Large deviations of reaction fluxes

Robert I. A. Patterson and D. R. Michiel Renger

February 9, 2018

Abstract

We study a system of interacting particles that randomly react to form new particles. The reaction flux is the rescaled number of reactions that take place in a time interval. We prove a dynamic large-deviation principle for the reaction fluxes under general assumptions that include mass-action kinetics. This result immediately implies the dynamic large deviations for the empirical concentration.

1 Introduction

Since Boltzmann's microscopic interpretation of entropy it is clear that thermodynamics is inherently related to large deviations. Onsager, in his papers [Ons31a, Ons31b] was able to extend this principle to the non-static regime - at least for reversible systems and close to equilibrium. More recently, it was shown that reversible stochastic particle systems induce a thermodynamically consistent gradient flow through their dynamical large deviations, see [ADPZ11, MPR14], and in particular [MPPR15] for an application to chemical reactions. This characterises dynamic behavior even far from equilibrium. However, a thermodynamically consistent representation of non-reversible particle systems remains one of the main open problems of non-equilibrium thermodynamics.

The difficulty in understanding irreversible particle systems lies in the occurrence of non-trivial fluxes, which is why flux large deviations are a commonly studied object, see [DDR04, BL10, BL12, Der07, BDSG$^+$05, BDSG$^+$06, BMN09] for examples covering Brownian motionss, random walkers and exclusion processes. In this work we apply the flux approach to reacting particles on a discrete state space.

Reacting particle system. We study a general network of reactions,

$$\sum_{y\in\mathcal{Y}} \alpha_y^{(r)} y \xrightarrow{\bar{k}^{(r)}} \sum_{y\in\mathcal{Y}} \beta_y^{(r)} y, \qquad r\in\mathcal{R}, \tag{1.1}$$

where $\mathcal{Y}$ is finite set of species, and $\mathcal{R}$ is a finite set of reactions, and $\bar{k}^{(r)}$ are the corresponding reaction rates. A typical choice of reaction rates is $\bar{k}^{(r)}(c) = \mathit{const} \times \prod_{y\in\mathcal{Y}} c_y^{\alpha_y}$; this is called *mass-action kinetics*, but we will consider a much more general class of rates.

For example, one could have the reactions

$$\mathsf{2H_2 + O_2} \xrightarrow{\bar{k}^{(\mathrm{fw})}} \mathsf{2H_2O}, \qquad \text{and} \qquad \mathsf{2H_2O} \xrightarrow{\bar{k}^{(\mathrm{bw})}} \mathsf{2H_2 + O_2}.$$

In this case the set of species is $\mathcal{Y} = \{\mathsf{H_2}, \mathsf{O_2}, \mathsf{H_2O}\}$, the set of reactions is $\mathcal{R} = \{\mathrm{fw}, \mathrm{bw}\}$, and $\bar{k}^{(\mathrm{fw})}, \bar{k}^{(\mathrm{bw})}$ are the reaction rates that depend on the concentration of the species in $\mathcal{Y}$. Furthermore, the species needed for the reactions can be grouped in the vectors $\alpha^{(\mathrm{fw})}, \alpha^{(\mathrm{bw})} = (2,1,0), (0,0,2)$, and similarly for the species resulting from the reactions $\beta^{(\mathrm{fw})}, \beta^{(\mathrm{bw})} = (0,0,2), (2,1,0)$. These vectors are called *complexes* or *stoichiometric* coefficients, the latter being Greek for "element counting".

The reaction networks described above are commonly modelled by the following microscopic particle system, see the survey [AK11] and the references therein. If at some given time t there are $N(t)$ particles of types $Y_1(t), \ldots, Y_{N(t)}(t)$ in the system with fixed volume V, then the empirical measure (or concentration) is defined as $C^{(V)}(t) := V^{-1}\sum_{i=1}^{N(t)} \mathbb{1}_{Y_i(t)}$.

With jump rate $k^{(r,V)}(C^{(V)}(t))$, also called *propensity*, a reaction r occurs, causing the concentration to jump to the new state $C^{(V)}(t) + \frac{1}{V}\gamma^{(r)}$, where $\gamma^{(r)} = \beta^{(r)} - \alpha^{(r)} \in \mathbb{R}^{\mathcal{Y}}$ is the *effective stoichiometric vector* (sometimes called *state change vector*) for reaction r and these are collected in a matrix $\Gamma := [\gamma^{(1)} \ldots, \gamma^{(R)}]$, which therefore maps rescaled reaction counts to changes in concentration. Since the propensities $k^{(r,V)}$ depend on the particles through the empirical concentration only, $C^{(V)}(t)$ is a Markov jump process in $\mathbb{R}^{\mathcal{Y}}$. The volume V controls the order of the (changing) number of particles in the system.

A classic result [Kur70, Kur72] says that the empirical measure $C^{(V)}(t)$ converges as $V \to \infty$ to the solution of the *reaction rate equation* $\dot{c}(t) = \sum_{r\in\mathcal{R}} \bar{k}^{(r)}\big(c(t)\big)$, where $V^{-1}k^{(V,r)} \to \bar{k}^{(r)}$ (in a way that we specify later).

Reaction Fluxes. More information is included in the *integrated empirical reaction flux*,

$$W^{(V,r)}(t) = \tfrac{1}{V}\#\big\{\text{reactions } r \text{ that occurred in time } (0,t]\big\}.$$

The pair $\big(C^{(V)}(t), W^{(V)}(t)\big)$ is then a Markov process in $\mathbb{R}_+^{\mathcal{Y}} \times \mathbb{R}_+^{\mathcal{R}}$ with generator

$$(\mathcal{Q}^{(V)}f)(c,w) = \sum_{r\in\mathcal{R}} k^{(r,V)}(c)\big(f(c + \tfrac{1}{V}\gamma^{(r)}, w + \tfrac{1}{V}\mathbb{1}_r) - f(c,w)\big). \tag{1.2}$$

As in the Kurtz limit, this pair converges to the solution of the system of ODEs

$$\begin{cases} \dot{c}(t) = \Gamma\dot{w}(t) = \sum_{r\in\mathcal{R}} \dot{w}^{(r)}(t)\gamma^{(r)},\\ \dot{w}(t) = \bar{k}\big(c(t)\big). \end{cases}$$

The first equation is a continuity equation, which also holds almost surely for the microscopic pair $(C^{(V)}, W^{(V)})$, for finite V.

Large deviations. The dynamic large-deviation principle for the concentrations $C^{(V)}$ have been proven in [Fen94, Léo95, DK95, SW95, SW05, DEW91, LL15, DRW16] under various assumptions. Large deviations for the pair $(C^{(V)}, W^{(V)})$ of concentrations and fluxes is, as far as we are aware, a relatively untred area. Formal large-deviation calculations for the reaction fluxes are found in [BMN09], a rigorous proof for the independent case was given in [Ren17], and a semigroup-based rigorous proof for a more general class of reaction fluxes can be found in [Kra17], still excluding mass-action kinetics. In our main result, we prove a dynamical large-deviation principle for the process $(C^{(V)}, W^{(V)})$, under initial distribution $(\mu^{(V)}, \delta_0)$, where we shall assume that $\mu^{(V)}$ satisfies a large-deviation principle with some rate functional $\mathcal{I}_0$. The precise statement reads:

Theorem 1.1. *Let $\mu^{(V)}$ satisfy a large-deviation principle with rate function $\mathcal{I}_0$, and let Assumptions 2.3 on $\mu^{(V)}$ and Assumption 2.2 on $k, \bar{k}$ hold. Then the process $(C^{(V)}(t), W^{(V)}(t))_{t=0}^T$ satisfies a large-deviation principle in $\mathrm{BV}(0,T;\mathbb{R}^{\mathcal{Y}} \times \mathbb{R}_+^{\mathcal{R}})$, equipped with the hybrid topology, with good rate functional $\mathcal{I}_0\big(c(0)\big) + \mathcal{J}(c,w)$, where*

$$\mathcal{J}(c,w) := \begin{cases} \int_0^T \mathcal{S}\big(\dot{w}(t) \mid \bar{k}(c(t)\big)\,\mathrm{d}t, & (c,w) \in W^{1,1}\big(0,T;\mathbb{R}^{\mathcal{Y}} \times \mathbb{R}_+^{\mathcal{R}}\big), \text{ and } \dot{c} = \Gamma\dot{w},\\ \infty, & \text{otherwise,} \end{cases} \tag{1.3}$$

with relative entropy

$$\mathcal{S}(j \mid \hat{j}) := \begin{cases} \sum_{r\in\mathcal{R}} s(j^{(r)} \mid \hat{j}^{(r)}), & \text{if } j \ll \hat{j},\\ +\infty & \text{otherwise,} \end{cases} \qquad \text{and}$$

$$s(j^{(r)} \mid \hat{j}^{(r)}) := \begin{cases} j^{(r)} \log\big(\frac{j^{(r)}}{\hat{j}^{(r)}}\big) - j^{(r)} + \hat{j}^{(r)}, & j^{(r)} > 0,\\ \hat{j}, & j^{(r)} = 0, \end{cases}$$

where $j \ll \hat{j}$ means that for all $r \in \mathcal{R}$ one has $\hat{j}^{(r)} = 0 \implies j^{(r)} = 0$.

The precise set of assumptions will be stated in Section 2.2. We choose to work in the hybrid topology on the space of paths of bounded variation rather than the commonly used Skorohod topology since it is in some sense natural for jump processes, and the compactness criteria are very simple; we will introduce and comment on this space, topology and σ-algebra in more detail in Section 2.1.

As an immediate consequence of Theorem 1.1, we obtain the large deviations for the concentrations:

Corollary 1.2. *Let $\mu^{(V)}$ satisfy a large-deviation principle with rate function $\mathcal{I}_0$, and let Assumptions 2.3 on $\mu^{(V)}$ and Assumption 2.2 on $k, \bar{k}$ hold. Then the process $C^{(V)}$ satisfies a large-deviation principle in $\mathrm{BV}(0,T;\mathbb{R}^{\mathcal{Y}})$, equipped with the hybrid topology, with good rate functional $\mathcal{I}_0\big(c(0)\big) + \mathcal{I}(c)$, where*

$$\mathcal{I}(c) := \inf_{\substack{w \in W^{1,1}(0,T;\mathbb{R}^{\mathcal{R}}_+):\\ \dot{c} = \Gamma \dot{w}}} \mathcal{J}(c,w).$$

Naturally, this result is consistent with the above mentioned articles, but now under a more general set of assumptions on the reaction rates. In particular, our assumptions allow for mass-action kinetics, as in [DRW16].

Initial conditions. Throughout the paper we consider two different initial conditions. The main statement, Theorem 1.1 holds if the initial condition is random and satisfies a large-deviation principle. We will assume continuity of this initial large-deviation rate functional, which is essential to approximate the rate functional by sufficiently regular paths. For some results we shall consider a deterministic initial condition $C^{(V)}(0) = \tilde{c}^{(V)}(0)$ such that $c^{(V)}(0) \to \tilde{c}(0) \in \mathbb{R}^{\mathcal{Y}}$ for some limit initial condition. Those results can then be extended to random initial conditions via a mixture argument [Big04]. For the integrated fluxes we set $W^{(V,r)}(0) = 0$ almost surely; we shall therefore always implicitly assume that any large-deviation rate blows up unless $w(0) = 0$.

Strategy and overview. Section 2 describes the setting of the paper: the topology used for the dynamic large deviations, the precise assumptions on the propensities, reaction rates and initial condition. We then discuss existence and convergence of the path measures, which serves as a prerequisite for the large-deviations. Section 3 is dedicated to the analysis of the rate functional. Most importantly, it is shown that the rate functional has an alternative formulation as a convex dual, and that the rate functional can be approximated by curves that are sufficiently regular to be able to perform a change-of-measure. In a sense, these approximation lemmas are the core of the large-deviation proof. We shall see that the fact that the rate functional has a relatively simple formulation makes these proofs rather direct (which would be much more cumbersome when proving the large deviations of the concentrations only). Finally, Section 4 is devoted to the proof of the large-deviation principle, Theorem 1.1. It will be shown that one can always construct sufficiently steep compact cones on which the path measures place all but exponentially vanishing probability. We then show the lower bound of the measures with the random initial conditions via a double tilting argument, exploiting the approximation lemmas. After this, the upper bound is proven under deterministic initial conditions, which implies the large-deviations upper bound by a mixture argument.

2 Setting

In this section we specify the setting that we will be used in the paper. More specifically, we first introduce the hybrid topology used in the large deviations, and the precise assumptions on the propensities, reaction rates and initial condition that we will need. Finally, we construct the Markov process and its corresponding limit.

2.1 The hybrid topology

For any path $(c,w) \in L^1(0,T;\mathbb{R}^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}})$, the essential pointwise variation is

$$\operatorname{epvar}(c,w) := \inf_{\substack{(\tilde{c},\tilde{w})=(c,w)\\ t-\text{a.e.}}} \sup_{0=t_1<\ldots<t_K=T} \sum_{k=1}^{K} \big| \big(\tilde{c}(t_{k+1}), \tilde{w}(t_{k+1})\big) - \big(\tilde{c}(t_k), \tilde{w}(t_k)\big) \big|,$$

and the space of paths of bounded variation is defined as:

$$\mathrm{BV}(0,T;\mathbb{R}^{\mathcal{Y}}\times\mathbb{R}^{\mathcal{R}}) := \big\{(c,w)\in L^1(0,T;\mathbb{R}^{\mathcal{Y}}\times\mathbb{R}^{\mathcal{R}}) : \mathrm{epvar}(c,w)<\infty\big\}.$$

Some key properties of paths of bounded variation include, see [AFP00]:

(i.) Left and right limits are well-defined, and one can (and we will) always take a càdlàg version. Wherever we write $(c(0),w(0))$, we implicitly mean the right limit $(c(0+),w(0+))$.

(ii.) Any path $(c(t),w(t))$ of bounded variation has a measure-valued derivative $\big(\dot c(\mathrm{d}t),\dot w(\mathrm{d}t)\big)$, and $\|(\dot c,\dot w)\|_{\mathrm{TV}}=\mathrm{epvar}(c,w)$.

(iii.) $\mathrm{BV}(0,T;\mathbb{R}^{\mathcal{Y}}\times\mathbb{R}^{\mathcal{R}})$ equipped with the norm $\|\cdot\|_{L^1}+\mathrm{epvar}(\cdot)$ is a Banach space, and it is isometrically isomorphic to the dual of a Banach space.

