Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Dynamical large deviations of countable reaction networks under a weak reversibility condition

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

submitted: June 30, 2016

Weierstraß-Institut
Mohrenstraße 39
10117 Berlin
Germany
E-Mail: robert.patterson@wias-berlin.de
d.r.michiel.renger@wias-berlin.de

No. 2273
Berlin 2016

2010 *Mathematics Subject Classification.* 60F10, 60J25, 80A30, 82C22.

2008 *Physics and Astronomy Classification Scheme.* 82.40.Bj, 82.60.-s, 82.20.Db, 82.20.Fd.

Key words and phrases. Chemical reaction networks, Markov processes, Large deviations.

The authors thank Mark Peletier and Alexander Mielke for their helpful comments and discussions. This research has been partially funded by Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1114.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

A dynamic large deviations principle for a countable reaction network including coagulation–fragmentation models is proved. The rate function is represented as the infimal cost of the reaction fluxes and a minimiser for this variational problem is shown to exist. A weak reversibility condition is used to control the boundary behaviour and to guarantee a representation for the optimal fluxes via a Lagrange multiplier that can be used to construct the changes of measure used in standard tilting arguments. Reflecting the pure jump nature of the approximating processes, their paths are treated as elements of a BV function space.

1 Introduction

Since the initial works of Kurtz [Kur70, Kur72], reacting particle systems have been a major object of study, see also the survey [AK11] and the references therein. In these works it is proven that the concentration of particles converges to a deterministic limit as the number of particles goes to infinity. The next question is then how to prove a dynamical large deviation principle corresponding to this limit. It is important to know these large deviations for a number of reasons. Firstly, the large deviations can often be used to improve efficiency of rare event simulations [Sie76]. Secondly, the large deviation rate encodes thermodynamic properties of the macroscopic system, like the free energy functional that drives the system and sometimes the dissipation potential. Together these completely characterise the macroscopic evolution as a function of thermodynamic driving forces. In fact, we studied such connections in the article [MPPR15], which was also the main motivation behind the current study. As we discuss below, there are a number of large deviation results for chemical reactions in the literature. However, applications to modern biochemistry and coagulation-fragmentation type reactions show a practical need to extend the theory to large and complex reaction networks.

In this paper we study a network of reactions,

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

where $\mathcal{Y}$ is a countable space of species, $\mathcal{R}$ is a countable space of reactions. The numbers $\alpha_y^{(r)}, \beta_y^{(r)}$ are called the *stoichiometric coefficients* or *complexes*, and the reaction rates $\bar{k}^{(r)}$ depend on the concentration of the species in $\mathcal{Y}$.

Microscopic model. The reaction networks described above are commonly modelled by the following microscopic particle system. 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. Since the propensities $k^{(r,V)}$ depend on the particles through the empirical concentration only, $C^{(V)}(t)$ is a Markov jump process in $\mathbb{R}^{\mathcal{Y}}$ with generator

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

Convergence to Reaction Rate Equation. We will assume that the initial concentration $C^{(V)}(0)$ is deterministic and converges to a fixed initial concentration $c(0)$ as $V \to \infty$. As such, the volume V controls the order of the (changing) number of particles in the system. We will assume further that the scaled propensities $V^{-1}k^{(r,V)}(c)$ converge to the macroscopic reaction rates $\bar{k}^{(r)}(c)$ from (1.1). Then under the appropriate conditions, the process $C^{(V)}(t)$ converges as $V \to \infty$ to the deterministic concentration $c(t)$ that solves the *Reaction Rate Equation* [Kur70],

$$\dot{c}(t) = \sum_{r\in\mathcal{R}} \bar{k}^{(r)}\big(c(t)\big)\gamma^{(r)}. \tag{1.3}$$

Large deviations. In our main result, Theorem 5.1, we prove that the processes $C^{(V)}(t)|_{t=0}^{T}$ satisfy a large deviation principle as $V \to \infty$, with rate functional

$$I(c) := \sup_{\xi\in C_b^1(0,T;\mathfrak{l}^\infty(\mathcal{Y}))} \int_0^T \xi(t)\cdot\dot{c}(\mathrm{d}t) - \int_0^T H\big(c(t),\xi(t)\big)\,\mathrm{d}t, \tag{1.4}$$

where

$$H(c,\xi) := \sum_{r\in\mathcal{R}} \bar{k}^{(r)}(c)\big(e^{\xi\cdot\gamma^{(r)}} - 1\big).$$

Spaces and topologies. In our setting, concentrations $C^{(V)}(t), c(t)$ will always lie in the space of summable sequences $\mathfrak{l}^1(\mathcal{Y})$. Here, the countability of $\mathcal{Y}$ (and $\mathcal{R}$) is particularly useful because, although the weak and strong topologies on $\mathfrak{l}^1(\mathcal{Y})$ differ, they have precisely the same convergent sequences. We also exploit the fact that $\mathfrak{l}^1(\mathcal{Y})$ is the dual of $\mathfrak{c}_0(\mathcal{Y})$ with Banach-Alaoglu arguments. We briefly mention that instead of $\mathfrak{l}^1(\mathcal{Y})$ one could also work with the more general Orlicz space of sequences for which the rate functional is finite, as in [Léo95].

For the paths in $\mathfrak{l}^1(\mathcal{Y})$ we take a slightly different space than the usual Cadlag/Skorohod space, namely the space of *functions of bounded variation*. In this space any path c has a measure-valued derivative $\dot{c}$, which is used in (1.4). We equip the path space with (what we call) the *hybrid topology*. This topology is weaker than the norm topology but stronger than the weak-* topology [HPR]. It turns out that some large deviation results, in particular the exponential tightness, have a very natural and explicit construction in this topology. Our large-deviation result also holds in the Skorohod space, however controlling the variation requires more complicated estimates.

Literature and techniques. In [Fen94] and [Léo95], dynamical large deviation principles are proved for particle systems with a countable, respectively continuous range of species and reactions. In these models as in ours, the reaction rates depend on the empirical measure of the entire particle system—there is a mean-field interaction. However, only jump-type unimolecular reactions are considered, meaning that each reaction only transforms one particle into one other. The model of [Fen94] includes independent jumps and unbounded mean-field interaction, but the mean-field interaction is of a very specific type. In the model of [Léo95] each particle is permanently assigned to a spatial lattice site, and the jump rates in an additional chemical coordinate are bounded. In the paper [DK95], the results of [Fen94, Léo95] are combined into a more general framework, still for unimolecular reactions. These works follow the ideas of [DG87], where one first proves a Sanov-type large deviation principle of the empirical measure on the space of paths, and then transforms back to the path of the empirical measure. However, the Sanov-type large-deviation principle is based on an explicit Radon-Nykodym derivative for the original process with respect to a process of independent particles, which fails for reactions involving multiple molecules.

In [SW95, Ch. 5] a dynamical large deviation principle is proven for systems with more general chemical reactions, but for a finite number of species and reactions (implying bounded propensities) and under the assumption that the propensities are bounded away from zero. This lower bound on the propensities controls the behaviour at the boundary of the state space. However, this assumption can be physically and mathematically undesirable, especially if one aims to generalise to a countable number of reactions. In the paper [SW05] the authors were able to relax this condition by requiring only local existence of a reaction with a propensity bounded away from zero.

The recent work [DEW91, DRW16] focusses on a specific form of the propensities $k^{(r,V)}$, in particular on a generalisation of mass-action kinetics. Their work covers propensities that vanish at the boundary, with a a few fairly weak assumptions, but restricted to the case of simultaneous unimolecular reactions on a finite state space.

The techniques that we use are most closely related to [SW95]. As in their model, we will assume that the total propensities are bounded from above. Since we work on a countable state space, this implies that the reaction rates cannot be bounded away from zero. In order to control the boundary behaviour we use a slightly different assumption to [SW05] and [DRW16]. Our assumption (see (3.3h)) is similar to what is sometimes called weak reversibility, see [AK11, Def. 1.10]. Whether a large deviation principle still holds without such condition remains an interesting and important open question.

Dual formulation of the rate functional. Following [SW95], an essential step is to prove that the rate functional I is equal to the dual formulation:

$$\widetilde{I}(c) := \begin{cases} \inf\limits_{\substack{u \in L^1(0,T;\mathfrak{l}^1_{\geq 0}(\mathcal{R})) \\ \Gamma u = \dot{c}}} \int_0^T \mathrm{Ent}_{c(t)}\big(u(t)\big)\,\mathrm{d}t, & c \in W^{1,1}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big), \\ \infty, & \text{otherwise,} \end{cases} \tag{1.5}$$

where

$$(\Gamma u)_y := \sum_{r\in\mathcal{R}} \gamma_y^{(r)} u^{(r)}, \qquad \mathrm{Ent}_c(u) := \begin{cases} \sum_{r\in\mathcal{R}} \bar{k}^{(r)}(c)\lambda_B\big(\frac{u^{(r)}}{\bar{k}^{(r)}(c)}\big), & \text{if } u^{(r)} \ll \bar{k}^{(r)}(c) \\ +\infty & \text{otherwise,} \end{cases} \tag{1.6}$$

and $\lambda_B(z) := z\log z - z + 1$, $\lambda_B(0) := 0$ is the Boltzmann function. The component $u^{(r)}(t)$ measures the amount of reaction r that takes place at time t. As such, the large deviation rate (1.5) can be seen as the large deviations for the empirical flow (see for example [BFG15]), transformed back to the large deviations for the empirical measure by a contraction principle. With this interpretation, one could consider the relation $\sum_{r\in\mathcal{R}}\gamma^{(r)}u^{(r)}=\dot{c}$ as a discrete-space counterpart of the continuity equation $-\operatorname{div}(cu)=\dot{c}$, common in mass transport theory and related large deviation results (see for example [DG87] and [AGS08]).

It turns out that with our weak reversibility condition, there always exists an optimal flow u which is non-zero. This flow can then be used to construct an optimal ξ in (1.4), which plays a crucial role in the proof of the lower bound.

Outline. In Section 2 we introduce the state space and path space. In Section 3 we introduce the precise particle system by a list of assumptions under which we will prove our results. In Section 4 we analyse the rate functionals (1.4) and (1.5). Finally, in Section 5 we prove the large deviation principle for the processes $\big(C^{(V)}\big)_{V>0}$.

2 Preliminaries

In Subsection 2.1 we introduce the state space along with some notation that we will need later. We then introduce the stoichiometric simplices as a subspace in which the processes remain, and prove a that these simplices are compact. In Subsection 2.2 we introduce the space of functions of bounded variation on a countable state space and the corresponding hybrid topology.

2.1 State space and stoichiometric simplices

As mentioned in the introduction, concentrations will be taken from the space

$$\mathfrak{l}^1(\mathcal{Y}):=\big\{c\in\mathbb{R}^{\mathcal{Y}}:|c|_1:=\textstyle\sum_{y\in\mathcal{Y}}|c_y|<\infty\big\}.$$

Test functions ξ are taken either from the space of bounded sequences, the space of vanishing sequences, or the space of sequences with finite support:

$$\begin{aligned}\mathfrak{l}^\infty(\mathcal{Y})&:=\big\{\xi:\mathcal{Y}\to\mathbb{R},|\xi|_\infty:=\sup_{y\in\mathcal{Y}}|\xi_y|<\infty\big\},\\ \mathfrak{c}_0(\mathcal{Y})&:=\big\{\xi\in\mathfrak{l}^\infty(\mathcal{Y}):\forall\epsilon>0\,\exists F\overset{\text{finite}}{\subset}\mathcal{Y}\text{ s.t. }\xi|_{F^c}<\epsilon\big\},\\ \mathfrak{c}_c(\mathcal{Y})&:=\big\{\xi\in\mathfrak{l}^\infty(\mathcal{Y}):\exists F\overset{\text{finite}}{\subset}\mathcal{Y}\text{ s.t. }\xi|_{F^c}=0\big\}.\end{aligned}$$

Similarly we will work with the space of summable or bounded sequences in $\mathfrak{l}^1(\mathcal{R})$ and $\mathfrak{l}^\infty(\mathcal{R})$. Recall that $\mathfrak{l}^\infty(\mathcal{Y})=\mathfrak{l}^1(\mathcal{Y})^*=\mathfrak{c}_0(\mathcal{Y})^{**}$. The usual dual pairing will be denoted by $\xi\cdot c:=\sum_{y\in\mathcal{Y}}\xi_y c_y$. Wherever we write a space with the subscript ≥ 0 or >0, we restrict to the elements with non-negative, respectively positive coordinates.

