\le \max\left(-M, 2\|A\|_{\mathbb{H}^{\infty,\infty}}|p| + C_t\right) && (\text{Monotony of max}, B \le 0) \\ &= \overline{f}(t,p) && (-M \le 0 \le C_t \le 2\|A\|_{\mathbb{H}^{\infty,\infty}}|p| + C_t). \end{aligned}$$

Consider the BSDEs:

$$\underline{P}_t := \int_t^T \underline{f}(s,\underline{P}_s)\mathrm{d}s, \tag{5.1}$$

$$P_t^{(M)} := \mathbb{E}_t\left[\Gamma + \int_t^T f^{(M)}(s,P_s^{(M)})\mathrm{d}s\right], \tag{5.2}$$

$$\overline{P}_t := \mathbb{E}_t\left[\Gamma + \int_t^T \overline{f}(s,\overline{P}_s)\mathrm{d}s\right]. \tag{5.3}$$

First, the BSDE (5.1) is actually an ODE with locally Lipschitz-continuous driver, and therefore, by Cauchy-Lipschitz theorem has a unique solution on some maximal interval $(\tau, T]$. The null function is clearly the unique solution of (5.1) and $\tau = -\infty$. This yields the well-posedness of $\underline{P}$ and the explicit expression $\underline{P}_t = 0, \forall t \in [0,T]$.

Second, we notice that (5.2) and (5.3) are BSDE with Lipschitz drivers, so that they are well-defined on $[0,T]$, according to [EPQ97, Theorem 5.1] or [ØZ12, Theorem 3.1, p. 705]. Notice that $\mathrm{d}\mathbb{P}\otimes\mathrm{d}t$-a.e., for any $p \in \mathbb{R}$,

$$0 \le \Gamma \quad ; \quad \underline{f}(t,p) \le f^{(M)}(t,p) \le \overline{f}(t,p).$$

By comparison theorem for BSDEs (see [ØZ12, Theorem 3.4, p. 710].), we get for all $M > 0$, $\mathrm{d}\mathbb{P} \otimes \mathrm{d}t$-a.e. on $[0, T]$:

$$0 = \underline{P}_t \leq P_t^{(M)} \leq \overline{P}_t.$$

Besides, $\overline{P}$ satisfies for some martingale $M \in \mathcal{M}_0$:

$$\begin{cases} -\mathrm{d}\overline{P}_t = \left(2\|A\|_{\mathbb{H}^{\infty,\infty}} \overline{P}_t + C_t\right) \mathrm{d}t - \mathrm{d}M_t, \\ \overline{P}_T = \Gamma. \end{cases} \tag{5.4}$$

Using the Integration by Parts formula in [Pro03, Corollary 2, p. 68] to $t \mapsto \overline{P}_t \exp\left(-2\|A\|_{\mathbb{H}^{\infty,\infty}}(T-t)\right)$ yields the explicit expression (3.5) and the estimation on $\|P\|_{\mathbb{H}^{\infty,\infty}}$. Now notice that for $M > M_0 := 2\|A\|_{\mathbb{H}^{\infty,\infty}}\|\overline{P}\|_{\mathbb{H}^{\infty,\infty}} + \|B\|_{\mathbb{H}^{\infty,\infty}}\|\overline{P}\|^2_{\mathbb{H}^{\infty,\infty}} + \|C\|_{\mathbb{H}^{\infty,\infty}}$, we have:

$$\forall p \in [0, \|\overline{P}\|_{\mathbb{H}^{\infty,\infty}}], \mathrm{d}\mathbb{P} \otimes \mathrm{d}t - a.e., \quad f^{(M)}(t, p) = 2A_t p + B_t p^2 + C_t. \tag{5.5}$$

Since $0 \leq P_t^{(M)} \leq \overline{P}_t$, $\mathrm{d}\mathbb{P} \otimes \mathrm{d}t$-a.e., we get, for $M > M_0$:

$$P_t^{(M)} = \mathbb{E}_t\left[\Gamma + \int_t^T \left(2A_s P_s^{(M)} + B_s \left(P_s^{(M)}\right)^2 + C_s\right) \mathrm{d}s\right].$$

This shows existence of solution of the Riccati BSDE (3.4).