Because of the last point, the space can also be equipped with a weak-* topology, which amounts to vague convergence of both the paths $(c^{(n)},w^{(n)})$ and its derivatives $(\dot c,\dot w)$, defined by pairing with test functions $(\phi,\psi)\in C_0(0,T;\mathbb{R}^{\mathcal{Y}}\times\mathbb{R}^{\mathcal{Y}})$. Naturally, weak-* compactness is simply characterised by norm-boundedness. Unfortunately, the weak-* topology is not metric, and hence difficult to use for stochastic analysis. Nevertheless, norm-boundedness is known to yield compactness in a slightly stronger topology [AFP00, Prop. 3.13], which we call the *hybrid* topology[1], defined through the convergence:

$$(c^{(n)},w^{(n)})\xrightarrow{\text{hybrid}}(c,w) \iff \|(c^{(n)},w^{(n)})-(c,w)\|_{L^1}\to 0 \qquad \text{and}$$
$$\big\langle(\phi,\psi),(\dot c^{(n)},\dot w^{(n)})\big\rangle\to\big\langle(\phi,\psi),(\dot c,\dot w)\big\rangle \quad \forall(\phi,\psi)\in C_0(0,T;\mathbb{R}^{\mathcal{Y}}\times\mathbb{R}^{\mathcal{R}}).$$

It turns out that the hybrid topology, although not metric, is 'perfectly normal', which implies that the corresponding Borel σ-algebra behaves nicely, and all probabilistic tools that we will need are valid, see [HPR16, Sec. 4].

2.2 The assumptions

We now state the set of assumptions under which we will prove our main result. A central role is played by the sets of concentrations that are reachable via chemical reactions:

Definition 2.1 (Stoichiometric simplex)**.** *Let* $c_0\in\mathbb{R}^{\mathcal{Y}}_+$

$$\mathscr{S}(c) := \{\tilde c = c+\Gamma w : w\in\mathbb{R}^{\mathcal{R}}_+, \tilde c\geq 0\} \tag{2.1}$$

$$\mathscr{S}_\epsilon(c) = \bigcup_{\tilde c:\ |c-\tilde c|\leq\epsilon}\mathscr{S}(\tilde c) \tag{2.2}$$

For vectors in $\mathbb{R}^{\mathcal{Y}}$ or $\mathbb{R}^{\mathcal{R}}$ we write $\geq$ for the partial ordering obtained by coordinate-wise inequalities. The set of assumptions on the propensities and reaction rates are the following:

Assumption 2.2 (Conditions on reaction rates)**.**

(i) $k^{(r,V)}(c)=0$ *whenever* $c_y<-V^{-1}\gamma^{(r)}_y$ *for at least one* $y\in\mathcal{Y}$*,*

(ii) $\sup_{c\in\mathscr{S}_\epsilon(c(0))}\sum_{r\in\mathcal{R}}\big|\frac{1}{V}k^{(V,r)}(c)-\bar k^{(r)}(c)\big|\to 0$ *for all* $\epsilon>0$ *and* $c(0)\in\mathbb{R}^{\mathcal{Y}}_+$*,*

(iii) $\bar k\in C^1(\mathbb{R}^{\mathcal{Y}}_+;\mathbb{R}^{\mathcal{R}}_+)$*,*

(iv) $\sup_{\tilde c\in\mathscr{S}_\epsilon(c)}|\bar k(\tilde c)|\vee|\nabla_c\bar k(c)|<\infty$ *for all* $c\in\mathbb{R}^{\mathcal{Y}}_+$ *and* $\epsilon>0$*,*

(v) $\bar k(\hat c)\geq\bar k(c)$ *for all* $\hat c\geq c$ *in* $\mathbb{R}^{\mathcal{Y}}_+$*,*

(vi) there exists a strictly increasing bijection $\psi:[0,1]\to[0,1]$ *such that*

$$\bar k^{(r)}(\delta c)\geq\psi(\delta)\bar k^{(r)}(c) \qquad \textit{for all } c\in\mathbb{R}^{\mathcal{Y}}_+,\delta>0 \textit{ and } r\in\mathcal{R}.$$

[1] The hybrid topology is usually called the weak-* topology. We name it differently to distinguish it from the functional analytically defined weak-* topology. The two topologies coincide on compact sets; in infinite dimensions the distinction becomes more subtle, see [HPR16].

The first assumption is needed to make sure that the stochastic model does not allow for negative concentrations. No assumptions related to boundedness or compactness of the stoichiometric simplices $\mathscr{S}(c(0))$ are required; the only assumption that is needed is (iv): that the reaction rates remain bounded on these simplices. Furthermore, the superhomogeneity assumption (vi) holds for most practical purposes, in particular for models with mass-action kinetics. We expect that the C^1-regularity can be relaxed to a locally Lipschitz condition, and that the monotonicity is only required in regions where the rates are small. Taken together (i) and (ii) imply that $c \geq 0$ is necessary in order to have $\bar{k}^{(r)}(c) > 0$.

The generality of the class of allowed reaction rates comes at the price of some regularity assumptions on the initial condition:

Assumption 2.3 (Sufficiently regular initial LDP)**.** *The initial measure $\mu^{(V)}$ satisfies a large-deviation principle in $\mathbb{R}_+^{\mathcal{Y}}$ with rate function $\mathcal{I}_0$ such that*

(i) *$\mathcal{I}_0$ is convex,*

(ii) *$\mathcal{I}_0$ is continuous,*

(iii) *$\mu^{(V)}$ converges in distribution to $\delta_{\tilde{c}(0)}$ for some $\tilde{c}(0) \in \mathbb{R}_+^{\mathcal{Y}}$,*

(iv) *$\mu^{(V)}$ is exponentially tight (and hence $\mathcal{I}_0$ is good),*

(v) *$\mathcal{I}_0$ satisfies the conditions of Varadhan's Integral Lemma[DZ87, Th. 4.3.1] for linear functions, i.e. for all $z \in \mathbb{R}^{\mathcal{Y}}$,*

  (a) $\lim_{M\to\infty} \limsup_{V\to\infty} \frac{1}{V} \log \int_{z\cdot c(0)\geq M} \mathrm{e}^{Vz\cdot c(0)}\, \mu^{(V)}\big(c(0)\big) = -\infty$, *or*

  (b) $\limsup_{V\to\infty} \frac{1}{V} \log \int \mathrm{e}^{Vaz\cdot c(0)}\, \mu^{(V)}\big(c(0)\big) < \infty$ *for some* $a > 1$.

(vi) $\partial\mathcal{I}_0\big(c(0)\big) \neq \varnothing$ *for all* $c(0) \in \mathbb{R}_+^{\mathcal{Y}}$.

Although this list of assumptions is a bit technical, we point out that most assumptions mean that $C^{(V)}(0)$ satisfy a 'sufficiently nice' large-deviation principle. For thermodynamic properties, one is mostly interested in the large deviations where the process starts from the invariant measure [Ren17, Sec. 4], which often satisfies a large-deviation principle with all the needed assumptions. The continuity of $\mathcal{I}_0$ will be exploited (and are essential) in the approximation lemmas 3.6,3.7,3.8 and 3.9, and the last assumption is a technical requirement that is needed to prove the large-deviation lower bound for the mixture.

2.3 Construction and convergence of the process

We denote by $\mathbb{P}^{(V)}$ the path measure of the process $\big(C^{(V)}(t), W^{(V)}(t)\big)$ with jump dynamics as captured in the generator (1.2) and initial distribution $\mu^{(V)} \times \delta_0$. This is well-defined, as Assumptions 2.2(ii) and (iv) imply that the jump rates are uniformly bounded on each stoichiometric simplex $\mathscr{S}(c)$, and hence (1.2) indeed generates a Markov process on $\mathrm{BV}(0,T;\mathbb{R}^{\mathcal{Y}}\times\mathbb{R}^{\mathcal{R}})$ (see [HPR16, Sect. 4] for a discussion of the Borel σ-algebra of the hybrid topology, and related properties).

For technical reasons we shall also consider the dynamics obtained by perturbing the jump rates using exponentials of $\zeta \in C_c(0,T;\mathbb{R}^{\mathcal{R}})$, leading to the time dependent generator

$$(\mathcal{Q}^{(V)}_{\zeta,t}\Phi)(c,w) := \sum_{r\in\mathcal{R}} k^{(V,r)}(c) e^{\zeta(t)\cdot\gamma^{(r)}} \big[\Phi(c + \tfrac{1}{V}\gamma^{(r)}, w + \tfrac{1}{V}\mathbb{1}_r) - \Phi(c,w)\big]. \tag{2.3}$$

Since the jump rates remain uniformly bounded under the perturbation, this generator also defines a path measure $\mathbb{P}^{(V)}_\zeta$ with initial condition $\mu^{(V)} \times \delta_0$.

In the interests of brevity we merely state the laws of large numbers for these measures, using the fact that the equations

$$\begin{cases} \dot{c}(t) = \Gamma\dot{w}(t), \\ \dot{w}(t) = \bar{k}^{(r)}\big(c(t)\big)e^{\zeta(t)\cdot\gamma^{(r)}}, \end{cases} \tag{2.4}$$

are well posed in $W^{1,1}(0,T;\mathbb{R}^{\mathcal{Y}}\times\mathbb{R}^{\mathcal{R}})$ for non-negative initial data; this may be checked by a Picard–Lindelöf argument. The basic ideas of the convergence proof go back to Kurtz [Kur70, Kur72].

Proposition 2.4. *Let $\zeta \in C_c(0,T;\mathbb{R}^{\mathcal{R}})$, Assumption 2.2 hold and suppose $\widetilde{\mu}^{(V)}$ converges narrowly to $\delta_{(\tilde{c}(0),0)}$. Then the laws $\widetilde{\mathbb{P}}^{(V)}_{\zeta}$ of the Markov processes with initial conditions $\widetilde{\mu}^{(V)}$ and dynamics given by* (2.3) *converge narrowly to the delta measure concentrated on the $(c,w) \in W^{1,1}(0,T;\mathbb{R}^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}})$ that is the unique solution to* (2.4) *with initial data $(\tilde{c}(0),0)$.*

Note that this result includes the cases of random initial conditions $\widetilde{\mu}^{(V)} = \mu^{(V)}$ as in Assumption 2.3, as well as the case of deterministic initial conditions $\widetilde{\mu}^{(V)} = \delta_{(\tilde{c}^{(V)}(0),0)}$ where $\tilde{c}^{(V)}(0) \to \tilde{c}(0)$. Note also that narrow convergence of probability measures on a metric space (convergence in distribution of the associated random variables) to a deterministic limit implies convergence in probability; this can readily be generalised to the hybrid topology on the space of bounded variation paths.

3 Analysis of the rate functional

A detailed knowledge of the properties of the rate function allows for a more concise presentation of the LDP, so these properties are developed here before we embark on the stochastic aspects of the proof. It will be practical to prove a dual, variational formulation of the rate functional:

$$\tilde{\mathcal{J}}(c,w) = \begin{cases} \sup\limits_{\zeta \in C^1_c(0,T;\mathbb{R}^{\mathcal{R}})} G(c,w,\zeta), & \text{if } \dot{c} = \Gamma \dot{w}, \\ \infty, & \text{otherwise}, \end{cases} \tag{3.1}$$

where

$$G(c,w,\zeta) := \int_0^T \big[\zeta(t)\cdot \dot{w}(\mathrm{d}t) - H\big(c(t),\zeta(t)\big)\big]\,\mathrm{d}t, \tag{3.2}$$

$$H(c,\zeta) := \sum_{r\in\mathcal{R}} \bar{k}^{(r)}(c)\big(\mathrm{e}^{\zeta^{(r)}} - 1\big). \tag{3.3}$$

Remark 3.1. $\tilde{\mathcal{J}}\colon \mathrm{BV}(0,T;\mathbb{R}^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}}_+) \to [0,\infty]$ is lower semincontinuous with respect to the hybrid topology on $\mathrm{BV}(0,T;\mathbb{R}^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}}_+)$ since for any $\zeta \in C_0(0,T;\mathbb{R}^{\mathcal{R}})$ the function $(c,w) \mapsto G(c,w,\zeta)$ is hybrid continuous. □

Remark 3.2. One can also rewrite the rate functional as a convex dual without restricting to pairs that satisfy the continuity equation:

$$\tilde{\mathcal{J}}(c,w) = \sup_{\substack{\xi \in C^1_c(0,T;\mathbb{R}^{\mathcal{Y}}) \\ \zeta \in C^1_c(0,T;\mathbb{R}^{\mathcal{R}})}} \int_0^T \zeta(t)\cdot\dot{w}(\mathrm{d}t) + \int_0^T \xi(t)\cdot\dot{c}(\mathrm{d}t) - \int_0^T H(c(t),\zeta(t))\,\mathrm{d}t.$$

A straight-forward calculation then shows that the rate functional reduces to (1.3) if the continuity equation is satisfied, and ∞ otherwise. The variation over the dual variable to $\dot{c}$ corresponds in some sense to zero-probability fluctuations in the continuity equation. Therefore it is more natural to omit that supremum, which also shortens notation considerably. □

3.1 Characterisation of the domain

This section is devoted to the proof that both formulations of the rate functional coincide. For the relative entropy formulation $\mathcal{J}$ of the rate functional, it is built into the definition (1.3) that $(c,w) \in W^{1,1}(0,T;\mathbb{R}^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}}_+)$ for finite $\mathcal{J}(c,w)$. The following Lemma says that the concentrations remain non-negative.

Lemma 3.3. *Let $(c,w) \in \mathrm{BV}(0,T;\mathbb{R}^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}}_+)$ and $c(0) \geq 0$. If $\mathcal{J}(c,w) < \infty$ then $c \geq 0$.*

Proof. Assume on the contrary that one may find t_1, y_1 such that $c_{y_1}(t_1) < 0$. By definition $\mathcal{J}(c,w) < \infty$ implies c_{y_1} (has a representative that) is absolutely continuous so one may take $0 \leq t_2 < t_1$ such that $0 \geq c_{y_1}(t_2) > c_{y_1}(t_1)$. This implies the existence of $r_1 \in \mathcal{R}$ such that $\gamma^{(r)}_y < 0$ and $\int_{t_2}^{t_1} \dot{w}^{(r_1)}(s)\mathrm{d}s > 0$ so (1.3) requires $\bar{k}^{(r_1)}(c(s)) > 0$ almost everywhere in $[t_2,t_1]$. However from Assumption 2.2 parts (i) and (ii) one sees that $\bar{k}^{(r_1)}(c(s)) = 0$ for all $s \in [t_2,t_1]$. □

In order to compare $\mathcal{J}$ to the variational formulation $\tilde{\mathcal{J}}$ we need to prove the same regularity result for $\tilde{\mathcal{J}}$:

Lemma 3.4. *Let $(c, w) \in \mathrm{BV}(0, T; \mathbb{R}_+^{\mathcal{Y}} \times \mathbb{R}_+^{\mathcal{R}})$. If $\tilde{\mathcal{J}}(c, w) < \infty$ then $(c, w) \in W^{1,1}(0, T; \mathbb{R}^{\mathcal{Y}} \times \mathbb{R}_+^{\mathcal{R}})$ and $c \geq 0$.*

Proof. Let $(c, w) \in \mathrm{BV}(0, T; \mathbb{R}_+^{\mathcal{Y}} \times \mathbb{R}_+^{\mathcal{R}})$ and $\tilde{\mathcal{J}}(c, w) < \infty$; the proof is carried out in three stages:

1. $\dot{w}$ is a non-negative measure,
2. $\dot{w}(\mathrm{d}t) = \dot{w}(t)\,\mathrm{d}t$ for some density $\dot{w} \in L^1(0, T; \mathbb{R}_+^{\mathcal{R}})$,
3. $\dot{c}(t) = \Gamma \dot{w}(t)$,
4. $c \geq 0$.