We will assume that $\|\Gamma\|:=\sup_{r\in\mathcal{R}}|\gamma^{(r)}|_1<\infty$, so that the linear operator $\Gamma:\mathfrak{l}^1(\mathcal{R})\to\mathfrak{l}^1(\mathcal{Y})$ from (1.5) is in fact bounded. We stress that the range $\operatorname{Ran}\Gamma:=\big\{\Gamma u:u\in\mathfrak{l}^1(\mathcal{R})\big\}\subset\mathfrak{l}^1(\mathcal{Y})$ allows for infinite combinations of the stoichiometric vectors. Equivalently, we can interpret the vectors $(\gamma^{(r)})_{r\in\mathcal{R}}$ as a Schauder basis for $\operatorname{Ran}\Gamma$ in $\mathfrak{l}^1(\mathcal{Y})$, but not as a Hamel basis in the sense of *finite* linear combinations.

Because the process $C^{(V)}(\cdot)$ does not necessarily conserve the number of particles, we need to introduce some structure that permits the use of tightness and compactness arguments. Such structure will be given by the stoichiometric simplexes set out below.

Given the effective stoichiometric vectors $\gamma^{(r)}$ and some fixed deterministic initial condition $c(0)$ in $\mathfrak{l}^1(\mathcal{Y})$, both the random process $C^{(V)}(t)$ and the deterministic limit $c(t)$ will remain in the *stoichiometric simplex* (sometimes also called the *non-negative stoichiometric compatibility class* [AK11, Def. 1.11]),

$$\mathfrak{l}^1_{c(0)}(\mathcal{Y}) := \mathfrak{l}^1_{\geq 0}(\mathcal{Y}) \cap \big(c(0) + \operatorname{Ran}\Gamma\big).$$

An important assumption on the stoichiometric simplex will be the existence of a quantity $m \in \mathbb{R}^{\mathcal{Y}}_{\geq 0}$ that is conserved $m \cdot \operatorname{Ran}\Gamma = 0$ and bounded away from zero. This quantity plays the role of a mass function so that this assumption is very physical. We stress that m does not have to lie in $\operatorname{Ker}\Gamma^T \subset \mathfrak{l}^\infty(\mathcal{Y})$ as it may not be bounded; on the contrary, we assume a growth condition on m. These assumptions will automatically imply compactness of the stoichiometric simplices (cf. [SW05, Ass. 2.1]):

Lemma 2.1 (Compact containment)**.** *Assume there exists an $m \in \mathbb{R}^{\mathcal{Y}}_{\geq 0}$ with*

- $m \cdot z = 0 \quad \forall z \in \operatorname{Ran}\Gamma$*,*
- $\inf_{y\in\mathcal{Y}} m_y =: \underline{m} > 0$*, and*
- *for any $\delta > 0$ there exists a finite subset $\mathcal{Y}_\delta \subset \mathcal{Y}$ such that $m_y > 1/\delta$ on $y \in \mathcal{Y}^c_\delta$.*

If $c(0) \in \mathfrak{l}^1_{\geq 0}(\mathcal{Y})$ with $m \cdot c(0) < \infty$ then the closure of the stoichiometric simplex $\operatorname{Cl}\big(\mathfrak{l}^1_{c(0)}(\mathcal{Y})\big)$ is compact in $\mathfrak{l}^1(\mathcal{Y})$.

Proof. We prove the relative compactness of $\mathfrak{l}^1_{c(0)}(\mathcal{Y})$. For any $c \in \mathfrak{l}^1_{c(0)}(\mathcal{Y})$, the quantity $m \cdot c \in m \cdot \big(c(0) + \operatorname{Ran}\Gamma\big) = m \cdot c(0) < \infty$ is conserved. From the boundedness from below of m we then deduce for any $c \in \mathfrak{l}^1_{c(0)}(\mathcal{Y})$ that $|c|_1 \leq \underline{m}^{-1}\sum_y m_y c_y = \underline{m}^{-1} m \cdot c(0) < \infty$, hence the set $\mathfrak{l}^1_{c(0)}(\mathcal{Y})$ is norm-bounded. By Banach-Alaoglu it is therefore relatively compact in the weak-$*$ topology, defined by pairings with $\mathfrak{c}_0(\mathcal{Y})$. We can then exploit the growth condition to improve the compactness to the weak topology, defined by pairings with $\mathfrak{l}^\infty(\mathcal{Y})$. Indeed, for any $\epsilon > 0$ and $c \in \mathfrak{l}^1_{c(0)}(\mathcal{Y})$, take $\delta < \epsilon/(m \cdot c(0))$. Then $\sum_{y\in\mathcal{Y}^c_\delta} c_y < \delta \sum_{y\in\mathcal{Y}^c_\delta} m_y c_y \leq \delta\, m \cdot c = \delta\, m \cdot c(0) < \epsilon$, so that $\mathfrak{l}^1_{c(0)}(\mathcal{Y})$ is tight, and weakly sequentially relatively compact by Prokhorov's Theorem. Recall that weak convergence of sequences in $\mathfrak{l}^1(\mathcal{Y})$ coincides with strong convergence [Con90, Prop. 5.2]. Hence the set $\mathfrak{l}^1_{c(0)}(\mathcal{Y})$ is strongly sequentially relatively compact, which is then also strongly topologically relatively compact since the strong topology is metrisable. □

2.2 Functions of bounded variation with the hybrid topology

We now describe the space and topology in which we prove the large-deviation principle. For any function $c \in L^1\big(0, T; \mathfrak{l}^1(\mathcal{Y})\big)$ the essential pointwise variation is defined as [AFP00]

$$\operatorname{epvar}(c) := \inf_{\hat{c}=c \text{ a.e.}} \sup_{0<t_1<\ldots<t_K<T} \sum_{k=1}^{K-1} |\hat{c}(t_{k+1}) - \hat{c}(t_k)|_1,$$

where the supremum ranges over all finite partitions of the interval $(0,T)$ and the infimum deals with the fact that L^1-functions are equivalence classes. The space of functions of bounded variation is then defined as

$$\mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big) := \Big\{c \in L^1\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big) : \mathrm{epvar}\big(c\big) < \infty\Big\},$$

with the corresponding norm $\|c\|_{\mathrm{BV}} := \|c\|_{L^1(0,T;\mathfrak{l}^1(\mathcal{Y}))} + \mathrm{epvar}\big(c\big)$.

An important feature of functions of bounded variation is that they always have a signed, measure-valued derivative $\dot{c} \in \mathrm{rca}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$, in the sense that [HPR]

$$\int_0^T \dot{\xi}(t)\cdot c(t)\,\mathrm{d}t = -\int_0^T \xi(t)\cdot \dot{c}(\mathrm{d}t) \qquad \text{for all } \xi \in C_0^1\big(0,T;\mathfrak{c}_0(\mathcal{Y})\big).$$

The total variation of this derivative is in fact the variation, i.e. $\mathrm{epvar}(c) = \|\dot{c}\|_{\mathrm{TV}}$, the total variation norm of the measure-valued derivative [HPR].

In order to prove exponential tightness, one needs to characterise compactness in the space of paths, which is generally difficult in a norm topology. Therefore one typically switches to a weak-$*$ topology. It can be shown that the space $\mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$ indeed has a predual, and that weak-$*$ convergence of functions $c^{(V)}$ is characterised by vague convergence of the measures $\mathrm{d}t \mapsto c^{(V)}(t)\,\mathrm{d}t$ and $\dot{c}^{(V)}$, see [AFP00, Rem. 3.12] and [HPR]. If the species space $\mathcal{Y}$ is finite, then this is equivalent to strong L^1-convergence of $c^{(V)}$ plus vague convergence of the measures $\dot{c}^{(V)}$. However, this equivalence is no longer true if $\mathcal{Y}$ is infinite. Instead, we thus get a topology that is stronger than the weak-$*$ topology and weaker than the norm topology. More precisely,

Definition 2.2. *We say that $c^{(V)}$ converges to c in $\mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$ in the* hybrid *topology, denoted by $c^{(V)} \rightrightarrows c$, if*

- $\|c^{(V)} - c\|_{L^1(0,T;\mathfrak{l}^1(\mathcal{Y}))} \to 0$, *and*
- $\int_0^T \xi(t)\cdot \dot{c}^{(V)}(\mathrm{d}t) \to \int_0^T \xi(t)\cdot \dot{c}(\mathrm{d}t)$ *for all test functions* $\xi \in C_0\big(0,T;\mathfrak{c}_0(\mathcal{Y})\big)$.

Beware of the logical mistake that the second condition follows from the first. Indeed, the strong L^1-convergence implies weak convergence of pairings with test functions $C_0^1\big(0,T;\mathfrak{c}_0(\mathcal{Y})\big)$, but not with test functions in $C_0\big(0,T;\mathfrak{c}_0(\mathcal{Y})\big)$.

We remark that the vague topology is not metrisible for infinite $\mathcal{Y}$ [Rud73, Th. 3.16 and Ex. 3.15], so that the hybrid topology is only characterised by convergence if one considers nets rather than sequences. Also because we work in a stronger topology than the weak-$*$ topology, we cannot use a Banach-Alaoglu argument to identify compact subsets. Instead we have the following result.

Theorem 2.3 ([HPR]). *For any $c \in \mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$, compact $K \subset \mathfrak{l}^1(\mathcal{Y})$, and $L > 0$, the set*

$$\Big\{\hat{c} \in \mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big) : \|\hat{c} - c\|_{\mathrm{BV}} \le L\Big\} \cap K^{(0,T)}, \quad \text{where}$$
$$K^{(0,T)} := \big\{c : (0,T) \to \mathfrak{l}^1(\mathcal{Y}) \text{ s.t. } c(t) \in K \ \forall t \in (0,T)\big\}$$

is (topologically and sequentially) compact in $\mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$ with the hybrid topology.

Remark 2.4. In fact, the compact set K is even allowed to depend on t, but we will not need that level of generality in the current paper. □

3 The empirical process

In Subsection 3.1 we specify the exact setting by introducing a set of conditions that we will assume throughout the paper. We show that the assumptions are satisfied by a subclass of models with mass-action kinetics, and prove that our microscopic models are well defined and bounded for all time. Next, in Subsection 3.2 we show the convergence of the process to the solution of the Reaction Rate Equation. This convergence will be used explictly in the proof of the large deviation principle.

3.1 Setting

Since we want to focus on stochastic fluctuations due to the dynamics, we will assume (without loss of generality) deterministic initial conditions:

Assumption 3.1. *For each* $V > 0$ *the initial condition is deterministic* $C^{(V)}(0) = c^{(V)}(0)$*, and* $c^{(V)}(0) \xrightarrow{\mathfrak{l}^1} c(0)$ *with* $\sup_{V>0} m \cdot c^{(V)}(0) < \infty$.

We show below that the processes $\big(C^{(V)}(t)\big)_{t\in(0,T),V>0}$, $\big(c(t)\big)_{t\in(0,T)}$ remain a.s. in the union

$$K := \mathrm{Cl}\Big(\mathfrak{l}^1_{c(0)}(\mathcal{Y}) \cup \bigcup_{V>0} \mathfrak{l}^1_{c^{(V)}(0)}(\mathcal{Y})\Big). \tag{3.1}$$

This set will be used in the compact containment condition in Theorem 2.3. By the following argument it is indeed compact:

Lemma 3.2 (Compact containment II)**.** *Under Assumptions 3.3a and 3.1, the set* K *is compact in* $\mathfrak{l}^1(\mathcal{Y})$.

Proof. This is a simple adaptation of the proof of Lemma 2.1, where we now take

$$\delta < \frac{\epsilon}{\sup_V m \cdot c^{(V)}(0)}. \tag{3.2}$$

□

The assumptions on the process under which we will prove our results are the following.

