Revisit of Macroscopic Dynamics for Some Non-equilibrium Chemical Reactions from a Hamiltonian Viewpoint

Yuan Gao[1] · **Jian-Guo Liu**[2]

Abstract

Most biochemical reactions in living cells are open systems interacting with environment through chemostats to exchange both energy and materials. At a mesoscopic scale, the number of each species in those biochemical reactions can be modeled by a random time-changed Poisson processes. To characterize macroscopic behaviors in the large number limit, the law of large numbers in the path space determines a mean-field limit nonlinear reaction rate equation describing the dynamics of the concentration of species, while the WKB expansion for the chemical master equation yields a Hamilton–Jacobi equation and the Legendre transform of the corresponding Hamiltonian gives the good rate function (action functional) in the large deviation principle. In this paper, we decompose a general macroscopic reaction rate equation into a conservative part and a dissipative part in terms of the stationary solution to the Hamilton–Jacobi equation. This stationary solution is used to determine the energy landscape and thermodynamics for general chemical reactions, which particularly maintains a positive entropy production rate at a non-equilibrium steady state. The associated energy dissipation law at both the mesoscopic and macroscopic levels is proved together with a passage from the mesoscopic to macroscopic one. A non-convex energy landscape emerges from the convex mesoscopic relative entropy functional in the large number limit, which picks up the non-equilibrium features. The existence of this stationary solution is ensured by the optimal control representation at an undetermined time horizon for the weak KAM solution to the stationary Hamilton–Jacobi equation. Furthermore, we use a symmetric Hamiltonian to study a class of non-equilibrium enzyme reactions, which leads to nonconvex energy landscape due to flux grouping degeneracy and reduces the conservative–dissipative decomposition to an Onsager-type strong gradient flow. This symmetric Hamiltonian implies that the transition paths between multiple steady states (rare events in biochemical reactions) is a modified time reversed least action path with associated path affinities and energy barriers. We illustrate this

✉ Yuan Gao
  gao662@purdue.edu

  Jian-Guo Liu
  jliu@math.duke.edu

1 Department of Mathematics, Purdue University, West Lafayette, IN, USA

2 Department of Mathematics and Department of Physics, Duke University, Durham, NC, USA

⚡ Springer

idea through a bistable catalysis reaction and compute the energy barrier for the transition path connecting two steady states via its energy landscape.

Keywords Conservative–dissipative decomposition · Positive entropy production rate · Time reversal · Transition path with energy barrier · Thermodynamic limit · Large deviation principle

Mathematics Subject Classification 80A30 · 35F21 · 70H33 · 49N99

1 Introduction

At a mesoscopic scale, chemical or biochemical reactions can be understood from a probabilistic viewpoint. A convenient way to stochastically describe chemical reactions is via random time-changed Poisson processes; c.f. [3]. Based on this, one can observe 'statistical properties' of chemical reactions in the thermodynamic limit. For instance, the law of large numbers gives the 'mean path' of a chemical reaction while the large deviation principle can give rough estimates on the probability of the occurrence in a vicinity of any path, particularly the transition path (the most probable path) between two stable states; c.f. [21]. In this paper, we investigate various macroscopic behaviors for general chemical reactions, including the conservative–dissipative decomposition for macroscopic dynamics, the passage from mesoscopic to macroscopic free energy dissipation relations and symmetric structures brought by Markov chain detailed balance in some enzyme reactions. The studies for non-equilibrium thermodynamics and metastability in biochemical oscillations was pioneered by Prigogine [58]. The coexistence of multiple stable steady states breaks the chemical version detailed balance (1.12) or complex balance (2.6) properties for chemical reactions and leads to bifurcations and transition paths. However, some enzyme reactions, most important non-equilibrium reactions in an open system to maintain metabolite concentrations, can still be characterized by a process with a Markov chain detailed balance (4.3). This mesoscopic reaction process yields the energy landscape ψ^{ss} of the chemical reaction and the corresponding macroscopic Hamiltonian is symmetric w.r.t $\nabla \psi^{ss}$; see (4.5). This enables us to study transition paths, energy barriers and gradient flow structures for a class of non-equilibrium dynamics with multiple steady states. Before we introduce the main results, we first review some backgrounds for the macroscopic limiting ODE from the large number limit of the mesoscopic stochastic processes and backgrounds for the Wentzel–Kramers–Brillouin (WKB) expansion, a corresponding Hamiltonian H and the good rate function in an associated large deviation principle.

Background for Large Number Process and Its Macroscopic Limiting ODE

Chemical reaction with $i = 1, \ldots, N$ species X_i and $j = 1, \ldots, M$ reactions can be kinematically described as

$$\text{reaction } j : \quad \sum_i \nu_{ji}^+ X_i \; \underset{k_j^-}{\overset{k_j^+}{\rightleftharpoons}} \; \sum_i \nu_{ji}^- X_i, \qquad (1.1)$$
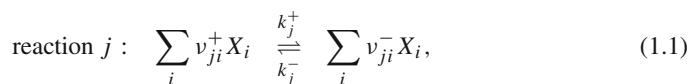
where nonnegative integers $\nu_{ji}^\pm \geq 0$ are stoichiometric coefficients and $k_j^\pm \geq 0$ are reaction rates for the j-th forward/backward reactions. Denote $\nu_{ji} := (\nu_{ji}^- - \nu_{ji}^+)$ as the net change

in molecular numbers for species X_i in the j-th forward reaction. The $M \times N$ matrix $\nu :=$ (ν_{ji}), $j = 1, \ldots, M$, $i = 1, \ldots, N$ as the Wegscheider matrix and ν^T is referred as the stoichiometric matrix [51]. The column vector $\vec{\nu}_j := \vec{\nu}_j^- - \vec{\nu}_j^+ := \left(\nu_{ji}^- - \nu_{ji}^+ \right)_{i=1:N} \in \mathbb{Z}^N$ is called the reaction vector for the j-th reaction. In this paper, all vectors $\vec{X} = (X_i)_{i=1:N} \in \mathbb{R}^N$ and $(\varphi_j)_{j=1:M}$, $(k_j)_{j=1:M} \in \mathbb{R}^M$ are column vectors. We remark the description (1.1) includes reactions both in a closed system and an open system. In an open system, $\vec{\nu}_j^+ = \vec{0}$ represents birth/death reaction with the corresponding reactant on the right hand side called chemostats and birth rate k_j^+, death rate k_j^-; see a bistable example in Sect. 4.4 for materials/energy exchange with environment.

Let the space of natural numbers $\mathbb{N}$ be the state space of the counting process $X_i(t)$ and let $X_i(t) \in \mathbb{N}$ for $i = 1, \ldots, N$ be the number of each species in those biochemical reactions. The random time-changed Poisson representation for chemical reactions (1.1) is, c.f. [3, 44],

$$\vec{X}(t) = \vec{X}(0) + \sum_{j=1}^{M} \vec{\nu}_j \left(\mathbb{1}_{\{\vec{X}(t_-) + \vec{\nu}_j \geq 0\}} Y_j^+ \left(\tilde{t}_j^+ \right) - \mathbb{1}_{\{\vec{X}(t_-) - \vec{\nu}_j \geq 0\}} Y_j^- \left(\tilde{t}_j^- \right) \right),$$

$$\tilde{t}_j^\pm(t) := \int_0^t \lambda_j^\pm(s) \, \mathrm{d}s,$$

(1.2)

where for the j-th reaction channel, $Y_j^\pm(t)$ are i.i.d. unit rate Poisson processes and $\mathbb{1}$ is the characteristic function indicating that there is no reaction if the next state $\vec{X}(t_-) \pm \vec{\nu}_j$ is negative for some component. Here and in the following a vector $\vec{x} \geq 0$ is understood as componentwisely nonnegative. The existence and uniqueness to the stochastic equation (1.2) was proved by [3, 44] in terms of the corresponding martingale problem. In (1.2), the intensity function $\lambda_j^\pm(s) = \varphi_j^\pm(\vec{X}(s))$ for the time clock $\tilde{t}_j^\pm(t)$ is usually chosen as the mesoscopic law of mass action (LMA)

$$\varphi_j^\pm(\vec{X}) = k_j^\pm V \prod_{\ell=1}^{N} \frac{X_\ell!}{V^{\nu_{j\ell}^\pm} \left(X_\ell - \nu_{j\ell}^\pm \right)!}.$$

(1.3)

Here $V \gg 1$ is the volume for species in the chemical reaction in a container. Because we assume chemical reactions in a container is independent of molecule position and the molecular number is proportional to the container volume, we call the limit for the large number of molecules as thermodynamic limit or macroscopic limit. Rescale the process (1.2) as $C_i^{\mathrm{v}} := \frac{X_i}{V}$ and denote the rescaled discrete state variable as $\vec{x}_{\mathrm{v}} := \frac{\vec{n}}{V}$, $\vec{n} \in \mathbb{N}^N$. Denote the forward and backward rescaled fluxes as $\tilde{\Phi}_j^\pm(\vec{x}_{\mathrm{v}}) := \frac{\varphi_j(\vec{n})}{V}$. Then the large number process C^{v} satisfies

$$C^{\mathrm{v}}(t) = C^{\mathrm{v}}(0) + \sum_{j=1}^{M} \frac{\vec{\nu}_j}{V} \left(\mathbb{1}_{\{C^{\mathrm{v}}(t_-) + \frac{\vec{\nu}_j}{V} \geq 0\}} Y_j^+ \left(V \int_0^t \tilde{\Phi}_j^+(C^{\mathrm{v}}(s)) \, \mathrm{d}s \right) \right.$$

$$\left. - \mathbb{1}_{\{C^{\mathrm{v}}(t_-) - \frac{\vec{\nu}_j}{V} \geq 0\}} Y_j^- \left(V \int_0^t \tilde{\Phi}_j^-(C^{\mathrm{v}}(s)) \, \mathrm{d}s \right) \right).$$

(1.4)

The 'no reaction' constraints above is to ensure that there is actually no jump if the number of some species will be negative in the container.

For a chemical reaction modeled by (1.2), denote the counting probability of $C^{\mathrm{v}}(t)$ as $p(\vec{x}_{\mathrm{v}}, t) = \mathbb{E}(\mathbb{1}_{\vec{x}_{\mathrm{v}}}(C^{\mathrm{v}}(t)))$, where $\mathbb{1}_{\vec{x}_{\mathrm{v}}}$ is the indicator function. Then $p(\vec{x}_{\mathrm{v}}, t)$ satisfies the chemical master equation (CME), c.f. [3]

$$\frac{\mathrm{d}}{\mathrm{d}t} p(\vec{x}_v, t) = (Q_v^* p)(\vec{x}_v, t) = V \sum_{j=1, \vec{x}_v - \frac{\vec{v}_j}{V} \geq 0}^{M} \left(\tilde{\Phi}_j^+(\vec{x}_v - \frac{\vec{v}_j}{V}) p(\vec{x}_v - \frac{\vec{v}_j}{V}, t) - \tilde{\Phi}_j^-(\vec{x}_v) p(\vec{x}_v, t) \right)$$

$$+ V \sum_{j=1, \vec{x}_v + \frac{\vec{v}_j}{V} \geq 0}^{M} \left(\tilde{\Phi}_j^-(\vec{x}_v + \frac{\vec{v}_j}{V}) p(\vec{x}_v + \frac{\vec{v}_j}{V}, t) - \tilde{\Phi}_j^-(\vec{x}_v) p(\vec{x}_v, t) \right).$$

(1.5)

Here Q_v^* is the transpose of the generator Q_v of process $C^v(t)$. Here we call the constraint $\vec{x}_v \pm \vec{v}_j h \geq 0$ the 'no reaction' boundary condition for CME which inherits from (1.4); see [24]. In [3, 43], this 'no reaction' restriction was omitted in the process C^v, thus to derive the master equation and generator including this 'no reaction' constraint, we give a pedagogical derivation in Appendix A. We refer to [28] for the existence and regularity of solutions to CME.

The mesoscopic jumping process C^v in (1.4) can be regarded as a large number interacting particle system. In the large number limit (thermodynamic limit), this interacting particle system can be approximately described by a mean field equation, i.e., a macroscopic nonlinear chemical reaction-rate equation. If the law of large numbers in the mean field limit holds, i.e., $p(\vec{x}_v, t) \to \delta_{\vec{x}(t)}$ for some $\vec{x}(t)$, then the limit $\vec{x}(t)$ describes the dynamics of the concentration of N species in the continuous state space $\mathbb{R}_+^N := \{\vec{x} \in \mathbb{R}^N; x_i \geq 0\}$ and is given by the following reaction rate equation (RRE), also known as chemical kinetic rate equation,

$$\frac{\mathrm{d}}{\mathrm{d}t} \vec{x} = \sum_{j=1}^{M} \vec{v}_j \left(\Phi_j^+(\vec{x}) - \Phi_j^-(\vec{x}) \right).$$

(1.6)

Here the macroscopic fluxes $\Phi_j^\pm$ satisfy the macroscopic LMA

$$\Phi_j^\pm(\vec{x}_v) = k_j^\pm \prod_{\ell=1}^{N} (x_\ell)^{v_{j\ell}^\pm},$$

(1.7)

which can be viewed as a large number approximation for the mesoscopic LMA (1.3). This RRE with LMA were first proposed by Guldberg and Waage in 1864. The limiting macroscopic LMA in RRE (1.6) is same as long as the mesoscopic LMA satisfies $\frac{\varphi_j(\vec{n})}{V} \approx \Phi_j^\pm(\vec{x}_v) = k_j^\pm \prod_{\ell=1}^{N} (x_\ell)^{v_{j\ell}^\pm}$. Indeed, Kurtz [43] proved the law of large numbers for the large number process $C^v(t)$; c.f., [3, Theorem 4.1]. Suppose $\Phi_j^\pm$ is local Lipschitz. If $C^v(0) \to \vec{x}(0)$ as $V \to +\infty$, then for any $\varepsilon > 0, t > 0$,

$$\lim_{V \to +\infty} \mathbb{P}\{ \sup_{0 \leq s \leq t} |C^v(s) - \vec{x}(s)| \geq \varepsilon \} = 0.$$

(1.8)

Thus we will also call the large number limiting ODE (1.6) as the macroscopic RRE. This gives a passage from the mesoscopic LMA (1.3) to the macroscopic one (1.7). In Appendix B, we give a pedagogical derivation for this mean field limit result (1.8) to include 'no reaction' boundary condition. We also refer to recent results in [49] which proves the evolutionary Γ-convergence from CME to the Liouville equation and thus starting from a deterministic state $\vec{x}_0$, [49, Theorem 4.7] recovers Kurtz's results on the mean field limit of CME.

If there exists a positive vector $\vec{m} \in \mathbb{R}_+^N$ (for instance due to the conservation of mass for each reaction j in a closed system) such that

$$\vec{v}_j \cdot \vec{m} = 0, \quad j = 1, 2, \ldots, M,$$

(1.9)

where $\vec{m} = (m_i)_{i=1:N}$ and m_i represents the molecular weight for the i-th species, then the Wegscheider matrix ν has a nonzero kernel, i.e., $\dim(\mathrm{Ker}(\nu)) \geq 1$ and we have a direct decomposition for the species space

$$\mathbb{R}^N = \mathrm{Ran}(\nu^T) \oplus \mathrm{Ker}(\nu). \tag{1.10}$$

Denote the stoichiometric space $G := \mathrm{Ran}(\nu^T)$. Given an initial state $\vec{X}_0 \in \vec{q} + G$, $\vec{q} \in \mathrm{Ker}(\nu)$, the dynamics of both mesoscopic (1.4) and macroscopic (1.6) states stay in the same space $G_q := \vec{q} + G$, called a stoichiometric compatibility class. We will see later the corresponding Hamiltonian/Lagrangian from WKB expansion for (1.5) are strictly convex in G while degenerate in $\mathrm{Ker}(\nu)$. Below we discuss the uniqueness of steady states within one stoichiometric compatibility class for the RRE detailed/complex balance case.

Denote a steady state to RRE (1.6) as $\vec{x}^{\mathrm{s}}$, which satisfies

$$\sum_{j=1}^{M} \vec{\nu}_j \left(\Phi_j^+(\vec{x}^{\mathrm{s}}) - \Phi_j^-(\vec{x}^{\mathrm{s}}) \right) = 0. \tag{1.11}$$

The detailed balance condition for RRE (1.6) is defined by Wegscheider 1901, Lewis 1925 as: (i) there exists a $\vec{x}^{\mathrm{s}} > 0$ (componentwise); and (ii) $\vec{x}^{\mathrm{s}}$ satisfies

$$\Phi_j^+(\vec{x}^{\mathrm{s}}) - \Phi_j^-(\vec{x}^{\mathrm{s}}) = 0, \quad \forall j. \tag{1.12}$$

We call RRE (1.6) detailed balanced if there exists such a detailed balanced state $\vec{x}^{\mathrm{s}}$. This immediately gives a necessary condition that both $k_j^{\pm} > 0$, i.e., the reaction is reversible. This concept of detailed balance for RRE (1.12) is commonly used in chemistry and biology, while it is different from the Markov chain detailed balance condition (4.3) for the mesoscopic jump process C^V. The latter is a more proper mathematical definition for the detailed balance condition and includes some non-equilibrium reactions which can not be characterized via the more constrained chemical version detailed balance (1.12); see Sect. 4. We will use both concepts in this paper, so we call (1.12) the detailed balance for RRE and (4.3) the Markov chain detailed balance for CME, respectively.

Under RRE detailed balance condition (1.12), all the positive steady solutions to (1.6) are detailed balanced and are characterized by $\vec{x}^{\mathrm{s}} e^{\vec{q}} > 0$ for some $\vec{q} \in \mathrm{Ker}(\nu)$. This is also true for a weaker condition called the complex balance condition (2.6). We summarize the well-known result on the uniqueness of the RRE detailed/complex balanced steady state discovered in [31, Theorem 6A] ([3, Theorem 3.5]) as Lemma 2.2 and the deficiency zero theorem proved by Horn and Feinberg [16, 31] is revisited in Sect. 2.2.1. Lemma 2.2 says that for each stoichiometric compatibility class, there is only one equilibrium steady state for a detailed/complex balanced RRE system.

On the contrary, non-equilibrium chemical reaction system has coexistent steady states and nonzero steady fluxes, so how to find transition paths between different non-equilibrium steady states and to compute the corresponding energy barriers for the macroscopic RRE (1.6) are the main goals of this paper. Particularly, in many biochemical reactions, such as heterogeneous catalytic oxidations and some enzyme reactions, the RRE detailed/complex balance conditions do not hold. We will use a symmetric Hamiltonian (1.26) to study some non-equilibrium reaction dynamics, which enables us to explore Onsager's strong form gradient flow structure (see (4.17)) and to compute the explicit transition path formula with the associated path affinity (see Proposition 4.6). Before describing our main results, let us further review some important properties for a Hamiltonian raising from WKB expansion of $p(\vec{x}_{\mathrm{v}}, t)$ below.

Background for WKB Expansion, Hamiltonian and Large Deviation Principle for Chemical Reactions

Besides the macroscopic trajectory $\vec{x}(t)$ given by the law of large numbers, WKB expansion for $p(\vec{x}_v, t)$ in CME (1.5) is another standard method [12, 27, 33, 38, 71], which builds up a more informative bridge between the mesoscopic dynamics and the macroscopic behaviors. We remark the WKB expansion has different names in different fields, such as the eikonal approximation, or the instanton technique, or the nonlinear semigroup, or the Cole–Hopf transformation.

To characterize the exponential asymptotic behavior, we assume there exists a continuous function $\psi(\vec{x}, t)$ such that $p(\vec{x}_v, t)$ has a WKB reformulation

$$p(\vec{x}_v, t) = e^{-V\psi(\vec{x}_v, t)}, \quad p(\vec{x}_v, 0) = p_0(\vec{x}_v). \tag{1.13}$$

The fluctuation on path space, i.e., the large deviation principle, can be computed through WKB expansion, and the good rate function for the large number process in a chemical reaction is rigorously proved by Agazzi et al. in [2, Theorem 1.6]; see explanations below. We know ψ satisfies

$$\partial_t \psi(\vec{x}_v, t) = -\frac{1}{V} e^{V\psi(\vec{x}_v, t)} Q_V^* e^{-V\psi(\vec{x}_v, t)} =: -\frac{1}{V} H_V^*(V\psi), \quad \psi(\vec{x}_v, 0) = -\frac{1}{V} \log p_0(\vec{x}_v). \tag{1.14}$$

By Taylor's expansion of $\psi(\vec{x} \pm \frac{\vec{v}_j}{V}, t)$ in (1.5) and (1.14), we obtain the following Hamilton–Jacobi equation (HJE) for the rescaled master equation (1.5) for ψ

$$\partial_t \psi(\vec{x}, t) = -\sum_{j=1}^{M} \left(\Phi_j^+(\vec{x}) \left(e^{\vec{v}_j \cdot \nabla \psi(\vec{x}, t)} - 1 \right) + \Phi_j^-(\vec{x}) \left(e^{-\vec{v}_j \cdot \nabla \psi(\vec{x}, t)} - 1 \right) \right); \tag{1.15}$$

see also derivations for (2.23) later. Define Hamiltonian $H(\vec{p}, \vec{x})$ on $\mathbb{R}^N \times \mathbb{R}^N$ as

$$H(\vec{p}, \vec{x}) := \sum_{j=1}^{M} \left(\Phi_j^+(\vec{x}) e^{\vec{v}_j \cdot \vec{p}} - \Phi_j^+(\vec{x}) + \Phi_j^-(\vec{x}) e^{-\vec{v}_j \cdot \vec{p}} - \Phi_j^-(\vec{x}) \right). \tag{1.16}$$

Then the HJE for $\psi(\vec{x}, t)$ can be recast as

$$\partial_t \psi + H(\nabla \psi, \vec{x}) = 0. \tag{1.17}$$

The WKB analysis above defines a Hamiltonian $H(\vec{p}, \vec{x})$, which contains almost all the information for the macroscopic dynamics. We remark this kind of WKB expansion was first used by Kubo et al. [38] for master equations for general Markov processes and later was applied to CME by Hu in [33]. In [12], Dykman et al. first derived the HJE (1.17) with the associated Hamiltonian H in (1.16) and reviewed the symmetry of Hamiltonian H in the RRE detailed balance case $H(\vec{p}, \vec{x}) = H(\log \frac{\vec{x}}{\vec{x}^s} - \vec{p}, \vec{x})$.