Let us now turn to uniqueness. Consider two solutions P and Q of the Riccati BSDE (3.4). By application of the comparison principle for BSDEs, we obtain with similar arguments as before:

$$\mathrm{d}\mathbb{P} \otimes \mathrm{d}t - a.e., \quad 0 \leq P_t \leq \overline{P}_t, \qquad 0 \leq Q_t \leq \overline{P}_t,$$

and therefore, for $M > M_0$, (5.5) shows that P and Q are both solutions of the BSDE (5.2), which has Lipschitz driver, hence a unique solution, according to [EPQ97, Theorem 5.1] or [ØZ12, Theorem 3.1, p. 705]. In particular, $P = Q$, which yields uniqueness of solutions of (3.4). □

5.2 Proof of Lemma 3.6

Proof. Define Π by (3.6) and define as well:

$$R_t := \exp\left(\int_0^t (P_s B_s + A_s)\mathrm{d}s\right)\Pi_t + \int_0^t (a_s P_s + b_s) \exp\left(\int_0^s (P_r B_r + A_r)\mathrm{d}r\right)\mathrm{d}s = \mathbb{E}_t\left[R_T\right],$$

$$S_t := \exp\left(-\int_0^t (P_s B_s + A_s)\mathrm{d}s\right).$$

Then $R \in \mathbb{H}^{\infty,2}$ and R is an $(\mathcal{F}_t)$-adapted càdlàg martingale. Then, apply integration by parts formula [Pro03, Corollary 2, p. 68] to the product SR, using the fact that S is continuous with finite variations. After reorganizing terms and using that R has countable jumps, we find:

$$\begin{cases} -\mathrm{d}\Pi_t = ((P_t B_t + A_t)\Pi_t + a_t P_t + b_t)\,\mathrm{d}t + S_t \mathrm{d}R_t, \\ \Pi_T = \eta, \end{cases}$$

and $\int_{0^+}^t S_s \mathrm{d}R_s$ is a càdlàg martingale in $\mathbb{H}^{\infty,2}$, see [Pro03, Theorem 20 p.63, Corollary 3 p.73, Theorem 29 p.75]. We then find that Π solves (3.7) and it is the unique solution of this BSDE, as the BSDE has a Lipschitz driver, according to [EPQ97, Theorem 5.1] or [ØZ12, Theorem 3.1, p. 705]. □

5.3 Proof of Theorem 3.7

Proof. 1. FBSDEs have been studied in the general case in [Zha17] and [MY99]. In the affine-linear case, the result is a consequence of [Yon06] and the Martingale Representation Theorem if the filtration is Brownian. However, for more general filtrations, the result is outside the scope of [Yon06] and we provide a proof for this case, restricting ourselves to one-dimensional control and state processes.

Consider the following auxiliary linear-quadratic stochastic control problem:

$$\left.\begin{aligned} \mathcal{J}^{quad}(u) &:= \mathbb{E}\left[\int_0^T \left(\tfrac{1}{2}u_t^2 + \tfrac{1}{2}C_t X_t^2 + b_t X_t\right)\mathrm{d}t + \tfrac{1}{2}\Gamma X_T^2 + \eta X_T\right] \\ \text{s.t. } X_t &= x + \int_0^t (A_s X_s + \sqrt{-B_s}u_s + a_s)\mathrm{d}s. \end{aligned}\right\} \longrightarrow \min_{u \in \mathbb{H}^{2,2}_{\mathcal{P}}}. \tag{5.6}$$

Our assumptions show that $\mathcal{J}^{quad}$ satisfies the hypothesis of Proposition 2.3 and therefore, it has a unique minimizer $u^* \in \mathbb{H}^{2,2}$.