Assumption 3.3. *Given effective stoichiometric coefficients* $(\gamma^{(r)})_{r\in\mathcal{R}} \in \mathfrak{c}_c(\mathcal{Y})$*, propensities and reaction rates* $(k^{(r,V)})_{r\in\mathcal{R},V>0}, (\bar{k}^{(r)})_{r\in\mathcal{R}} : \mathfrak{l}^1_{\geq 0}(\mathcal{Y}) \to \mathbb{R}^{\mathcal{R}}_{\geq 0}$*, and deterministic initial condition* $c(0) \in \mathfrak{l}^1_{\geq 0}(\mathcal{Y})$*, we will further assume that:*

- *there exists an* $m \in \mathbb{R}^{\mathcal{Y}}_{\geq 0}$ *with*
  - $m \cdot z = 0 \quad \forall z \in \mathrm{Ran}\,\Gamma$, (3.3aa)
  - $\inf_{y\in\mathcal{Y}} m_y =: \underline{m} > 0$*, and* (3.3ab)
  - *for any* $\delta > 0$ *there exists a finite subset* $\mathcal{Y}_\delta \subset \mathcal{Y}$ *such that* $m_y > 1/\delta$ *on* $y \in \mathcal{Y}_\delta^c$*,* (3.3ac)
- $m \cdot c(0) < \infty$, (3.3b)
- $1 \leq \inf_{r\in\mathcal{R}} |\gamma^{(r)}|_1 \leq \sup_{r\in\mathcal{R}} |\gamma^{(r)}|_1 =: \|\Gamma\| < \infty$, (3.3c)

- $k^{(r,V)}(c) = 0$ *for* $c < -V^{-1}\gamma^{(r)}$ *(component wise),* (3.3d)
- $\lim_{V\to\infty} \sup_{c\in K} \sum_{r\in\mathcal{R}} \left|\frac{1}{V} k^{(r,V)}(c) - \bar{k}^{(r)}(c)\right| = 0,$ (3.3e)
- $\sup_{c\in K} \sum_{r\in\mathcal{R}} \bar{k}^{(r)}(c) < \infty,$ (3.3f)
- *the propensities* $\bar{k}^{(r)}(c)$ *are continuous, and the sum* $\sum_{r\in\mathcal{R}} \bar{k}^{(r)}(c)$ *is Lipschitz continuous in* $c \in \mathfrak{l}^1_{c(0)}(\mathcal{Y})$, (3.3g)
- *for every* $r \in \mathcal{R}$ *and any* $c \in l^1(\mathcal{Y})$ *with* $k^{(r)}(c) > 0$ *there is a (finitely supported)* $v \in \mathfrak{c}_{c,>0}(\mathcal{R})$ *such that* $\Gamma v = 0$ *and* $k^{(\hat{r})}(c) > 0$ *for* $\hat{r} \in \operatorname{supp} v$. (3.3h)

The most restrictive assumptions in this list are in our opinion the summability (3.3f), the local Lipschitzness (3.3g) and the weak reversibility (3.3h). The latter states that for every reaction with non-zero rate there is a reaction - or a finite chain of reactions - with non-zero rate that undoes this reaction. In some sense we need this assumption to deal with the boundary of the stoichiometric simplex. More precisely, we will prove the existence of a maximiser in (1.4), which, without Assumption (3.3h), even fails in the case of one reaction, see Remark 4.10.

The summability (3.3f) of the reaction rates imply that right-hand side of the Reaction Rate Equation (1.3) is a bounded operator. Together with the Lipschitz property (3.3g) this guarantees that the Reaction Rate Equation (1.3) has a unique solution for initial condition $c(0)$. This setting of bounded and Lipschitz operators can possibly be generalised to accretive and monotone ("one-sided Lipschitz") operators [GGZ74], which we will not pursue in this work.

The typical setting that we have in mind is the case of *mass-action kinetics*. In that case the reaction rates are of the form $\bar{k}^{(r)}(c) := \kappa^{(r)} c^{\alpha^{(r)}} := \kappa^{(r)} \prod_{y\in\mathcal{Y}} c_y^{\alpha_y^{(r)}}$, and the propensities are as commonly used [Kur72, MPPR15] in the chemical master equation: $k^{(r,V)}(c) = \kappa^{(r)} V^{1-|\alpha|_1} (cV)!/(cV-\alpha)!$.

Proposition 3.4 (Mass-action kinetics)**.** *Let* $\bar{k}^{(r)}(c) := \kappa^{(r)} c^{\alpha^{(r)}} := \kappa^{(r)} \prod_{y\in\mathcal{Y}} c_y^{\alpha_y^{(r)}}$ *for some rate constants* $(\kappa^{(r)})_{r\in\mathcal{R}} \subset \mathbb{R}_{\geq 0}$ *and stoichiometric coefficients* $\alpha^{(r)} \in \mathfrak{l}^1(\mathcal{Y}) \cap \mathbb{N}_0^{\mathcal{Y}}$*. Assume that* $\omega := \sup_{c\in\mathfrak{l}^1_{c(0)}(\mathcal{Y})} |c|_\infty \geq 1$*, that* $\sup_{r\in\mathcal{R}} |\alpha^{(r)}|_\infty < \infty$ *and that* $\sum_{r\in\mathcal{R}} \kappa^{(r)} \omega^{|\alpha^{(r)}|_1} < \infty$*. Then Assumptions* (3.3f) *and* (3.3g) *are satisfied.*

Proof. Recall from Lemma 2.1 that ω is finite. The summability of the rates (3.3f) is immediate from the assumption.

For the Lipschitz condition (3.3g) we first prove a Lipschitz estimate for one reaction r, i.e. for any $\hat{c}, c \in \mathfrak{l}^1_{c(0)}(\mathcal{Y})$,

$$\left|\hat{c}^{\alpha^{(r)}} - \hat{c}^{\alpha^{(r)}}\right| \leq |\alpha^{(r)}|_\infty \omega^{|\alpha^{(r)}|_1 - 1} |\hat{c} - c|_1. \tag{3.2}$$

Let $d := \#\{y \in \mathcal{Y} : \alpha_y^{(r)} \neq 0\}$. Then $d < \infty$ since we assumed that $\alpha^{(r)} \subset \mathfrak{l}^1(\mathcal{Y}) \cap \mathbb{N}_0^{\mathcal{Y}}$. We now proceed by induction on d. For $d = 1$, take the $y \in \mathcal{Y}$ for which $\alpha_y^{(r)} \neq 0$, and recall the well-known local Lipschitz estimate for the exponential function:

$$\left|\hat{c}^{\alpha^{(r)}} - \hat{c}^{\alpha^{(r)}}\right| = \left|\hat{c}_y^{\alpha_y^{(r)}} - \hat{c}_y^{\alpha_y^{(r)}}\right| = \left|(\hat{c}_y - c_y) \sum_{i=0}^{\alpha_y^{(r)}-1} \hat{c}_y^i c_y^{\alpha_y^{(r)}-1-i}\right| \leq \alpha_y^{(r)} \omega^{\alpha_y^{(r)}-1} |\hat{c}_y - c_y|.$$

Now assume that (3.2) holds for $d=n$. Then for $d=n+1$ we have (for ease of notation assuming $\mathcal{Y}=\mathbb{N}$),

$$\begin{aligned}\big|\hat c^{\alpha^{(r)}}-\hat c^{\alpha^{(r)}}\big| &\leq \hat c_{n+1}^{\alpha^{(r)}_{n+1}}\big|\textstyle\prod_{y\leq n}\hat c_y^{\alpha^{(r)}_y}-\prod_{y\leq n}c_y^{\alpha^{(r)}_y}\big|+\big(\prod_{y\leq n}c_y^{\alpha^{(r)}}\big)\big|\hat c_{n+1}^{\alpha^{(r)}_{n+1}}-c_{n+1}^{\alpha^{(r)}_{n+1}}\big|\\ &\leq \hat c_{n+1}^{\alpha^{(r)}_{n+1}}\big(\sup_{y\leq n}\alpha^{(r)}_y\big)\omega^{\sum_{y\leq n}\alpha^{(r)}_y-1}\big(\textstyle\sum_{y\leq n}|\hat c_y-c_y|\big)\\ &\qquad\qquad+\big(\textstyle\prod_{y\leq n}c_y^{\alpha^{(r)}}\big)\alpha^{(r)}_{n+1}\omega^{\alpha^{(r)}_{n+1}}|\hat c_{n+1}-c_{n+1}|\\ &\leq |\alpha^{(r)}|_\infty\omega^{|\alpha^{(r)}|_1-1}|\hat c-c|_1.\end{aligned}$$

This proves the estimate (3.2) for any $d\geq 0$.

Then, summing (3.2) over the reactions:

$$\begin{aligned}\big|\textstyle\sum_{r\in\mathcal{R}}\kappa^{(r)}\hat c^{\alpha^{(r)}}-\sum_{r\in\mathcal{R}}\kappa^{(r)}\hat c^{\alpha^{(r)}}\big| &\leq |\hat c-c|_1\sum_{r\in\mathcal{R}}\kappa^{(r)}|\alpha^{(r)}|_\infty\omega^{|\alpha^{(r)}|_1-1}\\ &\leq |\hat c-c|_1\big(\sup_{r\in\mathcal{R}}|\alpha^{(r)}|_\infty\big)\sum_{r\in\mathcal{R}}\kappa^{(r)}\omega^{|\alpha^{(r)}|_1-1},\end{aligned}$$

where the Lipschitz constant is finite by assumption. □

We conclude this section by showing the approximating processes are well defined and somewhat regular:

Proposition 3.5 (Properties of the processes for finite V)**.** *Under Assumptions 3.3 and 3.1, the processes $t\mapsto C^{(V)}(t)$ remain in K almost surely, they are non-explosive and hence well-defined on any $(0,T)$, and they define corresponding probability measures $\mathbb{P}^{(V)}\in\mathcal{P}\big(\mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)\big)$. Moreover, each $C^{(V)}(t)$ is almost surely bounded in $\mathfrak{l}^\infty(\mathcal{Y})$ uniformly in t, and even uniformly in V.*

Proof. Non-negativity is preserved by (3.3d). Moreover, the assumptions imply that the propensities are almost surely bounded as follows:

$$\frac{1}{V}\sum_{r\in\mathcal{R}}k^{(r,V)}\big(C^{(V)}(t)\big)\overset{\text{(3.3e)}}{\leq}\sum_{r\in\mathcal{R}}\bar k^{(r)}\big(C^{(V)}(t)\big)+\epsilon\leq\sup_{c\in\mathfrak{l}^1_{c(0)}(\mathcal{Y})}\sum_{r\in\mathcal{R}}\bar k^{(r)}(c)+\epsilon\overset{\text{(3.3f)}}{<}\infty.$$

Therefore, $C^{(V)}(t)$ has almost surely only a finite number of jumps in the finite time interval $(0,T)$, where each jump lies in $\operatorname{Ran}\Gamma\subset\mathfrak{l}^1(\mathcal{Y})$. This shows that the process is non-explosive, it remains in $\mathfrak{l}^1_{c(0)}(\mathcal{Y})$ a.s., and the paths are a.s. of bounded variation. Finally, by Lemma 3.2, the set K is compact in $\mathfrak{l}^1(\mathcal{Y})$ and therefore bounded in $\mathfrak{l}^\infty(\mathcal{Y})$, which proves the uniform bound. □

3.2 Law of Large numbers

Before moving onto the preparatory work for the large deviations we assert the well-posedness of the limit equation and give the basic convergence result for our processes. Since this paper focusses on large deviation we only sketch the proof of the uniqueness.

Proposition 3.6 (Uniqueness for the Limit Equation)**.** *Let Assumptions 3.3 hold, then* (1.3) *is well-posed for initial conditions $c(0)\in\mathfrak{l}^1_{\geq 0}(\mathcal{Y})$ satisfying $m\cdot c(0)<\infty$.*

Sketch of the proof. By (3.3aa) and the combination of (3.3e) with (3.3d) we see that for values on the boundary of $\mathfrak{l}^1_{c(0)}(\mathcal{Y})$ the derivative cannot point out of this simplex. We can thus use (3.3g) to get the local existence and uniqueness of a solution to (1.3) via a Picard fixed point argument and (3.3f) to see that uniqueness (and existence) hold for all time. A detailed presentation of the Picard argument in a more complex setting can be found in [Pat16]. □

Proposition 3.7 (Functional Law of Large Numbers)**.** *Let Assumptions 3.3 and 3.1 hold. Then as* $V \to \infty$*,*

$$\mathbb{P}^{(V)} \rightharpoonup \delta_c,$$

where $c \in \mathrm{BV}\big(0, T; \mathfrak{l}^1(\mathcal{Y})\big)$ *is the unique solution of the Reaction Rate Equation* (1.3) *with initial condition* $c(0)$*.*

Proof. In Theorem 5.1 below, we will prove the exponential tightness of the distributions of the $C^{(V)}$ on $\mathrm{BV}\big(0, T; \mathfrak{l}^1(\mathcal{Y})\big)$, which implies tightness. Therefore every sequence has a convergent subsequence [HPR].