Equivalent to WKB reformulation for CME, one can define Varadhan's nonlinear semigroup [18, 73] for process $C^v(t)$ via the WKB reformulation for the backward equation

$$u(\vec{x}_v, t) = \frac{1}{V} \log \mathbb{E}^{\vec{x}_v} \left(e^{V u_0(C_t^v)} \right) =: (S_t u_0)(\vec{x}_v) \tag{1.18}$$

and as $V \to +\infty$, with the same Hamiltonian H, $u(\vec{x}, t)$ satisfies

$$\partial_t u - H(\nabla u, \vec{x}) = 0. \tag{1.19}$$

Comparing with (1.17), the limiting HJE after WKB expansion for forward and backward equation only has a sign difference in the time derivative. The rigorous convergence from the Varadhan's nonlinear semigroup (1.18) to the viscosity solution of HJE (1.19) was proved in [24] by reformulating (1.18) as a monotone scheme to HJE (1.19). Two difficulties brought by the 'no reaction' boundary condition and the polynomial growth rate for the coefficients $\Phi(\vec{x})_j^{\pm}$ in Hamiltonian when constructing unique viscosity solution were overcame in [24] by constructing upper/semicontinuous envelopes which inherit the 'no reaction' constraint and by constructing barriers to control far field values. Based on this convergence and the Lax–Oleinik's representation for the viscosity solution to HJE (1.19)

$$u(\vec{x}, t) = \sup_{\vec{y} \in \mathbb{R}^N} \left(u_0(\vec{y}) - I_{\vec{x},t}(\vec{y}) \right), \quad I_{\vec{x},t}(\vec{y}) = \inf_{\gamma(0)=\vec{x}, \gamma(t)=\vec{y}} \int_0^t L(\dot{\gamma}(s), \gamma(s)) \, \mathrm{d}s,$$

(1.20)

[24] verified the Varadhan's inverse lemma for the large deviation principle [10]. Here L is the convex conjugate of H and $I_{\vec{x},t}(\vec{y})$ is the least action between fixed initial point $\vec{x}$ and ending point $\vec{y}$ at time t. The Lax–Oleinik's representation for $u(\vec{x}, t)$ can be interpreted as a deterministic optimal control problem with terminal profit u_0 at t and running cost given by the least action $I_{\vec{x},t}$. Combining this convergence with the exponential tightness of $C^v(t)$ at single times, the large deviation principle for the random variable $C^v(t)$ at any time t with good rate function $I_{\vec{x}_0,t}(\cdot)$ was then proved. The sample path large deviation principle, which requires further the exponential tightness in the path space, is more involved and we refer to Agazzi et al. [2, Theorem 1.6]. The relation between the Hamiltonian H and the rate function in the large deviation principle for a general Markov process was introduced by Fleming and Sheu [19]; see also [18]. In Sects. 2.3 and 2.4, we summarize key properties for $H(\vec{p}, \vec{x})$ and its convex conjugate $L(\vec{s}, \vec{x})$ and their relations to the macroscopic RRE, the HJE for the phase function ψ, and also the good rate function in the large deviation principle. The solution $\vec{x}(t)$ to RRE (1.6) is shown to be a least action curve with zero action cost $Act(\vec{x}(\cdot)) = 0$; see Lemma 2.5. Indeed, $\vec{x}(t)$ is a curve following the Hamiltonian dynamics with zero momentum $\vec{p} = 0$.

Main Results

In Sect. 3, we utilize the dynamic and stationary solutions to HJE (1.17) to study the characterization and decomposition of RRE. We first recast the RRE (1.6) as a bi-characteristic of HJE

$$\frac{\mathrm{d}}{\mathrm{d}t} \vec{x} = \nabla_p H(\vec{0}, \vec{x}), \quad \vec{p} \equiv \vec{0}.$$

(1.21)

This directly gives the characterization of the macroscopic RRE trajectory $\vec{x}(t)$, i.e.,

$$\vec{x}(t) = \mathrm{argmin}_{\vec{x}} \, \psi(\vec{x}, t), \quad \text{for all } t \in [0, T].$$

(1.22)

Second, the stationary solution $\psi^{ss}(\vec{x})$ to HJE (1.17) plays the role of a free energy, by which, we decompose the RRE as a conservative part and a dissipation part

$$\dot{\vec{x}} = W(\vec{x}) - K(\vec{x}) \nabla \psi^{ss}(\vec{x}),$$
$$W(\vec{x}) := \int_0^1 \nabla_p H(\theta \nabla \psi^{ss}(\vec{x}), \vec{x}) \, \mathrm{d}\theta, \quad K(\vec{x}) := \int_0^1 (1 - \theta) \nabla_{pp}^2 H(\theta \nabla \psi^{ss}(\vec{x}), \vec{x}) \, \mathrm{d}\theta.$$

(1.23)

Here the conservative part is orthogonal to $\nabla \psi^{ss}$, i.e. $\langle W(\vec{x}), \nabla \psi^{ss}(\vec{x}) \rangle = 0$ and the dissipation part is expressed using Onsager's nonnegative definite response operator $K(\vec{x})$ for $\dot{\vec{x}}$ w.r.t generalized force $\nabla \psi^{ss}(\vec{x})$; see details in Theorem 3.2. This orthogonal decomposition yields that any increasing function of the energy landscape $\phi(\psi^{ss})$ serves as the Lyapunov function of RRE; see (3.17). GENERIC formalism and anti-symmetric structures for RRE with additional mass conservation law are also discussed in Sect. 3.2.

Third, in Sect. 3.3, we use the above conservative–dissipative decomposition for RRE to derive the thermodynamic relations for general chemical reactions, i.e., we express the total entropy production rate as the adiabatic and nonadiabated entropy production rate

$$
\begin{aligned}
T\dot{S}_{\text{tot}} &= T\dot{S}_{na} + T\dot{S}_a \geq 0, \\
T\dot{S}_{na} &= k_{\text{B}}T \langle K(\vec{x})\nabla\psi^{ss}(\vec{x}), \ \nabla\psi^{ss}(\vec{x}) \rangle \geq 0, \\
T\dot{S}_a &= k_{\text{B}}T \sum_j \Big(\text{KL}(\Phi_j^+(\vec{x}(t)) || \Phi_j^-(\vec{x}(t))e^{-\vec{\nu}_j \cdot \nabla\psi^{ss}}) \\
&\quad + \text{KL}(\Phi_j^-(\vec{x}(t)) || \Phi_j^+(\vec{x}(t))e^{\vec{\nu}_j \cdot \nabla\psi^{ss}}) \Big) \geq 0,
\end{aligned}
\tag{1.24}
$$

where $\text{KL}(\vec{x}||\vec{x}^*) := \sum_i \left(x_i \ln \frac{x_i}{x_i^*} - x_i + x_i^* \right)$ is the relative entropy; see Proposition 3.7. Particularly, as $t \to +\infty$ and $\vec{x}(t)$ goes to a non-equilibrium steady state (NESS) $\vec{x}^s$, remaining a strictly positive entropy production rate is an important feature of a non-equilibrium chemical reaction [39]

$$
T\dot{S}_a \to k_{\text{B}}T \sum_j \left(\Phi_j^+(\vec{x}^s) - \Phi_j^-(\vec{x}^s) \right) \log \frac{\Phi_j^+(\vec{x}^s)}{\Phi_j^-(\vec{x}^s)} > 0.
\tag{1.25}
$$

Fourth, for general non-equilibrium RRE and the corresponding CME (1.5), we also derive a ϕ-divergence energy dissipation law based on the Q_V-matrix structure and a Bregman's divergence in Proposition 3.8. Particularly, if there exists a positive invariant measure π_v for mesoscopic CME, taking $\phi(p) = p \log \frac{p}{\pi_V}$, in the large number limit, the corresponding mesoscopic energy dissipation relation converges to the macroscopic energy dissipation relation (3.20) in terms of the energy landscape $\psi^{ss}(\vec{x})$. We emphasis that the mesoscopic energy functional $F(p) = \sum_{\vec{x}_v} \phi\left(\frac{p(\vec{x}_v)}{\pi(\vec{x}_v)} \right) \pi(\vec{x}_v)$ is always convex w.r.t. p. However, since $\phi(u)$ is convex, the nonlinear weight $\pi_v(\vec{x}_v)$ in $F(p)$ drastically pick up the complicated non-convex energy landscape for chemical reactions from $\pi_v(\vec{x}_v) \approx e^{-V\psi^{ss}(\vec{x})}$ in the large number limit. After the concentration of the measure in the large number limit, a non-convex energy landscape ψ^{ss} emerges. Notice there is no such a transition from convex functional to a nonconvex function under the RRE detailed balance assumption because the corresponding probability flux is only monomial. We point out the above emerged polynomial grouped probability flux (4.1) and the non-convex energy landscape are only linked to non-equilibrium in the specific context of chemical reactions. For general equilibrium models in statistical physics, non-convex energy landscape is common, for instance the Lagenvin dynamics with non-convex potential and Ising model of ferromagnetism.

In terms of the mesoscopic chemical reaction jumping process, the proper mathematical definition for detailed balance is there exists a positive invariant measure π_v satisfying the Markov chain detailed balance (4.3), which naturally includes the above grouped probability fluxed (4.1). In the large number limit $V \to +\infty$, this Markov chain detailed balance gives raise to a symmetry for the macroscopic Hamiltonian

$$
H(\vec{p}, \vec{x}) = H(\nabla\psi^{ss}(\vec{x}) - \vec{p}, \vec{x}), \quad \forall \vec{x}, \vec{p};
\tag{1.26}
$$

see Proposition 4.1. Taking $\vec{p} = \vec{0}$, we know $\psi^{ss}(\vec{x})$ is the stationary solution to HJE (1.17). Applying (1.13) and (1.17), we formally have $\psi^{ss}(\vec{x}) = -\lim_{V \to +\infty} \frac{\log \pi(\vec{x}_v)}{V}$. Rigorously, an upper semicontinuous viscosity solution ψ^{ss} to the stationary HJE was constructed from $\pi(\vec{x}_v)$ in [24] in the Barron-Jensen's sense [9].

The first consequence of this symmetric Hamiltonian (1.26) is the RRE becomes an Onsager's type strong gradient flow in terms of $\psi^{ss}(\vec{x})$. That is to say the conservative part $W(\vec{x})$ vanishes in the RRE decomposition (3.15); see Proposition 4.4.

The second consequence of this symmetric Hamiltonian is the time reversal symmetry in terms of the Lagrangian upto a null Lagrangian

$$L(\vec{s}, \vec{x}) - L(-\vec{s}, \vec{x}) = \vec{s} \cdot \nabla \psi^{ss}(\vec{x}), \quad \forall \vec{x}, \vec{s}. \tag{1.27}$$

This symmetric relation was first dated back to Morpurgo et al. in [53] for Hamiltonian dynamics in classical mechanics with a quadratic Hamiltonian. The quadratic form Hamiltonian $H(\vec{p}, \vec{x}) = \vec{p} \cdot (\vec{p} - \nabla U)$ from the WKB expansion of the Langevin dynamics is also symmetric w.r.t $\vec{p} = \frac{1}{2} \nabla U$, so the classical Freidlin–Wentzell theory [21] shows the most probable path connecting two steady states $\vec{x}^A, \vec{x}^B$ of U (assumed to be double well with Morse index 1) is piesewisely given by an 'uphill' least action curve which starts from $\vec{x}^A$, passes through a saddle point $\vec{x}^C$ and then matches with a 'downhill' least action curve from $\vec{x}^C$ to $\vec{x}^B$. The 'uphill' least action curve with nonzero action is exactly the time reversal of the zero-cost least action curve from $\vec{x}^C$ to $\vec{x}^A$. This symmetric relation (1.27) was systematically studied in [50], which established the relation between generalized gradient flow and the large deviation principle. The symmetry in the Hamiltonian (1.26) was also used in [6] for the macroscopic fluctuation theory; see recent developments in [37] for the fluctuation symmetry and the associated GENERIC formalism. In Proposition 4.6, with the symmetric Hamiltonian condition (1.26), (i) the 'uphill' least action curve connecting a stable steady state and a saddle point is still the time reversal $\vec{x}^R(t)$ of the 'downhill' curve $\vec{x}(t)$ while the corresponding momentum is reversed with an additional control force $\nabla \psi^{ss}$; (ii) the difference of Lagrangians between the forward and reversed curve is a null Lagrangian (1.27), so the difference of the action cost between the path and the time revered path depends only on the starting/end positions; (iii) the steady solution $\psi^{ss}(\vec{x})$ to HJE (1.17) defines the energy landscape for the chemical reaction and the path affinity is given by the difference between the values of ψ^{ss} at the starting/end positions

$$Act(\vec{x}^R(\cdot)) - Act(\vec{x}(\cdot)) = \psi^{ss}(\vec{x}_T^R) - \psi^{ss}(\vec{x}_0^R). \tag{1.28}$$

The globally defined energy landscape ψ^{ss} coincides with the quasipotential [21] upto a constant if the least action curve stays within a stable basin of attraction of a steady state.

The third consequence is we can use the symmetric Hamiltonian to study a class of non-equilibrium enzyme reactions. Notice the symmetric Hamiltonian condition brought by the mathematical definition of Markov chain detailed balance (4.3) is more general than the constrained chemical version detailed balance condition (1.12). Although the Markov chain detailed balance is a basic mathematical concept, it includes the grouped probability flux representing a nonzero steady flux in each reaction channel. The resulting symmetric Hamiltonian can be used to describe a class of non-equilibrium reactions including enzyme catalyzed reactions. In Sect. 4.4, a simplified Schlögl catalysis reaction was studied in detail, where three features for non-equilibrium chemical reactions: multiple steady states, nonzero steady state fluxes and positive entropy production rates at non-equilibrium steady states (NESS) are shown.

The stationary solution $\psi^{ss}(\vec{x})$ to HJE serves as the energy landscape of chemical reactions, facilitates the conservative–dissipative decomposition for RRE, and also determines both the energy barrier and thermodynamics of chemical reactions. For a detailed/complex balanced RRE, we simply have $\psi^{ss}(\vec{x}) = \mathrm{KL}(\vec{x}||\vec{x}^s)$; see Lemma 3.6. For general chemical reactions, we first discuss viscosity solutions to stationary HJE by the dynamic programming method [72, Theorem 2.41]. This is equivalent to an optimal control interpretation in an undefined time horizon; see Sect. 5.1. By Maupertuis's principle for an undefined time horizon (see (4.38)), we regard $\vec{p}$ as a control variable, then in terms of the Hamiltonian, the most probable path is solved by a constrained optimal control problem (see (5.1))

$$v(\vec{y};\ \vec{x}^{\mathrm{A}}, c) = \inf_{T, \vec{p}} \int_0^T \left(\vec{p} \cdot \nabla_p H(\vec{p}, \vec{x}) - H(\vec{p}, \vec{x}) + c \right)\, \mathrm{d}t,$$
$$\text{s.t. } \dot{\vec{x}} = \nabla_p H(\vec{p}, \vec{x}),\ t \in (0, T), \quad \vec{x}_0 = \vec{x}^{\mathrm{A}},\ \vec{x}_T = \vec{y}. \tag{1.29}$$

Here and afterwards, we use notation $\nabla_p H$ as the vector $\{\partial_{p_i} H\}_{i=1:N}$. The critical energy level c is zero for the Hamiltonian in chemical reactions. Let $\vec{x}^{\mathrm{A}}$ be a steady state of RRE, then the corresponding critical mañé potential $v(\vec{y};\ \vec{x}^{\mathrm{A}}, 0)$ gives a viscosity solution to the steady HJE [34]. However, to obtain a unique viscosity solution given by the energy landscape ψ^{ss} in a chemical reaction, some notation of selection principle in the weak KAM solutions needs to be imposed [25].

In general, a standard diffusion approximation can be obtained via the Kramers–Moyal expansion for the CME, which is equivalent to the quadratic approximations near $\vec{p} = 0$ for the Hamiltonian; see Sect. 5.2. However, this diffusion approximation only valid for a transition near the 'downhill' solution to the RRE. The 'uphill' transition path starting from a stable steady state ending at a saddle point is a rare transition in the large deviation regime and the energy barrier can not be computed by the above diffusion approximation. We refer to [14] for quantified analysis of the failure of the diffusion approximation via the Kramers–Moyal expansion (a.k.a 'system size expansion' by van Kampen [74]). Based on the strong gradient formulation (4.17), another drift-diffusion approximation (5.13) is proposed as a good quadratic approximation near not only the 'downhill' solution to the macroscopic RRE but also the 'uphill' least action curve. This diffusion approximation shares the same energy landscape and same symmetric Hamiltonian structure w.r.t. $\nabla \psi^{ss}$ and satisfies a fluctuation–dissipation relation with an invariant measure $\pi = e^{-V\psi^{ss}}$.

State of the Art

The WKB expansion is a classical and powerful tool to understand the exponential asymptotics that quantifies the fluctuations in many physical problems; see Kubo et al. [38] for a WKB expansion of a general stochastic process and see Doi–Peliti formalism [13, 55]. WKB analysis also initials physical studies of the large deviation(fluctuation) behaviors for stochastic models of chemical reactions from a Hamiltonian viewpoint, c.f. [4, 12, 33, 46, 71]. Particularly, in the large number limit of chemical reactions modeled by the CME, Dykman et al. [12] first derived HJE (1.17) with the associated Hamiltonian H and studied the symmetry of the Hamiltonian in a detailed balanced chemical reaction system. We also refer to a recent review article [4] using WKB approximations to study various large deviation behaviors such as population extinction/fixation, genetic switches and biological invasions.

The concept and the exponential asymptotics for the reaction rate in terms of the activation energy (energy barrier) for transitions between two states in a chemical reaction was pioneered by Arrhenius 1889 while the celebrated work by Kramers explicitly estimated it for a Langevin

dynamics. At the mathematical analysis level, the large deviation principle with the associated Lagrangian/Hamiltonian for general stochastic processes and the transition path (the most probable path) connecting two stable states were pioneered by Freidlin and Wentzell in late 1960s, c.f. [21]. The central idea of the Freidlin–Wentzell theory is that the steady solution $\psi^{ss}(\vec{s})$ to the HJE defines a quasipotential which quantifies the maximum probability or the energy barrier for a transition, i.e., an exit problem in the basin of attraction. We also refer to [18–20, 40, 41] for the optimal control and nonlinear semigroup viewpoint, which connect least action problems with HJEs. For chemical reactions with RRE detailed balance (1.12), the quasipotential is given by $\psi^{ss}(\vec{x}) = \text{KL}(\vec{x}||\vec{x}^s)$ [12], while for general large number process including non-equilibrium dynamics, [2] proved the large deviation principle for C^v with the associated good rate function. In [24], the large deviation principle at single times was proved via the convergence from the Varadhan's nonlinear semigroup to the Lax–Oleinik representation of the viscosity solution to HJE. Moreover, an upper semicontinuous viscosity solution in the Barron-Jensen's sense [9] to the stationary HJE was also obtained in [24] by using a positive detailed balanced invariant measure to C^v.

With the RRE detailed/complex balance condition, characterization and uniqueness of all steady states for the macroscopic RRE was proved in [16, 31]. On the contrary, thermodynamic relations, dissipation structures and computations for transition paths in non-equilibrium stochastic dynamics are challenging problems due to coexistent steady states and nonzero steady fluxes sustained by environment, whose studies were pioneered by Prigogine [58] from the Brussels School. We refer to [27, 39, 61, 62, 64, 67] and the references therein for thermodynamics relations, particularly the adiabatic/nonadiabatic decomposition for the nonzero entropy production rate in non-equilibrium biochemical reactions. In [45], Lazarescu et al. used a biased Hamiltonian H for chemical reaction based on time-averaged observations to study the first order phase transitions, particularly for metastable models in an open system with non-equilibrium steady states. However, it is not clear whether the biased Hamiltonian provides the most probable path (the least action path for the original Hamiltonian). Using a linear response relation with a susceptibility $\chi(\rho)$ between the current and the external field generating the fluctuation, a comprehensive review by Bertin et al. [7] discussed the macroscopic fluctuation theory including the time reversal, symmetry of Hamiltonians, fluctuation theorems at a macroscopic scale for various physical models. The macroscopic fluctuation theory was first developed by Bertini et al., [6]; see further mathematical analysis and variational structure including the density-flux pair large deviation principle in [57, 59, 65]. Without the quadratic approximation of the Hamiltonian, the calculations of transition paths, and the symmetry for fully nonlinear Hamiltonians in non-equilibrium reactions were not discussed in [6, 7]. Indeed, there were many studies for the failure in computing the correct energy barrier of transition paths using a simple diffusion approximation from the Kramers–Moyal expansion of the CME; c.f. Doering et al. [14] for the extinction problem in a birth-death stochastic population model.