The function $\mathcal{J}^{quad}$ also satisfies the assumptions of first order sufficient optimality conditions (see second point of Theorem 2.6), so that, if we define $(X^*, Y^*) \in \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ by:

$$\begin{cases} X_t^* = x + \int_0^t (A_s X_s^* + \sqrt{-B_s}u_s^* + a_s)\mathrm{d}s, \\ Y_t^* = \mathbb{E}_t\left[\Gamma X_T^* + \eta + \int_t^T (C_s X_s^* + A_s Y_s^* + b_s)\,\mathrm{d}s\right], \end{cases}$$

we have

$$u_t^* + \sqrt{-B_t}Y_{t-}^* = 0.$$

By eliminating u^* using the last equation and using the fact that the Lebesgue integral is left unchanged by changing the value of the integrand on a countable set, this shows that (X^*, Y^*) satisfies the FBSDE:

$$\begin{cases} X_t = x + \int_0^t (A_s X_s + B_s Y_s + a_s)\,\mathrm{d}s, \\ Y_t = \mathbb{E}_t\left[\Gamma X_T + \eta + \int_t^T (C_s X_s + A_s Y_s + b_s)\,\mathrm{d}s\right], \end{cases}$$

Let us turn to uniqueness. Consider two solutions (X^1, Y^1) and (X^2, Y^2) of the above FBSDE. Then (u^1, X^1, Y^1) and (u^2, X^2, Y^2) with $u_t^i = -\sqrt{-B}Y_{t-}^i$ for $i = 1, 2$ are both solutions of:

$$\begin{cases} X_t = x + \int_0^t (A_s X_s + \sqrt{-B_s}u_s + a_s)\mathrm{d}s, \\ Y_t = \mathbb{E}_t\left[\Gamma X_T + \eta + \int_t^T (C_s X_s + A_s Y_s + b_s)\,\mathrm{d}s\right], \\ u_t + \sqrt{-B_t}Y_{t-} = 0. \end{cases}$$

Hence u^1 and u^2 are both solutions of the first order conditions characterizing minimizers of $\mathcal{J}^{quad}$ by Theorem 2.6 and by Proposition (2.3), $u^1 = u^2$. This shows $X^1 = X^2$, then $Y^1 = Y^2$, hence the existence and uniqueness of a solution of the FBSDE.

2. By our previous results, P and Π are well defined in $\mathbb{H}^{\infty,\infty}$ and $\mathbb{H}^{\infty,2}$ respectively. Then X given in (3.8) solves an affine-linear ODE and the assumption on the coefficients show that it is well-defined (non-explosion) and given by:

$$\forall t \in [0, T], \quad X_t = x \exp\left(\int_0^t (A_s + B_s P_s)\mathrm{d}s\right) + \int_0^t (B_s \Pi_s + a_s) \exp\left(\int_s^t (A_r + B_r P_r)\mathrm{d}r\right)\mathrm{d}s.$$

The estimates on X and $Y = PX + \Pi$ in the spaces $\mathbb{H}^{\infty,2}$ come directly from that and Lemma 3.6. Let us now prove that (X, Y) is a solution of the affine-linear FBSDE, which will conclude the proof, by uniqueness of the solution of such FBSDE, by the previous point. Using $Y = PX + \Pi$, it is easy to show that X satisfies:

$$X_t = x + \int_0^t (A_s X_s + B_s Y_s + a_s)\,\mathrm{d}s.$$

It remains to show that Y satisfies the BSDE:

$$Y_t = \mathbb{E}_t\left[\Gamma X_T + \eta + \int_t^T (C_s X_s + A_s Y_s + b_s)\,\mathrm{d}s\right].$$

To do that, use the fact that $Y = PX + \Pi$ by definition so that $Y_T = P_T X_T + \Pi_T = \Gamma X_T + \eta$ so that the terminal condition is verified. Introduce $M^{(P)}$ in $\mathcal{M}_0^2 \cap \mathbb{H}^{\infty,\infty}$ and $M^{(\Pi)}$ in $\mathcal{M}_0^2$ such that:

$$\begin{cases} -\mathrm{d}P_t = (2A_t P_T + B_t P_t^2 + C_T)\mathrm{d}t - \mathrm{d}M_t^{(P)}, \\ P_T = \Gamma, \end{cases}$$

and:

$$\begin{cases} -\mathrm{d}\Pi_t = ((P_t B_t + A_t)\Pi_t + a_t P_t + b_t)\,\mathrm{d}t - \mathrm{d}M_t^{(\Pi)}, \\ P_T = \eta. \end{cases}$$