It remains to show that every subsequence of these distributions converges to the common limit point δ_c. To this aim, suppose that we have passed to a subsequence $C^{(V)}$ that converges in distribution to some random variable C taking values in $\mathrm{BV}\big(0, T; \mathfrak{l}^1(\mathcal{Y})\big)$. We proceed using a Martingale approach. Take an arbitrary $\Phi \in C^1_b\big(\mathfrak{l}^1(\mathcal{Y})\big)$. As the process $C^{(V)}(t)$ is generated by $Q^{(V)}$, defined in (1.2), the random variable

$$M^{(\Phi,V)}(t) := \Phi\left(C^{(V)}(t)\right) - \Phi\left(c^{(V)}(0)\right) - \int_0^t \left(\mathcal{Q}^{(V)}\Phi\right)\left(C^{(V)}(s)\right) \mathrm{d}s \tag{3.3}$$

is a martingale. Define the continuous mapping $M^{(\Phi)}\colon \mathrm{BV}\big(0, T; \mathfrak{l}^1(\mathcal{Y})\big) \to \mathrm{BV}\big(0, T; \mathbb{R}\big)$ (continuity is with respect to the hybrid topology on both spaces) given by

$$\left(M^{(\Phi)}\hat{c}\right)(t) := \Phi\left(\hat{c}(t)\right) - \int_0^t \left(Q\Phi\right)\left(\hat{c}(s)\right) \mathrm{d}s, \qquad \text{where}$$

$$(\mathcal{Q}\Phi)(\hat{c}) := \sum_{r\in\mathcal{R}} \bar{k}^{(r)}(\hat{c})\, \nabla\Phi(\hat{c}) \cdot \gamma^{(r)}$$

will be the limit generator. We can now write

$$\left(M^{(\Phi)}C^{(V)}\right)(t) - \Phi\left(c^{(V)}(0)\right) = M^{(\Phi,V)}(t) + \int_0^t \left[\, \left(Q^{(V)}\Phi\right)\left(C^{(V)}(s)\right) - \left(Q\Phi\right)\left(C^{(V)}(s)\right) \right] \mathrm{d}s. \tag{3.4}$$

We now pass to the limit in this expression. From the continuity of $M^{(\Phi)}$ and Φ in combination with Assumption 3.1 the left-hand side of (3.4) converges in distribution to

$$\Phi\left(C(t)\right) - \Phi\left(c(0)\right) - \int_0^t \left(Q\Phi\right)\left(C(s)\right) \mathrm{d}s, \tag{3.5}$$

which will be a.s. 0 once we show that the right-hand side of (3.4) converges in distribution to the identically 0 path. To this end check (either directly or via the quadratic variation using the BDG inequalities [Kal02, Th. 26.12]) that for any $t \geq 0$

$$\mathbb{E}\left[\sup_{s\leq t} \left(M^{(\Phi,V)}(s)\right)^2\right] \sim \mathcal{O}\left(\|\Phi\|^2_\infty / V\right) \tag{3.6}$$

and so by Chebyshev $M^{(\Phi,V)}$ converges in distribution to 0. Also note that uniformly on $\mathfrak{l}^1_{c(0)}(\mathcal{Y})$

$$\big|(\mathcal{Q}\Phi)(c) - (\mathcal{Q}^{(V)}\Phi)(c)\big| \sim \mathcal{O}\left(\|\Phi\|_\infty /V\right). \tag{3.7}$$

We therefore get that (3.5) is a.s. 0. By Proposition 3.6 this can only be true if $C = c$ a.s. □

We conclude this section with an extension to slightly more general processes. As usual in large-deviation theory, we consider time-inhomogeneous generators, exponentially perturbed by $\xi \in C_b\big(0,T;\mathfrak{l}^\infty(\mathcal{Y})\big)$,

$$(\mathcal{Q}^{(V)}_{\xi(t)}\Phi)(c) := \sum_{r\in\mathcal{R}} k^{(r,V)}(c)e^{\xi(t)\cdot\gamma^{(r)}}\big(\Phi(c+\tfrac{1}{V}\gamma^{(r)}) - \Phi(c)\big),$$

$$(\mathcal{Q}_{\xi(t)}\Phi)(c) := \sum_{r\in\mathcal{R}} \bar{k}^{(r)}(c)e^{\xi(t)\cdot\gamma^{(r)}}\,\nabla\Phi(c)\cdot\gamma^{(r)}. \tag{3.8}$$

Lemma 3.8. *Fix $\xi \in L^\infty\big(0,T;\mathfrak{l}^\infty(\mathcal{Y})\big)$ and let Assumptions 3.3 and 3.1 hold. Then Lemma 3.5 and Propositions 3.6 & 3.7 remain true for the perturbed processes, that is, $\mathcal{Q}^{(V)}_\xi$ and $\mathcal{Q}_\xi$ define probability measures $\mathbb{P}^{(V)}_\xi$ and $\mathbb{P}_\xi$ on $\mathrm{BV}\big(0,T;\mathfrak{l}^1_{\geq 0}(\mathcal{Y})\big)$, and as $V\to\infty$,*

$$\mathbb{P}^{(V)}_\xi \rightharpoonup \mathbb{P}_\xi = \delta_c, \tag{3.9}$$

where $c \in \mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$ is the unique solution to the perturbed equation

$$\dot{c}(t) = \sum_{r\in\mathcal{R}} \bar{k}^{(r)}\big(c(t)\big)e^{\xi(t)\cdot\gamma^{(r)}}\gamma^{(r)}$$

with initial condition $c(0)$.

Proof. This is immediate from the fact that the perturbation factors $e^{\xi(t)\cdot\gamma^{(r)}}$ are uniformly bounded. □

4 Analysis of the rate functional

This section is devoted to a number of properties of the rate functional I that are needed in the proof of the large deviations. In Subsection 4.1 we show that the rate functional has the form of an action $I(c) = \int_0^T L\big(c(t),\dot{c}(t)\big)\,\mathrm{d}t$, with

$$L(c,s) := \sup_{\xi\in\mathfrak{l}^\infty(\mathcal{Y})} \xi\cdot s - H(c,\xi), \tag{4.1}$$

using $(c,s) \in \mathfrak{l}^1_{c(0)}(\mathcal{Y}) \times \mathfrak{l}^1(\mathcal{Y})$ as placeholder variables for $\big(c(t),\dot{c}(t)\big)$. In Subsection 4.2 we show that the function L admits the dual formulation

$$\widetilde{L}(c,s) := \inf_{u\in\mathfrak{l}^1_{\geq 0}(\mathcal{R}):\,\Gamma u = s} \mathrm{Ent}_c(u), \tag{4.2}$$

where Ent_c is defined by (1.6). We point out that this formulation is basically a countable inf-convolution, which arises naturally from the convex dual L of a countable sum H, see for example [Roc70, Th. 16.4]. It turns out that the infimum can be taken out of the integral, which leads to the dual formulation $\widetilde{I}$ from (1.5). Finally, in Subsection 4.3 we show that the infima in (4.2) and (1.5) are attained, and from the minimiser we construct a maximiser for the problem (1.4).

From now on we implicitly work under Assumptions 3.1 and 3.3.

4.1 The action formulation of the rate functional

Before we prove the action form, we first show that the time derivative in this formulation is well-defined.

Proposition 4.1. *Let $c \in \mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$ with $I(c) < \infty$. Then*

(a) c is continuous;

(b) c is absolutely continuous. Hence, we can identify the derivative $\dot c(\mathrm{d}t) = \dot c(t)\,\mathrm{d}t$ with some $\dot c \in L^1\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$ and in fact $c \in \mathrm{AC}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big) \cap \mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big) = W^{1,1}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$.

Proof. *(a)* Using [HPR], we can take a representation of c that has no jumps at the boundary points $\{0,T\}$. Without loss of generality we can assume that c remains in one stoichiometric subspace $\mathfrak{l}^1_{c(0)}(\mathcal{Y})$. Now assume on the contrary that c has a jump discontinuity at, say $t_0 \in (0,T)$. Since the right and left limits of t_0 exist by [HPR], we have as a consequence:

$$\frac{1}{\delta}\int_{t_0-\delta}^{t_0} |c(t) - c(t_0^-)|_1\,\mathrm{d}t \xrightarrow{\delta\to0} 0 \qquad \text{and} \qquad \frac{1}{\delta}\int_{t_0}^{t_0+\delta} |c(t) - c(t_0^+)|_1\,\mathrm{d}t \xrightarrow{\delta\to0} 0. \tag{4.3}$$

Take a $g \in C_c^1(-1,1;[0,1])$ with $g(0) = 1$, and define, for any $A > 0$ and $\delta > 0$ sufficiently small:

$$\xi^{(\delta,A)}(t) := A\, g\big(\tfrac{t-t_0}{\delta}\big)\,\mathrm{sgn}\big(c(t_0^+) - c(t_0^-)\big),$$

where $\mathrm{sgn}\colon \mathfrak{l}^1(\mathcal{Y}) \to \mathfrak{l}^\infty(\mathcal{Y})$ is the coordinate-wise sign function. For this sequence we have, on the one hand,

$$\begin{aligned}\int_0^T \xi^{(\delta,A)}(t)\cdot \dot c(\mathrm{d}t) = &-\frac{A}{\delta}\int_{t_0}^{t_0+\delta} \dot g\big(\tfrac{t-t_0}{\delta}\big)\,\mathrm{sgn}\big(c(t_0^+) - c(t_0^-)\big)\cdot\big(c(t) - c(t_0^+)\big)\,\mathrm{d}t\\ &-\frac{A}{\delta}\int_{t_0}^{t_0+\delta} \dot g\big(\tfrac{t-t_0}{\delta}\big)\,\mathrm{sgn}\big(c(t_0^+) - c(t_0^-)\big)\cdot c(t_0^+)\,\mathrm{d}t\\ &-\frac{A}{\delta}\int_{t_0-\delta}^{t_0} \dot g\big(\tfrac{t-t_0}{\delta}\big)\,\mathrm{sgn}\big(c(t_0^+) - c(t_0^-)\big)\cdot\big(c(t) - c(t_0^-)\big)\,\mathrm{d}t\\ &-\frac{A}{\delta}\int_{t_0-\delta}^{t_0} \dot g\big(\tfrac{t-t_0}{\delta}\big)\,\mathrm{sgn}\big(c(t_0^+) - c(t_0^-)\big)\cdot c(t_0^-)\,\mathrm{d}t.\end{aligned}$$

Because $\delta^{-1}\int_{t_0-\delta}^{t_0}\dot g\big(\frac{t-t_0}{\delta}\big)\,\mathrm{d}t = 1$ and $\delta^{-1}\int_{t_0}^{t_0+\delta}\dot g\big(\frac{t-t_0}{\delta}\big)\,\mathrm{d}t = -1$, the second and fourth term are exactly $A\,\mathrm{sgn}\big(c(t_0^+) - c(t_0^-)\big)\cdot\big(c(t_0^+) - c(t_0^-)\big) = A\,|c(t_0^+) - c(t_0^-)|_1$. The first term vanishes as $\delta \to 0$; this follows from (4.3) and the estimate

$$\frac{A}{\delta}\int_{t_0}^{t_0+\delta} \big|\dot g\big(\tfrac{t-t_0}{\delta}\big)\,\mathrm{sgn}\big(c(t_0^+) - c(t_0^-)\big)\cdot\big(c(t) - c(t_0^+)\big)\big|\,\mathrm{d}t \le \|\dot g\|_\infty \frac{A}{\delta}\int_{t_0}^{t_0+\delta}\big|c(t) - c(t_0^+)\big|_1\,\mathrm{d}t.$$

Similarly the third term vanishes. We therefore find that

$$\int_0^T \xi^{(\delta,A)}(t)\cdot\dot c(\mathrm{d}t) \xrightarrow[\delta\to0]{} A\,|c(t_0^+) - c(t_0^-)|_1.$$

On the other hand, we see that for fixed A,

$$\int_0^T H\big(c(t), \xi^{(\delta,A)}(t)\big)\,\mathrm{d}t \le \sup_{c\in \mathfrak{l}^1_{c(0)}(\mathcal{Y})} \Big(\sum_{r\in\mathcal{R}} \bar{k}^{(r)}(c)\Big) \int_{t_0-\delta}^{t_0+\delta} \big(e^{A\|\Gamma\|}-1\big)\,\mathrm{d}t \xrightarrow[\delta\to 0]{} 0$$

because of Assumptions (3.3c) and (3.3f).

Putting the pieces together, we find

$$I(c) \ge \lim_{A\to\infty}\lim_{\delta\to 0} \int_0^T \xi^{(\delta,A)}(t)\cdot \dot{c}(\mathrm{d}t) - \int_0^T H\big(c(t),\xi^{(\delta,A)}(t)\big)\,\mathrm{d}t$$
$$\ge \lim_{A\to\infty} A|c(t_0^+)-c(t_0^-)|_1 = \infty,$$

which contradicts the assumption. Therefore, c must be continuous.