For the first point note that the existence of $\dot{w}$ as a (signed) vector measure of finite total variation follows from [HPR16, Thrm. 2.13]. Suppose now that there is some $r \in \mathcal{R}$ and a measurable set $A \subset (0, T)$ such that $\dot{w}^{(r)}(A) < 0$. Using the Hahn decomposition and the regularity of Borel measures on the metric space $(0, T)$ ([Bog07, Thrm. 7.1.7] or [Kal02, Lem. 1.34]) one has the existence of a closed $B \subset A$ with $\dot{w}^{(r)}(B) < 0$. Define $\zeta_n \in C_c^1(0, T; \mathbb{R}_-^{\mathcal{R}})$ by

$$\zeta_n^{(r')}(t) = \begin{cases} 0 & r' \neq r \\ -n\varphi(t) & r' = r \end{cases}$$

for some $\varphi \in C_c^1(0, T; [0, 1])$ such that $\mathbb{1}_B \leq \varphi \leq \mathbb{1}_A$. On can now check that $\lim_n G(c, w, \zeta_n) = +\infty$, which contradicts $\tilde{\mathcal{J}} < \infty$ so there cannot be any r for which $\dot{w}^{(r)}$ takes negative values.

For the absolute continuity suppose that there is an $r \in \mathcal{R}$ and a measurable set $A \subset (0, T)$ such that $\dot{w}^{(r)}(A) = \delta > 0$, but $|A| = 0$, where we write $|\cdot|$ for Lebesgue measure. By the regularity result already mentioned in this proof we have the existence of closed sets F_n and open sets G_n such that $F_n \subset A \subset G_n$ with $\dot{w}^{(r)}(G_n \setminus F_n) < \frac{1}{n}$ and $|G_n| \leq \frac{1}{n}$. Define $\zeta_n \in C_c^1(0, T; \mathbb{R}_+^{\mathcal{R}})$ by

$$\zeta_n^{(r')}(t) = \begin{cases} 0 & r' \neq r \\ -\log |G_n|\, \varphi_n(t) & r' = r \end{cases}$$

for some $\varphi \in C_c^1(0, T; [0, 1])$ such that $\mathbb{1}_{F_n} \leq \varphi \leq \mathbb{1}_{G_n}$ to get a contradiction as in the proof that $\dot{w} \geq 0$. The Radon-Nikodym theorem thus allows us with a little abuse of notation to write $\dot{w}^{(r)}(\mathrm{d}t) = \dot{w}^{(r)}(t)\mathrm{d}t$ for $\dot{w}^{(r)} \in L^1(0, T; \mathbb{R})$.

The proof that $c(t) \geq 0$ is the same as in Lemma 3.3, where now we have on the non-null set $B \subset (0, T)$,

$$\mathcal{J}(c, w) \geq \sup_{\zeta^{(r)} \in C_c^1(B)} \int_B \zeta^{(r)}(t) \cdot \dot{w}^{(r)}(t) - 0 = \infty.$$

□

Proposition 3.5. $\mathcal{J} = \tilde{\mathcal{J}}$.

Proof. Let $(c, w) \in \mathrm{BV}(0, T; \mathbb{R}_+^{\mathcal{Y}} \times \mathbb{R}_+^{\mathcal{R}})$ (possibly with $\mathcal{J}(c, w) = \infty$). If $(c, w) \notin W^{1,1}(0, T; \mathbb{R}_+^{\mathcal{Y}} \times \mathbb{R}_+^{\mathcal{R}})$ then by Lemma 3.4 both $\tilde{\mathcal{J}}(c, w) = \infty = \mathcal{J}(c, w)$. Now assume that $(c, w) \in W^{1,1}(0, T; \mathbb{R}_+^{\mathcal{Y}} \times \mathbb{R}_+^{\mathcal{R}})$. We can then write $G(c, w, \zeta) = \sum_{r \in \mathcal{R}} \int_0^T g^{(r)}\big(c(t), \dot{w}^{(r)}(t), \zeta^{(r)}(t)\big)\,\mathrm{d}t$ where

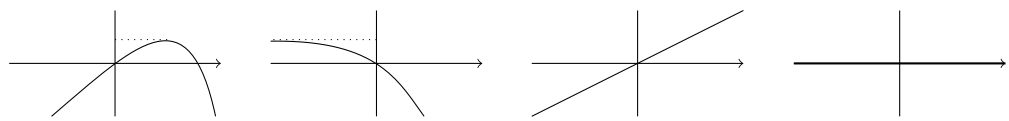
$$g^{(r)}(c, j^{(r)}, \zeta^{(r)}) := \zeta^{(r)} j^{(r)} - \bar{k}^{(r)}(c)\big(\mathrm{e}^{\zeta^{(r)}} - 1\big).$$

We now show that

$$\begin{aligned} \mathcal{J}(c, w) &= \sup_{\zeta : (0,T) \to \mathbb{R}^{\mathcal{R}}} G(c, w, \zeta) = \sup_{\zeta \in L^\infty(0,T;\mathbb{R}^{\mathcal{R}})} G(c, w, \zeta) \\ &= \sup_{\zeta \in C_b^1(0,T;\mathbb{R}^{\mathcal{R}})} G(c, w, \zeta) = \sup_{\zeta \in C_c^1(0,T;\mathbb{R}^{\mathcal{R}})} G(c, w, \zeta) = \tilde{\mathcal{J}}(c, w). \end{aligned} \tag{3.4}$$

The first equality in (3.4) can be calculated directly through the pointwise supremum. For the second equality, we construct, for each $t \in (0,T)$ and $r \in \mathcal{R}$, an explicit (pointwise) maximising sequence $\zeta_n^{(r)}(t)$ for $\sup_{\zeta^{(r)}} g^{(r)}\big(c(t), \dot w^{(r)}(t), \zeta^{(r)}(t)\big)$ as, see Figure 1,

$$\zeta_n^{(r)}(t) := \begin{cases} \log \frac{w^{(r)}(t)}{\bar k^{(r)}(c(t))} \wedge n, & \bar k^{(r)}(c(t)) > 0 \text{ and } \dot w^{(r)}(t) > 0, \\ -n, & \bar k^{(r)}(c(t)) > 0 \text{ and } \dot w^{(r)}(t) = 0, \\ n, & \bar k^{(r)}(c(t)) = 0 \text{ and } \dot w^{(r)}(t) > 0, \\ 0, & \bar k^{(r)}(c(t)) = 0 \text{ and } \dot w^{(r)}(t) = 0. \end{cases}$$



(a) $\bar k^{(r)}(c) > 0, j^{(r)} > 0$ (b) $\bar k^{(r)}(c) > 0, j^{(r)} = 0$ (c) $\bar k^{(r)}(c) = 0, j^{(r)} > 0$ (d) $\bar k^{(r)}(c) = 0, j^{(r)} = 0$

Figure 1: The function $\zeta^{(r)} \mapsto g^{(r)}(c, j^{(r)}, \zeta^{(r)})$.

Then each $\zeta_n \in L^\infty(0,T;\mathbb{R}^{\mathcal{R}})$ and $g^{(r)}\big(c(t), \dot w^{(r)}(t), \hat\zeta^{(r)}(t)\big)$ is non-decreasing in n and non-negative. Moreover, $g^{(r)}\big(c(t), \dot w^{(r)}(t), \hat\zeta^{(r)}(t)\big)$ converges pointwise in $t \in (0,T)$ and $r \in \mathcal{R}$ as $n \to \infty$ to the pointwise supremum. Hence by monotone convergence

$$\lim_{n\to\infty} \sum_{r\in\mathcal{R}} \int_0^T g^{(r)}\big(c(t), \dot w^{(r)}(t), \hat\zeta_n^{(r)}(t)\big) = \sum_{r\in\mathcal{R}} \int_0^T \sup_{\zeta^{(r)}} g^{(r)}\big(c(t), \dot w^{(r)}(t), \zeta^{(r)}(t)\big)\,\mathrm{d}t.$$

This shows that the pointwise supremum on the left of (3.4) can be taken over $L^\infty(0,T;\mathbb{R}^{\mathcal{R}})$.

For the third equality in (3.4) it suffices to show that for any $\zeta \in L^\infty(0,T;\mathbb{R}^{\mathcal{R}})$ the integrand can be approximated by a sequence in $C_b^2(0,T;\mathbb{R}^{\mathcal{R}})$. For an arbitrary $\zeta \in L^\infty(0,T;\mathbb{R}^{\mathcal{R}})$ consider the convolutions with smoothing kernels θ_δ for $\delta > 0$ that weakly converges to the Dirac measure at 0 as $\delta \to 0$. In the convolutions we extended the function ζ to zero outside the integral $(0,T)$. Since $\zeta \in L^1(\mathbb{R};\mathbb{R}^{\mathcal{R}})$ this sequence $\zeta * \theta_\delta$ converges strongly in $L^1(\mathbb{R};\mathbb{R}^{\mathcal{R}})$ to ζ as $\delta \to 0$, see [Eva02, App. C.4]. By a partial converse of the Dominated Convergence Theorem [Bré83, Th. IV.9], after passing to a subsequence $\zeta_n(t)^{(r)} := (\zeta * \theta_{\delta_n}^{(r)})(t)$ converges pointwise t-almost everywhere. Then the exponential $-\bar k\big(c(t)\big)\big(\mathrm{e}^{\zeta_n^{(r)}(t)} - 1\big)$ integrand part of $G(c, w, \zeta_n)$ also converges pointwise for almost every t. Moreover, we can bound

$$\|\bar k\|_\infty \ge -\bar k\big(c(t)\big)\big(\mathrm{e}^{\zeta_n^{(r)}(t)} - 1\big) \ge -\|\bar k\|_\infty \mathrm{e}^{\|\zeta_n\|_{L^\infty(0,T;\mathbb{R}^{\mathcal{R}})}} \ge -\|\bar k\|_\infty \mathrm{e}^{\|\zeta\|_{L^\infty(0,T;\mathbb{R}^{\mathcal{R}})}},$$

and hence by dominated convergence

$$-\sum_{r\in\mathcal{R}} \int_0^T \bar k\big(c(t)\big)\big(\mathrm{e}^{\zeta_n^{(r)}(t)} - 1\big)\,\mathrm{d}t \to -\sum_{r\in\mathcal{R}} \int_0^T \bar k\big(c(t)\big)\big(\mathrm{e}^{\zeta^{(r)}(t)} - 1\big)\,\mathrm{d}t.$$

Clearly the linear part $\sum_{r\in\mathcal{R}} \int_0^T \zeta_n^{(r)}(t)\dot w^{(r)}(t)\,\mathrm{d}t$ of $G(c, w, \zeta_n)$ converges to $\sum_{r\in\mathcal{R}} \int_0^T \zeta^{(r)}(t)\dot w^{(r)}(t)\,\mathrm{d}t$, and so $G(c, w, \zeta_n) \to G(c, w, \zeta)$. This proves the third equality in (3.4).

For the fourth equality, take any $\zeta \in C_b^1(0,T;\mathbb{R}^{\mathcal{R}})$, and approximate with $\zeta\eta_\delta \in C_c^1(0,T;\mathbb{R}^{\mathcal{R}})$ where

$$\eta_\delta(t) := \begin{cases} 0, & t \in (0,\delta] \cup [T-\delta, T), \\ 1, & t \in [2\delta, T-2\delta], \\ \text{smooth between 0 and 1}, & t \in [\delta, 2\delta] \cup [T-2\delta, T-\delta]. \end{cases} \tag{3.5}$$

Then, as $\delta \to 0$,

$$G(c, w, \zeta\eta_\delta) = \int_0^T \dot w(t)\zeta(t)\eta_\delta(t)\,\mathrm{d}t - \sum_{r\in\mathcal{R}} \int_0^T \bar k^{(r)}\big(c(t)\big)\big(\mathrm{e}^{\zeta_\delta(t)} - 1\big)\,\mathrm{d}t \to G(c, w, \zeta\eta_\delta) \to G(c, w, \zeta),$$

where for the linear part we use that $\zeta\eta_\delta \to \zeta$ weakly-* in $L^\infty(0,T;\mathbb{R}^{\mathcal{R}})$, and for the nonlinear part we use dominated convergence. ∎

3.2 Approximation by regular curves

A common challenge in proving a large-deviations lower bound for a Markov process is to approximate any curve of finite rate by curves for which one can perform a change-of-measure. In the setting of our paper, this set of sufficiently regular curves will be defined as:

$$\mathscr{A} := \Big\{(c,w) \in \mathrm{BV}\big(0,T;\mathbb{R}_+^{\mathcal{Y}}\times\mathbb{R}_+^{\mathcal{R}}\big)\cap \mathrm{AC}\big(0,T;\mathbb{R}^{\mathcal{Y}}\times\mathbb{R}^{\mathcal{R}}\big): \\ \zeta := \log\frac{\dot w}{\bar k(c)} \in C_c^1\big(0,T;\mathbb{R}^{\mathcal{R}}\big), \quad \dot c = \Gamma\dot w, \quad c,w,\dot w\geq 0, \quad w(0)=0\Big\}. \tag{3.6}$$

Observe that this set requires compactly supported perturbations, whereas the change-of-measure Theorem A.3 only requires boundedness. However, the compact support will be needed to control the end point in the tilting arguments, Lemmas 4.6 and 4.7.