Then, use the integration by parts formula in Protter [Pro03, Corollary 2, p. 68], combined with the fact that X is continuous with finite variations. We get:

$$\begin{aligned} -\mathrm{d}Y_t &= -(\mathrm{d}P_t)X_t - P_t(\mathrm{d}X_t) - \mathrm{d}\Pi_t \\ &= (2A_t P_T + B_t P_t^2 + C_T)X_t\mathrm{d}t - X_t\mathrm{d}M_t^{(P)} - P_t(A_t X_t + B_t Y_t + a_t)\mathrm{d}t + ((P_t B_t + A_t)\Pi_t + a_t P_t + b_t)\,\mathrm{d}t - \mathrm{d}M_t^{(\Pi)} \\ &= (A_t(P_t X_t + \Pi_t) + B_t P_t(P_t X_t + \Pi_t - Y_t) + C_t X_t + b_t)\,\mathrm{d}t - X_t\mathrm{d}M_t^{(P)} - \mathrm{d}M_t^{(\Pi)} \\ &= (A_t Y_t + C_t X_t + b_t)\mathrm{d}t - X_t\mathrm{d}M_t^{(P)} - \mathrm{d}M_t^{(\Pi)}. \end{aligned}$$

Using the fact that the last two terms are the increments of true martingales in $\mathcal{M}_0^2$ (as $X \in \mathbb{H}^{\infty,2}$ and $M^{(P)} \in \mathcal{M}_0^2 \cap \mathbb{H}^{\infty,\infty}$), this concludes the proof.

□

5.4 Proof of Corollary 3.8

Proof. It just remains to prove the estimates:

$$\begin{aligned} &\|P^u\|_{\mathbb{H}^{\infty,\infty}} \le C, \\ &\|\Pi^{u,w}\|_{\mathbb{H}^{\infty,2}} \le C\|w\|_{\mathbb{H}^{2,2}}, \\ &\|(\nabla^2\mathcal{J}(u))^{-1}(w)\|_{\mathbb{H}^{2,2}} \le C\|w\|_{\mathbb{H}^{2,2}}, \end{aligned}$$

for some constant C independent of u and w. Using the assumption of bounded second-order derivative and the fact that l is strongly convex in u (implying that l''_{uu} is uniformly bounded from below by a non-negative constant), we get for a constant C independent from u and w:

$$\begin{aligned} &\|A^u\|_{\mathbb{H}^{\infty,\infty}} + \|B^u\|_{\mathbb{H}^{\infty,\infty}} + \|C^u\|_{\mathbb{H}^{\infty,\infty}} + \|\Gamma^u\|_{\mathbb{L}_\Omega^\infty} \le C, \\ &\|a^{u,w}\|_{\mathbb{H}^{\infty,\infty}} + \|b^{u,w}\|_{\mathbb{H}^{\infty,\infty}} \le C\|w\|_{\mathbb{H}^{\infty,\infty}}. \end{aligned}$$

We get the bounds on $\|P^u\|_{\mathbb{H}^{\infty,\infty}}$ and $\|\Pi^{u,w}\|_{\mathbb{H}^{\infty,2}}$ by using the estimates on $\|P\|_{\mathbb{H}^{\infty,\infty}}$ and $\|\Pi\|_{\mathbb{H}^{\infty,2}}$ obtained in Lemmas **??** and 3.6, with $\eta = 0$ and $x = 0$. The bound on $\|(\nabla^2\mathcal{J}(u))^{-1}(w)\|_{\mathbb{H}^{2,2}}$ is then obtained using the strong convexity of l with respect to u and using the expression (3.2). □