Under the RRE detailed balance assumption, the macroscopic RRE has several gradient flow structures in terms of free energy $\text{KL}(\vec{x}||\vec{x}^s)$ [49, 54]. Particularly, a De Giorgi type generalized gradient flow structure brought by the symmetry in the Hamiltonian is closely related to the good rate function in the large deviation principle; see systematical studies in [50]. Recently, [48, 49] recovered the macroscopic RRE for chemical reactions via the evolutionary Γ-convergence techniques in [47, 70] in the gradient flow regime. The symmetric Hamiltonian was also used in Kraaij et al. [37] to study the fluctuation symmetry. This symmetry criteria reduces a pre-GENERIC system to a GENERIC formalism [37]. In general, the energetic decomposition for a dynamics is not unique and has different gradient flow structures with associated fluctuation estimates; c.f., [60].

The remaining part of this paper is organized as follows. In Sect. 2, we provide preliminaries for the RRE, WKB expansion and properties for the Hamiltonian and the Lagarangian. In Sect. 3, we study dynamic solutions, steady solution $\psi^{ss}(\vec{x})$ to the HJE. Using the stationary solution, we propose a conservative–dissipative decomposition for general non-equilibrium RRE (see Sect. 3.2) and also give a decomposition for its thermodynamics (see Sect. 3.3). The associated energy dissipation laws at both mesoscopic and macroscopic level with the passage from one to another is given in Sect. 3.4. In Sect. 4, we use a symmetric Hamiltonian to study a class of non-equilibrium enzyme reactions, which yields (i) an Onsager-type strong form of gradient flow and (ii) a modified time reversed curve serves as the transition paths between coexistent steady states. Bistable Schlögl example is discussed in Sect. 4.4. In Sect. 5, we clarify the existence of the stationary solution to HJE via an optimal control representation in an undefined time horizon and give a diffusion approximation for transition path computations that satisfies the fluctuation–dissipation relation and the same symmetric Hamiltonian. Pedagogical derivations for CME, the generator and the mean field limit RRE after including 'no reaction' boundary condition are given in Appendix.

2 Preliminaries: Macroscopic RRE, WKB Expansion and Large Deviation

As a preparation for our main results, in this section, we review some terminologies for the large number limiting RRE (1.6) and collect some preliminary lemmas for existence, uniqueness, characterization of steady states in a detailed/complex balanced RRE system. The associated Hamiltonian $H(\vec{p}, \vec{x})$ and HJE from the WKB expansion are also revisited. Moreover, the convex conjugate $L(\vec{s}, \vec{x})$ of $H(\vec{p}, \vec{x})$ gives the rate function in the large deviation principle for the large number process, which allows us to study the transition path for a non-equilibrium system in later sections. Most of the results in this section was known while we provide brief proofs for completeness.

2.1 Terminologies for the Macroscopic RRE and RRE Detailed/Complex Balance Conditions

Recall the forward and backward fluxes $\Phi_j^{\pm}$ satisfying LMA (1.7) and RRE (1.6). Using the $N \times M$ stoichiometric matrix ν^T and the reaction rate vector

$$\vec{r}(\vec{x}) = \left(r_j(\vec{x})\right)_{j=1:M} = (\Phi_j^+(\vec{x}) - \Phi_j^-(\vec{x}))_{j=1:M}, \tag{2.1}$$

we represent RRE (1.6) in a matrix form

$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{x} = \nu^T \vec{r} =: \vec{R}(\vec{x}), \tag{2.2}$$

where $\vec{R}(\vec{x})$ is called the production rate. Denote the range of matrix ν^T as $\mathrm{Ran}(\nu^T)$, i.e., the span of the column vectors $\{\vec{\nu}_j\}$ of ν^T. Then we know the production rate

$$\dot{\vec{x}}(t) \in \mathrm{Ran}(\nu^T) \subset \mathbb{R}^N. \tag{2.3}$$

Motivated by this, we define the subspace $G = \{\vec{x} \in \mathbb{R}^N; \ \vec{x} \in \mathrm{Ran}(\nu^T)\}$ which is known as the stoichiometric space. Recall (1.9), i.e., $\nu \vec{m} = \vec{0}$, which implies the conservation of total mass, i.e., $\frac{\mathrm{d}}{\mathrm{d}t}(\vec{m} \cdot \vec{x}) = \vec{m} \cdot \nu^T \vec{r} = 0$. Therefore the Wegscheider matrix ν always has a nonzero kernel, i.e., $\dim(\mathrm{Ker}(\nu)) \geq 1$.

We have the following lemma on existence and uniqueness of dynamic solution to (1.6).

Lemma 2.1 *Assume $v \in \mathbb{R}^{M \times N}$ is the Wegscheider matrix satisfying* (1.9). *Consider RRE* (1.6) *with flux $\Phi_j^{\pm}$ satisfying* (1.7). *We have*

(i) *The region $R_+^N := \{\vec{x} \in \mathbb{R}^N ; x_i \geq 0\}$ is an invariant region;*
(ii) *For any initial data $\vec{x}_0 \geq 0$, there exists a unique global-in-time bounded solution to* (1.6) *satisfying*

$$\frac{\mathrm{d}}{\mathrm{d}t} \vec{x}(t) \cdot \vec{m} = 0. \tag{2.4}$$

The statement (i) can be directly verified by proving $\frac{\mathrm{d}}{\mathrm{d}t} x_i \geq 0$ at any $x_i = 0$ using case by case arguments. The statement (ii) is a consequence of (1.9) and the standard ODE theory.

2.1.1 Detailed Balance and Complex Balance for the Macroscopic RRE

Recall the macroscopic RRE (1.6) and the RRE detailed balance condition (1.12) is equivalent to

$$\log k_j^+ - \log k_j^- = \vec{v}_j \cdot \log \vec{x}^{\mathrm{s}} \tag{2.5}$$

due to LMA (1.7).

Denote the complex space as the collection of distinct reaction vectors $\mathcal{C} := \left(\vec{v}_j^{\pm} \right)_{j=1:M}$. Then the complex balance condition means for each complex $\vec{\eta} \in \mathcal{C}$, all the reactant contributions in the flux equals all the product contributions in flux. Precisely, a strictly positive (componentwisely) state $\vec{x}_c^{\mathrm{s}} > 0$ is called complex balanced [31] if

$$\sum_{j, \vec{v}_j^+ = \vec{\eta}} \left(\Phi_j^-(\vec{x}_c^{\mathrm{s}}) - \Phi_j^+(\vec{x}_c^{\mathrm{s}}) \right) + \sum_{j, \vec{v}_j^- = \vec{\eta}} \left(\Phi_j^+(\vec{x}_c^{\mathrm{s}}) - \Phi_j^-(\vec{x}_c^{\mathrm{s}}) \right) = 0. \tag{2.6}$$

One can directly verify state $\vec{x}_c^{\mathrm{s}} > 0$ satisfying (2.6) is a steady state to RRE (1.6). Indeed, at $\vec{x}_c^{\mathrm{s}}$, recast the RHS of (1.6) as flux difference

$$\sum_j \vec{v}_j \left(\Phi_j^+(\vec{x}_c^{\mathrm{s}}) - \Phi_j^-(\vec{x}_c^{\mathrm{s}}) \right) = \sum_j \vec{v}_j^+ \left(\Phi_j^-(\vec{x}_c^{\mathrm{s}}) - \Phi_j^+(\vec{x}_c^{\mathrm{s}}) \right) + \sum_j \vec{v}_j^- \left(\Phi_j^+(\vec{x}_c^{\mathrm{s}}) - \Phi_j^-(\vec{x}_c^{\mathrm{s}}) \right). \tag{2.7}$$

The first term in the summation represents that for the reactant (aka substrate) complex $\vec{v}_j^+$ in the jth-forward reaction, the net flux is $\Phi_j^-(\vec{x}_c^{\mathrm{s}}) - \Phi_j^+(\vec{x}_c^{\mathrm{s}})$. So we can re-classify this summation w.r.t distinct reactant complex $\vec{v}_j^+ = \vec{\eta}, \vec{\eta} \in \mathcal{C}$

$$\sum_j \vec{v}_j^+ \left(\Phi_j^-(\vec{x}_c^{\mathrm{s}}) - \Phi_j^+(\vec{x}_c^{\mathrm{s}}) \right) = \sum_{\vec{\eta} \in \mathcal{C}} \vec{\eta} \sum_{j: \vec{v}_j^+ = \vec{\eta}} \left(\Phi_j^-(\vec{x}_c^{\mathrm{s}}) - \Phi_j^+(\vec{x}_c^{\mathrm{s}}) \right). \tag{2.8}$$

Similarly, the second term in the summation represents that the reactant complex $\vec{v}_j^-$ in the jth-backward reaction, the net flux is $\Phi_j^+(\vec{x}_c^{\mathrm{s}}) - \Phi_j^-(\vec{x}_c^{\mathrm{s}})$. Therefore we can choose to re-classify this summation w.r.t distinct reactant complex $\vec{v}_j^- = \vec{\eta}, \vec{\eta} \in \mathcal{C}$

$$\sum_j \vec{v}_j^- \left(\Phi_j^+(\vec{x}_c^{\mathrm{s}}) - \Phi_j^-(\vec{x}_c^{\mathrm{s}}) \right) = \sum_{\vec{\eta} \in \mathcal{C}} \vec{\eta} \sum_{j: \vec{v}_j^- = \vec{\eta}} \left(\Phi_j^+(\vec{x}_c^{\mathrm{s}}) - \Phi_j^-(\vec{x}_c^{\mathrm{s}}) \right). \tag{2.9}$$

Combining the above two ways of rearrangements for the summation in chemical channel j, (2.7) becomes

$$\sum_j \vec{v}_j \left(\Phi_j^+(\vec{x}_c^s) - \Phi_j^-(\vec{x}_c^s) \right)$$

$$= \sum_{\vec{\eta} \in \mathcal{C}} \vec{\eta} \left(\sum_{j:\vec{v}_j^+ = \vec{\eta}} \left(\Phi_j^-(\vec{x}_c^s) - \Phi_j^+(\vec{x}_c^s) \right) + \sum_{j:\vec{v}_j^- = \vec{\eta}} \left(\Phi_j^+(\vec{x}_c^s) - \Phi_j^-(\vec{x}_c^s) \right) \right) = \vec{0},$$

(2.10)

where we used the complex balance condition (2.6).

2.2 Characterization of RRE Steady State for the Detailed/Complex Balance Case

Now we investigate all the steady states of RRE (1.6), i.e.,

$$S_e := \{ \vec{x}^e \in \mathbb{R}_+^N; \ \vec{R}(\vec{x}^e) = \sum_j \vec{v}_j \left(\Phi_j^+(\vec{x}^e) - \Phi_j^-(\vec{x}^e) \right) = \vec{0} \}.$$

(2.11)

First, we show uniqueness of positive steady states for the detailed balanced RRE. From (1.7), we have

$$\vec{v}_j \cdot \log \frac{\vec{x}}{\vec{x}^s} = \log \left(\prod_{i=1}^N \left(\frac{x_i}{x_i^s} \right)^{v_{ji}} \right) = \log \left(\frac{\Phi_j^-(\vec{x}) \, \Phi_j^+(\vec{x}^s)}{\Phi_j^+(\vec{x}) \, \Phi_j^-(\vec{x}^s)} \right).$$

(2.12)

If $\vec{x}^e > 0$, then from (2.12), we have

$$0 = \log \frac{\vec{x}^e}{\vec{x}^s} \cdot \vec{R}(\vec{x}^e) = \sum_j \left(\Phi_j^+(\vec{x}^e) - \Phi_j^-(\vec{x}^e) \right) \log \left(\frac{\Phi_j^-(\vec{x}^e) \, \Phi_j^+(\vec{x}^s)}{\Phi_j^+(\vec{x}^e) \, \Phi_j^-(\vec{x}^s)} \right).$$

(2.13)

Under RRE detailed balance condition (1.12), the above equation implies $\Phi_j^+(\vec{x}^e) = \Phi_j^-(\vec{x}^e)$ and thus $\vec{x}^e$ also satisfies RRE detailed balance. Notice (1.7) and (1.12) implies identity

$$\vec{v}_j \cdot \log \frac{\vec{x}}{\vec{x}^s} = \log \left(\frac{\Phi_j^-(\vec{x})}{\Phi_j^+(\vec{x})} \right).$$

(2.14)

We know $\log \frac{\vec{x}^e}{\vec{x}^s} \in \text{Ker}(v)$. Given the stoichiometric space G and $\vec{q} \in \text{Ker}(v)$, $\vec{q} + G$ is called one stoichiometric compatibility class. Then it is easy to verify that if $\vec{x}^e$ and $\vec{x}^s$ are in the same stoichiometric compatibility class, then $\vec{x}^e = \vec{x}^s$. Indeed, from $\log \frac{\vec{x}^e}{\vec{x}^s} \in \text{Ker}(v)$ and $\vec{x}^e - \vec{x}^s \in G$, we know

$$\log \frac{\vec{x}^e}{\vec{x}^s} \cdot \left(\vec{x}^e - \vec{x}^s \right) = 0,$$

(2.15)

which implies $\vec{x}^e = \vec{x}^s$.

This uniqueness of steady states in one stoichiometric compatibility class still holds for the complex balanced system, with a slight modification of the above proof. We conclude the following well-known result on the uniqueness of steady state; c.f., [31, Theorem 6A], [3, Theorem 3.5].

Lemma 2.2 *Assume there exists a strictly positive steady state $\vec{x}_1^s$ satisfying complex balance* (2.6). *Then for any $\vec{q} \in \mathrm{Ker}(\nu)$, there exists a unique steady states $\vec{x}_*^s$ in the space $\{\vec{x} \in \vec{q} + G; \ \vec{x} > 0\}$. Moreover, $\vec{x}_*^s$ satisfies complex balance condition* (2.6), *and is characterized by*

$$(\vec{x}_*^s)_i = (\vec{x}_1^s)_i e^{q_i} > 0. \tag{2.16}$$

As a consequence, if $\vec{x}_1^s$ satisfies RRE detailed balance (1.12), thus it also satisfies (2.6). So (2.16) still holds and this unique steady state $\vec{x}_*^s$ in the space $\{\vec{x} \in \vec{q} + G; \ \vec{x} > 0\}$ is RRE detailed balanced. For both the detailed/complex balanced RRE system, $\vec{x}^s$ constructs a Lyapunov function for (1.6), known as the relative entropy $\mathrm{KL}(\vec{x}||\vec{x}^s) = \sum_i \left(x_i \ln \frac{x_i}{x_i^s} - x_i + x_i^s \right)$. Since $\mathrm{KL}(\vec{x}||\vec{x}^s)$ is strictly convex for $\{\vec{x} \in \vec{q} + G; \ \vec{x} > 0\}$, so one also have local stability of the RRE detailed/complex balanced steady state $\vec{x}_*^s$. To obtain global stability of $\vec{x}_*^s$, a necessary condition (see [69]) is that there shall be no equilibrium on the boundary of $\mathbb{R}_+^N$ for the positive stoichiometric compatibility class, i.e., $\{\vec{x} \in \vec{q} + G; \ \vec{x} > 0\}$. We refer to [5] for more detailed conditions to obtain global stability.

2.2.1 Deficiency Zero Theorem

The complex balance condition is an important property for balance between the product complex and the reactant complex. It also motivates a more important index theorem based only on the graph structure of the reaction networks. Recall the complex space $\mathcal{C} = \{\vec{v}_j^{\pm}\}_{j=1:M}$ and species $X = \{X_i\}_{i=1:N}$. A reaction network, denoted as $(X, \mathcal{C}, \mathcal{R})$, is a directed graph with nodes given by the complexes $\mathcal{C}$ and directed edges given by reactions $\mathcal{R} = \{\vec{v}_j^+ \to \vec{v}_j^-\}$. Each connected subgraph (regarded as undirected subgraph) is called a linkage class and denote the total number of the linkage classes of the reaction graph as ℓ. Denote the total number of distinct complex as n_c and denote the rank of ν as s. Then the deficiency of the reaction network is $\delta := n_c - \ell - s \geq 0$. In [16], Feinberg proved that a deficiency zero network, i.e., $\delta = 0$ is equivalent condition for that the equilibrium for (1.6) is complex balanced. Therefore, the equilibrium of the RRE (1.6) can be characterized using the deficiency zero theorem, which relies only on the network structure of $(X, \mathcal{C}, \mathcal{R})$. More precisely, we call the reaction network is weakly reversible if for any path connecting from complex $\mathcal{C}_i$ to complex $\mathcal{C}_j$, one can always find a path connecting from complex $\mathcal{C}_j$ to complex $\mathcal{C}_i$. Then the deficiency zero theorem proved by Horn and Feinberg [16, 31] states that if a chemical reaction network with LMA satisfies (i) $\delta = 0$ and (ii) weakly reversibility, then there is a unique positive steady state in each stoichiometric compatibility class; see also [17, Theorem 7.1.1].

2.3 WKB Expansion and Hamilton–Jacobi Equation

In this section, we use the WKB analysis of the CME for $p(\vec{x}_v, t)$ to study the exponential asymptotic behavior. We will investigate some good properties of the resulting HJE and the associated Hamiltonian $H(\vec{p}, \vec{x})$ defined in (1.16). Recall the large number process $C^v(t)$ in (1.4), which is also denoted as C_t for simplicity. For fixed V, recall Q_v^* defined in (1.5). Then for any continuous test function $f(\vec{x}_v)$, we have

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_{\vec{x}_v} f(\vec{x}_v) p(\vec{x}_v, t) = \sum_{\vec{x}_v} (Q_v f)(\vec{x}_v) p(\vec{x}_v, t). \tag{2.17}$$

Here Q_v is the duality of Q_v^*, see explicit definition in (A.12) after including 'no reaction' boundary condition.

Denote

$$w(\vec{x}_v, t) = \mathbb{E}^{\vec{x}_v}\left(f(C_t)\right), \tag{2.18}$$

then $w(\vec{x}_v, t)$ satisfies the backward equation

$$\partial_t w = Q_v w, \quad w(\vec{x}_v, 0) = f(\vec{x}_v). \tag{2.19}$$

We refer to [24] for the well-posedness of the backward equation after including 'no reaction' boundary condition.

Assume there exists a smooth enough function $u(\vec{x}, t)$ such that at $\vec{x} = \vec{x}_v$, we have WKB reformulation

$$w(\vec{x}_v, t) = e^{Vu(\vec{x}_v, t)}. \tag{2.20}$$

We obtain

$$\partial_t u(\vec{x}_v, t) = \frac{1}{V} e^{-Vu(\vec{x}_v, t)} Q_v e^{Vu(\vec{x}_v, t)} =: \frac{1}{V} H_V(Vu), \quad u(\vec{x}_v, 0) = \frac{1}{V} \log f(\vec{x}_v). \tag{2.21}$$

In summary,

$$u(\vec{x}_v, t) = \frac{1}{V} \log w(\vec{x}_v, t) = \frac{1}{V} \log \mathbb{E}^{\vec{x}_v}\left(f(C_t)\right) = \frac{1}{V} \log \mathbb{E}^{\vec{x}_v}\left(e^{Vu_0(C_t)}\right) =: (S_t u_0)(\vec{x}_v) \tag{2.22}$$

is the so-called Varadhan's nonlinear semigroup [18, 73] for process C_t.

For any $\vec{x} \in \mathbb{R}_+^N$, let $\vec{x}_V = \frac{\vec{n}}{V} \to \vec{x}$ as $V \to +\infty$. Then the after WKB reformulation at $\vec{x}_v \geq 0$ gives

$$Q_v e^{Vu(\vec{x}_v, t)} = V \sum_{j=1, \vec{x}_v + \frac{\vec{v}_j}{V} \geq 0}^{M} \Phi_j^+(\vec{x}_v) \left(e^{Vu(\vec{x}_v + \frac{\vec{v}_j}{V})} - e^{Vu(\vec{x}_v)}\right)$$

$$+ V \sum_{j=1, \vec{x}_v - \frac{\vec{v}_j}{V} \geq 0}^{M} \Phi_j^-(\vec{x}_v) \left(e^{Vu(\vec{x}_v - \frac{\vec{v}_j}{V})} - e^{Vu(\vec{x}_v)}\right)$$

For $\vec{x}_v \notin \mathbb{R}_+^N$, one can define a zero extension for $\tilde{\Phi}_j^\pm(\vec{x}_v)$; see [24]. Using Taylor's expansion w.r.t $\frac{\vec{v}_j}{V}$, we obtain HJE for u

$$\partial_t u(\vec{x}, t) = \sum_{j=1}^{M}\left(\Phi_j^+(\vec{x})\left(e^{\vec{v}_j \cdot \nabla u(\vec{x}, t)} - 1\right) + \Phi_j^-(\vec{x})\left(e^{-\vec{v}_j \cdot \nabla u(\vec{x}, t)} - 1\right)\right). \tag{2.23}$$

Similarly, starting from the froward equation (1.5), one can obtain the HJE (1.17) for $\psi(\vec{x}, t)$.