This section is dedicated to the proof of the required approximation result using a sequence of four approximation lemmas. We repeatedly exploit the lower semi-continuity of $\mathcal{J}$ to show that if $\lim_{\delta\searrow 0}(c_\delta,w_\delta)=(c,w)$ in the hybrid topology, then $\liminf_{\delta\searrow 0}\mathcal{J}(c_\delta,w_\delta)\geq\mathcal{J}(c,w)$.

Lemma 3.6 (Approximation I)**.** *Let $\mu^{(V)}$ satisfy Assumption 2.3 and $\bar k$ satisfy Assumptions 2.2*(iii),(iv),(v) *and* (vi)*. Given $(c,w)\in\mathrm{BV}\big(0,T;\mathbb{R}_+^{\mathcal{Y}}\times\mathbb{R}_+^{\mathcal{R}}\big)$ such that $\mathcal{J}(c,w)<\infty$, there exists a sequence $(c_\delta,w_\delta)_\delta\subset\mathrm{BV}\big(0,T;\mathbb{R}_+^{\mathcal{Y}}\times\mathbb{R}_+^{\mathcal{R}}\big)$ such that:*

(i) $c_\delta(0)\to c(0)$ *and* $(c_\delta,w_\delta)\rightrightarrows(c,w)$ *as* $\delta\to 0$*,*

(ii) $\mathcal{I}_0(c_\delta(0))+\mathcal{J}(c_\delta,w_\delta)\to\mathcal{I}_0(c(0))+\mathcal{J}(c,w)$ *as* $\delta\to 0$*,*

(iii) $\inf_{t\in(0,T),r\in\mathcal{R}}\bar k^{(r)}(c_\delta(t))>0$ *for any* $\delta>0$*,*

Proof. Without loss of generality we may assume that for each reaction r there exists a concentration $\hat c^{(r)}\in\mathbb{R}_y^+$ for which $\bar k^{(r)}(\hat c^{(r)})>0$. Set $\hat c=\sum_{r\in\mathcal{R}}\hat c^{(r)}$, so that by the assumed monotonicity,

$$\min_{r\in\mathcal{R}}\bar k^{(r)}(\hat c)>0.$$

For $\delta>0$ define

$$c_\delta(t) := \delta\hat c+(1-\delta)c(0)+\Gamma w_\delta(t), \qquad \text{and} \qquad w_\delta(t):=(1-\delta)w,$$

so that $c_\delta(t)=\delta\hat c+(1-\delta)c(t)\geq 0$.

The limits (i) are trivial. The lower bound (iii) follows by the monotonicity and superhomogeneity Assumptions 2.2(v) and (vi):

$$\inf_{t\in(0,T),r\in\mathcal{R}}\bar k^{(r)}\big(c_\delta(t)\big)\geq\min_{r\in\mathcal{R}}\bar k^{(r)}(\delta\hat c)\geq\psi(\delta)\min_{r\in\mathcal{R}}\bar k^{(r)}(\hat c)>0. \tag{3.7}$$

For the limits (ii), the convergence of $\mathcal{I}_0(c_\delta(0))$ follows by Assumption 2.3. Since $\liminf_{\delta\searrow 0}\mathcal{J}(c_\delta,w_\delta)\geq\mathcal{J}(c,w)$ it remains to check $\limsup_{\delta\searrow 0}\mathcal{J}(c_\delta,w_\delta)\leq\mathcal{J}(c,w)$. Using the fact that $\bar k^{(r)}(c_\delta(t))\geq\psi(1-\delta)\bar k^{(r)}(c(t))$ by the same argument as (3.7) above, we can rewrite and estimate:

$$s\big(\dot w_\delta^{(r)}(t)\mid\bar k^{(r)}(c_\delta(t))\big)=s\big(\dot w_\delta^{(r)}(t)\mid\bar k^{(r)}(c(t))\big)+\dot w_\delta^{(r)}(t)\log\left(\frac{\bar k^{(r)}(c(t))}{\bar k^{(r)}(c_\delta(t))}\right)-\bar k^{(r)}(c(t))+\bar k^{(r)}(c_\delta(t)) \tag{3.8}$$

$$\leq(1-\delta)s\big(\dot w^{(r)}(t)\mid\bar k^{(r)}(c(t))\big)+\delta\bar k^{(r)}(c(t)) \\ +(1-\delta)\dot w^{(r)}(t)\log\frac{1-\delta}{\psi(1-\delta)}+\big|\bar k^{(r)}(c(t))-\bar k^{(r)}(c_\delta(t))\big|. \tag{3.9}$$

Summing over r and integrating over t shows that, for δ sufficiently small,

$$\mathcal{J}(c_\delta,w_\delta)\leq(1-\delta)\mathcal{J}(c,w)+\delta T\sup_{\tilde c\in\mathscr{S}_{2\delta}(c(0))}|\bar k(\tilde c)|+(1-\delta)\log\frac{1-\delta}{\psi(1-\delta)}\|\dot w\|_{L^1}+\sqrt{|\mathcal{R}|}\,\mathrm{Lip}(\bar k)\,\|c_\delta-c\|_{L^1}.$$

Using Assumption 2.2(iv) it follows that all but the first term on the right-hand side vanish as $\delta\searrow 0$ and the result is established. □

For smoothing purposes we make use of convolutions with the heat kernels $\theta_\epsilon\colon\mathbb{R}\to\mathbb{R}_+; t\mapsto\exp(-t^2/2\epsilon)/\sqrt{2\pi\epsilon}$.

Lemma 3.7 (Approximation II). *Let $\mu^{(V)}$ satisfy Assumption 2.3 and $\bar{k}$ satisfy Assumptions 2.2*(iii) *and* (iv). *Given $(c,w) \in \mathrm{BV}\big(0,T;\mathbb{R}_+^{\mathcal{Y}} \times \mathbb{R}_+^{\mathcal{R}}\big)$ such that $\mathcal{J}(c,w) < \infty$ and $\inf_{t\in(0,T),r\in\mathcal{R}} \bar{k}^{(r)}(c(t)) > 0$, there exists a sequence $(c_\delta,w_\delta)_\delta \subset C_b^\infty(0,T;\mathbb{R}_+^{\mathcal{Y}} \times \mathbb{R}_+^{\mathcal{R}})$ such that:*

(i) $c_\delta(0) \to c(0)$ *and* $(c_\delta,w_\delta) \rightrightarrows (c,w)$ *as* $\delta \to 0$,

(ii) $\mathcal{I}_0(c_\delta(0)) + \mathcal{J}(c_\delta,w_\delta) \to \mathcal{I}_0(c(0)) + \mathcal{J}(c,w)$ *as* $\delta \to 0$,

(iii) $\inf_{t\in(0,T),r\in\mathcal{R}} \bar{k}^{(r)}(c_\delta(t)) > 0$ *for any sufficiently small* $\delta > 0$.

Proof. Define

$$c_\delta(t) := c(0) + (w*\theta_\delta)(0) + \Gamma w_\delta(t) \qquad \text{and} \qquad w_\delta(t) := (w*\theta_\delta)(t) - (w*\theta_\delta)(0),$$

where in the convolutions we extend w constantly to $w(0)$ and $w(T)$ outside the interval $(0,T)$. Observe that the definition is sound in the sense that $c_\delta(t) = (c*\theta_\delta)(t) \geq 0$ and $w_\delta, \dot{w}_\delta \geq 0$. Since $(c,w) \in W^{1,1}(0,T;\mathbb{R}^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}})$ by Lemma 3.4, the desired convergence (i) of the sequence can be shown by adapting the results in [Eva02, App. C.4] to mollifiers with non-compact support. Similarly to the proof of Proposition 3.5, we pass to a (relabelled) subsequence such that in fact $\dot{w}_\delta(t) \to \dot{w}(t)$ pointwise in almost every $t \in (0,T)$.

To show the lower bound (iii), observe that since c is continuous (on the compact interval $[0,T]$) by Lemma 3.4, $c_\delta \to c$ uniformly, see [Eva02, App. C.4]. Moreover, by the continuity of $\bar{k}$, we know there exists a $\tau > 0$ such that for $\delta < \tilde{\delta}$ and any $\tilde{c} \in \mathbb{R}_+^{\mathcal{Y}}$ with $|\tilde{c} - c| < \tau$ there holds $|\bar{k}(\tilde{c}) - \bar{k}(c)| < \frac{1}{2}\inf_{t\in(0,T),r\in\mathcal{R}} \bar{k}^{(r)}(c(t))$. From the uniform convergence of c_δ we get the existence of a $\tilde{\delta} > 0$ such that for any $\delta < \tilde{\delta}$ and any $t \in [0,T]$, there holds $|c(t) - c_\delta(t)| < \tau$. Therefore, for any $\delta < \tilde{\delta}$ and so $|\bar{k}(c_\delta(t)) - \bar{k}(c(t))| < \frac{1}{2}\inf_{t\in(0,T),r\in\mathcal{R}} \bar{k}^{(r)}(c(t))$, from which we deduce the lower bound (iii):

$$\bar{k}(c_\delta(t)) \geq \bar{k}(c(t)) - \tfrac{1}{2} \inf_{t\in(0,T),r\in\mathcal{R}} \bar{k}^{(r)}(c(\tilde{t})) \geq \tfrac{1}{2} \inf_{\tilde{t}\in(0,T),r\in\mathcal{R}} \bar{k}^{(r)}(c(\tilde{t})). \tag{3.10}$$

The convergence $\mathcal{I}_0\big(c_\delta(0)\big) = \mathcal{I}_0\big(c(0) + (w*\theta_\delta)(0)\big) \to \mathcal{I}_0\big(c(0)\big)$ follows by Assumption 2.3. For the convergence of $\mathcal{J}(c_\delta,w_\delta)$, we can bound the integrand, similarly as in (3.8),

$$0 \leq s\big(\dot{w}_\delta^{(r)}(t) \mid \bar{k}^{(r)}(c_\delta(t))\big) \leq s\big(\dot{w}_\delta^{(r)}(t) \mid \bar{k}^{(r)}(c(t))\big) + a\dot{w}_\delta^{(r)}(t) + \big|\bar{k}^{(r)}(c(t)) - \bar{k}^{(r)}(c_\delta(t))\big|, \tag{3.11}$$

where

$$\log\Big(\frac{\bar{k}^{(r)}(c(t))}{\bar{k}^{(r)}(c_\delta(t))}\Big) \overset{(3.10)}{\leq} \log\Big(\frac{2\sup_{t\in(0,T)}\bar{k}(c(t))}{\inf_{t\in(0,T)}\bar{k}(c(t))}\Big) =: a \in [0,\infty).$$

By the assumed continuity of the reaction rates $s\big(\dot{w}_\delta^{(r)}(t)|k^{(r)}(c_\delta(t))\big) \to s\big(\dot{w}^{(r)}(t)|k^{(r)}(c(t))\big)$ pointwise in $t \in (0,T)$. If we can prove that, after summing over $\mathcal{R}$ and integrating over $(0,T)$, the right-hand side in (3.11) converges to a finite integral, then $\mathcal{J}(c_\delta,w_\delta) \to \mathcal{J}(c,w)$ by a generalisation of the Dominated Convergence Theorem, see [LL01, Th. 1.8 & following remark].

Naturally the last two terms converge:

$$\sum_{r\in\mathcal{R}} \int_0^T \Big[a\dot{w}_\delta^{(r)}(t) + \big|\bar{k}^{(r)}(c(t)) - \bar{k}^{(r)}(c_\delta(t))\big|\Big]\,\mathrm{d}t \to a\|\dot{w}\|_{L^1} < \infty.$$

The convergence of the entropic part can be proven analogue to [Ren17, Lem. 4.11]. By lower semicontinuity,

$$\liminf_{\delta\to 0} \sum_{r\in\mathcal{R}} \int_0^T s\big(\dot{w}_\delta^{(r)}(t)|\bar{k}^{(r)}(c(t))\big)\,\mathrm{d}t \geq \sum_{r\in\mathcal{R}} \int_0^T s\big(\dot{w}^{(r)}(t)|\bar{k}^{(r)}(c(t))\big)\,\mathrm{d}t.$$

On the other hand, by Jensen's inequality,

$$\begin{aligned}\sum_{r\in\mathcal{R}} \int_0^T s\big(\dot{w}_\delta^{(r)}(t)|\bar{k}^{(r)}(c(t))\big)\,\mathrm{d}t &\leq \sum_{r\in\mathcal{R}} \int_0^T \Big(s\big(\dot{w}^{(r)}(\cdot)|\bar{k}^{(r)}(c(t))\big) * \theta_\epsilon\Big)(t)\,\mathrm{d}t \\ &= \sum_{r\in\mathcal{R}} \int_0^T \underbrace{(\dot{w}^{(r)}\log\dot{w}^{(r)})}_{\in L^1} * \theta_\epsilon(t) - (\dot{w}^{(r)} * \theta_\epsilon)(t)\underbrace{\big(1 + \log\bar{k}^{(r)}(c(t))\big)}_{\in L^\infty} + \bar{k}^{(r)}\big(c(t)\big) \\ &\to \sum_{r\in\mathcal{R}} \int_0^T s\big(\dot{w}^{(r)}(t)|\bar{k}^{(r)}(c(t))\big)\,\mathrm{d}t,\end{aligned}$$

again by [Eva02, App. C.4]. Therefore the summed and integrated right-hand side of (3.11) indeed converges to a finite integral, which concludes the proof of claim (ii).

□

Lemma 3.8 (Approximation III)**.** *Let $\mu^{(V)}$ satisfy Assumption 2.3 and $\bar{k}$ satisfy Assumptions 2.2*(iii),(iv), (v) *and* (vi). *Given $(c,w) \in C_b^\infty\big(0,T;\mathbb{R}_+^\mathcal{Y}\times\mathbb{R}_+^\mathcal{R}\big)$ such that $\mathcal{J}(c,w)<\infty$ and $\inf_{t\in(0,T),r\in\mathcal{R}}\bar{k}^{(r)}(c(t))>0$, there exists a sequence $(c_\delta,w_\delta)_\delta\subset C_b^\infty\big(0,T;\mathbb{R}_+^\mathcal{Y}\times\mathbb{R}_+^\mathcal{R}\big)$ such that:*

(i) $c_\delta(0)\to c(0)$ *and* $(c_\delta,w_\delta)\rightrightarrows(c,w)$ *as* $\delta\to 0$*,*

(ii) $\mathcal{I}_0(c_\delta(0))+\mathcal{J}(c_\delta,w_\delta)\to\mathcal{I}_0(c(0))+\mathcal{J}(c,w)$ *as* $\delta\to 0$*,*

(iii) $\inf_{t\in(0,T),r\in\mathcal{R}}\bar{k}^{(r)}(c_\delta(t))>0$ *for any* $\delta>0$*,*

(iv) $\inf_{t\in(0,T),r\in\mathcal{R}}\dot{w}_\delta^{(r)}(t)>0$ *for any* $\delta>0$*,*

(v) $\zeta_\delta:=\log\frac{\dot{w}_\delta}{\bar{k}(c_\delta)}\in C_b^1(0,T;\mathbb{R}^\mathcal{R})$.