(b) Suppose that c is not absolutely continuous: there exists an $\epsilon > 0$ such that for any $\delta > 0$ there is a sequence of disjoint open intervals $\bigcup_{n=1}^N (a_n, b_n) \subset (0,T)$ such that

$$\sum_{n=1}^N (b_n - a_n) < \delta \qquad \text{and} \qquad \sum_{n=1}^N |c(b_n)-c(a_n)|_1 \ge \epsilon.$$

For arbitrary $A > 0$ and $\delta > 0$, construct a sequence of test functions $\xi^{(\delta,A)} \in C_b^1(0,T;\mathfrak{l}^\infty(\mathcal{Y}))$ such that $\xi^{(\delta,A)}(t) = A\,\mathrm{sgn}\big(c(b_n)-c(a_n)\big)$ for $t \in (a_n,b_n)$. By *(a)* the curve c has no jumps, therefore we can extend the function $\xi^{(\delta,A)}$ smoothly between the intervals such that $0 = \int_{b_n}^{a_{n+1}} \xi^{(\delta,A)}(t)\cdot\dot{c}(\mathrm{d}t) = \int_0^{a_1}\xi^{(\delta,A)}(t)\cdot\dot{c}(\mathrm{d}t) = \int_{b_N}^T \xi^{(\delta,A)}(t)\cdot\dot{c}(\mathrm{d}t)$ for all $n = 1,\dots,N-1$. With this construction we get $[c,\xi^{(\delta,A)}] = A\sum_{n=1}^N |c(b_n)-c(a_n)| \ge A\epsilon$. On the other hand,

$$H\big(c(t),\xi^{(\delta,A)}(t)\big) \le \sum_{r\in\mathcal{R}} \bar{k}^{(r)}\big(c(t)\big)\big(e^{A|\gamma^{(r)}|_1}-1\big),$$

which is finite by Assumptions (3.3c) and (3.3f), because c remains in $\mathfrak{l}^1_{c(0)}(\mathcal{Y})$ else $I=\infty$. After plugging this test function into the definition we therefore get:

$$I(c) \ge A\epsilon - \delta\big(e^{A\|\Gamma\|}-1\big) \sup_{c\in\mathfrak{l}^1_{c(0)}(\mathcal{Y})} \sum_{r\in\mathcal{R}} \bar{k}^{(r)}(c),$$

First taking $\delta \to 0$ and then $A \to \infty$ shows that $I = \infty$, which is a contradiction, and hence the path c must be absolutely continuous. It follows that the L^1-valued derivative exists, see for example [AGS08, Rem. 1.1.3]. □

We now show that the supremum over $C_b^1(0,T;\mathfrak{l}^\infty(\mathcal{Y}))$ in (1.4) can be taken over two different spaces. Later on, we will see that the first space is sufficiently regular to apply a Girsanov argument, and that the second space is sufficiently large to guarantee existence of the maximiser. To shorten notation we introduce the functional

$$G(c,\xi) := \int_0^T \xi(t)\cdot\dot{c}(\mathrm{d}t) - \int_0^T H\big(c(t),\xi(t)\big)\mathrm{d}t. \tag{4.4}$$

Proposition 4.2. *For any $c \in \mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$,*

$$I(c) = \sup_{\xi \in C_b^2(0,T;\mathfrak{l}^\infty(\mathcal{Y}))} G(c,\xi), \qquad \textit{and} \tag{4.5}$$

$$= \sup_{\xi \in L^\infty(0,T;\mathfrak{l}^\infty(\mathcal{Y}))} G(c,\xi) \qquad \textit{whenever } I(c)<\infty. \tag{4.6}$$

Proof. The first equality follows from the fact that the functional $\xi \mapsto G(c,\xi)$ is continuous in $C_b^1\big(0,T;\mathfrak{l}^\infty(\mathcal{Y})\big)$, equipped with the uniform norm.

We prove the second equality via a number of approximations, more precisely:

$$\begin{aligned} I(c) &= \sup_{\xi \in C_b(0,T;\mathfrak{l}^\infty(\mathcal{Y}))} G(c,\xi) = \sup_{\xi \in C_b(0,T;\mathfrak{c}_0(\mathcal{Y}))} G(c,\xi) \\ &= \sup_{\xi \in L^\infty(0,T;\mathfrak{c}_0(\mathcal{Y}))} G(c,\xi) = \sup_{\xi \in L^\infty(0,T;\mathfrak{l}^\infty(\mathcal{Y}))} G(c,\xi). \end{aligned}$$

The relaxation to $C_b\big(0,T;\mathfrak{l}^\infty(\mathcal{Y})\big)$ is again true by the above mentioned continuity. For the second equality, pick an ordering of $\mathcal{Y}$ and approximate any arbitrary $\xi \in C_b\big(0,T;\mathfrak{l}^\infty(\mathcal{Y})\big)$ by $\xi^{(n)}_y(t) := \xi_y(t)\mathbb{1}_{\{y\le n\}}$. Then by dominated convergence $\lim_{n\to\infty} G(c,\xi^{(n)}) = G(c,\xi)$, which shows that the supremum can also be taken over $C_b\big(0,T;\mathfrak{c}_0(\mathcal{Y})\big)$.

Observe that the space $\mathfrak{c}_0(\mathcal{Y})$ is Polish, so that we can now use a Lusin approximation. More precisely, by Lusin's Theorem [Bog07, Th. II.7.14.26], we can construct, for any arbitrary $\xi \in L^\infty\big(0,T;\mathfrak{c}_0(\mathcal{Y})\big)$, a sequence $\xi^{(n)} \in C_b\big(0,T;\mathfrak{c}_0(\mathcal{Y})\big)$ with $\|\xi^{(n)}\|_\infty \le \|\xi\|_\infty$ such that $\xi^{(n)} \equiv \xi$ on some compact $A_n \subset (0,T)$ with Lebesgue measure at least $T-1/n$. Then again by dominated convergence we have $\lim_{n\to\infty} G(c,\xi^{(n)}) = G(c,\xi)$, so that the supremum can indeed be taken over $L^\infty\big(0,T;\mathfrak{c}_0(\mathcal{Y})\big)$.

For the final equality, pick a $\xi \in L^\infty\big(0,T;\mathfrak{l}^\infty(\mathcal{Y})\big)$, define the cut-off approximation as before $\xi^{(n)}_y(t) := \xi_y(t)\mathbb{1}_{\{y\le n\}}$, and use dominated convergence to see that that the supremum can indeed be taken over $L^\infty\big(0,T;\mathfrak{l}^\infty(\mathcal{Y})\big)$. □

We now seek to move the supremum inside the integral in (1.4). To show this we need the following simple lemma.

Lemma 4.3. *The supremum in* (4.1) *can be taken over a countable set which does not depend on* (c,s).

Proof. Write $\mathfrak{c}_c(\mathcal{Y}) \cap \mathbb{Q}^{\mathcal{Y}} \subset \mathfrak{l}^\infty(\mathcal{Y})$ for the set of rational sequences indexed by $\mathcal{Y}$ with at most finitely many non-zero values. For any $\xi \in \mathfrak{l}^\infty(\mathcal{Y})$ we can take a sequence $\xi^{(n)} \in \mathfrak{c}_c(\mathcal{Y}) \cap \mathbb{Q}^{\mathcal{Y}}$ such that $\big|\xi^{(n)}_y\big| \le |\xi_y|$ for all $y \in \mathcal{Y}$ and $\lim_{n\to\infty}\xi^{(n)}_y = \xi_y$. Then for fixed c and s by dominated convergence arguments $\lim_{n\to\infty} \xi^{(n)}\cdot s - H(c,\xi^{(n)}) = \xi\cdot s - H(c,\xi)$. □

We now show that I can be rewritten as an action functional:

Proposition 4.4.

$$I(c) = \begin{cases} \displaystyle\int_0^T L\left(c(t),\dot c(t)\right)\mathrm{d}t, & c \in W^{1,1}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big), \\ \infty, & \textit{otherwise}. \end{cases}$$

Proof. By Proposition (4.1) we can assume that $c \in W^{1,1}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$ whenever $I(c) < \infty$. Observe that $t \mapsto \xi \cdot \dot c(t) - H(c(t),\xi)$ is measurable for any ξ, and hence L is measurable in t, and the time integral is well-defined. Clearly we have the inequality $I(c) \leq \int_0^T L\big(c(t),\dot c(t)\big)\,\mathrm{d}t$.

To prove the other inequality, take a countable set $\{\zeta^{(n)} \colon n \in \mathbb{N}\}$ from Lemma 4.3 over which to take the supremum in (4.1) and add the zero vector if it is missing. This means that, given $\epsilon > 0$ and for each t individually we can always find n such that $\zeta^{(n)} \cdot \dot c(t) - H\left(c(t),\zeta^{(n)}\right) \geq \max\left(L(c,s) - \epsilon, 0\right)$ and thus define inductively the (measurable) sets,

$$A_{n,\epsilon} := \Big\{t \in (0,T) \colon \max\left\{0, L\left(c(t),\dot c(t)\right) - \epsilon\right\} \leq \zeta^{(n)} \cdot \dot c(t) - H(c,\zeta^{(n)})\Big\} \setminus \bigcup_{i=1}^{n-1} A_{i,\epsilon}.$$

Then for $a > 0$ we can define measurable functions $\xi^{(\epsilon)}, \xi^{(a,\epsilon)} \in L^\infty\left(0,T;\mathfrak{l}^\infty(\mathcal{Y})\right)$ by

$$\xi^{(\epsilon)}(t) := \sum_{n=1}^{\infty} \zeta^{(n)} \mathbb{1}_{A_{n,\epsilon}}(t), \qquad \text{and} \qquad \xi^{(a,\epsilon)}(t) := \xi^{(\epsilon)}(t) \mathbb{1}_{\left\{|\xi^{(\epsilon)}(t)| \leq a\right\}}.$$

If $0 < a < a' < \infty$ we now have

$$\begin{aligned} 0 &\leq \xi^{(a,\epsilon)}(t) \cdot \dot c(t) - H\left(c(t),\xi^{(a,\epsilon)}(t)\right) \\ &\leq \xi^{(a',\epsilon)}(t) \cdot \dot c(t) - H\left(c(t),\xi^{(a',\epsilon)}(t)\right) \\ &\leq \xi^{(\epsilon)}(t) \cdot \dot c(t) - H\left(c(t),\xi^{(\epsilon)}(t)\right) \end{aligned}$$

so by monotone convergence

$$\begin{aligned} &\lim_{a\to\infty} \int_0^T \left[\xi^{(a,\epsilon)}(t) \cdot \dot c(t) - H\left(c(t),\xi_{a,\epsilon}(t)\right)\right] \mathrm{d}t \\ &\qquad = \int_0^T \left[\xi^{(\epsilon)}(t) \cdot \dot c(t) - H\left(c(t),\xi^{(\epsilon)}(t)\right)\right] \mathrm{d}t \geq \int_0^T L\left(c(t),\dot c(t)\right) \mathrm{d}t - \epsilon T \end{aligned}$$

and we can send $\epsilon \to 0$ to see that

$$\int_0^T L\left(c(t),\dot c(t)\right) \mathrm{d}t \leq \sup_{\xi \in L^\infty(0,T;\mathfrak{l}^\infty(\mathcal{Y}))} \int_0^T \left[\xi(t) \cdot \dot c(t) - H\big(c(t),\xi(t)\big)\right] \mathrm{d}t.$$

□

4.2 The dual formulation of the rate functional

We first establish that the action functions (4.1) and (4.2) are equal.

Proposition 4.5. $L \equiv \widetilde{L}$.