2.3.1 Properties of Hamiltonian H

Recall the matrix form of the macroscopic RRE

$$\frac{d}{dt}\vec{x} = v^T \vec{r} =: \vec{R}(\vec{x}), \tag{2.24}$$

where $\nu \in \mathbb{R}^{M \times N}$ is a constant matrix. Recall the mass conservation law of chemical reactions (1.9) and direct decomposition (1.10), which always satisfies

$$\dim \left(\text{Ran}(\nu^T) \right) < N. \tag{2.25}$$

It motivates that for the WKB expansion and the corresponding relations with the rate function L in the large deviation principle, we will see L make sense in a 'more accurate' subspace G.

Lemma 2.3 *Hamiltonian $H(\vec{p}, \vec{x})$ in (1.16) is degenerate in the sense that*

$$H(\vec{p}, \vec{x}) = H(\vec{p}_1, \vec{x}), \tag{2.26}$$

where $\vec{p}_1 \in \text{Ran}(\nu^T)$ is the direct decomposition of $\vec{p}$ such that

$$\vec{p} = \vec{p}_1 + \vec{p}_2, \quad \vec{p}_1 \in \text{Ran}(\nu^T), \ \vec{p}_2 \in \text{Ker}(\nu). \tag{2.27}$$

Proof From the direct decomposition (1.10), we have (2.27). Thus $0 = \vec{\nu}_j \cdot \vec{p}_2$, which implies (2.26). $\qquad \square$

Lemma 2.4 *$H(\vec{p}, \vec{x})$ defined in (1.16) is strictly convex for $\vec{p} \in G$.*

Proof We compute the Hessian of H in $G \times \mathbb{R}^N$. For any $\vec{\alpha} \in G$,

$$\frac{\mathrm{d}^2}{\mathrm{d}\varepsilon^2}\Big|_{\varepsilon=0} H(\vec{p} + \varepsilon\vec{\alpha}, \vec{x}) = \sum_{j=1}^{M} \left(\vec{\nu}_j \cdot \vec{\alpha} \right)^2 \left(\Phi_j^+(\vec{x}) e^{\vec{\nu}_j \cdot \vec{p}} + \Phi_j^-(\vec{x}) e^{-\vec{\nu}_j \cdot \vec{p}} \right) \geq 0$$

and the equality holds if and only if $\nu\vec{\alpha} = \vec{0}$. Since $\vec{\alpha} \in G = \text{Ran}(\nu^T)$, there exists a vector $\vec{\beta} \in \mathbb{R}^M$ such that $\vec{\alpha} = \nu^T \vec{\beta}$. Thus the equality above holds if and only if

$$0 = \vec{\beta}^T \nu\vec{\alpha} = \vec{\beta}^T \nu\nu^T \vec{\beta}, \tag{2.28}$$

which implies $\vec{\alpha} = \vec{0} \in G$. $\qquad \square$

2.4 The Convex Conjugate $L(\vec{s}, \vec{x})$ Gives the Rate Function in Large Deviation Principle

Let us first introduce the convex conjugate function L and the associated action functional. Since H defined in (1.16) is convex w.r.t $\vec{p}$, we compute the convex conjugate of H via the Legendre transform. For any $\vec{s} \in \mathbb{R}^N$, define

$$L(\vec{s}, \vec{x}) := \sup_{\vec{p} \in \mathbb{R}^N} \left(\langle \vec{p}, \vec{s} \rangle - H(\vec{p}, \vec{x}) \right) = \langle \vec{p}^*, \vec{s} \rangle - H(\vec{p}^*, \vec{x}) \tag{2.29}$$

where $\vec{p}^*(\vec{s}, \vec{x})$ solves

$$\vec{s} = \nabla_p H(\vec{p}^*, \vec{x}) = \sum_j \vec{\nu}_j \left(\Phi_j^+ e^{\vec{\nu}_j \cdot \vec{p}^*} - \Phi_j^- e^{-\vec{\nu}_j \cdot \vec{p}^*} \right). \tag{2.30}$$

Recall here notation $\nabla_p H$ is a vector $\left(\partial_{p_i} H \right)_{i=1:N}$. Thus

$$L(\vec{s}, \vec{x}) = \vec{s} \cdot \vec{p}^*(\vec{s}, \vec{x}) - H(\vec{p}^*(\vec{s}, \vec{x}), \vec{x}). \tag{2.31}$$

Define the action functional as

$$Act(\vec{x}(\cdot)) = \int_0^T L(\dot{\vec{x}}(t), \vec{x}(t)) \, dt. \tag{2.32}$$

Then we have the following lemma.

Lemma 2.5 *For L function defined in* (2.29), *we know*

(i) $L(\vec{s}, \vec{x}) \geq 0$ *and*

$$L(\vec{s}, \vec{x}) = \begin{cases} \max_{\vec{p} \in G} \{\vec{s} \cdot \vec{p} - H(\vec{p}, \vec{x})\}, & \vec{s} \in G \\ +\infty, & \vec{s} \notin G; \end{cases} \tag{2.33}$$

moreover, L is strictly convex in G;

(ii) *For the action functional* $Act(\vec{x}(\cdot))$ *in* (2.32) *the least action* $\vec{x}(t)$ *satisfies the Euler-Lagrange equation*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\vec{x}}}(\dot{\vec{x}}(t), \vec{x}(t))\right) = \frac{\partial L}{\partial \vec{x}}(\dot{\vec{x}}(t), \vec{x}(t)), \tag{2.34}$$

which is equivalent to the Hamiltonian dynamics with H defined in (1.16)

$$\frac{d}{dt}\vec{x} = \nabla_p H(\vec{p}, \vec{x}), \quad \frac{d}{dt}\vec{p} = -\nabla_x H(\vec{p}, \vec{x}); \tag{2.35}$$

(iii) $\vec{x}(t)$ *is the solution to RRE* (1.6) *if and only if* $Act(\vec{x}(\cdot)) = 0$.

Proof (i) First, from (3.1), $H(\vec{0}, \vec{x}) \equiv 0$ thus we know $L(\vec{s}, \vec{x}) \geq 0$.

Second, from Lemma 2.3, we know for $\vec{s} \in \mathbb{R}^N$,

$$\begin{aligned} L(\vec{s}, \vec{x}) &= \sup_{\vec{p} \in \mathbb{R}^N} (\langle (\vec{p}, \vec{s}) - H(\vec{p}, \vec{x}) \rangle \\ &= \sup_{\vec{p} \in \mathbb{R}^N} (\langle \vec{p}_1, \vec{s} \rangle + \langle \vec{p}_2, \vec{s} \rangle - H(\vec{p}_1, \vec{x})), \end{aligned} \tag{2.36}$$

where $\vec{p}_1 \in G$ and $\vec{p}_2 \in \text{Ker}(\nu)$ are direct decomposition of $\vec{p}$. Therefore, for $\vec{s} \notin G$,

$$L(\vec{s}, \vec{x}) \geq \sup_{\vec{p}_2 \in \text{Ker}(\nu), \vec{p}_1 = \vec{0}} (\langle \vec{p}_1, \vec{s} \rangle + \langle \vec{p}_2, \vec{s} \rangle - H(\vec{p}_1, \vec{x})) = \sup_{\vec{p}_2 \in \text{Ker}(\nu)} \langle \vec{p}_2, \vec{s} \rangle = +\infty. \tag{2.37}$$

On the other hand, for $\vec{s} \in G$,

$$L(\vec{s}, \vec{x}) = \sup_{\vec{p}_1 \in G} (\langle \vec{p}_1, \vec{s} \rangle - H(\vec{p}_1, \vec{x})). \tag{2.38}$$

From the definition of H, we know H has a lower bound and is exponentially coercive. Indeed,

$$\lim_{|\vec{p}| \to +\infty} H(\vec{p}, \vec{x}) \geq \lim_{|\vec{p}| \to +\infty} \sum_j \left(\min(\Phi_j^+, \Phi_j^-) e^{(\vec{v}_j \cdot \hat{\vec{p}})|\vec{p}|} - \Phi_j^+(\vec{x}) - \Phi_j^-(\vec{x}) \right) = +\infty. \tag{2.39}$$

Therefore, the sup in (2.37) can be achieved and we conclude (2.33).

Third, we show that strict convexity of H in G implies strict convexity of L in G. For any $\vec{s}_1, \vec{s}_2 \in G$, from (2.38) above, there exist $\vec{p}_1, \vec{p}_2 \in G$ such that $\vec{s}_1 = \nabla_p H(\vec{p}_1, \vec{x})$, $\vec{s}_2 = \nabla_p H(\vec{p}_2, \vec{x})$. Then we have

$$(\vec{s}_1 - \vec{s}_2) \cdot (\nabla_s L(\vec{s}_1, \vec{x}) - \nabla_s L(\vec{s}_2, \vec{x})) = (\nabla_p H(\vec{p}_1, \vec{x}) - \nabla_p H(\vec{p}_2, \vec{x})) \cdot (\vec{p}_1 - \vec{p}_2) > 0 \tag{2.40}$$

due to the strict convexity of H in G.

(ii) Let $\vec{x}(t)$ be the least action such that

$$\vec{x}(\cdot) = \arg \min_{\vec{x}(0)=\vec{x}_0, \vec{x}(T)=\vec{b}} Act(\vec{x}(\cdot)). \tag{2.41}$$

Then $\vec{x}(t)$ satisfies the Euler-Lagrange equation (2.34). From (2.30), (2.31), we know given $\vec{s}, \vec{x}$

$$\frac{\partial L}{\partial \vec{x}}(\vec{s}, \vec{x}) = -\nabla_x H(\vec{p}^*(\vec{s}, \vec{x}), \vec{x}). \tag{2.42}$$

Thus for $\vec{p} = \frac{\partial L}{\partial \vec{s}}$, the Hamiltonian dynamics (2.35) holds.

(iii) First, let $\vec{x}(t)$ is the solution to RRE (1.6) with initial data $\vec{x}_0$ and set $\vec{p}(0) = \vec{0}$. Then

$$\vec{p} \equiv \vec{0}, \quad \frac{d}{dt}\vec{x} = \nabla_p H(\vec{0}, \vec{x}). \tag{2.43}$$

This corresponds to a least action $\vec{x}(t)$ such that $Act(\vec{x}(\cdot)) = 0$.

Second, assume $\vec{x}$ is a least action such that $Act(\vec{x}(\cdot)) = 0$, then $L(\dot{\vec{x}}(t), \vec{x}(t)) \equiv 0$ for all $t \in [0, T]$ and the Hamiltonian dynamics (2.35) holds. It is sufficient to prove the following two cases.

Case(I), if there exists t^* such that $\vec{p}(t^*) = \vec{q}$ for some $\vec{q} \in \text{Ker}(\nu)$, then $\vec{p}(t) \equiv \vec{q} \in \text{Ker}(\nu)$ because $\dot{\vec{p}} = -\nabla_x H(\vec{q}, \vec{x}) = \vec{0}$. Thus

$$\frac{d}{dt}\vec{x} = \nabla_p H(\vec{q}, \vec{x}) = \nabla_p H(\vec{0}, \vec{x}) = \sum_{j=1}^{M} \vec{\nu}_j \left(\Phi_j^+(\vec{x}) - \Phi_j^-(\vec{x}) \right) \tag{2.44}$$

implies $\vec{x}(t)$ is the solution to RRE (1.6).

Case (II), if $\vec{p}(t) \notin \text{Ker}(\nu)$ for all $t \in [0, T]$, then we know $\vec{\nu}_j \cdot \vec{p} \neq 0$. Then from (2.30) and Lemma 2.4, we have

$$\dot{\vec{x}}(t) = \vec{s} = \nabla_p H(\vec{p}(t), \vec{x}(t)) \neq \nabla_p H(\vec{0}, \vec{x}(t)). \tag{2.45}$$

However, from the strict convexity of L, $L(\vec{s}, \vec{x}) > 0$ for $\vec{s} \neq \nabla_p H(\vec{0}, \vec{x}(t))$, which contradicts with $Act(\vec{x}(\cdot)) = 0$.

Thus we conclude (iii). □

As mentioned in the introduction, Lax–Oleinik's representation (1.20), to which the Varadhan's nonlinear semigroup converges, shows that the function L defined in (2.29) actually gives the good rate function for the large derivation principle for the large number process $C^\nu(t)$ at single times. Precisely,

Theorem 2.6 [24] *Let C^ν be the large number process defined in (1.4) with generator Q_ν. Assume $C^\nu(0) = \vec{x}_0^\nu$ satisfying $\vec{x}_0^\nu \to \vec{x}_0$ in $\mathbb{R}^N$. Then at each time t, the random variable*

$C^v(t)$ satisfies the large deviation principle in $\mathbb{R}_+^N$ with the good rate function $I_{x_0,t}(\vec{y})$ defined in (1.20). That is, for any open set $\mathcal{O} \subset \mathbb{R}_+^N$, it holds

$$\liminf_{V \to +\infty} \frac{1}{V} \log \mathbb{P}_{\vec{x}_0^V}\{C^v(t) \in \mathcal{O}\} \geq -\inf_{\vec{x} \in \mathcal{O}} I_{\vec{x}_0,t}(\vec{x}) \qquad (2.46)$$

while for any closed set $\mathcal{C} \subset \mathbb{R}_+^N$, it holds

$$\limsup_{V \to +\infty} \frac{1}{V} \log \mathbb{P}_{\vec{x}_0^V}\{C^v(t) \in \mathcal{C}\} \leq -\inf_{\vec{x} \in \mathcal{C}} I_{\vec{x}_0,t}(\vec{x}). \qquad (2.47)$$

The sample path large deviation principle in the path space $D([0, T]; \mathbb{R}_+^N)$, i.e., the space of càdlàg functions, which is proved by Agazzi et al. [2], is more difficult and significant. Under some mild assumptions ensuring the existence of solution $C^v(t)$, [2] proved the sample path large derivation principle for $\{C^v(t)\}$, as restated below.

Theorem 2.7 (Theorem 1.6, [2]) *Let C^v be the large number process defined in (1.4) with generator Q_v defined in (2.17). Assume $C^v(0) = \vec{x}_0^V$ satisfying $\vec{x}_0^V \to \vec{x}_0$ in $\mathbb{R}^N$. Then the sample path $C^v(t)$, $t \in [0, T]$ satisfies the large deviation principle in $D([0, T]; \mathbb{R}_+^N)$ with the good rate function*

$$A_{x_0,T}(\vec{x}(\cdot)) := \begin{cases} \int_0^T L(\dot{\vec{x}}(t), \vec{x}(t))\, dt & \text{if } \vec{x}(0) = \vec{x}_0, \ \vec{x}(\cdot) \in AC([0, T]; \mathbb{R}^N), \\ +\infty & \text{otherwise.} \end{cases} \qquad (2.48)$$

That is, for any open set $\mathcal{E} \subset D([0, T]; \mathbb{R}_+^N)$, it holds

$$\liminf_{V \to +\infty} \frac{1}{V} \log \mathbb{P}_{\vec{x}_0^V}\{C^v(t) \in \mathcal{E}\} \geq -\inf_{\vec{x} \in \mathcal{E}} A_{\vec{x}_0,T}(\vec{x}(\cdot)), \qquad (2.49)$$

while for any closed set $\mathcal{G} \subset D([0, T]; \mathbb{R}_+^N)$, it holds

$$\limsup_{V \to +\infty} \frac{1}{V} \log \mathbb{P}_{\vec{x}_0^V}\{C^v(t) \in \mathcal{G}\} \leq -\inf_{\vec{x} \in \mathcal{G}} A_{\vec{x}_0,T}(\vec{x}(\cdot)), \qquad (2.50)$$

where $D([0, T]; \mathbb{R}_+^N)$ is Skorokhod space and $AC([0, T]; \mathbb{R}^N)$ is space of absolute continuous curves.

The sample path large deviation principle Theorem 2.7 covers the above single time result in Theorem 2.6. Indeed, for any fixed open set $\mathcal{O} \subset \mathbb{R}_+^N$, one takes special open set $\mathcal{E} \subset D([0, T]; \mathbb{R}_+^N)$ as $\mathcal{E} = \{\vec{x}(\cdot) \in D([0, T]; \mathbb{R}_+^N); \ \vec{x}(t) \in \mathcal{O}\}$. Then

$$\inf_{\vec{x} \in \mathcal{E}} A_{\vec{x}_0,T}(\vec{x}(\cdot)) = \inf_{\vec{y} \in \mathcal{O}} \left(\inf_{\vec{x}(\cdot) \in D([0,T];\mathbb{R}_+^N), \ \vec{x}(0)=\vec{x}_0, \ \vec{x}(t)=\vec{y}} \int_0^T L(\dot{\vec{x}}(s), \vec{x}(s))\, ds \right) = \inf_{\vec{y} \in \mathcal{O}} I_{\vec{x}_0,t}(\vec{y}).$$

Here in the last equality, the least action from 0 to T is the combination of the least action from 0 to t and a zero-cost action for t to T. However, [24] gives an alternative proof for the simple case in Theorem 2.6 using semigroup approach.

3 The Dynamic and Steady Solution to Hamilton–Jacobi Equation

In this section, we first study a general Hamiltonian $H(\vec{p}, \vec{x})$ and its HJE. (i) As a result of the law of large numbers $p(\vec{x}_v, t) \approx \delta_{\vec{x}(t)}$ in the large number limit, the minimizer of the dynamic solution to the HJE gives the deterministic macroscopic path, which is a solution to

the corresponding large number limiting ODE; see Proposition 3.1. (ii) The steady solution $\psi^{ss}(\vec{x})$ to the HJE gives a Lyapunov function and a conservative–dissipative decomposition for the macroscopic RRE; see Theorem 3.2. The thermodynamics for detailed/complex balanced RRE and also for general RRE will be discussed in Sect. 3.3 at the mesoscopic scale and in Sect. 3.4 after the passage from mesoscopic scale to macroscopic scale.

3.1 Kurtz's Limiting ODE as the Minimizer of HJE Solution ψ and Its Lyapunov Function

We further observe the following special properties for H

$$H(\vec{0}, \vec{x}) \equiv 0, \quad \text{and thus } \nabla_x H(\vec{0}, \vec{x}) \equiv 0. \tag{3.1}$$

In the following theorem, we will show the solution to the Kurtz limiting ODE (1.6) is the minimizer of HJE solution ψ. Indeed, without any symmetry assumptions, we will prove a general theorem that the HJE solution $\psi(\vec{x}, t)$ from the WKB expansion can always characterize the ODE path given by the law of large numbers, and the steady solution $\psi^{ss}(\vec{x})$ yields a Lyapunov function to RRE (1.6).

Proposition 3.1 *Let $\psi_0(\vec{x})$ be the initial data to HJE (1.17) with a generic Hamiltonian $H(\vec{p}, \vec{x})$ satisfying (3.1). Assume $\min_{\vec{x}} \psi_0(\vec{x}) = 0$ and assume $\psi_0(\vec{x})$ is smooth, strictly convex with a linear growth at the far field, i.e.*

$$c_1|\vec{x}| \leq \psi_0(\vec{x}) \leq c_2|\vec{x}|, \quad as \ |\vec{x}| \to +\infty. \tag{3.2}$$

Then

(i) *there exists a unique local-in-time strictly convex classical solution $\psi(\vec{x}, t)$, $t \in [0, T]$ to (1.17);*

(ii) *the global viscosity solution to (1.17) is given by the Lax–Oleinik semigroup (a.k.a the optimal control formulation)*

$$\psi(\vec{x}, t) = \inf_{\gamma(\cdot) \in AC([0,t]),\, \gamma(t) = \vec{x}} \int_0^t L(\dot{\gamma}(\tau), \gamma(\tau))\, d\tau + \psi_0(\gamma(0)), \quad t \in [0, +\infty); \tag{3.3}$$

(iii) *the solution to ODE $\frac{d}{dt}\vec{x} = \nabla_p H(\vec{0}, \vec{x})$ with initial data $\vec{x}_0 = \operatorname{argmin}_{\vec{x}} \psi_0(\vec{x})$ is the minimizer of $\psi(\vec{x}, t)$, i.e.,*

$$\vec{x}^*(t) = \operatorname{argmin}_{\vec{x}} \psi(\vec{x}, t), \quad for \ all \ t \in [0, T]. \tag{3.4}$$

Along the ODE solution, $\psi(\vec{x}(t), t) \equiv 0$. That is to say the trajectory of the corresponding Hamiltonian dynamics with $\vec{p} \equiv \vec{0}$ gives the mean path in the sense of the weak law of large numbers

$$\lim_{V \to +\infty} \mathbb{E}(\varphi(C_t^V)) = \varphi(\vec{x}^*(t)). \tag{3.5}$$

We remark the rigorous proof for the large deviation principle of process C^v shall be done via the convergence of the WKB reformulation for backward equation, i.e., the Varahdan's nonlinear semigroup, to the viscosity solution to the corresponding HJE (1.19) [24]. Then the concentration of measure gives the mean field limit equation $\frac{d}{dt}\vec{x} = \nabla_p H(\vec{0}, \vec{x})$. However, in this proposition, we use the WKB reformulation for the forward equation because the forward equation is more intuitive for computing the probability. Then the formal convergence from this WKB reformulation for the forward equation to the Lax–Oleinik semigroup representation of the viscosity solution to (1.17) yields the mean field limit equation.