Proof. Let $\beta^{(r)},\alpha^{(r)}\in\mathbb{R}_+^\mathcal{Y}$ be the positive and negative parts of $\gamma^{(r)}$, i.e. $\gamma^{(r)}=\beta^{(r)}-\alpha^{(r)}$. For $0<\delta<1$ define

$$c_\delta(t):=(1-\delta)c(0)+\delta T\sum_{r\in\mathcal{R}}\alpha^{(r)}+\Gamma w_\delta(t)\qquad\text{and}\qquad w_\delta(t):=(1-\delta)w(t)+\delta t,$$

so that $c_\delta(t)=(1-\delta)c(t)+\delta\sum_{r\in\mathcal{R}}\big[(T-t)\alpha^{(r)}+t\beta^{(r)}\big]\geq 0$ and $\dot{w}_\delta(t)=(1-\delta)\dot{w}(t)+\delta\geq\delta>0$. Hence the sequence is admissable, and property (iv) holds by construction. Again, the hybrid convergence (i) is trivial, and the monotonicity and superhomogeneity, Assumptions 2.2(v), (vi) imply the same estimate as (3.7), which shows that the bound (iii) is indeed retained.

The convergence $\mathcal{I}_0\big(c_\delta(0)\big)\to\mathcal{I}_0\big(c(0)\big)$ follows from the continuity of $\mathcal{I}_0$ and (i). As in the previous lemmas it is sufficient to show $\limsup_{\delta\searrow 0}\mathcal{J}(c_\delta,w_\delta)\leq\mathcal{J}(c,w)$ in order to establish (ii). We can again derive estimate (3.9), where the terms $\dot{w}_\delta^{(r)}(t)\log\bar{k}^{(r)}(c(t))/\bar{k}^{(r)}(c_\delta(t))$ and $\bar{k}^{(r)}(c(t))-\bar{k}^{(r)}(c_\delta(t))$ can be dealt with in exactly the same manner as in the proof of Lemma 3.6. It thus remains to show convergence of the integral $\sum_{r\in\mathcal{R}}\int_0^T s\big(\dot{w}_\delta^{(r)}(t)|\bar{k}^{(r)}(c(t))\big)\,\mathrm{d}t$. By the convexity of s in its first argument, we get for $0<\delta<1$,

$$\begin{aligned}s\big(\dot{w}_\delta^{(r)}(t)|\bar{k}^{(r)}(c(t))\big)&\leq(1-\delta)s\big(\dot{w}^{(r)}(t)|\bar{k}^{(r)}(c(t))\big)+\delta s\big(1|\bar{k}^{(r)}(c(t))\big)\\&\leq s\big(\dot{w}^{(r)}(t)|\bar{k}^{(r)}(c(t))\big)-\delta\log\bar{k}^{(r)}(c(t))+\delta\bar{k}^{(r)}(c(t)).\end{aligned}$$

Since the last two terms are bounded from below and above it follows that $\limsup_{\delta\searrow 0}\mathcal{J}(c_\delta,w_\delta)\leq\mathcal{J}(c,w)$.

Finally we can prove (v) for any $\delta>0$. Since the curve (c_δ,w_δ) is smooth we only need to prove boundedness of the functions

$$\zeta_\delta^{(r)}(t)=\log\frac{\dot{w}_\delta^{(r)}(t)}{\bar{k}(c_\delta(t))},\qquad\text{and}\qquad\dot{\zeta}_\delta^{(r)}(t)=\frac{\ddot{w}_\delta^{(r)}(t)}{\dot{w}_\delta^{(r)}(t)}-\frac{\nabla_c\bar{k}(c_\delta(t))\cdot\dot{c}_\delta(t)}{\bar{k}(c_\delta(t))}.$$

This follows from the boundedness away from zero of $\dot{w}_\delta$ and $\bar{k}(c_\delta)$, together Assumption 2.2(iv).

□

Lemma 3.9 (Approximation IV)**.** *Let $\bar{k}$ satisfy Assumptions* (iii),(iv). *Given $(c,w)\in C_b^\infty\big(0,T;\mathbb{R}_+^\mathcal{Y}\times\mathbb{R}_+^\mathcal{R}\big)$ such that $\mathcal{J}(c,w)<\infty$ and $\zeta=\log\dot{w}/\bar{k}(c)\in C_b^1(0,T;\mathbb{R}^\mathcal{R})$, there exists a sequence $(c_\delta,w_\delta)\in C_c^1\big(0,T;\mathbb{R}_+^\mathcal{Y}\times\mathbb{R}_+^\mathcal{R}\big)$ such that:*

(i) $c_\delta(0)\equiv c(0)$ *and* $(c_\delta,w_\delta)\rightrightarrows(c,w)$ *as* $\delta\to 0$*,*

(ii) $\mathcal{I}_0(c_\delta(0))+\mathcal{J}(c_\delta,w_\delta)\to\mathcal{I}_0(c(0))+\mathcal{J}(c,w)$ *as* $\delta\to 0$.

Proof. Given (c,w) with $\zeta=\log\dot{w}/\bar{k}\in C_b^1(0,T;\mathbb{R}^\mathcal{R})$, we approximate $\zeta_\delta:=\zeta\eta_\delta$ where η_δ is the usual compactly supported function (3.5). Clearly (c,w) satisfies the perturbed equation below, and we define, for each $\delta > 0$ the path (c_δ, w_δ) as the solution of the second perturbed equation:

$$\begin{cases} \dot c(t) = \Gamma \dot w(t), \\ \dot w(t) = \bar k^{(r)}\big(c(t)\big) e^{\zeta^{(r)}(t)}, \end{cases} \qquad \begin{cases} \dot c_\delta(t) = \Gamma \dot w_\delta(t), \\ \dot w_\delta(t) = \bar k^{(r)}\big(c_\delta(t)\big) e^{\zeta^{(r)}_\delta(t)}, \end{cases}$$

both under the same initial conditions $(c(0), 0)$.

Let us now introduce the matrix norm,

$$\|\Gamma\| := \max_{r\in\mathcal{R}} \left|\gamma^{(r)}\right|.$$

To prove convergence (i) we first estimate for any $0 \le t \le T$,

$$\begin{aligned} &|\dot w_\delta(t) - \dot w(t)| \\ &\le \sum_{r\in\mathcal{R}} \left| \bar k^{(r)}\big(c_\delta(t)\big)e^{\zeta_\delta^{(r)}(t)} - \bar k^{(r)}\big(c(t)\big)e^{\zeta_\delta^{(r)}(t)} \right| + \sum_{r\in\mathcal{R}} \left| \bar k^{(r)}\big(c(t)\big)e^{\zeta_\delta^{(r)}(t)} - \bar k^{(r)}\big(c(t)\big)e^{\zeta^{(r)}(t)} \right| \\ &\le \mathrm{Lip}(\bar k)\|\Gamma\| e^{\|\zeta\|_{L^\infty}} \int_0^t |\dot w_\delta(\hat t) - \dot w(\hat t)|\, d\hat t + \left(\sup_{\hat c\in\mathscr{S}(c(0))} \sum_{r\in\mathcal{R}} \bar k^{(r)}(\hat c) \right) \max_{r\in\mathcal{R}} \left| e^{\zeta_\delta^{(r)}(t)} - e^{\zeta^{(r)}(t)} \right|. \end{aligned}$$

From (3.5) one sees that $\zeta_\delta(t) = \zeta(t)$ except on two intervals each with length no more than 2δ. Gronwall's inequality yields

$$|\dot w_\delta(t) - \dot w(t)| \le \left(\sup_{\hat c\in\mathscr{S}(c(0))} \sum_{r\in\mathcal{R}} \bar k^{(r)}(\hat c) \right) \underbrace{\int_0^t |e^{\zeta_\delta(\hat t)} - e^{\zeta(\hat t)}|\, d\hat t}_{\le 4\delta \exp\left(\|\zeta\|_{L^\infty}\right)} e^{\mathrm{Lip}(\bar k)\|\Gamma\| e^{\|\zeta\|_{L^\infty}} t} \tag{3.12}$$

and so $w_\delta \to w$ in $W^{1,\infty}(0,T;\mathbb{R}_+^{\mathcal{R}})$, and by boundedness of the operator Γ also $c_\delta \to c$ in $W^{1,\infty}(0,T;\mathbb{R}_+^{\mathcal{Y}})$.

For the convergence (i) we only need to prove convergence of the dynamic rate $\mathscr{J}$: the initial conditions are identical. Indeed, by dominated convergence together with (3.12) and $\zeta_\delta \le \zeta$,

$$\mathscr{J}(c_\delta, w_\delta) = G(c_\delta, w_\delta, \zeta_\delta) \to G(c, w, \zeta) = \mathscr{J}(c_\delta, w_\delta).$$

□

Corollary 3.10. *Let $\mu^{(V)}$ satisfy Assumption 2.3 and $\bar k$ satisfy Assumptions 2.2*(iii), (iv), (v) *and* (vi). *Given $(c, w) \in \mathrm{BV}\big(0, T; \mathbb{R}_+^{\mathcal{Y}} \times \mathbb{R}_+^{\mathcal{R}}\big)$ such that $\mathscr{J}(c, w) < \infty$, there exists a sequence $(c_\delta, w_\delta)_\delta \subset \mathscr{A}$ such that:*

(i) $c_\delta(0) \to c(0)$ *and* $(c_\delta, w_\delta) \rightrightarrows (c, w)$ *as* $\delta \to 0$,

(ii) $\mathcal{I}_0(c_\delta(0)) + \mathscr{J}(c_\delta, w_\delta) \to \mathcal{I}_0(c(0)) + \mathscr{J}(c, w)$ *as* $\delta \to 0$.

4 Large deviations

We approach the proof of the main result, Theorem 1.1 with a fairly classical tilting approach with a twist. In Section 4.1 we prove exponential tightness, in Section 4.2 we prove the large deviations lower bound under initial distribution $\mu^{(V)}$, exploiting the approximation arguments from Section 3.2. In Section 4.3 we first prove the weak upper bound (i.e. on compact sets) for the conditional path measures, and then for the path measures under initial distribution $\mu^{(V)}$ again. The exponential tightness then guarantees that the lower bound also holds on closed sets, and that the rate functional is lower semicontinuous [DZ87, Lem. 1.2.18].

4.1 Exponential tightness

By a standard Chernoff argument, the balls

$$\mathscr{B}_m^{\mathrm{TV}} := \Big\{(c,w) \in \mathrm{BV}(0,T;\mathbb{R}_+^{\mathcal{Y}} \times \mathbb{R}_+^{\mathcal{R}}) : \|(c,w)\|_{L^1} + \|(\dot c,\dot w)\|_{\mathrm{TV}} \leq m\Big\} \tag{4.1}$$

can be used for the exponential tightness. However, in order to control the initial condition in the large-deviations upper bound we work with the cones (for some initial condition $\tilde c(0)$):

$$\mathscr{C}_{m,\epsilon} := \Big\{(c,w) \in \mathscr{B}_m^{\mathrm{TV}} : \big|\big(c(t),w(t)\big) - \big(\tilde c(0),0\big)\big| \leq \epsilon + tm \quad t\text{–a.e.}\Big\},$$

Lemma 4.1. *For any $m,\epsilon>0$ the cone $\mathscr{C}_{m,\epsilon}$ is hybrid-compact.*

Proof. The cone $\mathscr{C}_{m,\epsilon}$ is contained in the total-variation ball $\mathscr{B}_m^{\mathrm{TV}}$ and is clearly L^1-bounded, so it is relatively compact as discussed in Section 2.1. We thus need to show that $\mathscr{C}_{m,\epsilon}$ is hybrid closed. To that aim, take a hybrid-convergent net $(c^{(\omega)},w^{(\omega)})_\omega \subset \mathscr{C}_{m,\epsilon}$ with limit (c,w). By the weak-* lower semicontinuity of the TV-norm it follows that $\|(\dot c,\dot w)\|_{\mathrm{TV}} \leq m$. Moreover, the pointwise bound implies that,

$$\int_A \Big(\big|(c^{(\omega)}(t),w^{(\omega)}(t)) - (\tilde c(0),0)\big| - \epsilon - mt\Big)\,\mathrm{d}t \leq 0 \qquad \forall \text{ measurable } A \subset (0,T),$$

hence, after taking the limit in ω,

$$\int_A \Big(\big|(c(t),w(t)) - (\tilde c(0),0)\big| - \epsilon - mt\Big)\,\mathrm{d}t \leq 0 \qquad \forall \text{ measurable } A \subset (0,T),$$

which is equivalent to the pointwise bound $|(c(t),w(t)) - (\tilde c(0),0)| \leq \epsilon + mt$ for the limit. □

We first show exponential tightness for the conditional measures.

Lemma 4.2 (Uniform Exponential tightness of conditional measures)**.** *Let $\zeta \in C_{\mathrm{c}}(0,T;\mathbb{R}^{\mathcal{R}})$ and assume $\bar k^{(r)}$ is bounded on stoichiometric simplices (Assumption 2.2*(iv)*). Fix any convergent sequence $\frac1V\mathbb{N}_0^{\mathcal{R}} \ni \tilde c^{(V)}(0) \to \tilde c(0) \in \mathbb{R}_+^{\mathcal{R}}$ and let $\tilde{\mathbb{P}}_\zeta^{(V)}$ be the law of the Markov process with generator $\mathcal{Q}_{\zeta,t}^{(V)}$ and initial distribution $\delta_{c^{(V)}(0)}$. Then for any ϵ and $\eta>0$ there exists an m (not depending on the choice $\tilde c(0)$) such that*

$$\frac1V \log \tilde{\mathbb{P}}_\zeta^{(V)}\big(\mathscr{C}_{m,\epsilon}^{\mathsf{c}}\big) \leq -\eta.$$

Proof. For a $\delta>0$ to be determined later, define the set (see Figure 2):

$$\Sigma_{\delta,\epsilon} := \Big\{(c,w) \in \mathrm{BV}(0,T;\mathbb{R}^{\mathcal{Y}} \times \mathbb{R}_+^{\mathcal{R}}) : \big|(c(t),w(t)) - (\tilde c(0),0)\big| \leq \sigma_{\delta,\epsilon}(t) \quad t-æ\Big\},$$

$$\sigma_{\delta,\epsilon}(t) := \epsilon + \sum_{l=1}^{\lfloor T/\delta\rfloor} \tfrac12 \epsilon l \mathbb{1}_{[l\delta,(l+1)\delta)}(t).$$

Then $\mathscr{C}_{\epsilon/(2\delta),\epsilon} \supset \Sigma_{\delta,\epsilon} \cap \mathscr{B}_{\epsilon/(2\delta)}^{\mathrm{TV}}$ and so it suffices to prove that for any $\eta>0$ we can find

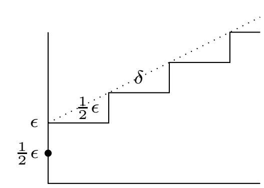


Figure 2: the function $\sigma_{\delta,\epsilon}(t) \leq \epsilon + \frac{\epsilon}{2\delta}t$.