6 Conclusion

In this paper, we extend the Newton method to the framework of stochastic control problems, which amounts to consider successive linearizations of the optimality system found by using the stochastic Pontryagin principle. We show that the computation of the Newton step amounts to solve a linear FBSDE with random coefficients (with some sign conditions), which in turn reduces to solving a Riccati BSDE and a linear BSDE. Then, an appropriate restriction of the space of processes is considered to obtained desirable regularity for the control problem, allowing to prove convergence results for the Newton method. To obtain a global convergence, an appropriate line-search which fits our infinite-dimensional setting is proposed. Global convergence of the Newton method combined with this adapted line-search is then proved theoretically. The Newton method is implemented on a problem of joint control of many identical batteries in order to maintain power balance on a given network. In particular, regression techniques are used in order to compute the solutions of the linear and non-linear BSDEs arising when computing the Newton step. So far, we have considered low dimensional problems: the regression steps are performed in $\mathbb{R}^2$ and the control and state variables are one-dimensional. In higher dimension, we expect a curse of dimensionality when solving the BSDEs using regression method. Other methods like Deep-learning could help solve the issue. However, the Newton method is iterative and training a network at each iteration seems computationally expensive. We also expect the Newton method to be applicable to other settings, like controlled diffusions for instance, which would change the form of the Riccati BSDEs arising when solving the successive linearizations of the FBSDE characterizing the optimal control. Other interesting perspectives to our work include designing appropriate stopping criteria in the Newton method implemented using regression techniques, or incorporate automatic tuning procedures for the hyper-parameters in regression steps in the algorithm.

References

[ANR17] I. F. Alves, C. Neves, and P. Rosário. "A general estimator for the right endpoint with an application to supercentenarian women's records". In: *Extremes* 20.1 (2017), pp. 199–237.

[Ang+19] A. Angiuli, C. Graves, H. Li, J.-F. Chassagneux, F. Delarue, and R. Carmona. "Cemracs 2017: numerical probabilistic approach to MFG". In: *ESAIM: Proceedings and Surveys* 65 (2019), pp. 84–113.

[Bad+18] J. Badosa, E. Gobet, M. Grangereau, and D. Kim. "Day-ahead probabilistic forecast of solar irradiance: a Stochastic Differential Equation approach". In: *Renewable Energy: Forecasting and Risk Management*. Ed. by P. Drobinski, M. Mougeot, D. Picard, R. Plougonven, and P. Tankov. Springer Proceedings in Mathematics & Statistics, 2018. Chap. 4, pp. 73–93.

[BZ+08] C. Bender, J. Zhang, et al. "Time discretization and Markovian iteration for coupled FBSDEs". In: *The Annals of Applied Probability* 18.1 (2008), pp. 143–177.

[BSS05] B. Bibby, I. Skovgaard, and M. Sorensen. "Diffusion-type models with given marginal distribution and autocorrelation function". In: *Bernoulli* 11.2 (2005), pp. 191–220.

[Bis76] J.-M. Bismut. "Linear quadratic optimal stochastic control with random coefficients". In: *SIAM Journal on Control and Optimization* 14.3 (1976), pp. 419–444.

[BZ03] J. F. Bonnans and H. Zidani. "Consistency of generalized finite difference schemes for the stochastic HJB equation". In: *SIAM Journal on Numerical Analysis* 41.3 (2003), pp. 1008–1021.

[BV04] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[Bre10] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Springer Science+Business Media, 2010.

[CL82] F. Chernousko and A. Lyubushin. "Method of successive approximations for solution of optimal control problems". In: *Optimal Control Applications and Methods* 3.2 (1982), pp. 101–114.

[EPQ97] N. El Karoui, S. Peng, and M. Quenez. "Backward stochastic differential equations in finance". In: *Math. Finance* 7.1 (1997), pp. 1–71.

[GT16a] E. Gobet and P. Turkedjiev. "Linear regression MDP scheme for discrete backward stochastic differential equations under general conditions". In: *Math. Comp.* 85.299 (2016), pp. 1359–1391.

[GT16b] E. Gobet and P. Turkedjiev. "Linear regression MDP scheme for discrete backward stochastic differential equations under general conditions". In: *Mathematics of Computation* 85.299 (2016), pp. 1359–1391.

[GŠ09] I. Gyöngy and D. Šiška. "On finite-difference approximations for normalized Bellman equations". In: *Applied Mathematics and Optimization* 60.3 (2009), p. 297.