Proof Step 1 Using the definition of the Hamiltonian in (1.16), we solve the following HJE by the characteristic method

$$\partial_t \psi + H(\nabla \psi, \vec{x}) = 0, \quad \psi(\vec{x}, 0) = \psi_0(\vec{x}). \tag{3.6}$$

Then constructing the characteristics $\vec{x}(t)$, $\vec{p}(t)$

$$\begin{aligned}\dot{\vec{x}} &= \nabla_p H(\vec{p}, \vec{x}), \quad \vec{x}(0) = \vec{x}_0, \\ \dot{\vec{p}} &= -\nabla_x H(\vec{p}, \vec{x}), \quad \vec{p}(0) = \nabla \psi_0(\vec{x}_0).\end{aligned} \tag{3.7}$$

From the assumptions on ψ_0, we know there exists T such that the characteristics $\vec{x}_1(t)$, $\vec{x}_2(t)$ starting from any initial data $(\vec{x}_1(0), \vec{p}_1(0))$, $(\vec{x}_2(0), \vec{p}_2(0))$ do not intersect. Thus upto $t \in [0, T]$, $\vec{x}(t)$, $\vec{p}(t)$ can be uniquely solved from (3.7). For any $t \in [0, T]$, we also know $\psi(\vec{x}, t)$ is convex. Then along characteristics, with $\vec{p}(t) = \nabla_x \psi(\vec{x}(t), t)$, we know $z(t) = \psi(\vec{x}(t), t)$ satisfies

$$\dot{z} = \nabla_x \psi(\vec{x}(t), t) \cdot \dot{\vec{x}} + \partial_t \psi(\vec{x}(t), t) = \vec{p} \cdot \nabla_p H(\vec{p}, \vec{x}) - H(\vec{p}, \vec{x}), \quad z(0) = \psi_0(\vec{x}_0). \tag{3.8}$$

Hence we can solve for $z(t) = \psi(\vec{x}(t), t)$. Then we know along the characteristic

$$\begin{aligned}\frac{d}{dt} H(\vec{x}(t), \vec{p}(t)) &= \dot{\vec{x}} \cdot \nabla_x H + \dot{\vec{p}} \cdot \nabla_p H = 0, \\ \dot{z} - \vec{p}(t) \cdot \nabla_p H(\vec{p}, \vec{x}) &= -H(\vec{p}, \vec{x}) = -H(\vec{p}_0, \vec{x}_0).\end{aligned} \tag{3.9}$$

The Lax–Oleinik formula for the global viscosity solution in conclusion (ii) is a direct application of the dynamic program principle (semigroup property of (3.3)); see [72, Theorem 2.22].

Step 2 Particularly, taking $\vec{x}_0$ as the minimizer of ψ_0 such that $\nabla_x \psi_0(\vec{x}_0) = \vec{0}$ and thus $\vec{p}(0) = \vec{0}$. Then from (3.1), we have

$$\vec{p}(t) \equiv \vec{0}, \quad \frac{d}{dt} \psi(\vec{x}(t), t) = \dot{z} \equiv -H(\vec{p}_0, \vec{x}_0) = 0 \tag{3.10}$$

and we obtain ODE

$$\frac{d}{dt} \vec{x}(t) = \nabla_p H(\vec{0}, \vec{x}(t)). \tag{3.11}$$

$$\vec{x}(t) = \operatorname{argmin}_{\vec{x}} \psi(\vec{x}, t). \tag{3.12}$$

Now we prove the trajectory in (3.12) is the mean path in the sense of a weak formulation of the law of large numbers. Recall the WKB expansion for the law of large number process C^V, i.e., $p(\vec{x}_v, t) = e^{-V\psi(\vec{x}_v, t)}$ for any $t > 0$. Then for any test function $\varphi(\vec{x})$, the expectation of φ satisfies

$$\mathbb{E}(\varphi(C_t^V)) = \sum_{\vec{x}_v} \varphi(\vec{x}_v) p(\vec{x}_v, t) = \frac{\sum_{\vec{x}_v} \varphi(\vec{x}_v) e^{-V\psi(\vec{x}_v, t)}}{\sum_{\vec{x}_v} e^{-V\psi(\vec{x}_v, t)}}, \quad t > 0. \tag{3.13}$$

Then by the Laplace principle, we have for any $t \in [0, T]$,

$$\lim_{V \to +\infty} \mathbb{E}(\varphi(C_t^V)) = \varphi(\vec{x}^*(t)), \quad \vec{x}^*(t) = \operatorname{argmin}_{\vec{x}} \psi(\vec{x}, t). \tag{3.14}$$

In other words, the process time marginal $C^V(t)$ converges to $\vec{x}^*(t)$ in law for any t. $\qquad \square$

3.2 Conservative–Dissipative Decomposition for the Macroscopic RRE

In this section, we study finer properties of the macroscopic RRE based on any smooth enough stationary solutions $\psi^{ss}(\vec{x})$ to HJE. We will first decompose the RRE as a conservative part and a dissipative part in Theorem 3.2. Then using another conservation law for the total mass in special chemical reactions, we explore the GENERIC formalism and bi-anti-symmetric structures in the decomposition.

Theorem 3.2 *Let $H(\vec{p}, \vec{x})$ be a convex Hamiltonian satisfying $H(\vec{0}, \vec{x}) = 0$. Assume $\psi^{ss}(\vec{x})$ is a steady solution to the corresponding HJE satisfying $H(\nabla \psi^{ss}(\vec{x}), \vec{x}) = 0$. Then we have the following conservative–dissipative decomposition for the macroscopic RRE $\frac{d}{dt}\vec{x} = \nabla_p H(\vec{0}, \vec{x})$*

$$\dot{\vec{x}} = W(\vec{x}) \, - \, K(\vec{x})\nabla \psi^{ss}(\vec{x}),$$

$$W(\vec{x}) := \int_0^1 \nabla_p H(\theta \nabla \psi^{ss}(\vec{x}), \vec{x}) \, d\theta, \quad K(\vec{x}) := \int_0^1 (1 - \theta)\nabla_{pp}^2 H(\theta \nabla \psi^{ss}(\vec{x}), \vec{x}) \, d\theta. \tag{3.15}$$

(i) $W(\vec{x})$ *is the conservative part satisfying*

$$\langle W(\vec{x}), \nabla \psi^{ss}(\vec{x}) \rangle = 0. \tag{3.16}$$

Thus we recast conservative part as $W(\vec{x}) = \mathcal{A}(\vec{x})\nabla \psi^{ss}(\vec{x})$, where $\mathcal{A}(\vec{x}) := \frac{W \otimes \nabla \psi^{ss}(\vec{x}) - \nabla \psi^{ss}(\vec{x}) \otimes W}{|\nabla \psi^{ss}(\vec{x})|^2}$ is an anti-symmetric operator.
(ii) $-K(\vec{x})\nabla \psi^{ss}(\vec{x})$ *is the dissipative part with a nonnegative definite operator $K(\vec{x})$. Thus any increasing function $\phi(\cdot)$ of $\psi^{ss}(\vec{x})$ is a Lyapunov function for the ODE with energy dissipation relation*

$$\frac{d}{dt}\phi(\psi^{ss}(\vec{x})) = \langle \dot{\vec{x}}, \, \phi'(\psi^{ss}(\vec{x}))\nabla \psi^{ss}(\vec{x}) \rangle = -\langle \phi'(\psi^{ss}(\vec{x}))K(\vec{x})\nabla \psi^{ss}(\vec{x}), \, \nabla \psi^{ss}(\vec{x}) \rangle \le 0. \tag{3.17}$$

Proof Notice $\psi^{ss}(\vec{x})$ is a steady solution to the HJE satisfying $H(\vec{0}, \vec{x}) = H(\nabla \psi^{ss}(\vec{x}), \vec{x}) = 0$.

We first recast the right-hand-side of the RRE as

$$\nabla_p H(\vec{0}, \vec{x}) = - \int_0^1 \partial_\theta \left((1 - \theta)\nabla_p H(\theta \nabla \psi^{ss}, \vec{x}) \right) d\theta$$

$$= - \int_0^1 (1 - \theta)\nabla_{pp}^2 H(\theta \nabla \psi^{ss}(\vec{x}), \vec{x}) \, d\theta \nabla \psi^{ss}(\vec{x}) \tag{3.18}$$

$$+ \int_0^1 \nabla_p H(\theta \nabla \psi^{ss}, \vec{x}) \, d\theta.$$

This gives the right-hand-side of (3.15) with the definitions $K(\vec{x})$, $W(\vec{x})$. For the first term in (3.15), taking inner product with $\nabla \psi^{ss}$ gives

$$\langle W(\vec{x}), \nabla \psi^{ss}(\vec{x}) \rangle = \int_0^1 \partial_\theta H(\theta \nabla \psi^{ss}(\vec{x}), \vec{x}) \, d\theta = H(\nabla \psi^{ss}(\vec{x}), \vec{x}) - H(0, \vec{x}) = 0 \tag{3.19}$$

due to $H(\vec{0}, \vec{x}) = H(\nabla \psi^{ss}(\vec{x}), \vec{x}) = 0$. From Lemmas 2.3 and 2.4, we know $\langle \nabla_{pp}^2 H(\vec{p}, \vec{x})\vec{p}, \vec{p} \rangle \ge 0$ and strictly positive in G, so K is nonnegative definite operator.

Second, from $\phi' \geq 0$ and the orthogonality (3.19), we conclude the energy dissipation (3.17). Particularly,

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi^{ss}(\vec{x}) = \langle \dot{\vec{x}}, \nabla\psi^{ss}\rangle = -\langle K(\vec{x})\nabla\psi^{ss}(\vec{x}), \nabla\psi^{ss}(\vec{x})\rangle \leq 0. \tag{3.20}$$

$\square$

Remark 3.3 We remark that in the chemical Langevin approximation, see for instance [23] and (5.9), the Lagrangian and Hamiltonian are both quadratic and thus the above decomposition becomes transparent. We point out the choice of the Hamiltonian for the RRE decomposition is not unique, which leads to different interpretations in energetics and kinetics; c.f., [60].

Remark 3.4 Although we have a family of Lyapunov functions $\phi(\psi^{ss}(\vec{x}(t)))$, we will see only the stationary solution ψ^{ss} is the energy landscape of the chemical reactions later in Sect. 4. The energy dissipation (3.20) can be regarded as the large number limit of the energy dissipation law for the mesoscopic master equation in terms of the natural relative entropy $\rho \log \frac{\rho}{\pi} - \rho + 1$; see Proposition 3.8 for a passage from mesoscopic to macroscopic in the large number limit. In other words, the RRE can be decomposed as an Onsager-type strong gradient flow in the direction of $\nabla\psi^{ss}$, and a conservative flow in the orthogonal direction of $\nabla\psi^{ss}$; both with the same free energy $\psi^{ss}(\vec{x})$. Thus (3.15) can be regarded as a conservative dynamics coupling with a dissipation structure. Due to the competition in the chemical reaction represented by Hamiltonian H between conservative force and dissipation in terms of a nonconvex energy landscape $\psi^{ss}(\vec{x})$, this system can exhibit complicated dynamic patterns such as limit cycles, oscillations, chaotic attractors and multi-stability, etc. The decomposition (3.15) is a "pre-GENERIC" formalism. The concept "pre-GENERIC" was proposed by Kraaij et al. [37] which replaces the Hamiltonian part $\mathcal{A}\nabla E$ in the original GENERIC formalism by a general orthogonal term W such that $\langle W, \nabla\psi^{ss}\rangle = 0$.

Below, we explore further the GENERIC formalism and two anti-symmetric structures for RRE by utilizing the additional mass conservation law for chemical reaction, i.e., for any $\vec{m} \in \mathrm{Ker}(\nu)$, we have $\frac{\mathrm{d}}{\mathrm{d}t}\vec{x}(t) \cdot \vec{m} = 0$.

3.2.1 GENERIC Formalism for RRE

Denote the conservative part in the decomposition (3.15) as

$$W(\vec{x}) := \int_0^1 \nabla_p H(\theta\nabla\psi^{ss}(\vec{x}), \vec{x})\,\mathrm{d}\theta. \tag{3.21}$$

For any $\vec{m} \in \mathrm{Ker}(\nu)$, we use the conserved mass as the role of the conserved energy functional in GENERIC formalism

$$E(\vec{x}) := \vec{m} \cdot \vec{x}. \tag{3.22}$$

Since $\nabla_p H(\vec{p}, \vec{x}) = \sum_j \vec{\nu}_j \left(\Phi_j^+ e^{\vec{\nu}_j \cdot \vec{p}} - \Phi_j^- e^{-\vec{\nu}_j \cdot \vec{p}}\right) \in G$, the conservative part W is orthogonal to $\nabla E = \vec{m}$

$$\langle W(\vec{x}), \vec{m}\rangle = 0, \quad \forall \vec{m} \in \mathrm{Ker}(\nu). \tag{3.23}$$

Therefore, W can be recast as

$$W(\vec{x}) = \frac{(W \otimes \vec{m} - \vec{m} \otimes W)\vec{m}}{|\vec{m}|^2} =: \mathcal{A}_1(\vec{x})\nabla E, \tag{3.24}$$

where $\mathcal{A}_1(\vec{x}) := \frac{(W \otimes \vec{m} - \vec{m} \otimes W)}{|\vec{m}|^2}$ is an anti-symmetric matrix satisfying

$$\mathcal{A}_1(\vec{x}) \nabla \psi^{ss}(\vec{x}) = \vec{0}. \tag{3.25}$$

Here we remark that since H is degenerate in $\mathrm{Ker}(\nu)$, without loss of generality, $\psi^{ss}(\vec{x})$ can be chosen within G or we only require $\langle W(\vec{x}), \nabla \psi^{ss}(\vec{x}) \rangle = 0$.

On the other hand, by Lemma 2.3, we know $\vec{m}^T K(\vec{x}) \vec{m} = 0$. Since $K(\vec{x})$ is nonnegative definite, we have

$$K(\vec{x}) \nabla E = 0. \tag{3.26}$$

In summary, for special Hamiltonian H in chemical reactions, the decomposition (3.15) for RRE can be recast as GENERIC formalism

$$\dot{\vec{x}} = \mathcal{A}_1(\vec{x}) \nabla E(\vec{x}) - K(\vec{x}) \nabla \psi^{ss}(\vec{x}). \tag{3.27}$$

3.2.2 Two Anti-symmetric Structures for the Conservative Part

Apart from the full decomposition, we study two anti-symmetric structures of the conservative part $W(\vec{x}) = \int_0^1 \nabla_p H(\theta \nabla \psi^{ss}(\vec{x}), \vec{x}) \, d\theta$ in (3.15) for RRE.

First, we observe there are two conservation laws for W: one is (3.23) and the other one is

$$\langle W(\vec{x}), \nabla \psi^{ss}(\vec{x}) \rangle = 0 \tag{3.28}$$

due to (3.19). Following the same idea for constructing the anti-symmetric operator in (3.24), given any conservation laws $\langle W, \vec{a} \rangle = 0$, $W(\vec{x})$ can always be recast as anti-symmetric matrix $\mathcal{A}$ using

$$W = \frac{(W \otimes \vec{a} - \vec{a} \otimes W) \vec{a}}{|\vec{a}|^2} =: \mathcal{A}(\vec{x}) \vec{a}. \tag{3.29}$$

Particularly, using the formula for H in (1.16),

$$
\begin{aligned}
W(\vec{x}) &= \sum_j \vec{\nu}_j \left(\Phi_j^+ \int_0^1 e^{\theta \vec{\nu}_j \cdot \nabla \psi^{ss}} \, d\theta - \Phi_j^- \int_0^1 e^{-\theta \vec{\nu}_j \cdot \nabla \psi^{ss}} \, d\theta \right) \\
&= \sum_j \left(\Phi_j^+ \left(e^{\vec{\nu}_j \cdot \nabla \psi^{ss}} - 1 \right) + \Phi_j^- \left(e^{-\vec{\nu}_j \cdot \nabla \psi^{ss}} - 1 \right) \right) \frac{(\vec{\nu}_j \otimes \nabla \psi^{ss}) \nabla \psi^{ss}}{\vec{\nu}_j \cdot \nabla \psi^{ss} |\nabla \psi^{ss}|^2} \\
&= \sum_j \left(\Phi_j^+ \left(e^{\vec{\nu}_j \cdot \nabla \psi^{ss}} - 1 \right) + \Phi_j^- \left(e^{-\vec{\nu}_j \cdot \nabla \psi^{ss}} - 1 \right) \right) \frac{(\vec{\nu}_j \otimes \nabla \psi^{ss} - \nabla \psi^{ss} \otimes \vec{\nu}_j)}{\vec{\nu}_j \cdot \nabla \psi^{ss} |\nabla \psi^{ss}|^2} \nabla \psi^{ss} \\
&=: \mathcal{A}_2(\vec{x}) \nabla \psi^{ss}.
\end{aligned}
\tag{3.30}
$$

That is to say, for chemical reactions with H in (1.16), the decomposition (3.15) can be recast as

$$\frac{d}{dt} \vec{x} = \nabla_p H(\vec{0}, \vec{x}) = \sum_{j=1}^M \vec{\nu}_j \left(\Phi_j^+(\vec{x}) - \Phi_j^-(\vec{x}) \right) =: \mathcal{A}_2(\vec{x}) \nabla \psi^{ss}(\vec{x}) - K(\vec{x}) \nabla \psi^{ss}(\vec{x}) \tag{3.31}$$

with explicit formulas

$$K(\vec{x}) := \sum_{j=1}^{M} \left(\Phi_j^+ \left(e^{\vec{v}_j \cdot \nabla \psi^{ss}} - 1 - \vec{v}_j \cdot \nabla \psi^{ss} \right) + \Phi_j^- \left(e^{-\vec{v}_j \cdot \nabla \psi^{ss}} - 1 + \vec{v}_j \cdot \nabla \psi^{ss} \right) \right) \frac{\vec{v}_j \otimes \vec{v}_j}{|\vec{v}_j \cdot \nabla \psi^{ss}|^2}$$

$$\mathcal{A}_2(\vec{x}) := \sum_{j=1}^{M} \left(\Phi_j^+ \left(e^{\vec{v}_j \cdot \nabla \psi^{ss}} - 1 \right) + \Phi_j^- \left(e^{-\vec{v}_j \cdot \nabla \psi^{ss}} - 1 \right) \right) \frac{(\vec{v}_j \otimes \nabla \psi^{ss} - \nabla \psi^{ss} \otimes \vec{v}_j)}{\vec{v}_j \cdot \nabla \psi^{ss} |\nabla \psi^{ss}|^2}.$$

$$(3.32)$$

It is easy to verify the positive symmetry of K, the anti-symmetry of $\mathcal{A}$ and

$$\langle K(\vec{x}) \nabla \psi^{ss}(\vec{x}), \nabla \psi^{ss}(\vec{x}) \rangle \geq 0, \quad \langle \mathcal{A}_2(\vec{x}) \nabla \psi^{ss}(\vec{x}), \nabla \psi^{ss}(\vec{x}) \rangle = 0.$$

We point out the decomposition (3.31) is analogous to the Landau–Lifshitz equation instead of the GENERIC formalism (3.27).

In summary, combining (3.24) and (3.30), two conservation laws leads to two anti-symmetric structures for the conservative part $W(\vec{x})$

$$W(\vec{x}) = \mathcal{A}_1(\vec{x}) \nabla E(\vec{x}) = \mathcal{A}_2(\vec{x}) \nabla \psi^{ss}(\vec{x}). \tag{3.33}$$

With the above characterization for relations between solutions to HJE (1.17) and RRE trajectory in Proposition 3.1, we summarize the corresponding relations and the decompositions for RRE in the following Corollary.

Corollary 3.5 *For Hamiltonian $H(\vec{p}, \vec{x})$ in (1.16) for a chemical reaction, the minimizer of the dynamic solution $\psi(\vec{x}, t)$ gives the unique solution $\vec{x}(t)$ to the RRE (1.6) with initial data $\vec{x}_0 = \min_{\vec{x}} \psi_0$. The steady solution $\psi^{ss}(\vec{x})$ gives a Lyapunov function to RRE (1.6). The GENERIC formalism for RRE reads as (3.27). Another decomposition for RRE reads as (3.31) with K and $\mathcal{A}_2$ defined in (3.32).*

As an example, we show that if RRE satisfies the detailed/complex balance condition, then it is well-known that a closed formula solution for the Lyapunov function is $\psi^{ss}(\vec{x}) = \mathrm{KL}(\vec{x}||\vec{x}^s)$; c.f. [69]. In the RRE detailed balanced case, using (2.14), we have dissipation relation $\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{KL}(\vec{x}(t)||\vec{x}^s) = \sum_i \dot{x}_i \log \frac{x_i}{x_i^s} = -\sum_j \left(\Phi_j^-(\vec{x}) - \Phi_j^+(\vec{x}) \right) \log \left(\frac{\Phi_j^-(\vec{x})}{\Phi_j^+(\vec{x})} \right) \leq 0$. For the complex balance case, we summarize the following equivalent characterization for complex balance condition (2.6). The proof will be given in Appendix C

Lemma 3.6 *Assume $\vec{x}^s > 0$ is a positive steady state to RRE (1.6). Then the following statements are equivalent:*

(i) *$\vec{x}^s$ satisfies complex balance condition (2.6);*
(ii) *The relative entropy $\psi^{ss}(\vec{x}) = \mathrm{KL}(\vec{x}||\vec{x}^s)$ is a steady solution to HJE (1.17);*
(iii) *The product of Poisson distribution $\pi_V(\vec{x}_v) = \Pi_{i=1}^{N} \frac{(V x_i^S)^{V x_i}}{(V x_i)!} e^{-V x_i^S}$ is a positive invariant measure $\pi_V(\vec{x}_v)$ to the mesoscopic CME (1.5).*

3.3 Thermodynamics of Chemical Reaction Systems

Suppose chemical reactions is in a large reservoir that has a constant temperature T and a constant pressure. In this section, we devote to study the thermodynamics for non-equilibrium chemical reactions. For non-equilibrium chemical reactions, the reaction affinity along j-th reaction pathway is introduced by Kondepudi and Prigogine [39]

$$\mathcal{A}_j = k_{\mathrm{B}} T \log \frac{\Phi_j^+(\vec{x})}{\Phi_j^-(\vec{x})},$$

where k_{B} is the Boltzmann constant. For the equilibrium reaction, $\mathcal{A}_j$ reduces to the difference of the Gibbs free energy along j-th reaction pathway (3.38), so this is a natural extension from equilibrium reactions. Then the total entropy production rate is

$$T\dot{S}_{\mathrm{tot}} := \sum_j \left(\Phi_j^+(\vec{x}) - \Phi_j^-(\vec{x}) \right) \mathcal{A}_j = k_{\mathrm{B}} T \sum_j \left(\Phi_j^+(\vec{x}) - \Phi_j^-(\vec{x}) \right) \log \frac{\Phi_j^+(\vec{x})}{\Phi_j^-(\vec{x})}. \quad (3.34)$$

We will first decompose the total entropy production rate into adiabatic and nonadiabatic contributions. Then we apply it to equilibrium reactions to check the consistency with classical thermodynamic relations based on the Gibbs theory. Particularly, as $t \to +\infty$, at non-equilibrium steady states (NESS), the non-equilibrium dynamics still maintain a positive total entropy production rate.