$m,\delta>0$ such that

$$\limsup_{V\to\infty} \tfrac1n \log \tilde{\mathbb{P}}^{(V)}(\Sigma_{\delta,\epsilon}^{\mathsf{c}}) \leq -\eta \qquad \textit{and} \tag{4.2}$$

$$\limsup_{V\to\infty} \tfrac1n \log \tilde{\mathbb{P}}^{(V)}(\mathscr{B}_m^{\mathrm{TV}\,\mathsf{c}}) \leq -\eta. \tag{4.3}$$

Observe that by the convergence of the initial condition, for V sufficiently large and $\tilde{\mathbb{P}}^{(V)}$-almost surely,

$$|(c(0), w(0)) - (\tilde{c}(0), 0)| \leq \tfrac{1}{2}\epsilon. \tag{4.4}$$

To prove (4.3), observe that the Markov jump process $\sum_{r\in\mathcal{R}} W^{(V,r)}(t)$ is bounded by a Poisson process $\frac{1}{V}N_{V\lambda}(t)$ with $\lambda := e^{\|\Gamma\|\|\zeta\|_\infty} \sup_{\mathscr{S}_{1/2\epsilon}(\tilde{c}(0))} \sum_{r\in\mathcal{R}} 1 + \bar{k}^{(r)} < \infty$ due to Assumptions 2.2(ii) and (iv). A standard Chernoff bound therefore yields

$$\begin{aligned}\tilde{\mathbb{P}}^{(V)}(\mathscr{B}_m^{\mathrm{TV}\mathsf{c}}) &= \tilde{\mathbb{P}}^{(V)}\big(\{\textstyle\sum_{r\in\mathcal{R}} W^{(V,r)}(T) > m\}\big)\\ &\leq \mathrm{Prob}\,\big(N_{V\lambda}(T) > Vm\big) \leq e^{V\lambda TV - nm} = e^{-V\eta},\end{aligned}$$

if we choose $m := \lambda T e + \eta$.

We now prove (4.2). Because of (4.4) we may assume that for any $(c,w) \in \Sigma_{\delta,\epsilon}^{\mathsf{c}}$ there exists an interval $(l\delta, (l+1)\delta)$ on which the process has jumped more than $\frac{1}{2}\epsilon$. Since the norm of each jump is bounded from below by $\frac{1}{V}$ (the W-coordinate always jumps at least that length) we can estimate:

$$\begin{aligned}\tilde{\mathbb{P}}^{(V)}(\Sigma_{\delta,\epsilon}^{\mathsf{c}}) &\leq \tilde{\mathbb{P}}^{(V)}\Big(\textstyle\bigcup_{l=1}^{\lfloor T/\delta\rfloor} \big\{\frac{1}{V}\sum_{r\in\mathcal{R}} VW^{(V,r)}((l+1)\delta) - VW^{(V,r)}(l\delta) > \frac{\epsilon}{2}\big\}\Big)\\ &\leq \sum_{l=1}^{\lfloor T/\delta\rfloor} \Big(\big\{\textstyle\frac{1}{V}\sum_{r\in\mathcal{R}} VW^{(V,r)}((l+1)\delta) - VW^{(V,r)}(l\delta) > \frac{\epsilon}{2}\big\}\Big)\\ &\leq \frac{T}{\delta}\,\mathrm{Prob}\,\big(N_{V\lambda}(\delta) > \tfrac{V\epsilon}{2}\big)\\ &\leq \frac{T}{\delta} e^{-Vs(\epsilon/2|\lambda\delta)},\end{aligned}$$

where the latter is found by first applying a Chernoff bound to $\mathrm{Prob}\,\big(aN_{V\lambda}(\delta) > \frac{aV\epsilon}{2}\big)$ for arbitrary $a > 0$ and then minimising over a. With the choice

$$\delta := \frac{\epsilon}{2\lambda}\exp\big(-\frac{2\eta}{\epsilon} - 1\big),$$

we find $s(\epsilon/2|\lambda\delta) = \eta + \frac{\epsilon}{2}e^{-2\eta/\epsilon-1} \geq \eta$ which proves (4.2). □

From [Big04, Prop. 6] we now immediately obtain exponential tightness under the initial distribution $\mu^{(V)}$:

Corollary 4.3 (Exponential tightness)**.** *Let $\zeta \in C_c(0,T;\mathbb{R}^{\mathcal{R}})$, assume $\bar{k}^{(r)}$ is bounded on stoichiometric simplices (Assumption 2.2*(iv) *and let $\mu^{(V)}$ be exponentially tight (Assumption 2.3*(iv)*). Then $\mathbb{P}_\zeta^{(V)}$ (under initial distribution $\mu^{(V)}$) is exponentially tight.*

Remark 4.4. The results in this section apply in particular to the $\mathbb{P}^{(V)} = \mathbb{P}_0^{(V)}$. □

4.2 Lower bound

Proposition 4.5. *Let $\mu^{(V)}$ satisfy Assumption 2.3 and $\bar{k}$ satisfy Assumptions 2.2. For any hybrid-open set $\mathscr{O} \subset \mathrm{BV}\big(0,T;\mathbb{R}^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}}\big)$,*

$$\liminf_{V\to\infty} \frac{1}{V}\log \mathbb{P}^{(V)}(\mathscr{O}) \geq -\inf_{(c,w)\in\mathscr{O}} \mathcal{I}_0(c(0)) + \mathcal{J}(c,w).$$

Proof. Recall the definition of the set $\mathscr{A}$ in (3.6). Choose an arbitrary hybrid-open set $\mathscr{O} \subset \mathrm{BV}(0,T;\mathbb{R}^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}})$. From Lemma 4.6 proven below, it follows that

$$\liminf_{V\to\infty} \frac{1}{V}\log \mathbb{P}^{(V)}(\mathscr{O}) \geq -\inf_{(c,w)\in\mathscr{O}\cap\mathscr{A}} \mathcal{I}_0(c(0)) + \mathcal{J}(c,w).$$

By Corollary 3.10 it then follows that

$$\inf_{(c,w)\in\mathscr{O}} \mathcal{I}_0(c(0)) + \mathcal{J}(c,w) = \inf_{(c,w)\in\mathscr{O}\cap\mathscr{A}} \mathcal{I}_0(c(0)) + \mathcal{J}(c,w).$$

□

For the lower bound it thus remains to prove the following:

Lemma 4.6. *Let Assumption 2.2 on the rates and Assumption 2.3 on the initial distribution hold. Let $\mathscr{O} \subset \mathrm{BV}\big(0,T;\mathbb{R}^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}}\big)$ be any hybrid-open set, and $\mathscr{A}$ be the set* (3.6)*. Then for any $(c,w) \in \mathscr{O} \cap \mathscr{A}$,*

$$\liminf_{V\to\infty} \frac{1}{V} \log \mathbb{P}^{(V)}(\mathscr{O}) \geq -\mathcal{I}_0\big(c(0)\big) - \mathcal{J}(c,w). \tag{4.5}$$

Proof. Take a pair $(c,w) \in \mathscr{O} \cap \mathscr{A}$, and let $\zeta := \log \dot{w}/\bar{k}(c) \in C_c^1\big(0,T;\mathbb{R}^{\mathcal{R}}\big)$ and $z \in \partial\mathcal{I}_0\big(c(0)\big)$, which is non-empty by Assumption 2.3. Without loss of generality we assume that $\mathcal{I}_0(c(0)) + \mathcal{J}(c,w) < \infty$.

We also define a perturbed initial distribution on $\mathbb{R}^{\mathcal{Y}}$ by setting:

$$\mu_z^{(V)}\big(d\tilde{c}(0)\big) = \mathrm{e}^{Vz\cdot\tilde{c}(0) - V\Lambda^{(V)}(z)}\mu^{(V)}\big(d\tilde{c}(0)\big), \qquad \text{with} \quad \Lambda^{(V)}(z) := \tfrac{1}{V}\log\int \mathrm{e}^{Vz\cdot\tilde{c}(0)}\mu^{(V)}\big(d\tilde{c}(0)\big).$$

By Assumption 2.3, we can apply Varadhan's Lemma [DZ87, Th. 4.3.1], and so, combined with the assumption $z \in \partial\mathcal{I}_0\big(c(0)\big)$,

$$\lim_{V\to\infty} \Lambda^{(V)}(z) = \sup_{\tilde{c}(0)\in\mathbb{R}_+^{\mathcal{Y}}} z\cdot\tilde{c}(0) - \mathcal{I}_0\big(\tilde{c}(0)\big) = z\cdot c(0) - \mathcal{I}_0\big(c(0)\big) =: \Lambda(z), \tag{4.6}$$

Since we assumed that $\mathcal{I}_0\big(c(0)\big)$ is finite, it follows that (at least for sufficiently large V), the value $\Lambda^{(V)}(z)$ is finite. Naturally $\mathrm{e}^{V\Lambda^{(V)}(z)}$ is simply a normalisation factor so that the perturbed $\mu_z^{(V)}$ is a probability measure. We can now define the perturbed path measure

$$\mathbb{P}_{\zeta,z}^{(V)}(\mathrm{d}w'\,\mathrm{d}c') := \int \mathbb{P}_\zeta^{(V)}\big(\mathrm{d}w'\,\mathrm{d}c' \mid c'(0) = \tilde{c}(0)\big)\mu_z^{(V)}\big(d\tilde{c}(0)\big).$$

The next step is to apply Theorem A.3 to see that

$$\log\frac{\mathrm{d}\mathbb{P}^{(V)}}{\mathrm{d}\mathbb{P}_{\zeta,z}^{(V)}}\big(c',w'\big) = -VG^{(V)}(c',w',\zeta) - Vz\cdot c'(0) + V\Lambda^{(V)}(z). \tag{4.7}$$

When checking the applicability of the results from the appendix one may take $K_{(c',w')} = \mathscr{S}(c')\times\mathbb{R}_+^{\mathcal{R}}$. To establish Assumption A.3 one observes that $\sup_{t\in(0,T)}|c'(t),w'(t)|$ is bounded by $|c'(0)|$ plus a constant times the number of jumps up to time T, and that under $\mathbb{P}^{(V)}$ the number of jumps is stochastically dominated by a Poisson random variable with finite expectation due to Assumption 2.2 parts (ii)&(iv).

We now apply a standard tilting argument with respect to this measure. We first introduce the sets, for some arbitrary small $\epsilon > 0$ (recall that (c,w) is already fixed),

$$\mathcal{G}_\epsilon^\zeta = \mathcal{G}_\epsilon^\zeta[c,w] := \Big\{(c',w') \in \mathrm{BV}\left(0,T;\mathbb{R}^{\mathcal{Y}}\times\mathbb{R}^{\mathcal{R}}\right) : \big|G(c',w',\zeta) - G(c,w,\zeta)\big| < \epsilon\Big\}, \quad \text{and} \tag{4.8}$$

$$\mathscr{B}_\epsilon = \mathscr{B}_\epsilon[c(0)] := \big\{c'(0) \in \mathbb{R}_+^{\mathcal{Y}} : |c'(0) - c(0)| < \epsilon\big\}, \tag{4.9}$$

and let $\pi_0[c'] := c'(0)$. Although $\mathcal{G}_\epsilon^\zeta$ is not restricted to the positive cone, the probabilities are of course concentrated on non-negative concentrations and fluxes. Using (4.7)

$$\frac{1}{V}\log\mathbb{P}^{(V)}\big(\mathscr{O}\big) \geq \inf_{\substack{(c',w')\in\\ \mathscr{O}\cap\mathcal{G}_\epsilon^\zeta\cap\pi_0^{-1}[\mathscr{B}_\epsilon]}} \Big[-z\cdot c'(0) + \Lambda^{(V)}(z) - G^{(V)}(c',w',\zeta)\Big] + \frac{1}{V}\log\mathbb{P}_{\zeta,z}^{(V)}\big(\mathscr{O}\cap\mathcal{G}_\epsilon^\zeta\big), \tag{4.10}$$

where

$$G^{(V)}(c',w',\zeta) := \int_0^T \big[\zeta(t)\cdot\dot{w}'(\mathrm{d}t) - \sum_{r\in\mathcal{R}}\tfrac{1}{V}k^{(V,r)}\big(c'(t)\big)\big(e^{\zeta^{(r)}}(t) - 1\big)\big]\,\mathrm{d}t, \tag{4.11}$$

The first term is bounded by $-z\cdot\tilde{c}(0) \geq -z\cdot c(0) - \epsilon$ by definition of $\mathscr{B}_\epsilon$; for the second term we use (4.6) so that

$$\big|\Lambda^{(V)}(z) - \Lambda(z)\big| < \epsilon. \tag{4.12}$$

for V sufficiently large. For the third term we estimate,

$$\begin{aligned}
&\big|G^{(V)}(c',w',\zeta) - G(c,w,\zeta)\big| \leq \\
&\qquad \sup_{(c',w')\in\pi_0^{-1}[\mathscr{B}_\epsilon]} \big|G^{(V)}(c',w',\zeta) - G(c',w',\zeta)\big| + \sup_{(c',w')\in\mathcal{G}_\epsilon^\zeta} \big|G(c',w',\zeta) - G(c,w,\zeta)\big| \\
&\qquad\qquad \leq T(\mathrm{e}^{\|\zeta\|_{L^\infty}}+1) \sup_{c'\in\mathscr{S}_\epsilon(c(0))} \sum_{r\in\mathcal{R}} \big|\tfrac{1}{V}k^{(V,r)}(c') - \bar{k}^{(r)}(c')\big| + \epsilon \\
&\qquad\qquad\qquad \leq \mathit{const}\,\epsilon,
\end{aligned} \tag{4.13}$$

for sufficiently large V because of Assumption 2.2(ii). For the last term in (4.10) we use Proposition 2.4 together with the Portemanteau Theorem:

$$\begin{aligned}
\liminf_{V\to\infty} \mathbb{P}^{(V)}_{\zeta,z}\big(\mathscr{O}\cap\mathcal{G}_\epsilon^\zeta\cap\pi_0^{-1}[\mathscr{B}_\epsilon]\big) &\geq \liminf_{V\to\infty} \mathbb{P}^{(V)}_{\zeta,z}\big(\mathscr{O}\cap\mathcal{G}_\epsilon^\zeta\big) - \limsup_{V\to\infty} \mathbb{P}^{(V)}_{\zeta,z}\big(\pi_0^{-1}[\mathscr{B}_\epsilon]^\mathsf{c}\big) \\
&= \liminf_{V\to\infty} \mathbb{P}^{(V)}_{\zeta,z}\big(\mathscr{O}\cap\mathcal{G}_\epsilon^\zeta\big) - \limsup_{V\to\infty} \mu_z^{(V)}\big(\mathscr{B}_\epsilon^\mathsf{c}\big) \geq 1,
\end{aligned}$$

which is valid since $\mathscr{O}\cap\mathcal{G}_\epsilon^\zeta$ is hybrid-open by the continuity of $(c,w)\mapsto G(c,w,\zeta)$ (recall $\zeta\in C_c^1(0,T;\mathbb{R}^\mathcal{R})$), and $\mathscr{B}_\epsilon^\mathsf{c}$ is closed in $\mathbb{R}^\mathcal{Y}$.