Proof. We prove equality of the convex duals, i.e. for any $\xi \in \mathfrak{l}^\infty(\mathcal{Y})$,

$$\begin{aligned} \sup_{s \in \mathfrak{l}^1(\mathcal{Y})} \xi \cdot s - \inf_{u \colon \Gamma u = s} \mathrm{Ent}_c(u) &= \sup_{u \in \mathfrak{l}^1_{\geq 0}(\mathcal{R})} \xi \cdot \Gamma u - \mathrm{Ent}_c(u) \\ &\leq \sum_{r\in\mathcal{R}} \bar k^{(r)}(c) \sup_{u^{(r)} \geq 0} \left(\frac{u^{(r)}}{\bar k^{(r)}(c)} \xi \cdot \gamma^{(r)} - \frac{u^{(r)}}{\bar k^{(r)}(c)} \log \frac{u^{(r)}}{\bar k^{(r)}(c)} - \frac{u^{(r)}}{\bar k^{(r)}(c)} + 1\right) \\ &= \sum_{r\in\mathcal{R}} \bar k^{(r)}(c) \big(e^{\xi\cdot\gamma^{(r)}} - 1\big) = H(c,\xi). \end{aligned}$$

In fact, the inequality above is an equality, since the pointwise maximiser $u \in \mathbb{R}^{\mathcal{R}}$, given by $u^{(r)} = \bar{k}^{(r)}(c)e^{\xi\cdot\gamma^{(r)}}$, lies in $\mathfrak{l}^1(\mathcal{R})$ because of Assumptions (3.3c), (3.3f) and $\xi \in \mathfrak{l}^\infty(\mathcal{Y})$. Since both $\widetilde{L}$ and L are convex and have the same convex dual they must be equal. □

Next, using the existence of the minimiser that we will prove in the next subsection, we get that the infimum can be pulled out of the time-integral, which leads to the dual formulation (1.5).

Proposition 4.6. $I = \tilde{I}$.

Proof. Using Propositions 4.4 and 4.5, we get

$$I(c) = \int_0^T L\big(c(t), \dot{c}(t)\big)\,\mathrm{d}t = \int_0^T \inf_{\substack{u(t)\in\mathfrak{l}^1_{\geq 0}(\mathcal{R})\\ \Gamma u(t)=\dot{c}(t)}} \mathrm{Ent}_{c(t)}\big(u(t)\big)\,\mathrm{d}t \qquad (4.7)$$

$$\leq \inf_{\substack{u\in L^1(0,T;\mathfrak{l}^1_{\geq 0}(\mathcal{R}))\\ \Gamma u=\dot{c}}} \int_0^T \mathrm{Ent}_{c(t)}\big(u(t)\big)\,\mathrm{d}t.$$

If the left-hand side is infinite, then so is the right-hand side. Now consider the case when the left-hand side is finite. Then from Proposition 4.7(b) it then follows that the minimiser $u(t)$ in $\widetilde{L}\big(c(t), \dot{c}(t)\big)$ exists t-almost everywhere, and by Proposition 4.8 this pointwise minimiser lies in fact in $L^1\big(0, T; \mathfrak{l}^1_{\geq 0}(\mathcal{R})\big)$. This shows that the above inequality is indeed an equality. □

4.3 Existence of minimisers and maximisers

We first study the minimisers in the definition of $\widetilde{L}$, that is, pointwise in t.

Proposition 4.7. *Let $(c, s) \in \mathfrak{l}^1_{c(0)}(\mathcal{Y}) \times \mathfrak{l}^1(\mathcal{Y})$ be such that $L(c, s) < \infty$. Then*

(a) the function $\mathrm{Ent}_c(u)$ has (strongly) compact sublevel sets in $\mathfrak{l}^1_{\geq 0}(\mathcal{R})$;

(b) there is a unique minimiser $u \in \mathfrak{l}^1_{\geq 0}(\mathcal{R})$ such that $L(c, s) = \mathrm{Ent}_c(u)$ and $\Gamma u = s$;

(c) $s \in \operatorname{Ran}\Gamma$;

(d) for the unique minimiser from (b) we have $k^{(r)}(c) > 0 \iff u^{(r)} > 0$.

We stress that (d) is the only place where we use the weak reversibility condition (3.3h).

Proof. For brevity we omit dependencies on (c, s).

(a) The proof of this statement was suggested to us by Alex Mielke. The function $\mathrm{Ent}(u)$ is lower semicontinuous, which follows from Fatou's Lemma together with the fact that $u_n \to u$ implies $u_n^{(r)} \to u^{(r)}$ for all $r \in \mathcal{R}$. It thus suffices to prove relative compactness of the sublevel sets.

Since the sum $\sum_{r\in\mathcal{R}} \bar{k}^{(r)}$ is finite, there is a sequence $b^{(r)}$ in $\mathbb{R}_{\geq 0}$ such that $\sup_{r\in\mathcal{R}} b^{(r)} = \infty$ but we still have $\sum_{r\in\mathcal{R}} \bar{k}^{(r)} b^{(r)} < \infty$. Without loss of generality we assume that $\inf_{r\in\mathcal{R}} \log b^{(r)} > 0$. By convexity of the Boltzmann function λ_B we have

$$\bar{k}^{(r)}\lambda_B(u^{(r)}/\bar{k}^{(r)}) \geq \bar{k}^{(r)}\lambda_B\big((\bar{k}^{(r)}b^{(r)})/\bar{k}^{(r)}\big) + (u^{(r)} - \bar{k}^{(r)}b^{(r)})\lambda_B'\big((\bar{k}^{(r)}b^{(r)})/\bar{k}^{(r)}\big)$$
$$= u^{(r)}\log b^{(r)} - \bar{k}^{(r)}b^{(r)} + \bar{k}^{(r)}.$$

Thus, in sublevels $\{u \in \mathfrak{l}^1_{\geq 0}(\mathcal{R}) : \mathrm{Ent}(u) \leq E\}$ we have the estimate

$$E \geq \mathrm{Ent}(u) \geq \sum_{r\in\mathcal{R}} u^{(r)} \log b^{(r)} - \underbrace{\sum_{r\in\mathcal{R}} \bar{k}^{(r)}(b^{(r)} - 1)}_{=:a}.$$

Take any sequence $(u_n)_{n\geq 1}$ in the sublevel set. The estimate above together with the fact that $\inf_{r\in\mathcal{R}} \log b^{(r)} > 0$ shows that $|u|_1 \leq \sum_{r\in\mathcal{R}} u_n^{(r)} \leq E + a$. By Banach-Alaoglu we can then extract a weak-* converging subsequence, meaning that (relabeling) $u_n^{(r)} \to u^{(r)}$ for all $r \in \mathcal{R}$. Moreover, by the above estimates (picking any ordering of the reaction space $\mathcal{R}$),

$$\begin{aligned} 0 \leq \sup_{n\geq 1} \sum_{r\geq R} u_n^{(r)} &\leq \frac{1}{\inf_{r\geq R} \log b^{(r)}} \sup_{n\geq 1} \sum_{r\geq R} u^{(r)} \log b^{(r)} \\ &\leq \frac{1}{\inf_{r\geq R} \log b^{(r)}} (E + a) \xrightarrow{R\to\infty} 0, \end{aligned}$$

which proves that the sequence $u_n^{(r)}/\bar{k}^{(r)}$ is uniformly integrable in the measure space $(\mathcal{R}, 2^{\mathcal{R}}, k)$. Therefore $\sum_{r\in\mathcal{R}} \bar{k}^{(r)} |u_n^{(r)}/\bar{k}^{(r)} - u^{(r)}/\bar{k}^{(r)}| \to 0$ by the Vitali Convergence Theorem, so that the (sub)sequence u_n converges strongly in $\mathfrak{l}^1(\mathcal{Y})$.

(b) Since $\mathrm{Ent}(u) < L(c, s) < \infty$ we can take a minimising sequence $(u_n)_{n\geq 1} \subset \{u \in \mathfrak{l}^1_{\geq 0}(\mathcal{R}) : \Gamma u = s\}$. By (a) there is subsequence that converges strongly, and the limit is a minimiser because of the lower semicontinuity. The limit lies in the feasible set as the mapping Γ is continuous. Finally, by the strict convexity of the functional the minimiser is unique.

(c) This follows immediately from Proposition 4.5.

(d) Clearly if $k^{(r)} = 0$ then $u^{(r)} = 0$ else $L(c, s) = \infty$. Now assume that there exists a $r \in \mathcal{R}$ for which $k^{(r)} > 0$ but $u^{(r)} = 0$. Take a $v \in \mathfrak{c}_{c,>0}(\mathcal{R})$ from Assumption (3.3h) such that $\Gamma v = 0$ and $\bar{k}^{(\hat{r})}(c) > 0$ for all $\hat{r} \in \operatorname{supp} v$. Then for any $\epsilon > 0$ we have $\Gamma(u + \epsilon v) = \Gamma u = s$, and so $u + \epsilon v$ it is also a feasible candidate for the minimiser. However,

$$\frac{d}{d\epsilon}\mathrm{Ent}(u + \epsilon v) = \sum_{\hat{r}\in\operatorname{supp} v} \log \frac{u^{(\hat{r})} + \epsilon v^{(r)}}{\bar{k}^{(r)}} \xrightarrow{\epsilon\to 0} \sum_{\hat{r}\in\operatorname{supp} v} v^{(\hat{r})} \log \frac{u^{(\hat{r})}}{\bar{k}^{(r)}} = -\infty,$$

because $v^{(r)} > 0$ and $u^{(r)} = 0$ but $k^{(r)} > 0$. Therefore there exists an $\epsilon > 0$ such that $\mathrm{Ent}(u + \epsilon v) < \mathrm{Ent}(u)$, which contradicts the assumption that u is a minimiser.

□

Next we prove existence and uniqueness of the pathwise minimiser.

Proposition 4.8. *Let $c \in W^{1,1}\big(0, T; \mathfrak{l}^1_{\geq 0}(\mathcal{Y})\big)$ such that $I(c) < \infty$. Then the pointwise minimiser $t \mapsto u(t)$ in* (4.7) *lies in $L^1\big(0, T; \mathfrak{l}^1_{\geq 0}(\mathcal{R})\big)$, and it is therefore also a pathwise minimiser in the expression $\tilde{I}$ from* (1.5)*. Moreover, the pathwise minimiser u is unique, and t-almost everywhere $u(t) > 0$.*

Proof. Naturally, $I(c)<\infty$ implies that, t-almost everywhere, we have $\widetilde{L}\big(c(t),\dot c(t)\big)<\infty$, and hence the pointwise minimiser $u(t)$ in $\widetilde{L}$ exists by Proposition 4.7(b). Then by Jensen's inequality we get

$$\begin{aligned}\infty > I(c) &= \sum_{r\in\mathcal{R}} \int_0^T \bar{k}^{(r)}\big(c(t)\big)\lambda_B\left(\frac{u^{(r)}(t)}{\bar{k}^{(r)}(c(t))}\right)\mathrm{d}t\\ &\geq \left(\sum_{r}\int_0^T \bar{k}^{(r)}\big(c(t)\big)\,\mathrm{d}t\right)\lambda_B\left(\frac{1}{\sum_{r\in\mathcal{R}}\int_0^T\bar{k}^{(r)}\big(c(t)\big)\,\mathrm{d}t}\sum_{r\in\mathcal{R}}\int_0^T u^{(r)}(t)\,\mathrm{d}t\right).\end{aligned}$$

We can assume that $0<\sum_{r\in\mathcal{R}}\int_0^T\bar{k}^{(r)}\big(c(t)\big)\,\mathrm{d}t<\infty$; if it were zero then $I(c)=\infty$, and the finiteness follows from Assumption (3.3f). By continuity of the Boltzmann function λ_B we get that $\|u\|_{L^1(0,T;\mathfrak{l}^1(\mathcal{R}))}<\infty$, and so the pointwise minimiser is indeed a pathwise minimiser in $L^1\big(0,T;\mathfrak{l}^1_{\geq 0}(\mathcal{R})\big)$. The uniqueness follows from strict convexity. The fact that $u(t)>0$ follows from Proposition 4.7(d). □

Using the pathwise minimiser of the dual formulation $\widetilde{I}$, we can now construct the pathwise maximiser in the definition of the rate functional I.