Recall that the energy landscape ψ^{ss} gives dissipation relation (3.20) in Proposition 3.1. The nonadiabatic entropy production rate representing the dissipation of the energy landscape ψ^{ss} is defined as

$$T\dot{S}_{na} := -k_{\mathrm{B}} T \frac{\mathrm{d}}{\mathrm{d}t} \psi^{ss}(\vec{x}(t)) = k_{\mathrm{B}} T \langle K(\vec{x}) \nabla \psi^{ss}(\vec{x}), \nabla \psi^{ss}(\vec{x}) \rangle \geq 0.$$

As $t \to +\infty$, this nonadiabatic entropy production rate goes to zero. However, as one of the most important features for non-equilibrium reaction, at NESS, the total entropy production rate is positive. Therefore, apart from the nonadiabatic entropy production rate $T\dot{S}_{na}$, the remaining part in $T\dot{S}_{\mathrm{tot}}$ is usually called the adiabatic entropy production rate

$$T\dot{S}_a = T\dot{S}_{\mathrm{tot}} - T\dot{S}_{na}$$

$$= k_{\mathrm{B}} T \sum_j \left(\Phi_j^+(\vec{x}(t)) - \Phi_j^-(\vec{x}(t)) \right) \log \left(\frac{\Phi_j^+(\vec{x}(t))}{\Phi_j^-(\vec{x}(t))} e^{\vec{\nu}_j \cdot \nabla \psi^{ss}} \right)$$

$$= k_{\mathrm{B}} T \sum_j \left(\mathrm{KL}(\Phi_j^+(\vec{x}(t)) || \Phi_j^-(\vec{x}(t)) e^{-\vec{\nu}_j \cdot \nabla \psi^{ss}}) + \mathrm{KL}(\Phi_j^-(\vec{x}(t)) || \Phi_j^+(\vec{x}(t)) e^{\vec{\nu}_j \cdot \nabla \psi^{ss}}) \right) \geq 0,$$

$$(3.35)$$

where in the last equality, we used $H(\nabla \psi^{ss}(\vec{x}), \vec{x}) = 0$. When $t \to +\infty$, we have

$$T\dot{S}_a \to k_{\mathrm{B}} T \sum_j \left(\Phi_j^+(\vec{x}^{\mathrm{s}}) - \Phi_j^-(\vec{x}^{\mathrm{s}}) \right) \log \frac{\Phi_j^+(\vec{x}^{\mathrm{s}})}{\Phi_j^-(\vec{x}^{\mathrm{s}})} \quad (3.36)$$

which is strictly positive for $\vec{x}^{\mathrm{s}}$ being NESS.

In summery, we have

Proposition 3.7 *The thermodynamic decomposition for the total entropy production rate* (3.34) *of a non-equilibrium chemical reaction is given by*

$$T\dot{S}_{tot} = T\dot{S}_{na} + T\dot{S}_a \geq 0,$$

$$T\dot{S}_{na} = k_B T \langle K(\vec{x}) \nabla \psi^{ss}(\vec{x}), \nabla \psi^{ss}(\vec{x}) \rangle \geq 0,$$

$$T\dot{S}_a = k_B T \sum_j \left(\mathrm{KL}(\Phi_j^+(\vec{x}(t)) || \Phi_j^-(\vec{x}(t)) e^{-\vec{\nu}_j \cdot \nabla \psi^{ss}}) \right. \quad (3.37)$$

$$\left. + \mathrm{KL}(\Phi_j^-(\vec{x}(t)) || \Phi_j^+(\vec{x}(t)) e^{\vec{\nu}_j \cdot \nabla \psi^{ss}}) \right) \geq 0,$$

where $K(\vec{x})$ is the nonnegative definite operator in (3.32) and $\psi^{ss}(\vec{x})$ is the stationary solution to HJE (1.17).

Now we review the thermodynamic decomposition for the equilibrium chemical reactions to see the above discussion is a natural generalization to non-equilibrium reactions. Assume the chemical potential μ_i satisfies the thermodynamic relation

$$\mu_i = \mu_i^G + k_B T \log x_i, \quad \mu_i^G = \mu_i^0 - k_B T \log x_i^0$$

where μ_i^G is a reference Gibbs free energy (physically called standard-state Gibbs free energy) and $\log x_i - \log x_i^0$ together is a dimensionless quantity. From LMA (1.7), we have

$$\log\left(\frac{\Phi_j^-(\vec{x})}{\Phi_j^-(\vec{x}^s)}\frac{\Phi_j^+(\vec{x}^s)}{\Phi_j^+(\vec{x})}\right) = \sum_i v_{ji}\left(\log x_i - \log x_i^s\right) = \sum_i v_{ji}\frac{\mu_i - \mu_i^{eq}}{k_B T}. \quad (3.38)$$

In the RRE detailed/complex balanced case, Rao and Esposito [64] introduced a Lyapunov function called Shear Lyapunov function $\psi^{ss}(\vec{x}) = G_{eq} + k_B T\, \mathrm{KL}(\vec{x}(t)||\vec{x}^s)$, where G_{eq} is the equilibrium Gibbs free energy with an additional linear combination of conservative quantities. Particularly, [64] also extend this relation to an open chemical reaction network, where chemostat species interacting with both internal species and environment are included. With the special Shear Lyapunov function, the decomposition (3.35) is reduced to

$$T\dot{S}_a = k_B T \sum_j \left(\Phi_j^+(\vec{x}(t)) - \Phi_j^-(\vec{x}(t))\right) \log\left(\frac{\Phi_j^+(\vec{x}(t))}{\Phi_j^-(\vec{x}(t))} e^{\vec{v}_j \cdot \log\frac{x_i}{x_i^s}}\right)$$

$$= k_B T \sum_j \left(\Phi_j^+(\vec{x}(t)) - \Phi_j^-(\vec{x}(t))\right) \log\left(\frac{\Phi_j^+(\vec{x}^s(t))}{\Phi_j^-(\vec{x}^s(t))}\right) \quad (3.39)$$

$$= T\dot{S}_{tot} - T\dot{S}_{na} = T\dot{S}_{tot} + k_B T \frac{d}{dt} \mathrm{KL}(\vec{x}(t)||\vec{x}^s),$$

where the entropy production rate from adiabatic contribution $T\dot{S}_a$ represents the chemical work rate performed by the chemostats when interacting with the environment; see [64, Eq. (84)].

We point out for general non-equilibrium reactions, $\psi^{ss}(\vec{x})$ is an asymptotically effective energy in the large number limit, which is different from the thermodynamic free energy (Kirkwood potential [36]). On the other hand, we also discuss the passage from the mesoscopic thermodynamics to the macroscopic thermodynamics in Sect. 3.4 below, where ψ^{ss} can be regarded as the large number limit of the mesoscopic relative entropy.

3.4 Energy Dissipation Law and Passage from Mesoscopic to Macroscopic Dynamics

For a general non-equilibrium RRE and the corresponding CME, we first derive the ϕ-divergence energy dissipation law based on the Q_V-matrix structure and a Bregman's divergence. This type of ϕ-divergence energy dissipation law was previously derived by [49] under the RRE detailed balance condition.

Based on this energy dissipation law for general non-equilibrium reactions, we take $\phi(\rho) = \rho \log\frac{\rho}{\pi_V}$, then as $V \to +\infty$, the corresponding mesoscopic energy dissipation relation converges to the macroscopic energy dissipation relation in terms of the energy landscape $\psi^{ss}(\vec{x})$. This shows the passage from a mesoscopic convex functional to a macroscopic non-convex function for general non-equilibrium chemical reactions. We also remark

that under the RRE detailed balance condition, [49] rigorously proved the evolutionary Γ-convergence in the generalized gradient flow setting for the passage from mesoscopic to macroscopic dynamics. However, in the RRE detailed balance, there is no such a transition from convex functional to a non-convex function; see Remark 3.9.

Proposition 3.8 *Assume there exists a positive invariant measure $\pi_V(\vec{x}_v)$ and the limit $\psi^{ss}(\vec{x}) := \lim_{V \to +\infty} -\frac{\log \pi_V(\vec{x}_v)}{V}$ exists. Then*

(i) *for any convex function ϕ, we have the mesoscopic energy dissipation relation*

$$\frac{d}{dt} \sum_{\vec{x}_v} \phi\left(\frac{p(\vec{x}_v)}{\pi(\vec{x}_v)}\right) \pi(\vec{x}_v) = -\sum_{\vec{x}_v, \vec{y}_v} Q(\vec{y}_v, \vec{x}_v) \pi(\vec{y}_v) D_\phi\left(\frac{p(\vec{y}_v)}{\pi(\vec{y}_v)}, \frac{p(\vec{x}_v)}{\pi(\vec{x}_v)}\right) \le 0,$$
(3.40)

where $D_\phi(y, x) := (y - x)^2 \int_0^1 (1 - \theta)\phi''(x + \theta(y - x)) \, d\theta \ge 0$. Particularly, taking $\phi(x) = x \log x - x + 1 \ge 0$

$$\frac{d}{dt} \sum_{\vec{x}_v} p(\vec{x}_v) \log \frac{p(\vec{x}_v)}{\pi(\vec{x}_v)} = -\sum_{\vec{x}_v, \vec{y}_v} Q(\vec{y}_v, \vec{x}_v) p(\vec{y}_v) \log \frac{p(\vec{y}_v)\pi(\vec{x}_v)}{\pi(\vec{y}_v)p(\vec{x}_v)} \le 0, \quad (3.41)$$

where Q is the Q_V-matrix in (1.5) for the mesoscopic jumping process C_V;

(ii) *Formally, as $V \to +\infty$, the mesoscopic dissipation law (3.41) converges to the macroscopic dissipation law (3.20) in the sense that*

$$\frac{1}{V} \sum_{\vec{x}_v} p(\vec{x}_v) \log \frac{p(\vec{x}_v)}{\pi(\vec{x}_v)} \to \psi^{ss}(\vec{x}^*),$$

$$\frac{1}{V} \sum_{\vec{x}_v, \vec{y}_v} Q(\vec{y}_v, \vec{x}_v) p(\vec{y}_v) \log \frac{p(\vec{y}_v)\pi(\vec{x}_v)}{\pi(\vec{y}_v)p(\vec{x}_v)} \to \langle K(\vec{x}^*)\nabla\psi^{ss}(\vec{x}^*), \nabla\psi^{ss}(\vec{x}^*)\rangle, \quad (3.42)$$

where $\vec{x}^*(t) = \arg\min_{\vec{x}} \psi(\vec{x}, t)$ is the mean path obtained in Proposition 3.1 such that $C^V(t)$ converges in law to $\vec{x}^*(t)$.

Proof First, recast the master equation for chemical reaction (1.5) as $\frac{dp^T}{dt} = p^T Q$. Since $\sum_{\vec{x}_v} Q(\vec{y}_v, \vec{x}_v) = 0$ and $\sum_{\vec{y}_v} Q(\vec{y}_v, \vec{x}_v)\pi(\vec{y}_v) = 0$, then for any functions $\phi(x), \psi(x)$,

$$\frac{d}{dt} \sum_{\vec{x}_v} \phi\left(\frac{p(\vec{x}_v)}{\pi(\vec{x}_v)}\right) \pi(\vec{x}_v) = \sum_{\vec{x}_v, \vec{y}_v} Q(\vec{y}_v, \vec{x}_v) p(\vec{y}_v) \phi'\left(\frac{p(\vec{x}_v)}{\pi(\vec{x}_v)}\right)$$

$$= \sum_{\vec{x}_v, \vec{y}_v} Q(\vec{y}_v, \vec{x}_v) \pi(\vec{y}_v) \frac{p(\vec{y}_v)}{\pi(\vec{y}_v)} \left(\phi'\left(\frac{p(\vec{x}_v)}{\pi(\vec{x}_v)}\right) - \phi'\left(\frac{p(\vec{y}_v)}{\pi(\vec{y}_v)}\right)\right)$$

$$= \sum_{\vec{x}_v, \vec{y}_v} Q(\vec{y}_v, \vec{x}_v) \pi(\vec{y}_v) \frac{p(\vec{y}_v)}{\pi(\vec{y}_v)} \left(\phi'\left(\frac{p(\vec{x}_v)}{\pi(\vec{x}_v)}\right) - \phi'\left(\frac{p(\vec{y}_v)}{\pi(\vec{y}_v)}\right)\right)$$

$$\qquad - \sum_{\vec{x}_v, \vec{y}_v} Q(\vec{y}_v, \vec{x}_v) \pi(\vec{y}_v) \left(\psi\left(\frac{p(\vec{x}_v)}{\pi(\vec{x}_v)}\right) - \psi\left(\frac{p(\vec{y}_v)}{\pi(\vec{y}_v)}\right)\right) \qquad (3.43)$$

$$= \sum_{\vec{x}_v, \vec{y}_v} Q(\vec{y}_v, \vec{x}_v) \pi(\vec{y}_v) \left(\frac{p(\vec{y}_v)}{\pi(\vec{y}_v)}\left(\phi'\left(\frac{p(\vec{x}_v)}{\pi(\vec{x}_v)}\right) - \phi'\left(\frac{p(\vec{y}_v)}{\pi(\vec{y}_v)}\right)\right)\right.$$

$$\qquad \left. - \left(\psi\left(\frac{p(\vec{x}_v)}{\pi(\vec{x}_v)}\right) - \psi\left(\frac{p(\vec{y}_v)}{\pi(\vec{y}_v)}\right)\right)\right).$$

Furthermore, take $\psi(x) := x\phi'(x) - \phi(x)$. Denote $y = \frac{p(\vec{y}_v)}{\pi(\vec{y}_v)}$ and $x = \frac{p(\vec{x}_v)}{\pi(\vec{x}_v)}$, then the dissipation can be rewritten as a Bregman's divergence

$$y[\phi'(x) - \phi'(y)] - [\psi(x) - \psi(y)] = (y - x)\phi'(x) + \phi(x) - \phi(y) =: -D_\phi(y, x). \tag{3.44}$$

Using the integral form of the reminder in Taylor expansion,

$$D_\phi(y, x) = (y - x)^2 \int_0^1 (1 - \theta)\phi''(x + \theta(y - x)) \, d\theta \geq 0. \tag{3.45}$$

This concludes (3.40). Take $\phi(x) = x \log x - x + 1 \geq 0$, the dissipation relation becomes (3.41).

Second, recall the change of variables in WKB expansion

$$\psi_V(\vec{x}_v, t) = -\frac{\log p(\vec{x}_v, t)}{V}, \quad \psi_V^{ss}(\vec{x}_v) = -\frac{\log \pi(\vec{x}_v)}{V}. \tag{3.46}$$

Then

$$\log \frac{p(\vec{x}_v - \frac{\vec{v}_j}{V}, t)}{p(\vec{x}_v, t)} = -V\left(\psi(\vec{x}_v - \frac{\vec{v}_j}{V}, t) - \psi(\vec{x}_v)\right) = \vec{v}_j \cdot \int_0^1 \nabla\psi(\vec{x}_v - \theta\frac{\vec{v}_j}{V}, t) \, d\theta. \tag{3.47}$$

Using the definition of Q-matrix, the dissipation relation (3.41) reads

$$\frac{d}{dt} \sum_{\vec{x}_v} p(\vec{x}_v, t) \left(\psi_V^{ss}(\vec{x}_v) - \psi_V(\vec{x}_v, t)\right)$$

$$= -\sum_{\vec{x}_v} \sum_{j=1}^M \left[\tilde{\Phi}_j^+(\vec{x}_v - \frac{\vec{v}_j}{V})p(\vec{x}_v - \frac{\vec{v}_j}{V}, t)\vec{v}_j \cdot \int_0^1 \nabla\left(\psi_V - \psi_V^{ss}\right)(\vec{x}_v - \theta\frac{\vec{v}_j}{V}) \, d\theta\right.$$

$$\left. - \tilde{\Phi}_j^-(\vec{x}_v + \frac{\vec{v}_j}{V})p(\vec{x}_v + \frac{\vec{v}_j}{V}, t)\vec{v}_j \cdot \int_0^1 \nabla\left(\psi_V - \psi_V^{ss}\right)(\vec{x}_v + \theta\frac{\vec{v}_j}{V}) \, d\theta\right] \tag{3.48}$$

Taking limit $V \to +\infty$, from Proposition 3.1, $C^V(t)$ converges in law to $\vec{x}^*(t) = \arg\min \psi(\vec{x}, t)$. Using the fact that $\psi(\vec{x}^*(t), t) = 0$, $\nabla\psi(\vec{x}^*(t), t) = 0$, the left-hand-side of (3.48) satisfies

$$\frac{1}{V} \sum_{\vec{x}_v} p(\vec{x}_v) \log \frac{p(\vec{x}_v)}{\pi(\vec{x}_v)} = \sum_{\vec{x}_v} p(\vec{x}_v, t) \left(\psi_V^{ss}(\vec{x}_v) - \psi_V(\vec{x}_v, t)\right) \to \psi^{ss}(\vec{x}^*(t)) - \psi(\vec{x}^*(t), t)$$
$$= \psi^{ss}(\vec{x}^*(t)). \tag{3.49}$$

Similarly, the right-hand-side of (3.48) satisfies

$$-\sum_{\vec{x}_v} \sum_{j=1}^M \tilde{\Phi}_j^+(\vec{x}_v - \frac{\vec{v}_j}{V})p(\vec{x}_v - \frac{\vec{v}_j}{V}, t)\vec{v}_j \cdot \int_0^1 \nabla\left(\psi_V - \psi_V^{ss}\right)(\vec{x}_v - \theta\frac{\vec{v}_j}{V}) \, d\theta$$

$$\to -\sum_{j=1}^M \Phi_j^+(\vec{x}^*(t))\vec{v}_j \cdot \left(\nabla\psi(\vec{x}^*(t), t) - \nabla\psi^{ss}(\vec{x}^*(t))\right) = \sum_{j=1}^M \Phi_j^+(\vec{x}^*(t))\vec{v}_j \cdot \nabla\psi^{ss}(\vec{x}^*(t)), \tag{3.50}$$

thus we arrive at

$$-\frac{1}{V} \sum_{\vec{x}_v, \vec{y}_v} Q(\vec{y}_v, \vec{x}_v) p(\vec{y}_v) \log \frac{p(\vec{y}_v)\pi(\vec{x}_v)}{\pi(\vec{y}_v)p(\vec{x}_v)}$$

$$\rightarrow \sum_{j=1}^{M}(\Phi_j^+(\vec{x}^*(t)) - \Phi_j^-(\vec{x}^*(t)))\vec{v}_j \cdot \nabla\psi^{ss}(\vec{x}^*(t)) = -\langle K(\vec{x})\nabla\psi^{ss}(\vec{x}^*(t)), \nabla\psi^{ss}(\vec{x}^*(t))\rangle. \tag{3.51}$$

Notice the uniqueness of weak convergence. Combining (3.41), (3.49) and (3.51), we conclude $\frac{d}{dt}\psi^{ss}(\vec{x}^*(t)) = -\langle K(\vec{x})\nabla\psi^{ss}(\vec{x}^*(t)), \nabla\psi^{ss}(\vec{x}^*(t))\rangle$, which is exactly the Lyapunov estimate (3.20) for RRE in Theorem 3.2. □

Remark 3.9 At the mesoscopic level, the energy functional $F(p) = \sum_{\vec{x}_v} \phi\left(\frac{p(\vec{x}_v)}{\pi(\vec{x}_v)}\right)\pi(\vec{x}_v)$ is also convex w.r.t. p. However, since $\phi(u)$ is convex, the nonlinear weight $\pi_v(\vec{x}_v)$ in $F(p)$ drastically pick up the complicated non-convex energy landscape for chemical reactions from $\pi_v(\vec{x}_v) \approx e^{-V\psi^{ss}(\vec{x})}$ in the large number limit. Therefore, it is natural that after the concentration of the measure in the large number limit, a non-convex energy landscape emerges. Notice also there is no such a transition from convex functional to a nonconvex function under the RRE detailed balance assumption because the corresponding probability flux is only monomial. This grouped probability flux including polynomials, which leads to non-convex energy landscape, are only linked to nonequilibrium in the specific context of chemical reactions. For general equilibrium models in statistical physics, non convex energy landscape is common, for instance the Lagenvin dynamics with non-convex potential and Ising model of ferromagnetism.