Putting all these estimates and convergence results together we find from (4.10) that

$$\liminf_{V\to\infty} \frac{1}{V}\log\mathbb{P}^{(V)}\big(\mathscr{O}\big) \geq -z\cdot c(0) + \Lambda^{(V)}(z) - G(c,w,\zeta) - \mathit{const}\,\epsilon = -\mathcal{I}_0\big(c(0)\big) - \mathcal{J}(c,w) - \mathit{const}\,\epsilon,$$

as ζ and z were chosen to make the final equality true, assuming convexity of $\mathcal{I}_0$. This proves the claim since ϵ was arbitrary. □

4.3 Upper bound

For the upper bound we work first with a deterministic initial condition and then an argument of Biggins' [Big04] to deduce the upper bound for the 'mixture'.

Lemma 4.7. *Let $\bar{k}$ satisfy Assumptions 2.2*(iii),(ii),(iv)*, and fix any convergent sequence $\tilde{c}^{(V)}(0)\to\tilde{c}(0)$ in $\mathbb{R}_+^\mathcal{Y}$. Let $\tilde{\mathbb{P}}^{(V)}$ be the law of the Markov process with deterministic initial condition $\tilde{c}^{(V)}(0)$ and the dynamics given by* (1.2)*. Then for any hybrid-compact set $\mathscr{K}\subset \mathrm{BV}\big(0,T;\mathbb{R}^\mathcal{Y}\times\mathbb{R}^\mathcal{R}\big)$,*

$$\limsup_{V\to\infty}\frac{1}{V}\log\tilde{\mathbb{P}}^{(V)}(\mathscr{K}) \leq -\inf_{\substack{(c,w)\in\mathscr{K}\\ c(0)=\tilde{c}(0)}} \mathcal{J}(c,w).$$

Proof. We use an adaptation of the usual covering technique as in the proof of the Gärtner-Ellis Theorem [DZ87, Th. 4.5.3]. Fix a convergent sequence $\tilde{c}^{(V)}(0)\to\tilde{c}(0)$ in $\mathbb{R}_+^\mathcal{Y}$, a hybrid-compact set $\mathscr{K}\subset\mathrm{BV}\big(0,T;\mathbb{R}^\mathcal{Y}\times\mathbb{R}^\mathcal{R}\big)$, and arbitrary $\epsilon>0$.

To control the initial condition, we use the compact cones $\mathscr{C}_{m,\epsilon}$ from the exponential tightness, Proposition 4.3. In this proof we will take $m>0$ such that $\limsup_{V\to\infty}\frac{1}{V}\log\tilde{\mathbb{P}}^{(V)}\big(\mathscr{C}_{m,\epsilon}^\mathsf{c}\big) < -1/\epsilon$. Note that for V sufficiently large, $\tilde{c}^{(V)}(0)\in\mathscr{B}_\epsilon[\tilde{c}(0)] = \pi_0\mathscr{C}_{m,\epsilon}$.

By Proposition 3.5 we can find, for any $(c,w)\in\mathrm{BV}\big(0,T;\mathbb{R}^\mathcal{Y}\times\mathbb{R}^\mathcal{R}\big)$ and $\epsilon>0$, a $\zeta[c,w]\in C_c^1(0,T;\mathbb{R}^\mathcal{R})$ such that $G(c,w,\zeta[c,w])\geq\mathcal{J}(c,w)-\epsilon$. Then the sets $\mathcal{G}_\epsilon^\zeta[c,w]$ from (4.8) form an open covering $\bigcup_{(c,w)\in\mathscr{K}}\mathcal{G}_\epsilon^{\zeta[c,w]}(c,w)\supset\mathscr{K}\cap\mathscr{C}_{m,\epsilon}$, and hence there exists a finite subset $(c^{(n)},w^{(n)})_{n=1,\ldots,N}\subset\mathscr{K}$ such that $\bigcup_{n=1,\ldots,N}\mathcal{G}_\epsilon^{\zeta[c^{(n)},w^{(n)}]}(c^{(n)},w^{(n)})\supset\mathscr{K}\cap\mathscr{C}_{m,\epsilon}$.

Let $\tilde{\mathbb{P}}_\zeta^{(V)}$ be the law of the Markov process with deterministic initial condition $\tilde{c}^{(V)}(0)$ and the dynamics given by the perturbed generator (2.3) For each $n=1,\ldots,N$ we find for sufficiently large V (here we only intersect with the initial balls in order to employ (4.13) in the end of this calculation),

$$\frac{1}{V}\log\tilde{\mathbb{P}}^{(V)}\big(\mathcal{G}_\epsilon^{\zeta[c^{(n)},w^{(n)}]}(c^{(n)},w^{(n)})\big)=\frac{1}{V}\log\tilde{\mathbb{P}}^{(V)}\big(\mathcal{G}_\epsilon^{\zeta[c^{(n)},w^{(n)}]}(c^{(n)},w^{(n)})\cap\pi_0^{-1}\mathscr{B}_\epsilon[\tilde{c}(0)]\big)$$

$$\le \sup_{(c,w)\in\mathcal{G}_\epsilon^{\zeta[c^{(n)},w^{(n)}]}(c^{(n)},w^{(n)})\cap\pi_0^{-1}\mathscr{B}_\epsilon[\tilde{c}(0)]}\frac{1}{V}\log\frac{d\tilde{\mathbb{P}}^{(V)}}{d\tilde{\mathbb{P}}^{(V)}_{\zeta[c^{(n)},w^{(n)}]}}(c,w)$$

$$+\underbrace{\frac{1}{V}\log\tilde{\mathbb{P}}^{(V)}_{\zeta[c^{(n)},w^{(n)}]}\big(\mathcal{G}_\epsilon^{\zeta[c^{(n)},w^{(n)}]}(c^{(n)},w^{(n)})\big)}_{\le 0}$$

$$\overset{\text{Th. A.3}}{\le}\sup_{(c,w)\in\mathcal{G}_\epsilon^{\zeta[c^{(n)},w^{(n)}]}(c^{(n)},w^{(n)})\cap\pi_0^{-1}\mathscr{B}_\epsilon[\tilde{c}(0)]}-G^{(V)}(c,w,\zeta[c^{(n)},w^{(n)}])$$

$$\overset{(4.13)}{\le}\sup_{(c,w)\in\mathcal{G}_\epsilon^{\zeta[c^{(n)},w^{(n)}]}(c^{(n)},w^{(n)})}-G(c,w,\zeta[c^{(n)},w^{(n)}])+\mathit{const}\,\epsilon$$

$$\overset{(4.8)}{\le}-G(c^{(n)},w^{(n)},\zeta[c^{(n)},w^{(n)}])+\mathit{const}\,\epsilon.$$

Because of the finiteness of the covering we can now use the Laplace Principle:

$$\begin{aligned}\limsup_{V\to\infty}\frac{1}{V}\log\tilde{\mathbb{P}}^{(V)}(\mathscr{K})&\le\limsup_{V\to\infty}\frac{1}{V}\log\big(\tilde{\mathbb{P}}^{(V)}(\mathscr{K}\cap\mathscr{C}_{m,\epsilon})+\tilde{\mathbb{P}}^{(V)}(\mathscr{C}^{\mathsf{c}}_{m,\epsilon})\big)\\
&\le\max_{n=1,\dots,N}\limsup_{V\to\infty}\frac{1}{V}\log\tilde{\mathbb{P}}^{(V)}(\mathcal{G}_\epsilon^{\zeta[c^{(n)},w^{(n)}]})\vee-\frac{1}{\epsilon}\\
&\le\max_{n=1,\dots,N}\big(-G(c^{(n)},w^{(n)},\zeta[c^{(n)},w^{(n)}])+\mathit{const}\,\epsilon\big)\vee-\frac{1}{\epsilon}\\
&\le\max_{n=1,\dots,N}\big(-\mathcal{J}(c^{(n)},w^{(n)})+\mathit{const}\,\epsilon\big)\vee-\frac{1}{\epsilon}\\
&\le\Big(-\inf_{(c,w)\in\mathscr{K}\cap\mathscr{C}_{m,\epsilon}}\mathcal{J}(c,w)+\mathit{const}\,\epsilon\Big)\vee-\frac{1}{\epsilon}\\
&\le\Big(-\inf_{\substack{(c,w)\in\mathscr{K}:\\ c(0)\in\mathscr{B}_\epsilon[\tilde{c}(0)]}}\mathcal{J}(c,w)+\mathit{const}\,\epsilon\Big)\vee-\frac{1}{\epsilon}.\end{aligned}$$

This proves the claim as ϵ was chosen arbitrarily. □

We can now deduce the large-deviations upper bound for the mixture:

Corollary 4.8. *Let $\mu^{(V)}$ satisfy Assumption 2.3 and $\bar{k}$ satisfy Assumptions 2.2. For any hybrid-compact set $\mathscr{K}\subset\mathrm{BV}\big(0,T;\mathbb{R}^{\mathcal{Y}}\times\mathbb{R}^{\mathcal{R}}\big)$,*

$$\liminf_{V\to\infty}\frac{1}{V}\log\mathbb{P}^{(V)}(\mathscr{K})\ge-\inf_{(c,w)\in\mathscr{K}}\mathcal{I}_0(c(0))+\mathcal{J}(c,w).$$

Proof. By Assumption 2.3 and Lemma 4.7 one can apply [Big04, Lemma 12] noting that the proof in [Big04] only uses the upper bound proved in Lemma 4.7 not a full LDP. □

A A change-of-measure result for linear test functionals on jump processes

Changes of measure are central to the proof of the large deviations principle presented in this work. This appendix arose out of the need to clarify under exactly what technical conditions [KL99, Appendix 1, Prop. 7.3] could be adapted to the setting of the present work, in particular so that functions of the form $x\mapsto\zeta\cdot x$ could be used since these are not bounded functions (although they are bounded linear operators). This boundedness restriction is avoided in [PR02], but functions used in the change of measure are no longer time dependent and the conditions are less explicit. Here the aim is to include unbounded, time dependent functions in the change of measure formula, but to give relatively explicit, sufficient conditions that can easily be checked using the model assumptions from the main part of the paper. In this endeavour the results are restricted to pure jump processes.

Let $\mathcal{X}$ be a Banach space, $T \in (0,\infty]$ and $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T)}\big)$ be a filtered probability space with canonical random variable $X : \Omega \to \Omega$, where

- Ω is a subset of the càdlàg functions $[0,T) \to \mathcal{X}$, with the convention $f(T) := f(T-)$ if $T < \infty$,
- $\mathcal{F}$ is the Borel σ-algebra generated by a separable topology on Ω and equal to the σ-algebra generated by the time evaluation functions $X \mapsto X(t)$.

Note that $T = \infty$ is allowed for now. The application in this paper is to the case $\Omega = \mathrm{BV}(0,T,\mathbb{R}^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}})$ with the hybrid topology, but this is not a necessary assumption.

We define the jump process through a given family of jump kernels $(\alpha_t(x,\cdot)_{t\in[0,T),x\in\mathcal{X}}$ where $\alpha_t(x,A)$ is the instantaneous jump rate at time t from $x \in \mathcal{X}$ into a measurable set $A \subset \mathcal{X}$, together with a given initial distribution μ. Let $\mathbb{P}$ be the law of this process, a probability measure on $(\Omega,\mathcal{F})$ and $\mathbb{E}$ the associated expectation operator.

We now define a class of test functions for which the associated propagators (a two-paramater semigroup of linear operators) are well-defined. To construct this set we will assume that there exists a family of measurable (not necessarily compact or bounded) subsets $(K_x)_x$ of $\mathcal{X}$ such that for all $x \in \mathcal{X}$:

- $x \in K_x$ and $\bigcup_{y\in K_x} K_y = K_x$,
- $\int_0^T \sup_{y\in K_x} \alpha_t(y,\mathcal{X})\mathrm{d}t < \infty$,
- $\sup_{t\in[0,T)} \sup_{y\in K_x} \alpha_t(y, \mathcal{X} \setminus K_x) = 0$.

This expresses the idea that the process started from x can never explode nor leave K_x. Then the propagators $(P_{s,t}f)(x) := \mathbb{E}\left[f(X(t))|X(s)=x\right]$ preserve the set

$$B_{\mathrm{K}}(\mathcal{X}) := \left\{ f : \mathcal{X} \to \mathbb{R} \text{ measurable, such that } \forall x \in \mathcal{X} \sup_{y\in K_x} |f(y)| < \infty \right\},$$

and satisfy $\frac{\mathrm{d}}{\mathrm{d}s}(P_{s,t}f)(x) = -(\mathcal{Q}_s P_{s,t} f)(x)$ with (time-dependent) generator

$$(\mathcal{Q}_t f)(x) := \int_X \left[f(y) - f(x)\right] \alpha_t(x, \mathrm{d}y).$$

We now make three additional assumptions under which the change-of-measure formula holds.