[HJW18] J. Han, A. Jentzen, and E Weinan. "Solving high-dimensional partial differential equations using deep learning". In: *Proceedings of the National Academy of Sciences* 115.34 (2018), pp. 8505–8510.

[HL20] J. Han and J. Long. "Convergence of the deep BSDE method for coupled FBSDEs". In: *Probability, Uncertainty and Quantitative Risk* 5.1 (2020), pp. 1–33.

[How60] R. A. Howard. "Dynamic Programming and Markov Processes." In: (1960).

[HP95] Y. Hu and S. Peng. "Solution of forward-backward stochastic differential equations". In: *Probability Theory and Related Fields* 103.2 (1995), pp. 273–283.

[Ji+20] S. Ji, S. Peng, Y. Peng, and X. Zhang. "Three algorithms for solving high-dimensional fully-coupled FBSDEs through deep learning". In: *IEEE Intelligent Systems* (2020).

[Kan48] L. V. Kantorovich. "Functional analysis and applied mathematics". In: *Uspekhi Matematicheskikh Nauk* 3.6 (1948), pp. 89–185.

[KŠS20a] B. Kerimkulov, D. Šiška, and Ł. Szpruch. "A modified MSA for stochastic control problems". In: *arXiv preprint arXiv:2007.05209* (2020).

[KŠS20b] B. Kerimkulov, D. Šiška, and L. Szpruch. "Exponential Convergence and Stability of Howard's Policy Improvement Algorithm for Controlled Diffusions". In: *SIAM Journal on Control and Optimization* 58.3 (2020), pp. 1314–1340.

[Kry08] N. V. Krylov. *Controlled Diffusion Processes*. Vol. 14. Springer Science & Business Media, 2008.

[MY99] J. Ma and J. Yong. *Forward-Backward Stochastic Differential Equations and their Applications*. A course on stochastic processes. Lecture Notes in Mathematics, 1702, Springer-Verlag, 1999.

[MPY94] J. Ma, P. Protter, and J. Yong. "Solving forward-backward stochastic differential equations explicitly—a four step scheme". In: *Probability theory and related fields* 98.3 (1994), pp. 339–359.

[NN94] Y. Nesterov and A. Nemirovskii. *Interior-point polynomial algorithms in convex programming*. SIAM, 1994.

[NW06] J. Nocedal and S. Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

[ØZ12] B. Øksendal and T. Zhang. "Backward stochastic differential equations with respect to general filtrations and applications to insider finance". In: *Communications on Stochastic Analysis* 6.4 (2012), p. 13.

[PT99] E. Pardoux and S. Tang. "Forward-backward stochastic differential equations and quasilinear parabolic PDEs". In: *Probability Theory and Related Fields* 114.2 (1999), pp. 123–150. ISSN: 1432-2064. DOI: 10.1007/s004409970001. URL: https://doi.org/10.1007/s004409970001.

[PW99] S. Peng and Z. Wu. "Fully coupled forward-backward stochastic differential equations and applications to optimal control". In: *SIAM Journal on Control and Optimization* 37.3 (1999), pp. 825–843.

[Pha09] H. Pham. *Continuous-time stochastic control and optimization with financial applications*. Vol. 61. Springer Science & Business Media, 2009.

[Pro03] P. E. Protter. *Stochastic Integration and Differential Equations*. second. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003. ISBN: 978-3-662-10061-5. DOI: 10.1007/978-3-662-10061-5_6. URL: https://doi.org/10.1007/978-3-662-10061-5_6.

[Yon06] J. Yong. "Linear forward-backward stochastic differential equations with random coefficients". In: *Probability theory and related fields* 135.1 (2006), pp. 53–83.

[Yon97] J. Yong. "Finding adapted solutions of forward–backward stochastic differential equations: Method of continuation". In: *Probability Theory and Related Fields* 107.4 (1997), pp. 537–572.

[Yon99] J. Yong. "Linear Forward—Backward Stochastic Differential Equations". In: *Applied Mathematics and Optimization* 39.1 (1999), pp. 93–119.

[Zha17] J. Zhang. “Backward stochastic differential equations”. In: *Backward Stochastic Differential Equations*. Springer, 2017, pp. 79–99.