4 Symmetric Hamiltonian: Strong Gradient Flow, Reversed Least Action Curve, Non-equilibrium Enzyme Reactions

In this section, we explore the symmetry in the mesoscopic CME and its macroscopic consequences. In Sect. 4.1, we first clarify that the Markov chain detailed balance implies a symmetric Hamiltonian; see (1.26). A proper mathematical Markov chain detailed balance condition for CME is a weaker condition than the more constrained chemical version of detailed balance, and thus includes a class of non-equilibrium enzyme reactions with three distinguished features: multiple steady states, nonzero steady state fluxes and positive entropy production rates at non-equilibrium steady states [39].

Then we study in detail two consequences of this symmetry. (I) We show the conservative part $W(\vec{x})$ vanishes in the conservative–dissipation decomposition for RRE (3.15); see Sect. 4.2. (II) We prove the 'uphill' least action path is a modified time reversed curve corresponding to the RRE, i.e., corresponding to a zero action 'downhill' path, and the associated path affinity is given by the difference of ψ^{ss}; see Proposition 4.6. We will call ψ^{ss} as the energy landscape since it is a Lyapunov function of RRE and will eventually give the energy barrier of a transition path. We also provide a mesoscopic interpretation of path affinity; see Sect. 4.3.

4.1 Markov Chain Detailed Balance Implies a Symmetric Hamiltonian

We first observe the Markov chain detailed balance condition (4.3) for the mesoscopic CME is different from the more constraint RRE detailed balance (1.12). The RRE detailed balance

is a very strong symmetric condition that implies the Markov chain detailed balance condition (4.3). But the latter one is a proper mathematical definition of detailed balance for a Markov process, which leads to a symmetric Hamiltonian in a reaction system.

Observe the jumping process with generator Q_V only distinct the same reaction vector $\vec{\xi}$ and then the summation in j shall be rearranged in terms of all j such that $\vec{v}_j = \pm\vec{\xi}$. Therefore, define the probability flux for the same reaction vector $\vec{\xi}$ as

$$\Phi_{\vec{\xi}}^+(\vec{x}) := \sum_{j:\vec{v}_j=\vec{\xi}} \Phi_j^+(\vec{x}) + \sum_{j:\vec{v}_j=-\vec{\xi}} \Phi_j^-(\vec{x}), \quad \Phi_{\vec{\xi}}^-(\vec{x}) := \sum_{j:\vec{v}_j=\vec{\xi}} \Phi_j^-(\vec{x}) + \sum_{j:\vec{v}_j=-\vec{\xi}} \Phi_j^+(\vec{x}).$$

(4.1)

With the grouped probability flux, CME (1.5) can be recast as

$$\frac{d}{dt}p(\vec{x}_v, t) = V \sum_{\xi, \vec{x}_v - \frac{\vec{\xi}}{V} \geq 0} \left(\tilde{\Phi}_{\vec{\xi}}^+ (\vec{x}_v - \frac{\vec{\xi}}{V}) p(\vec{x}_v - \frac{\vec{\xi}}{V}, t) - \tilde{\Phi}_{\vec{\xi}}^-(\vec{x}_v) p(\vec{x}_v, t) \right)$$

$$+ V \sum_{\xi, \vec{x}_v + \frac{\vec{\xi}}{V} \geq 0} \left(\tilde{\Phi}_{\vec{\xi}}^-(\vec{x}_v + \frac{\vec{\xi}}{V}) p(\vec{x}_v + \frac{\vec{\xi}_j}{V}, t) - \tilde{\Phi}_{\vec{\xi}}^-(\vec{x}_v) p(\vec{x}_v, t) \right),$$

(4.2)

where the $\tilde{\Phi}_{\vec{\xi}}^\pm$ has the same definition as $\Phi_{\vec{\xi}}^\pm$ in (4.1) but replacing $\Phi_j^\pm$ by $\tilde{\Phi}_j^\pm$. For any $\vec{x}_v$ and $\vec{\xi}$, the proper mathematical definition for the Markov chain detailed balance for CME (1.5) means that there exists a positive invariant measure $\pi(\vec{x}_v)$ to CME (1.5) such that the total forward probability steady flux from $\vec{x}_v$ to $\vec{x}_v + \frac{\vec{\xi}}{V} \geq 0$ equals the total backward one

$$\tilde{\Phi}_{\vec{\xi}}^-(\vec{x}_v + \frac{\vec{\xi}}{V})\pi(\vec{x}_v + \frac{\vec{\xi}}{V}) = \tilde{\Phi}_{\vec{\xi}}^+(\vec{x}_v)\pi(\vec{x}_v), \quad \forall\vec{\xi}.$$

(4.3)

On the other hand, a commonly used detailed balance condition in biochemistry is the more constrained chemical version of detailed balance for each reaction channel, c.f., [77, Ch7, Lemma 3.1], [1, Theorem 4.5], [61, (7.30)],

$$\tilde{\Phi}_j^-(\vec{x}_v + \frac{\vec{v}_j}{V})\pi(\vec{x}_v + \frac{\vec{v}_j}{V}) = \tilde{\Phi}_j^+(\vec{x}_v)\pi(\vec{x}_v), \quad \forall j.$$

(4.4)

This is also known as Whittle's Markov chain detailed balance [35]. It is well known that the mesoscopic Whittle's Markov chain detailed balance (4.4) is equivalent to the macroscopic RRE detailed balance (1.12). Indeed, Whittle use the product of Poisson distributions with intensity $V\vec{x}^s$ to construct a detailed balanced invariant measure $\pi_V(\vec{x}_v)$ satisfying (4.4). While from (4.4), it is nontrivial to obtain a detailed balance steady state $\vec{x}^s$; see [77, Ch7, Lemma 3.1].

The proposition below shows that Markov chain detailed balance condition (4.3) gives raise a symmetry in the Hamiltonian $H(\vec{p}, \vec{x})$ (see (1.26))

$$H(\vec{p}, \vec{x}) = H(\nabla\psi^{ss}(\vec{x}) - \vec{p}, \vec{x}), \quad \forall\vec{x}, \vec{p}.$$

(4.5)

Taking $\vec{p} = \vec{0}$, we see ψ^{ss} is a steady solution to HJE (1.17).

Proposition 4.1 *Assume there exists a positive invariant measure $\pi_V(\vec{x}_v)$ satisfying the Markov chain detailed balance condition (4.3) for mesoscopic effective stochastic process. Assume the macroscopic energy landscape $\psi^{ss}(\vec{x}) := \lim_{V\to+\infty} -\frac{\log\pi_V(\vec{x}_v)}{V}$ exists, then the*

Hamiltonian $H(\vec{p}, \vec{x})$ for macroscopic RRE satisfies the symmetry (1.26) w.r.t. $\psi^{ss}(\vec{x})$, or equivalently

$$e^{\vec{\xi} \cdot \nabla \psi^{ss}(\vec{x})} \Phi_{\xi}^{+}(\vec{x}) = \Phi_{\xi}^{-}(\vec{x}). \tag{4.6}$$

Proof Let the equilibrium $\pi_V(\vec{x}_v)$ to mesoscopic effective stochastic process be $\pi_V(\vec{x}_v) = e^{-V \psi_V^{ss}(\vec{x}_v)}$. Then the Markov chain detailed balance condition (4.3) implies

$$\frac{\pi_V(\vec{x}_v)}{\pi_V(\vec{x}_v + \frac{\vec{\xi}}{V})} \tilde{\Phi}_{\xi}^{+}(\vec{x}_v) = \tilde{\Phi}_{\xi}^{-}(\vec{x}_v + \frac{\vec{\xi}}{V}). \tag{4.7}$$

Since as $V \to +\infty$, $\vec{x}_v \to \vec{x}$ and $\psi_V^{ss}(\vec{x}_v) \to \psi^{ss}(\vec{x})$, then

$$\frac{\pi_V(\vec{x}_v)}{\pi_V(\vec{x}_v + \frac{\vec{\xi}}{V})} = e^{\vec{\xi} \cdot \int_0^1 \nabla \psi_V^{ss}(\vec{x}_v + \theta \frac{\vec{\xi}}{V}) \, d\theta} \to e^{\vec{\xi} \cdot \nabla \psi^{ss}(\vec{x})}. \tag{4.8}$$

Then taking limit in (4.7), we obtain (4.6). Using the notation for probability flux $\Phi_{\xi}^{\pm}(\vec{x})$ in (4.1), the Hamiltonian becomes

$$H(\vec{p}, \vec{x}) := \sum_{\vec{\xi}} \left(\Phi_{\xi}^{+}(\vec{x}) e^{\vec{\xi} \cdot \vec{p}} - \Phi_{\xi}^{+}(\vec{x}) + \Phi_{\xi}^{-}(\vec{x}) e^{-\vec{\xi} \cdot \vec{p}} - \Phi_{\xi}^{-}(\vec{x}) \right). \tag{4.9}$$

Since $\vec{p}$ in (1.26) is arbitrary, we rearrange and take out common factor $e^{\vec{\xi} \cdot \vec{p}}$. Then the even symmetry of H in (4.13) is equivalent to for any $\vec{p}$

$$
\begin{aligned}
&H(\nabla \psi^{ss}(\vec{x}) - \vec{p}, \vec{x}) - H(\vec{p}, \vec{x}) \\
&= \sum_{\vec{\xi}} \left(\Phi_{\xi}^{+}(\vec{x}) e^{\vec{\xi} \cdot \nabla \psi^{ss}(\vec{x})} e^{-\vec{\xi} \cdot \vec{p}} + \Phi_{\xi}^{-}(\vec{x}) e^{-\vec{\xi} \cdot \nabla \psi^{ss}(\vec{x})} e^{\vec{\xi} \cdot \vec{p}} - \Phi_{\xi}^{+}(\vec{x}) e^{\vec{\xi} \cdot \vec{p}} - \Phi_{\xi}^{-}(\vec{x}) e^{-\vec{\xi} \cdot \vec{p}} \right) \\
&= \sum_{\vec{\xi}} \left(\Phi_{\xi}^{+}(\vec{x}) e^{\vec{\xi} \cdot \nabla \psi^{ss}(\vec{x})} - \Phi_{\xi}^{-}(\vec{x}) \right) \left(e^{-\vec{\xi} \cdot \vec{p}} - e^{-\vec{\xi} \cdot \nabla \psi^{ss}(\vec{x})} e^{\vec{\xi} \cdot \vec{p}} \right).
\end{aligned} \tag{4.10}
$$

This means the coefficients of exponential function $e^{\vec{\xi} \cdot \vec{p}}$ for each distinct $\vec{\xi}$ must be same, hence (4.6) is equivalent to (1.26). □

We remark the existence of $\psi^{ss}(\vec{x})$ constructed from a positive invariant measure $\pi_V(\vec{x}_v)$ satisfying (4.3) was proved in [24] in the sense of an upper semicontinuous (USC) viscosity solution to stationary HJE following Barron–Jensen's definition [9] for USC viscosity solution. However, the uniqueness and selection principle for those USC viscosity solution is still open.

Corollary 4.2 *Let H be the Hamiltonian for a chemical reaction defined in (1.16) and $\vec{x}^s$ be any steady states for RRE (1.6). Then a necessary condition for the symmetry of H (1.26) is that for each reaction vector $\vec{\xi}$ in the chemical reaction*

$$\Phi_{\xi}^{+}(\vec{x}^s) = \Phi_{\xi}^{-}(\vec{x}^s). \tag{4.11}$$

Proof If $\vec{x}^s$ is a steady state for RRE (1.6), we know $\nu \nabla \psi^{ss}(\vec{x}^s) = \vec{0}$. Otherwise, take $\vec{x}_0 = \vec{x}^s$ as initial data, from the estimate (3.20),

$$0 \equiv \frac{d \psi^{ss}(\vec{x}(t))}{dt} = -\langle \nabla \psi^{ss}(\vec{x}), K(\vec{x}) \nabla \psi^{ss} \rangle < 0, \tag{4.12}$$

due to H is strictly convex in G (see Lemma 2.4). This contradiction shows $\nabla \psi^{ss}(\vec{x}^s) \in \mathrm{Ker}(\nu)$ thus $\nu \nabla \psi^{ss}(\vec{x}^s) = \vec{0}$. Therefore, evaluating (4.6) at $\vec{x}^s$ yields (4.11). □

Remark 4.3 As a slight generalization, an even-symmetry of the Hamiltonian w.r.t $\frac{\vec{q}}{2}$ is

$$H(\vec{p}, \vec{x}) = H(\vec{q}(\vec{x}) - \vec{p}, \vec{x}), \quad \forall \vec{x}, \vec{p} \tag{4.13}$$

for some function $\vec{q}(\vec{x})$. This is equivalent to

$$L(\vec{s}, \vec{x}) - L(-\vec{s}, \vec{x}) = \vec{s} \cdot \vec{q}(\vec{x}). \tag{4.14}$$

A Hamiltonian which is quadratic in terms of the momentum $\vec{p}$ is a special cases of (1.26). For an irreversible drift-diffusion process, $dx = -\vec{q} \, dt + \sqrt{2\varepsilon} \, dB$, the corresponding Hamiltonian $H(\vec{p}, \vec{x}) = \vec{p} \cdot (\vec{p} - \vec{q})$ satisfies even-symmetry (4.13). Another example in electromagnetism is the even-symmetry for momentum $\vec{p}$ in Hamiltonian w.r.t the magnetic vector potential [53].

In the following subsections, we study two consequences for a symmetric Hamiltonian.

(I) Under symmetric assumption (1.26), we provide an Onsager's strong form of gradient flow structure in terms of the energy landscape ψ^{ss} in Sect. 4.2. That is to say, the conservative part $W(\vec{x})$ vanish in the previous conservative–dissipation decomposition for RRE (3.15).

(II) The symmetric Hamiltonian (1.26) is equivalent to the time reversal symmetry in the Lagrangian L upto a null Lagrangian (see (1.27))

$$L(\vec{s}, \vec{x}) - L(-\vec{s}, \vec{x}) = \vec{s} \cdot \nabla \psi^{ss}(\vec{x}), \quad \forall \vec{x}, \vec{s}. \tag{4.15}$$

Here $\vec{s} \cdot \nabla \psi^{ss}(\vec{x})$ is a null Lagrangian, whose Euler-Lagrange equation vanishes. Denote the time reversed curve of $\vec{x}(t)$ as $\vec{x}^R(t) = \vec{x}(T-t)$ with $\vec{x}^R(T) = \vec{x}_T^R = \vec{x}^C$ and $\vec{x}^R(0) = \vec{x}_0^R = \vec{x}^A$. Then take $\vec{s} = \dot{\vec{x}}^R(t)$ in (4.15) and integrate w.r.t time t from 0 to T leads to the action cost identity

$$Act(\vec{x}^R(\cdot)) - Act(\vec{x}(\cdot)) = \psi^{ss}(\vec{x}_T^R) - \psi^{ss}(\vec{x}_0^R); \tag{4.16}$$

see Proposition 4.6. We point out the time reversed least action path is an application of the Freidlin–Wentzell theory [21] for a general exit problem to the chemical reactions while it also gives the most probable path connecting two steady states $\vec{x}^A, \vec{x}^B$. As a well-known application of due to symmetric Hamiltonian, for a Langevin dynamics with a potential form drift $-\nabla U$, the Freidlin–Wentzell theory [21] shows the 'uphill' least action curve with nonzero action is exactly the time reversal of the 'downhill' least action curve. The associated Hamiltonian for this Langevin dynamics is $H(\vec{p}, \vec{x}) = \vec{p} \cdot (\vec{p} - \nabla U)$, which is symmetric $H(\vec{p}, \vec{x}) = H(\nabla U - \vec{p}, \vec{x}), \quad \forall \vec{x}, \vec{p}$.

The idea of using this kind of symmetric Hamiltonian to find the time reversed least action curve for some classical mechanics was first discovered by Morpurgo et al. [53]. In the RRE detailed balanced case, the symmetric property w.r.t $\frac{1}{2} \nabla KL(\vec{x}, \vec{x}^s)$ of the Hamiltonian was first studied in [12], Dykman et al. With a symmetric Hamiltonian, the corresponding generalized gradient flow was first studied in Mielke et al. [50], where the residual of the gradient flow was connected with the rate function in the large deviation principle. The symmetric Hamiltonian was also used in Bertin [6] to study the fluctuation symmetry; see also a comprehensive review [7] on the macroscopic fluctuation theory and recent development in [37, 59, 65].

4.2 Onsager's Strong Form of Gradient Flow in Terms of Energy Landscape ψ^{ss}

In this section, under the symmetric assumption (1.26) for Hamiltonian, we derive a strong form of gradient flow formulation, where the steady solution ψ^{ss} to the HJE serves as a free

energy. This gradient flow immediately gives vanishing of the conservative part $W(\vec{x}) = 0$ in RRE decomposition (3.15).

Proposition 4.4 *Under the symmetric assumption* (1.26)*, the RRE* (1.6) *becomes a strong gradient flow in terms of* $\psi^{ss}(\vec{x})$

$$\dot{\vec{x}} = -K(\vec{x})\nabla\psi^{ss}(\vec{x}), \quad K(\vec{x}) = \int_0^1 \frac{1}{2}\nabla_{pp}^2 H(\theta\nabla\psi^{ss}(\vec{x}))\,\mathrm{d}\theta. \tag{4.17}$$

Particularly, for chemical reaction with RRE detailed balance (1.12)*,* (4.17) *reduces to*

$$\dot{\vec{x}} = -K(\vec{x})\nabla\,\mathrm{KL}(\vec{x}||\vec{x}^s), \quad K(\vec{x}) = \sum_{j=1}^M \Lambda\left(\Phi_j^+(\vec{x}), \Phi_j^-(\vec{x})\right)\left(\vec{v}_j \otimes \vec{v}_j\right), \tag{4.18}$$

where $\Lambda(x, y) := \frac{x-y}{\log x - \log y}$ *is the logarithmic mean.*

Proof Recall the decomposition for RRE in (3.15). When $H(\vec{p}, \vec{x})$ satisfies the symmetric condition (1.26), we have

$$\nabla_p H(\nabla\psi^{ss}(\vec{x}) - \vec{p}, \vec{x}) = -\nabla_p H(\vec{p}, \vec{x}), \quad \forall\vec{p}. \tag{4.19}$$

Taking $\vec{p} = \theta\nabla\psi^{ss}$, then

$$\nabla_p H((1-\theta)\nabla\psi^{ss}(\vec{x}), \vec{x}) = -\nabla_p H(\theta\nabla\psi, \vec{x}). \tag{4.20}$$

Then integrating w.r.t θ implies

$$\int_0^1 \nabla_p H(\theta\nabla\psi^{ss}(\vec{x}), \vec{x})\,\mathrm{d}\theta = 0$$

and thus the RRE is simply a strong gradient flow (4.17). Furthermore, recall symmetric nonnegative operator $K(\vec{x}) = \int_0^1 (1-\theta)\nabla_{pp}^2 H(\theta\nabla\psi^{ss}(\vec{x}))\,\mathrm{d}\theta$. From (4.19), we have the symmetry

$$\int_0^1 (1-\theta)\nabla_{pp}^2 H(\theta\nabla\psi^{ss}(\vec{x}))\,\mathrm{d}\theta = \int_0^1 \theta\nabla_{pp}^2 H(\theta\nabla\psi^{ss}(\vec{x}))\,\mathrm{d}\theta.$$

Thus

$$K(\vec{x}) = \int_0^1 (1-\theta)\nabla_{pp}^2 H(\theta\nabla\psi^{ss}(\vec{x}))\,\mathrm{d}\theta = \int_0^1 \frac{1}{2}\nabla_{pp}^2 H(\theta\nabla\psi^{ss}(\vec{x}))\,\mathrm{d}\theta. \tag{4.21}$$

Particularly, let $\vec{x}^s$ be a steady solution to (1.6) satisfying RRE detailed balance condition (1.12). Then we know $\psi^{ss}(\vec{x}) = \mathrm{KL}(\vec{x}||\vec{x}^s)$ and $\vec{v}_j \cdot \log\frac{\vec{x}}{\vec{x}^s} = \log\frac{\Phi_j^-(\vec{x})}{\Phi_j^+(\vec{x})}$. Thus the K-matrix in (4.21) reduces to

$$K(\vec{x}) = \sum_{j=1}^M \frac{\left(\Phi_j^-(\vec{x}) - \Phi_j^+(\vec{x})\right)\vec{v}_j \otimes \vec{v}_j}{\vec{v}_j \cdot \nabla\psi^{ss}(\vec{x})} = \sum_{j=1}^M \Lambda\left(\Phi_j^+(\vec{x}), \Phi_j^-(\vec{x})\right)\left(\vec{v}_j \otimes \vec{v}_j\right). \tag{4.22}$$

$\square$

The above formula (4.18) can also be written as $\sum_{j=1}^M \Phi_j(\vec{x}^s)\Lambda\left(\frac{\Phi_j^+(\vec{x})}{\Phi_j(\vec{x}^s)}, \frac{\Phi_j^-(\vec{x})}{\Phi_j(\vec{x}^s)}\right)\left(\vec{v}_j \otimes \vec{v}_j\right)$, which exactly recovers the well-known strong gradient flow represented by the logarithmic mean $\Lambda(x, y) = \frac{x-y}{\log x - \log y}$ [30] ([49, Theorem 2.2]) for RRE detailed balance case.