- there is a $\gamma > 0$ such that $|y-x| \le \gamma$ for all $x \in \mathcal{X}, t \in (0,T)$, and $\alpha_t(x,\cdot)$-æ y, (A.1)
- $\lim_{n\to\infty} \mathbb{P}(\tau_n < t) = 0$ for all $t \in (0,T), n \in \mathbb{N}$, where $\tau_n := \inf\{t\colon \alpha_t(X(t),\mathcal{X}) \ge n\}$, (A.2)
- $\mathbb{E}\left[Z^\beta(t)\right] < \infty$ for all $t \in (0,T), \beta > 0$, where (A.3)

$$Z^\beta(t) := \exp(\beta|X(0)|) + \exp(\beta|X(t)|) \\ + \int_0^t \exp(\beta|X(s)|)\,\beta|X(s)|\,\mathrm{d}s + \int_0^t \exp(\beta|X(s)|)\,\beta\alpha_s(X(s),\mathcal{X})\,\mathrm{d}s.$$

The next result is a variation on [KL99, Appendix 1, Lem. 5.1]:

Proposition A.1. *Let $f\colon [0,T) \times \mathcal{X} \to \mathbb{R}$ be bounded, absolutely continuous in t and measurable in x, with measurable, uniformly bounded derivative $\partial_t f(t,x)$. Then under Assumptions* (A.2) *&* (A.1)*,*

$$M^f(t) := f(t,X(t)) - f(0,X(0)) - \int_0^t \big((\partial_s + \mathcal{Q}_s) f\big)(s,X(s))\,\mathrm{d}s$$

is a Martingale in the filtration $(\mathcal{F}_t)_{t\ge0}$ generated by $X(t)$.

Proof. In the case that f does not depend on time and $\sup_{t,x} \alpha_t(x,\mathcal{X}) < \infty$ the result follows from [EK86, Ch. 4 Sect. 7]. The additional term ∂_s is added for time-dependent test functions due to a chain rule. By approximating by the process stopped at τ_n and using Assumptions (A.2)&(A.3) one can remove the boundedness assumption on α. ◻

Lemma A.2. *Under Assumptions* (A.1), (A.2) *&* (A.3)*, the conclusion of Proposition A.1 is valid when* $f(t,x) = \zeta(t) \cdot x$ *and when* $f(t,x) = e^{\zeta(t)\cdot x}$*, in both cases for* $\zeta \in C^1_{\mathrm{b}}([0,T);\mathcal{X}^*)$ *where* $\mathcal{X}^*$ *is the Banach dual of* $\mathcal{X}$.

Proof. The exponential case is proved here; the linear case is similar. Let $\theta_n \in C^\infty(\mathbb{R})$ be such that $\theta_n(y) = y$ for $y \le n$, $\theta_n \le n+1$ and $0 \le \theta_n' \le 1$. Take an arbitrary $\zeta \in C^1_{\mathrm{b}}([0,\infty);\mathcal{X}^*)$ and set $f_n(t,x) = \exp(\theta_n(\zeta(t)\cdot x))$ so that Proposition A.1 can be applied to $M^{f_n}(t)$.

It follows from the definitions that for all t and x,

$$\lim_n f_n(t,x) = f(t,x) \quad \text{and} \quad \lim_n \partial_t f_n(t,x) = \partial_t f(t,x).$$

Because of Assumption (A.1) $\mathcal{Q}_t f$ is well defined and one can prove by dominated convergence that $\lim_n (\mathcal{Q}_t f_n)(t,x) = (\mathcal{Q}_t f)(t,x)$ for all t,x. Preparatory to further applications of dominated convergence we estimate

$$\begin{aligned} f_n(t,x) &\le \exp\left(\|\zeta\|_\infty |x|\right), \\ |\partial_t f_n(t,x)| &\le \exp\left(\|\zeta\|_\infty |x|\right) \left\|\dot{\zeta}\right\|_\infty |x|, \qquad \text{and} \\ |(\mathcal{Q}_t f_n)(t,x)| &\le \exp\left(\|\zeta\|_\infty |x|\right) \left(\exp\left(\|\zeta\|_\infty \gamma\right) + 1\right) \alpha_t(x,\mathcal{X}). \end{aligned}$$

With these estimates and Assumption (A.3) one checks $\lim_n M^{f_n}(t) = M^f(t)$ almost surely. Again using Assumption (A.3) one can find a $\beta > 0$ such that $\left|M^{f_n}(t)\right| \le Z^\beta(t)$ almost surely. By the conditional expectation form of the dominated convergence theorem, for $s < t$,

$$M^f(s) = \lim_n M^{f_n}(s) = \lim_n \mathbb{E}\left[M^{f_n}(t)|\mathcal{F}_s\right] = \mathbb{E}\left[\lim_n M^{f_n}(t)|\mathcal{F}_s\right] = \mathbb{E}\left[M^f(t)|\mathcal{F}_s\right].$$

□

Finally, for the exponential change of measure we will need a bounded time interval.

Theorem A.3. *Let* $T < \infty$*,* $\zeta \in C^1_b(0,T;\mathbb{R}^{\mathcal{R}})$*, and let Assumptions* (A.1), (A.2) *and* (A.3) *all hold. Suppose* $\mathbb{P}_\zeta$ *is the law of some process with paths in* Ω *and having initial distribution* μ*. Under* $\mathbb{P}_\zeta$*,* X *is a Markov process with generator*

$$(\mathcal{Q}_{\zeta,t} f)(x) = \int_{\mathcal{X}} [f(y) - f(x)]\, e^{\zeta(t)\cdot y - \zeta(t)\cdot x} \alpha_t(x, \mathrm{d}y)$$

if and only if

$$\log \frac{d\mathbb{P}_\zeta}{d\mathbb{P}}(X) = \zeta(T)\cdot X(T) - \zeta(0)\cdot X(0) - \int_0^T e^{-\zeta(t)\cdot X(t)} \left(\partial_t + \mathcal{Q}_t\right) e^{\zeta(t)\cdot X(t)} \mathrm{d}t. \tag{A.4}$$

Proof. We only need to show the direction " $\Longleftarrow$ "; the converse then follows immediately from the uniqueness of the generator. To this end define $\widehat{\mathbb{P}}_\zeta$ by (A.4) and let the associated expectation operator be $\widehat{\mathbb{E}}_\zeta$. We sketch a number of steps, similar to [KL99, Appendix 1, Sect. 7] and [PR02], by which it is shown that under $\widehat{\mathbb{P}}_\zeta$ X is Markov with generator $\mathcal{Q}_{\zeta,t}$.

1. Define for $t \in (0,T)$, the process

$$E(t) := \exp\left(\zeta(t)\cdot X(t) - \zeta(0)\cdot X(0) - \int_0^t e^{-\zeta(s)\cdot X(s)} \left(\partial_s + \mathcal{Q}_s\right) e^{\zeta(s)\cdot X(s)} \mathrm{d}s\right)$$

   and recall $E(T) = \lim_{t\nearrow T} E(t)$. By Lemma A.2 above, $E(t)$ is a strictly positive, mean-one $\mathbb{P}$-Martingale. One then shows that $\left.\frac{\mathrm{d}\widehat{\mathbb{P}}_\zeta}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t} = E(t)$ and $\left.\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\widehat{\mathbb{P}}_\zeta}\right|_{\mathcal{F}_t} = \frac{1}{E(t)}$.
2. For any $Y \in L^1(\Omega,\mathcal{F})$, using the definition of conditional expectation and the results from the previous point, it follows that $\widehat{\mathbb{E}}_\zeta\left[Y|\mathcal{F}_t\right] = \mathbb{E}\left[Y E(T)/E(t)|\mathcal{F}_t\right]$.
3. Next one can use the result from point 2 to show via conditional expectations under $\mathbb{P}$ and the $\mathbb{P}$-Markov property that for $t \ge s$ and any bounded and measurable $f: \mathcal{X} \to \mathbb{R}$, we have $\widehat{\mathbb{E}}_\zeta\left[f(X(t))|\mathcal{F}_s\right] = \widehat{\mathbb{E}}_\zeta\left[f(X(t))|\sigma(X(s))\right]$, and so X is $\widehat{\mathbb{P}}_\zeta$-Markov.
4. Finally, the propagators $(P^\zeta_{s,t} f)(x) := \widehat{\mathbb{E}}_\zeta\left[f(X(t))|X(s) = x\right]$ then satisfy $\frac{\mathrm{d}}{\mathrm{d}s}(P_{s,t} f)(x) = -(\mathcal{Q}_{\zeta,s} P_{s,t} f)(x)$. This implies that under $\widehat{\mathbb{P}}_\zeta$, X has the same finite dimensional distributions as the process with generator $\mathcal{Q}_{\zeta,t}$ and thus $\widehat{\mathbb{P}}_\zeta = \mathbb{P}_\zeta$.

□

Acknowledgements

This research has been funded by the Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1114 "Scaling Cascades in Complex Systems", Project C08.

References

[ADPZ11] S. Adams, N. Dirr, M. A. Peletier, and J. Zimmer. From a large-deviations principle to the Wasserstein gradient flow: a new micro-macro passage. *Communications in Mathematical Physics*, 307(3):791–815, 2011.

[AFP00] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford university press, Oxford, UK, 2000.

[AK11] D.F. Anderson and T.G Kurtz. *Design and Analysis of Biomolecular Circuits: Engineering Approaches to Systems and Synthetic Biology*, chapter Continuous time Markov chain models for chemical reaction networks, pages 1–44. Springer, 2011.

[BDSG+05] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim. Current fluctuations in stochastic lattice gases. *Phys. Rev. Lett.*, 94(3):030601, Jan 2005.

[BDSG+06] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim. Non equilibrium current fluctuations in stochastic lattice gases. *Journal of Statistical Physics*, 123(2):237–276, April 2006.

[Big04] J.D. Biggings. Large deviations for mixtures. *Electronic Communications in Probability*, 9:60–71, 2004.

[BL10] T. Bodineau and M. Lagouge. Current large deviations in a driven dissipative model. *Journal of Statistical Physics*, 139(2):201–218, April 2010.

[BL12] T. Bodineau and M. Lagouge. Large deviations of the empirical currents for a boundary-driven reaction diffusion model. *Ann. Appl. Probab.*, 22(6):2282–2319, December 2012.

[BMN09] M. Baiesi, C. Maes, and K. Netočný. Computation of current cumulants for small nonequilibrium systems. *Journal of Statistical Physics*, 135(1):57–75, 2009.

[Bog07] V.I. Bogachev. *Measure theory Vol. I and II.* Springer, Berlin, Germany, 2007.

[Bré83] H. Brézis. *Analyse fonctionelle, Théorie et applications (In French, English translation available).* Mason, Paris, France, 1983.

[DDR04] B. Derrida, B. Douçot, and P.-E. Roche. Current fluctuations in the one-dimensional symmetric exclusion process with open boundaries. *Journal of Statistical Physics*, 115(3):717–748, May 2004.

[Der07] B. Derrida. Non-equilibrium steady states: fluctuations and large deviations of the density and of the current. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(7):P07023, 2007.

[DEW91] P. Dupuis, R.S. Ellis, and A. Weiss. Large deviations for Markov processes with discontinuous statistics, I: General upper bounds. *The Annals of Probability*, 19(3):1280–1297, 1991.

[DK95] B. Djehiche and I. Kaj. The rate function for some measure-valued jump processes. *The annals of probability*, 23(3):1414–1438, 1995.

[DRW16] P. Dupuis, K. Ramanan, and W. Wu. Large deviation principle for finite-state mean field interacting particle systems. arXiv:1601.06219, 2016.

[DZ87] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic modelling and applied probability.* Springer, New York, NY, USA, 2nd edition, 1987.

[EK86] S.N. Ethier and T.G. Kurtz. *Markov processes – characterization and convergence.* John Wiley & Sons, Hoboken, NJ, USA, 1986.

[Eva02] L.C. Evans. *Partial Differential Equations*. American Mathematical Society, Providence, RI, USA, 2002.

[Fen94] S. Feng. Large deviations for empirical process of mean-field interacting particle system with unbounded jumps. *The Annals of Probability*, 22(4):1679–2274, 1994.

[HPR16] M. Heida, R.I.A. Patterson, and D.R.M. Renger. The space of bounded variation with infinite-dimensional codomain. Technical Report 2353, Weierstrass Institute, Berlin, 2016.

[Kal02] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[KL99] C. Kipnis and C. Landim. *Scaling limits of interacting particle systems*. Springer, Berlin-Heidelberg, Germany, 1999.

[Kra17] R. Kraaij. Flux large deviations of weakly interacting jump processes via well-posedness of an associated Hamilton-Jacobi equation. `https://arxiv.org/abs/1711.00274`, 2017.

[Kur70] T.G. Kurtz. Solutions of ordinary differential equations as limits of pure jump processes. *Journal of Applied Probability*, 7(1):49–58, 1970.

[Kur72] T. G. Kurtz. The relationship between stochastic and deterministic models for chemical reactions. *The Journal of Chemical Physics*, 57(7):2976–2978, 1972.

[Léo95] C. Léonard. Large deviations for long range interacting particle systems with jumps. *Annales de l'Institut Henri Poincaré, section B*, 31(2):289–323, 1995.

[LL01] E.H. Lieb and M. Loss. *Analysis*. American Mathematical Society, Providence, RI, USA, 2nd edition, 2001.

[LL15] T. Li and F. Lin. Large deviations for two scale chemical kinetic processes. *Communications in Mathematical Sciences*, 15(1):123–163, 2015.

[MPPR15] A. Mielke, R.I.A. Patterson, M.A. Peletier, and D.R.M. Renger. Non-equilibrium thermodynamic principles for nonlinear chemical reactions and systems with coagulation and fragmentation. *WIAS Preprint*, 2165, 2015.

[MPR14] A. Mielke, M.A. Peletier, and D.R.M. Renger. On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion. *Potential Analysis*, 41(4):1293–1327, 2014.

[Ons31a] L. Onsager. Reciprocal relations in irreversible processes I. *Phys. Rev.*, 37(4):405–426, Feb 1931.

[Ons31b] L. Onsager. Reciprocal relations in irreversible processes II. *Phys. Rev.*, 38(12):2265–2279, Dec 1931.

[PR02] Zbigniew Palmowski and Tomasz Rolski. A technique for exponential change of measure for markov processes. *Bernoulli*, 8(6):767–785, 12 2002.

[Ren17] D.R.M. Renger. Flux large deviations of independent and reacting particle systems, with implications for macroscopic fluctuation theory. WIAS Preprint No. 2375, 2017.

[SW95] A. Shwartz and A. Weiss. *Large deviations for performance analysis: queues, communications, and computing*. Chapman & Hall, London, UK, 1995.

[SW05] A. Shwartz and A. Weiss. Large deviations with diminishing rates. *Mathematics of Operations Research*, 30(2):281–310, 2005.