Proposition 4.9. *Let $c\in\mathrm{BV}\big(0,T;\mathfrak{l}^1_{\geq 0}(\mathcal{Y})\big)$ such that $I(c)<\infty$. Then there exists a maximiser $\xi\in L^\infty\big(0,T;\mathfrak{l}^\infty(\mathcal{Y})\big)$ in the maximisation problem* (4.6)*. Moroever this maximiser is related to the minimiser $u\in L^1\big(0,T;\mathfrak{l}^1(\mathcal{R})\big)$ from Proposition 4.8 through $u^{(r)}(t)=\bar{k}^{(r)}\big(c(t)\big)e^{\xi(t)\cdot\gamma^{(r)}}$.*

Proof. By Proposition (4.8) there exists a unique minimiser $u\in L^1\big(0,T;\mathfrak{l}^1_{\geq 0}(\mathcal{R})\big)$ of the constraint minimisation problem (4.6). The constraint $t\mapsto \Gamma u(t)=t\mapsto\dot c(t)$ lies in $L^1\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$, and since u is non-zero and the constraint is linear, the minimiser is a regular point. Then [Lue69, §9.3, Th. 1] there exists a Lagrange multiplier $\xi\in L^\infty\big(0,T;\mathfrak{l}^\infty(\mathcal{Y})\big)$ such that u is a stationary point of $\hat u\mapsto\mathrm{Ent}(\hat u)+\langle\xi,\dot c-\Gamma\hat u\rangle$. This leads to the stability equation:

$$0=\log\frac{u^{(r)}(t)}{\bar{k}^{(r)}\big(c(t)\big)}-\xi(t)\cdot\gamma^{(r)}\qquad\text{for all } r\in\mathcal{R}\text{ and almost every } t\in(0,T),$$

from which we deduce that the minimiser has the form $u^{(r)}(t)=\bar{k}^{(r)}\big(c(t)\big)e^{\xi(t)\cdot\gamma^{(r)}}$. Plugging this representation back in the dual formulation yields, using 4.2 and Propositions 4.6,

$$\begin{aligned}&\sup_{\hat\xi\in L^\infty(0,T;\mathfrak{l}^\infty(\mathcal{Y}))}\int_0^T\Big[\hat\xi(t)\cdot\dot c(t)-H\big(c(t),\hat\xi(t)\big)\Big]\,\mathrm{d}t=I(c)=\widetilde{I}(c)\\ &\quad=\int_0^T\mathrm{Ent}_{c(t)}\Big(\bar{k}^{(r)}\big(c(t)\big)e^{\xi(t)\cdot\gamma^{(r)}}\Big)\,\mathrm{d}t=\int_0^T\Big[\xi(t)\cdot\dot c(t)-H\big(c(t),\xi(t)\big)\Big]\,\mathrm{d}t.\end{aligned}$$

We then see that the Lagrange multiplier for the minimisation problem is a maximiser of the maximisation problem 4.6. □

Remark 4.10. Without the weak reversibility condition 3.3h, the minimiser u might not be a regular point, in which case the Lagrange Multiplier Theorem cannot be used. In that case, the maximiser ξ may not exist at all, even for the case of one reaction $|\mathcal{R}|=1$. Indeed for one reaction, the large deviation rate of a constant path $c(t)\equiv c(0)$ is $I(c)=\sup_\xi -\bar{k}\big(c(0)\big)\int_0^T(e^{\xi(t)}-1)\,\mathrm{d}t=\bar{k}\big(c(0)\big)<\infty$, but a maximising ξ does not exist. □

5 Large deviations

In this section we prove the main result of this paper.

Theorem 5.1. *Let Assumptions 3.3 and 3.1 be satisfied. The sequence* $\big(C^{(V)}(\cdot)\big)_{V>0}$ *satisfies a large-deviation principle in* $\mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$ *in the hybrid topology with unique good rate functional* I*, defined by* (1.4).

Proof. In Lemma 5.2 we first prove that the sequence is exponentially tight. In Lemma 5.5 we prove that for any hybrid-open set $\mathcal{O}\subset \mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$,

$$\liminf_{V\to\infty}\frac{1}{V}\log \mathbb{P}^{(V)}(\mathcal{O})\geq -\inf_{c\in\mathcal{O}} I(c),$$

and in Lemma 5.6 we show that for any hybrid-compact set $\mathcal{K}\subset \mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$,

$$\limsup_{V\to\infty}\frac{1}{V}\log \mathbb{P}^{(V)}(\mathcal{K})\leq -\inf_{c\in\mathcal{K}} I(c).$$

Hence, the process satisfies a weak large deviation principle, which together with exponential tightness yields the full large deviation principle with a good rate functional [DZ87, Lem. 1.2.18]. The uniqueness follows from the fact that the hybrid topology is finer than the L^1-topology, and therefore Hausdorff and regular [DZ87, Lem. 4.1.4]. □

5.1 Exponential tightness

The exponential tightness follows naturally from the compactness Theorem 2.3.

Lemma 5.2. *The measures* $(\mathbb{P}^{(V)})_{V\geq 0}$ *are exponentially tight in* $\mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$ *with the hybrid topology.*

Proof. For $\eta>0$ we define the set of curves,

$$\mathcal{B}_\eta := \big\{c\in \mathrm{BV}\big(0,T;\mathfrak{l}^1_{\geq 0}(\mathcal{Y})\big) : \mathrm{var}(c)\leq \eta\|\Gamma\|+1\big\}, \tag{5.1}$$

Denote by $\Lambda^{(r,V)}$ the number reactions r that take place during the time interval $(0,T)$, and let $\mathcal{N}^{(V)}$ be a Poisson process with intensity κV where

$$\kappa = 1+\sup_{c\in K}\sum_{r\in\mathcal{R}}\bar{k}^{(r)}(c),$$

which is finite by (3.3e) and (3.3f). Then for any sufficiently large V,

$$\begin{aligned}
\mathbb{P}^{(V)}\big(\mathcal{B}^c_\eta\big) &= \mathrm{Prob}\Big(\textstyle\sum_{r\in\mathcal{R}}\Lambda^{(r,V)}\frac{|\gamma^{(r)}|_1}{V}>\eta\|\Gamma\|\Big)\\
&\leq \mathrm{Prob}\Big(\textstyle\sum_{r\in\mathcal{R}}\Lambda^{(r,V)}>V\eta\Big)\\
&\leq \mathrm{Prob}\Big(\mathcal{N}^{(V)}_T>V\eta\Big)\\
&\overset{\text{(Chernoff)}}{\leq} \exp\left(V\left(\kappa Te-\eta\right)\right).
\end{aligned}$$

Let K be defined by (3.1). Since $C^{(V)}(t)$ remains in K almost surely by Proposition 3.5, we find

$$\limsup_{V\to\infty} \frac{1}{V} \log \mathbb{P}^{(V)}\Big(\big(K^{(0,T)} \cap \mathcal{B}_\eta\big)^c\Big) \leq \kappa T e - \eta.$$

The set $K^{(0,T)} \cap \mathcal{B}_\eta$ is bounded in $L^1\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$ and therefore hybrid-compact by Lemma 3.2 and Theorem 2.3. □

5.2 The lower bound

We first prove the lower bound for a set of sufficiently regular curves:

$$\mathcal{A} := \Big\{ c \in \mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big) : \dot{c}(t) = \sum_{r\in\mathcal{R}} \bar{k}^{(r)}\big(c(t)\big) e^{\xi(t)\cdot\gamma^{(r)}} \gamma^{(r)} \quad \text{for some } \xi \in C^2_b\big(0,T;\mathfrak{c}_0(\mathcal{Y})\big) \Big\}.$$

Lemma 5.3. *Let $\mathcal{O} \subset \mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$ be any hybrid-open set. Then for any $c \in \mathcal{O} \cap \mathcal{A}$,*

$$\liminf_{V\to\infty} \frac{1}{V} \log \mathbb{P}^{(V)}(\mathcal{O}) \geq -I(c). \tag{5.2}$$

Proof. To ease notation, let us introduce the functional

$$\begin{aligned} G^{(V)}(c,\xi) &:= \int_0^T \Big[\xi(t)\cdot\dot{c}(t) - H^{(V)}\big(c(t),\xi(t)\big)\Big]\,\mathrm{d}t, \quad \text{where} \\ H^{(V)}(c,\xi) &:= \frac{1}{V}\sum_{r\in\mathcal{R}} k^{(r,V)}(c)(e^{\xi\cdot\gamma^{(r)}} - 1). \end{aligned}$$

We can assume that $I(c) < \infty$, else the claim is trivial. Take a $c \in \mathcal{O} \cap \mathcal{A}$ with a corresponding $\xi \in C^2_b\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$. For any $\epsilon > 0$ we define the set

$$\mathcal{U}_\epsilon(c) := \left\{ \hat{c} \in \mathrm{BV}\left(0,T;\mathfrak{l}^1(\mathcal{Y})\right) : G(\hat{c},\xi) < G(c,\xi) + \epsilon \right\}.$$

By standard tilting arguments and the Girsanov Theorem [KL99, A1, Th. 7.3],

$$\begin{aligned} \frac{1}{V}\log \mathbb{P}^{(V)}\big(\mathcal{O}\big) &\geq \frac{1}{V}\log \mathbb{P}^{(V)}\big(\mathcal{O}\cap\mathcal{U}_\epsilon(c)\big) \\ &\geq \frac{1}{V}\log \mathbb{P}^{(V)}_\xi\big(\mathcal{O}\cap\mathcal{U}_\epsilon(c)\big) + \frac{1}{V}\log \mathbb{P}^{(V)}\text{-}\operatorname*{ess\,inf}_{\hat{c}\in\mathcal{O}\cap\mathcal{U}_\epsilon(c)} \frac{d\mathbb{P}^{(V)}}{d\mathbb{P}^{(V)}_\xi}(\hat{c}) \\ &\geq \frac{1}{V}\log \mathbb{P}^{(V)}_\xi\big(\mathcal{O}\cap\mathcal{U}_\epsilon(c)\big) - \sup_{\hat{c}\in\mathcal{O}\cap\mathcal{U}_\epsilon(c)} G^{(V)}(\hat{c},\xi). \end{aligned} \tag{5.3}$$

To bound the first term in (5.3), recall from Lemma 3.8 that $\mathbb{P}^{(V)}_\xi \rightharpoonup \delta_c$. Moreover, because of the continuity of $c \mapsto G(c,\xi)$ the set $\mathcal{U}_\epsilon(c)$ is hybrid-open. Then by [HPR] we can use a generalised Portmanteau Theorem:

$$\liminf_{V\to\infty} \mathbb{P}^{(V)}_\xi\big(\mathcal{O}\cap\mathcal{U}_\epsilon\big) \geq \delta_c\big(\mathcal{O}\cap\mathcal{U}_\epsilon\big) = 1,$$

so that the first term in (5.3) vanishes. To bound the second term we use the definition of the set $\mathcal{U}_\epsilon(c)$ to get:

$$\limsup_{V\to\infty} \sup_{\hat{c}\in\mathcal{O}\cap\mathcal{U}_\epsilon(c)} G^{(V)}(\hat{c},\xi) \\ \leq G(c,\xi) + \epsilon + \limsup_{V\to\infty} \sup_{\hat{c}\in\mathcal{O}\cap\mathcal{U}_\epsilon(c)} \int_0^T \left[H\big(\hat{c}(t),\xi(t)\big) - H^{(V)}\big(\hat{c}(t),\xi(t)\big)\right] \mathrm{d}t.$$

Using Assumption (3.3e) the final term vanishes as $V\to\infty$ so that for arbitrary $\epsilon>0$

$$\liminf_{V\to\infty} \frac{1}{V} \log \mathbb{P}^{(V)}(\mathcal{O}) \geq -G(c,\xi) - \epsilon. \tag{5.4}$$

Therefore, the same inequality is also true with $\epsilon = 0$. The desired lower bound (5.2) holds as ξ maximises $G(\xi,c)$. □

In order to improve the lower bound to any open set we need an approximation argument. For this we will use the following Lemma.

Lemma 5.4. *Let $c \in \mathrm{BV}\big(0,T;\mathfrak{l}^1_{c(0)}(\mathcal{Y})\big)$ with $I(c)<\infty$. Then there is a sequence $(\xi^{(n)})_{n\in\mathbb{N}} \subset C_b^2\big(0,T;\mathfrak{l}^\infty(\mathcal{Y})\big)$ and corresponding solutions $c^{(n)} \in W^{1,1}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$ to*

$$\dot{c}^{(n)}(t) = \sum_{r\in\mathcal{R}} \bar{k}^{(r)}\big(c^{(n)}(t)\big) e^{\xi^{(n)}(t)\cdot\gamma^{(r)}}\gamma^{(r)}, \qquad c^{(n)}(0) = c(0), \tag{5.5}$$

such that

(1) $c^{(n)} \to c$ in the $W^{1,1}$-norm topology and thus also in the BV-norm and -hybrid topologies,

(2) $\lim_{n\to\infty} I(c^{(n)}) = I(c)$.

Proof. Let $u^{(r)}(t) = \bar{k}^{(r)}\big(c(t)\big)e^{\xi(t)\cdot\gamma^{(r)}}$ be the minimiser from Proposition 4.8, where $\xi \in L^\infty\big(0,T;\mathfrak{l}^\infty(\mathcal{Y})\big)$ is the maximiser from Proposition 4.9. Extend ξ to zero outside the interval $(0,T)$, and define the mollified function

$$\xi_y^{(n)}(t) := (\theta_n * \xi_y)(t), \qquad \text{where} \qquad \theta_n(t) := \sqrt{\frac{n}{4\pi}}e^{-\frac{nt^2}{4}}.$$

This sequence has the properties

- $\xi^{(n)}(t)\cdot\gamma^{(r)} \xrightarrow{n\to\infty} \xi(t)\cdot\gamma^{(r)}$ for all $r\in\mathcal{R}$ and almost every $t\in(0,T)$, and (5.6)
- $\sup_{n\in\mathbb{N}} \operatorname{ess\,sup}_{t\in(0,T)} \sup_{r\in\mathcal{R}} \xi^{(n)}(t)\cdot\gamma^{(r)} \leq \|\xi\|_\infty \|\Gamma\| < \infty.$ (5.7)

To prove *(1)*, let

$$a^{(n)}(t) := \sum_{r\in\mathcal{R}} \bar{k}^{(r)}\left(c(t)\right)\left|e^{\xi^{(n)}(t)\cdot\gamma^{(r)}} - e^{\xi(t)\cdot\gamma^{(r)}}\right| \leq 2e^{\|\xi\|_\infty\|\Gamma\|} \sup_{c\in\mathfrak{l}^1_{c(0)}(\mathcal{Y})} \sum_{r\in\mathcal{R}} \bar{k}^{(r)}(c).$$

This estimate is independent of t, and finite by Assumptions (3.3c) and (3.3f). Moreover, by assumption each term in the sum vanishes, and so by Fubini and dominated convergence we find that $\lim_{n\to\infty}\int_0^T a^{(n)}(t)\,\mathrm{d}t = 0$. One now has

$$\begin{aligned}\left|\dot c(t) - \dot c^{(n)}(t)\right|_1 &\leq \left|\sum_{r\in\mathcal{R}} \Big(u^{(r)}(t) - \bar k^{(r)}\big(c(t)\big)e^{\xi^{(n)}(t)\cdot\gamma^{(r)}}\Big)\,\gamma^{(r)}\right|_1 \\ &\quad + \left|\sum_{r\in\mathcal{R}} e^{\xi^{(n)}(t)\cdot\gamma^{(r)}}\left(\bar k^{(r)}\left(c(t)\right) - \bar k^{(r)}\left(c^{(n)}(t)\right)\right)\gamma^{(r)}\right|_1 \\ &\leq a^{(n)}(t)\|\Gamma\| + e^{\|\xi\|_\infty\|\Gamma\|}\operatorname{Lip}_{\sum_r \bar k^{(r)}}\|\Gamma\|\left|c(t)-c(t)\right|_1, \end{aligned} \tag{5.8}$$

which is finite by Assumptions (3.3c) and (3.3g). Thus by Gronwall's inequality we find that

$$\left|c(t) - c^{(n)}(t)\right|_1 \leq \|\Gamma\|\int_0^t a^{(n)}(s)e^{\|\xi\|_\infty\leq\|\Gamma\|(t-s)}\,\mathrm{d}s \leq \|\Gamma\|e^{\|\xi\|_\infty\leq\|\Gamma\|T}\int_0^T a^{(n)}(s)\,\mathrm{d}s \to 0$$

as proven above. Therefore $c^{(n)} \to c$ in the L^∞ and L^1 norms and, returning to (5.8), the same convergences also hold for the time derivatives.

To prove *(2)*, observe that

$$\begin{aligned} I(c^{(n)}) &= \sup_{\zeta\in L^\infty(0,T;\mathfrak{l}^\infty(\mathcal{Y}))}\sum_{r\in\mathcal{R}}\int_0^T \bar k^{(r)}\big(c^{(n)}(t)\big)\big(e^{\xi^{(n)}(t)\cdot\gamma^{(r)}}\zeta(t)\cdot\gamma^{(r)} - e^{\zeta(t)\cdot\gamma^{(r)}} + 1\big)\,\mathrm{d}t \\ &\leq \sum_{r\in\mathcal{R}}\int_0^T \bar k^{(r)}\big(c^{(n)}(t)\big)\big(e^{\xi^{(n)}(t)\cdot\gamma^{(r)}}\xi^{(n)}(t)\cdot\gamma^{(r)} - e^{\xi^{(n)}(t)\cdot\gamma^{(r)}} + 1\big)\,\mathrm{d}t \\ &\leq \big(\sup_{c\in\mathfrak{l}^1_{c(0)}(\mathcal{Y})}\textstyle\sum_{r\in\mathcal{R}}\bar k^{(r)}(c)\big)T\big(\|\xi\|_\infty\|\Gamma\|e^{\|\xi\|_\infty\|\Gamma\|}+1\big) < \infty. \end{aligned} \tag{5.9}$$

Hence again by Proposition 4.9 the maximiser ζ in the supremum above is attained; by a straightforward calculation it follows that this maximiser must be $\zeta = \xi^{(n)}$. We can thus write

$$\begin{aligned} |I(c) - I(c^{(n)})| &\leq \int_0^T \Big|L\big(c(t),\dot c(t)\big) - L\big(c^{(n)}(t),\dot c^{(n)}(t)\big)\Big|\,\mathrm{d}t \\ &\leq \int_0^T\sum_{r\in\mathcal{R}}\bar k^{(r)}\left(c(t)\right)\Big|\lambda_{\mathrm{B}}\Big(e^{\xi(t)\cdot\gamma^{(r)}}\Big) - \lambda_{\mathrm{B}}\Big(e^{\xi^{(n)}(t)\cdot\gamma^{(r)}}\Big)\Big|\,\mathrm{d}t \\ &\quad + \int_0^T\sum_{r\in\mathcal{R}}\lambda_{\mathrm{B}}\left(e^{\xi^{(n)}(t)\cdot\gamma^{(r)}}\right)\left|\bar k^{(r)}\left(c(t)\right) - \bar k^{(r)}\left(c^{(n)}(t)\right)\right|\,\mathrm{d}t. \end{aligned}$$

We finally show that this difference vanishes, which proves the claim. Indeed, by assumption and continuity all integrand terms vanish. Moreover the integrand terms are dominated by

$$\bar k^{(r)}\big(c(t)\big)\max\{1,\lambda_B(e^{\|\xi\|_\infty\|\Gamma\|})\} + \max\{1,\lambda_B(e^{\|\xi\|_\infty\|\Gamma\|})\}\big(\bar k^{(r)}\big(c(t)\big) + \bar k^{(r)}\big(c^{(r)}(t)\big),$$

which is summable over $r\in\mathcal{R}$ and integrable over $t\in(0,T)$ by Assumption (3.3f). The claim then follows by dominated convergence. □

The general lower bound now follows immediately from Proposition 5.3 and Lemma 5.4.

Lemma 5.5. *For any hybrid-open set $\mathcal{O}\subset\mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$,*

$$\liminf_{V\to\infty}\frac{1}{V}\log\mathbb{P}^{(V)}(\mathcal{O}) \geq -\inf_{c\in\mathcal{O}} I(c).$$

5.3 The upper bound

We conclude the large deviation proof with the upper bound.

Lemma 5.6. *For any hybrid-compact set* $\mathcal{K} \subset \mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$,

$$\limsup_{V\to\infty} \frac{1}{V}\log \mathbb{P}^{(V)}(\mathcal{K}) \leq -\inf_{c\in\mathcal{K}} I\big(c\big).$$

Proof. The argument mainly follows a standard covering technique as in the proof of the Gärtner-Ellis Theorem [DZ87, Th. 4.5.3]. Fix an arbitrary hybrid-compact set $\mathcal{K}$ and an $\epsilon > 0$. Because of (4.5), we can find, for any given $c \in \mathrm{BV}\big(0,T;\mathfrak{l}^1(\mathcal{Y})\big)$, a corresponding $\xi^{(c)} \in C^2_b\big(0,T;\mathfrak{l}^\infty(\mathcal{Y})\big)$ such that $G(c,\xi^{(c)}) \geq \inf_{c\in\mathcal{K}} I(c) - \epsilon$. Then by hybrid-continuity of $\hat{c} \mapsto G(\hat{c},\xi^{(c)})$ the sets

$$\mathcal{V}_\epsilon(c) := \big\{\hat{c} \in \mathfrak{l}^1(\mathcal{Y})\colon G(\hat{c},\xi^{(c)}) > G\left(c,\xi^{(c)}\right) - \epsilon\big\} \tag{5.10}$$

form an open covering. By compactness, there is a finite covering $\bigcup_{n=1,\ldots,N}\mathcal{V}_\epsilon(c^{(n)}) \supset \mathcal{K}$. Observe that the functions $\xi^{(c^{(n)})}$ are sufficiently regular to apply a Girsanov transformation [KL99, A1,Th. 7.3]. Hence for each n we find, similarly to (5.3),

$$\begin{aligned}
\limsup_{V\to\infty}\frac{1}{V}\log\mathbb{P}^{(V)}\big(\mathcal{V}_\epsilon(c^{(n)})\big) &\leq \limsup_{V\to\infty}\frac{1}{V}\underbrace{\log\mathbb{P}^{(V)}_{\xi^{(c^{(n)})}}\big(\mathcal{V}_\epsilon(c^{(n)})\big)}_{\leq 0} \\
&\qquad + \frac{1}{V}\log\mathbb{P}^{(V)}\text{-}\operatorname*{ess\,sup}_{\hat{c}\in\mathcal{V}_\epsilon(c^{(n)})}\frac{d\mathbb{P}^{(V)}}{d\mathbb{P}^{(V)}_{\xi^{(c^{(n)})}}}(\hat{c}) \\
&\leq \limsup_{V\to\infty} -\inf_{\hat{c}\in\mathcal{V}_\epsilon(c^{(n)})} G^{(V)}(\hat{c},\xi^{(c^{(n)})}) \\
&= -\inf_{\hat{c}\in\mathcal{V}_\epsilon(c^{(n)})} G(\hat{c},\xi^{(c^{(n)})}),
\end{aligned}$$

since $\lim_{V\to\infty} G^{(V)}(\hat{c},\xi^{(c^{(n)})}) = G(\hat{c},\xi^{(c^{(n)})})$ uniformly in $\hat{c}$, due to assumption (3.3e).

We can now exploit the finiteness of the covering to use the Laplace Principle on the compact set:

$$\begin{aligned}
\limsup_{V\to\infty}\frac{1}{V}\log\mathbb{P}^{(V)}(\mathcal{K}) &\leq \max_{n=1,\ldots,N}\limsup_{V\to\infty}\frac{1}{V}\log\mathbb{P}^{(V)}\big(\mathcal{V}_\epsilon(c^{(n)})\big) \\
&\leq \max_{n=1,\ldots,N} -\inf_{\hat{c}\in\mathcal{V}_\epsilon(c^{(n)})} G(\hat{c},\xi^{(c^{(n)})}) \\
&\overset{(5.10)}{\leq} -\inf_{c\in\mathcal{K}} I(c) + 2\epsilon.
\end{aligned}$$

Since ϵ was arbitrary, this proves the claim. □

References

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

[AGS08] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics. ETH Zürich. Birkhauser, Basel, Switzerland, 2nd edition, 2008.

[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.

[BFG15] L. Bertini, A. Faggionato, and D. Gabrielli. Large deviations of the empirical flow for continuous time Markov chains. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 51(3):867–900, 2015.

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

[Con90] J.B. Conway. *A course in functional analysis*. Springer, Berlin, Germany, 2nd edition, 1990.

[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.

[DG87] D.A. Dawson and J. Gärtner. Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics*, 20(4):247–308, 1987.

[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.

[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.

[GGZ74] H. Gajewski, K. Gröger, and K. Zacharias. *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*. Akademie Verlag, Berlin, DDR, 1974.

[HPR] M. Heida, R.I.A. Patterson, and D.R.M. Renger. The space of bounded variation with infinite-dimensional codomain. In progress.

[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.

[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'I.H.P., section B*, 31(2):289–323, 1995.

[Lue69] D.G. Luenberger. *Optimization by vector space methods*. John Wiley & Sons, New York, N.Y., U.S.A., 1969.

[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.

[Pat16] R.I.A. Patterson. Properties of the solutions of delocalised coagulation and inception problems with outflow boundaries. *J. Evoln. Eqns.*, in press, 2016.

[Roc70] R.T. Rockafellar. *Convex analysis*. Princeton university press, Princeton, N.J., U.S.A., 1970.

[Rud73] W. Rudin. *Functional Analysis*. McGraw-Hill, New York, NY, USA, 1973.

[Sie76] D. Siegmund. Importance sampling in the Monte Carlo study of sequential tests. *Annals of Statistics*, 4(4):673–684, 1976.

[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.