Remark 4.5 To fit into more general biochemical reactions such as gene switch [66], we give a slightly more general symmetric condition for H so that the strong gradient flow structure still holds. Assume there exists $\alpha(\vec{x}) > 0$ such that

$$H(\vec{p}, \vec{x}) = H(\nabla \psi^{ss}(\vec{x}) - \alpha(\vec{x})\vec{p}, \vec{x}) \tag{4.23}$$

and H still satisfies $H(\vec{0}, \vec{x}) = H(\nabla \psi^{ss}(\vec{x}), \vec{x}) = 0$. Then we have

$$\frac{d}{dt}\vec{x} = \nabla_p H(\vec{0}, \vec{x}) = -\alpha(\vec{x}) \nabla_p H(\nabla \psi^{ss}(\vec{x}(t)), \vec{x}(t))$$

$$= -\alpha(\vec{x}) \left(\nabla_p H(\vec{0}, \vec{x}(t)) + 2K(\vec{x}) \nabla \psi^{ss}(\vec{x}) \right), \tag{4.24}$$

which yields a gradient flow structure

$$\frac{d}{dt}\vec{x} = \nabla_p H(\vec{0}, \vec{x}) = -\frac{2\alpha(\vec{x})}{1 + \alpha(\vec{x})} K(\vec{x}) \nabla \psi^{ss}(\vec{x}). \tag{4.25}$$

4.3 Reversed Least Action Curve and Path Affinity Described by Energy Landscape ψ^{ss}

In order to study the transition path between two states $\vec{x}^A$ and $\vec{x}^B$, we now characterize the time reversed solution to RRE (1.6). Let $\vec{x}(t)$ be the 'downhill' solution to RRE (1.6) with $\vec{x}(0) = \vec{x}^C$ and $\vec{x}(T) = \vec{x}^A$ for some finite time T. Notice this requires $\vec{x}^A$, $\vec{x}^C$ are not steady states to RRE (1.6). However, $\vec{x}^A$, $\vec{x}^C$ can be in a small neighborhood of steady states and then taking time goes to infinity gives the transition path between two stable states passing through a saddle point. Then by Proposition 2.5, we know $\vec{x}(t)$ is a least action solution with action cost $Act(\vec{x}^R(\cdot)) = 0$ in (2.32). We define the time reversed curve for $\vec{x}(t)$, $\vec{p}(t)$ by

$$\vec{x}^R(t) = \vec{x}(T - t), \quad \vec{p}^R(t) = \vec{p}(T - t), \quad 0 \le t \le T. \tag{4.26}$$

Then we know $\vec{x}^R$ satisfies

$$\dot{\vec{x}}^R = \sum_{j=1}^M \vec{v}_j \left(\Phi_j^-(\vec{x}^R) - \Phi_j^+(\vec{x}^R) \right), \quad \vec{x}_0^R = \vec{x}(T) = \vec{x}^A, \ \vec{x}_T^R = \vec{x}(0) = \vec{x}^C. \tag{4.27}$$

The following Proposition 4.6 states that the time reversed solution $\vec{x}^R$ with a modified reversed momentum

$$\vec{p}^{MR}(t) = \nabla \psi^{ss}(\vec{x}^R(t)) - \vec{p}^R(t) \tag{4.28}$$

is a 'uphill' least action solution from $\vec{x}_0^R = \vec{x}^A$ to $\vec{x}_T^R = \vec{x}^C$ but with a non-zero action $Act(\vec{x}^R(\cdot))$.

Proposition 4.6 *Given a Hamiltonian $H(\vec{p}, \vec{x})$ satisfying (1.26), suppose $L(\vec{s}, \vec{x})$ is its convex conjugate. Let $\vec{x}(t)$, $\vec{p}(t)$ be a least action solution for the action functional $Act(\vec{x}(\cdot)) = \int_0^T L(\dot{\vec{x}}, \vec{x}) \, dt$ starting from $\vec{x}(0) = \vec{x}^C$ and ending at $\vec{x}(T) = \vec{x}^A$. Then for the time reversed solution $\vec{x}^R(t)$, $\vec{p}^R(t)$ defined in (4.26), we know*

(i) *the modified time reversed solution $\vec{x}^R(t)$, $\vec{p}^{MR}(t) = \nabla \psi^{ss}(\vec{x}^R(t)) - \vec{p}^R(t)$ is a least action curve starting from $\vec{x}_0^R = \vec{x}(T) = \vec{x}^A$, ending at $\vec{x}_T^R = \vec{x}(0) = \vec{x}^C$ and satisfies the Hamiltonian dynamics*

$$\frac{d}{dt}\vec{x}^R = \nabla_p H(\vec{p}^{MR}, \vec{x}^R), \quad \frac{d}{dt}\vec{p}^{MR} = -\nabla_x H(\vec{p}^{MR}, \vec{x}^R); \tag{4.29}$$

(ii) *the corresponding action cost for the least action curve $\vec{x}^R(t)$ is given by*

$$Act(\vec{x}^R(\cdot)) = Act(\vec{x}(\cdot)) + \psi^{ss}(\vec{x}_T^R) - \psi^{ss}(\vec{x}_0^R). \tag{4.30}$$

Proof First, recall the definition of H in (1.16) which satisfies (1.26). Then regarding $\vec{p}, \vec{x}$ as independent variables in $H(\nabla\psi^{ss}(\vec{x}) - \vec{p}, \vec{x}) \equiv H(\vec{p}, \vec{x})$, taking derivatives, we directly have following identities

$$\nabla_p H(\vec{p}, \vec{x}) \equiv -\nabla_p H(\nabla\psi^{ss}(\vec{x}) - \vec{p}, \vec{x}), \tag{4.31}$$

$$\nabla_x H(\vec{p}, \vec{x}) \equiv \nabla_x H(\nabla\psi^{ss}(\vec{x}) - \vec{p}, \vec{x}^R) + \nabla^2\psi^{ss}(\vec{x})\nabla_p H(\nabla\psi^{ss}(\vec{x}) - \vec{p}, \vec{x}). \tag{4.32}$$

Second, from (4.31), we have

$$\frac{d}{dt}\vec{x}^R(t) = -\nabla_p H(\vec{p}^R(t), \vec{x}^R(t)) = \nabla_p H(\vec{p}^{MR}(t), \vec{x}^R(t)). \tag{4.33}$$

Third, by the definition of modified reversed momentum $\vec{p}^{MR}$ and (4.32), we have

$$\begin{aligned}
\frac{d}{dt}\vec{p}^{MR}(t) &= -\frac{d}{dt}\vec{p}^R(t) + \nabla^2\psi^{ss}(\vec{x}^R(t))\dot{\vec{x}}^R(t) = \dot{\vec{p}}\Big|_{T-t} \\
&\quad +\nabla^2\psi^{ss}(\vec{x}^R(t))\nabla_p H(\vec{p}^{MR}(t), \vec{x}^R(t)) \\
&= -\nabla_x H(\vec{p}^R(t), \vec{x}^R(t)) + \nabla^2\psi^{ss}(\vec{x}^R(t))\nabla_p H(\vec{p}^{MR}(t), \vec{x}^R(t)) \\
&= -\nabla_x H(\vec{p}^{MR}(t), \vec{x}^R(t)),
\end{aligned} \tag{4.34}$$

Fourth, combining (4.33) and (4.34), we know $(\vec{x}^R(t), \vec{p}^{MR}(t))$ solves a Hamiltonian dynamics. Notice for any Hamiltonian trajectory $(\vec{x}(t), \vec{p}(t))$, the Lagrangian can be expressed as

$$L(\dot{\vec{x}}(t), \vec{x}(t)) = \vec{p}(t) \cdot \dot{\vec{x}}(t) - H(\vec{p}(t), \vec{x}(t)). \tag{4.35}$$

From (1.26), one can directly compute the action cost along $\vec{x}^R$

$$\begin{aligned}
Act(\vec{x}^R(\cdot)) &= \int_0^T L(\dot{\vec{x}}^R(t), \vec{x}^R(t))\, dt = \int_0^T \left(\vec{p}^{MR}(t) \cdot \dot{\vec{x}}^R(t) - H(\vec{p}^{MR}(t), \vec{x}^R(t))\right) dt \\
&= \int_0^T \left(\nabla\psi^{ss}(\vec{x}^R) \cdot \frac{d}{dt}\vec{x}^R(t) - \vec{p}^R(t) \cdot \nabla_p H(\vec{p}^{MR}, \vec{x}^R) - H(\vec{p}^{MR}(t), \vec{x}^R(t))\right) dt.
\end{aligned} \tag{4.36}$$

where we used (4.28). Then by (4.31), we obtain

$$\begin{aligned}
Act(\vec{x}^R(\cdot)) &= \int_0^T \left(\nabla\psi^{ss}(\vec{x}^R) \cdot \frac{d}{dt}\vec{x}^R(t) + \vec{p}^R(t) \cdot \nabla_p H(\vec{p}^R, \vec{x}^R) - H(\vec{p}^R(t), \vec{x}^R(t))\right) dt \\
&= \int_0^T L(\dot{\vec{x}}(T-t), \vec{x}(T-t))\, dt + \int_0^T \frac{d}{dt}\psi^{ss}(\vec{x}^R(t))\, dt \\
&= Act(\vec{x}(\cdot)) + \psi^{ss}(\vec{x}_T^R) - \psi^{ss}(\vec{x}_0^R),
\end{aligned} \tag{4.37}$$

which concludes (4.30). □

The proof for the minimum cost only relies on the observation for null Lagrange $L(\dot{\vec{x}}, \vec{x}) - L(-\dot{\vec{x}}, \vec{x})$ in (1.27). The statement (ii) for the reversed action cost can be understood as a path affinity describing in which direction the chemical reaction (or a general nonlinear dynamics) proceed. Precisely, this affinity is given by $Act(\vec{x}^R(\cdot)) - Act(\vec{x}(\cdot)) = \psi^{ss}(\vec{x}_T^R) - \psi^{ss}(\vec{x}_0^R)$. In the case of $\vec{p}^R \equiv \vec{0}$, the 'downhill' path corresponds to the solution to RRE (1.6) with action cost $Act(\vec{x}(\cdot)) = 0$. In this case, the reversed action cost is

$$Act(\vec{x}^R(\cdot)) = \psi^{ss}(\vec{x}_T^R) - \psi^{ss}(\vec{x}_0^R) = \psi^{ss}(\vec{x}^C) - \psi^{ss}(\vec{x}^A). \tag{4.38}$$

Usually, the steady solution $\psi^{ss}(\vec{x})$ to HJE is known as the quasipotential $V(\vec{x}^{\mathrm{C}}; \vec{x}^{\mathrm{A}})$ upto a constant for the exit problem in the Freidlin–Wentzell theory [21] in the sense that for $\vec{x}^{\mathrm{C}}$ in the basin of attraction

$$\psi^{ss}(\vec{x}^{\mathrm{C}}) - \psi^{ss}(\vec{x}^{\mathrm{A}}) = V(\vec{x}^{\mathrm{C}}; \vec{x}^{\mathrm{A}}) = \inf_{T>0,\ \vec{x}^{\mathrm{R}}(0)=\vec{x}^{\mathrm{A}},\ \vec{x}^{\mathrm{R}}(T)=\vec{x}^{\mathrm{C}}} \int_0^T L(\dot{\vec{x}}^{\mathrm{R}}(t), \vec{x}^{\mathrm{R}}(t))\, \mathrm{d}t.$$

(4.39)

Below we explain why the modified least action path with associated action cost in (4.38) for a given fixed time T coincides with the quasipotential $\psi^{ss}(\vec{x})$ in (4.39), where the time T is also a variable to minimize. On the one hand, the assumption in Proposition 4.6 that there exists a forward least action curve from $\vec{x}(0) = \vec{x}^{\mathrm{C}}$ to $\vec{x}(T) = \vec{x}^{\mathrm{A}}$ already gives the curve trajectory with a fixed reaching time T. On the other hand, as proved above, the symmetric property ensures that the modified time reversal yields exactly the 'uphill' least action curve with the same trajectory and the same reaching time T, but only along a reversed time order. That is to say, the associated action cost in (4.38) with the reaching time T also implies the optimal time in (4.39) is exactly the reaching time T in the 'downhill' solution to the RRE. The minimization problem (4.39) is also called Maupertuis's principle of least action in an undefined time horizon and we will formulate it as an optimal control problem in a undefined (infinite) time horizon in Sect. 5 without the detailed balance.

The quasipotential is a generalized potential function to quantify the energy barrier for the transitions starting from one stable state to another one. The concept of energy barrier (a.k.a activation energy) was initialed by Arrhenius in 1889 who related the transition rate $\mathcal{K}$ to the free energy difference via Arrhenius's law $\mathcal{K} \propto e^{-\frac{\delta E}{\varepsilon}}$ with a noise parameter ε indicating the thermal energy. Using a Langevin dynamics starting from $\vec{x}$ in a basin of attraction, Kramers estimated the mean first passage time $\tau(\vec{x}) = \mathbb{E}^{\vec{x}}(\tau_C)$, i.e., the expectation of the stopping time defined as the first hitting time on the boundary of the basin of attraction, which gives an explicit formula for transition rate $\mathcal{K} = \frac{1}{\tau(\vec{x})} \propto e^{-\frac{\delta E}{\varepsilon}}$. At a rigorously mathematical level, the large deviation theory for general stochastic processes gives the estimate for the reaction rate by computing the probability of the exit problem from the basin of attraction via the good rate functional $\mathbb{P}\{C^V \in \Gamma\} \approx e^{-\frac{\inf_{\vec{x}\in\Gamma} Act(\vec{x}(\cdot))}{\varepsilon}}$; see comprehensive studies in the Freidlin–Wentzell theory [21] and precise statement in [2, Theorem 1.6] for chemical reactions. Therefore, the least action cost computed in (4.38) gives the energy barrier for the transition path problem in chemical reactions and this is why we call ψ^{ss} the energy landscape for RRE (1.6).

In Proposition 4.6, the assumption that there exists a forward least action curve from $\vec{x}(0) = \vec{x}^{\mathrm{C}}$ to $\vec{x}(T) = \vec{x}^{\mathrm{A}}$ already limits the curve within the stable basin of $\vec{x}^{\mathrm{A}}$ and not passing beyond the separatrix (boundary of the basin). Within the basin of attraction of stable state $\vec{x}^{\mathrm{A}}$, the globally defined energy landscape ψ^{ss} coincides with the quasipotential upto a constant. On the other hand, the transition path connecting $\vec{x}^{\mathrm{A}}$ to $\vec{x}^{\mathrm{B}}$ and passing through some saddle point $\vec{x}^{\mathrm{C}}$ is one of the most important scientific questions. In this case, the most probable path is piecewise defined by finding the least action curve from $\vec{x}^{\mathrm{A}}$ to $\vec{x}^{\mathrm{C}}$ and then from $\vec{x}^{\mathrm{C}}$ to $\vec{x}^{\mathrm{B}}$; see Sect. 5.1. The energy barrier for the rare transition shall be computed piecewise, for instance, the energy barrier for transition $\vec{x}^{\mathrm{A}}$ to $\vec{x}^{\mathrm{B}}$ is given by the 'uphill' action cost $\psi^{ss}(\vec{x}^{\mathrm{C}}) - \psi^{ss}(\vec{x}^{\mathrm{A}})$ plus zero action cost for the 'downhill' curve. In practice, given the energy landscape ψ^{ss}, there are many methods such as string method [75] to find the saddle point $\vec{x}^{\mathrm{C}}$.

4.3.1 Mesoscopic Interpretation of Path Affinity

Denote $(\Omega, \mathcal{F}, \mathbb{P})$ as a probobility space. For the large number process $C^V(t)$, define a new random variable on the Skorokhod space $D([0, T]; \mathbb{R}_+^N)$ through the froward trajectory $\vec{x}^V(\cdot)$ as

$$Z_v := \log \frac{d\mathbb{P}^V_{[0,T]}}{d\mathbb{P}^V_{[T,0]}}(\vec{x}^V(\cdot)) \tag{4.40}$$

where $\vec{x}^V_R(t) := \vec{x}^V(T - t) =: (\mathcal{R} \circ \vec{x}^V)(t)$ is the time reversed trajectory and $\mathcal{R}$ is the reversed operator. Here $\mathbb{P}^V_{[0,T]}$ is the probability measure on path space $D([0, T]; \mathbb{R}_+^N)$ defined via pushforward of $\mathbb{P}$ by $\vec{x}^V(\cdot)$, i.e., $\vec{x}^V(\cdot)_\# \mathbb{P}$ and with the reversed operator $\mathcal{R}$, $\mathbb{P}^V_{[T,0]}$ is the probability measure on path space $D([T, 0]; \mathbb{R}_+^N)$ defined via pushforward $(\mathcal{R} \circ \vec{x}^V)_\# \mathbb{P}$. The ratio $\frac{d\mathbb{P}^V_{[0,T]}}{d\mathbb{P}^V_{[T,0]}}(\vec{x}^V(\cdot))$ is the Radon-Nikodym derivative. This Z_v is known as the fluctuating entropy production rate of any forward trajectory $\vec{x}^V(\cdot)$, with respect to its reversed trajectory. Z_v was also historically introduced by Onsager as a dissipation function [15, 78].

For simplicity in presentation, we assume there is probability density (with the same notations as the distribution) for $d\mathbb{P}^V_{[0,T]}$ then $d\mathbb{P}^V_{[0,T]}(\vec{x}^V) = \mathbb{P}^V_{[0,T]}(\vec{x}^V) dx$ and similarly we have $d\mathbb{P}^V_{[T,0]}(\vec{x}^V) = \mathbb{P}^V_{[0,T]}(\mathcal{R} \circ \vec{x}^V) dx$. Fix starting point $\vec{x}_0 = \vec{x}^B$ and ending point $\vec{x}_T = \vec{x}^A$. We define a subset $\Gamma \subset D([0, T]; \mathbb{R}_+^N)$ as all trajectories $\vec{x}(\cdot)$ starting from $\vec{x}_0$ and ending at $\vec{x}_T$. Then by the large deviation principle in [2, Theorem 1.6] (Theorem 2.6 above), we have

$$\lim_{V \to +\infty} \frac{1}{V} \log \mathbb{P}^V_{[0,T]}(\vec{x}^V(\cdot) \in \Gamma) = - \min_{\vec{x}(\cdot) \in \Gamma \cap AC([0,T];\mathbb{R}_+^N)} \int_0^T L(\dot{\vec{x}}(t), \vec{x}(t)) \, dt = Act(\vec{x}^*(\cdot)), \tag{4.41}$$

where min is achieved at an interior point $\vec{x}^*(\cdot)$, i.e., the least action curve among Γ satisfying Euler–Lagrange equation (2.34). On the other hand, for $\mathbb{P}^V_{[T,0]}$ defined on reversed trajectory above, we have

$$\mathbb{P}^V_{[T,0]}\{\vec{x}^V \in \Gamma\} = \mathbb{P}^V_{[0,T]}\{\vec{x}^V_R \in \Gamma\} = \mathbb{P}^V_{[0,T]}\{\vec{x}^V \in \Gamma^R\}, \tag{4.42}$$

where $\Gamma^R \subset D([T, 0]; \mathbb{R}_+^n)$ is the set of any trajectories $\vec{x}(\cdot)$ starting from $\vec{x}_T$ and ending at $\vec{x}_0$. Then we have

$$\begin{aligned}
\lim_{V \to +\infty} \frac{1}{V} \log \mathbb{P}^V_{[T,0]}(\vec{x}^V(\cdot) \in \Gamma) &= \lim_{V \to +\infty} \frac{1}{V} \log \mathbb{P}^V_{[0,T]}(\vec{x}^V(\cdot) \in \Gamma^R) \\
&= - \min_{\vec{x}(\cdot) \in \Gamma^R \cap AC} \int_0^T L(\dot{\vec{x}}(t), \vec{x}(t)) \, dt \\
&= - \min_{\vec{x}(\cdot) \in \Gamma \cap AC} \int_0^T L(\dot{\vec{x}}^R(t), \vec{x}^R(t)) \, dt \\
&= - \min_{\vec{x}(\cdot) \in \Gamma \cap AC} \int_0^T L(-\dot{\vec{x}}(t), \vec{x}(t)) \, dt = Act(\vec{x}^*_R),
\end{aligned} \tag{4.43}$$

where $\vec{x}^*_R(\cdot)$ is the least action curve among Γ^R satisfying Euler–Lagrange equation (2.34).

Under symmetric assumption (4.13) for H, Proposition 4.6 tells us $\vec{x}^*_R$ is exactly the time reversal of $\vec{x}^*$, i.e., $\vec{x}^*_R(t) = \vec{x}^*(T - t)$. Thus plugging the least action curve $\vec{x}^*$ into (4.41)

and (4.43) and taking difference, we use the relation (1.27) to derive

$$
\lim_{V \to +\infty} \frac{1}{V} \log \mathbb{P}^V_{[0,T]}(\vec{x}^V(\cdot) \in \Gamma) - \lim_{V \to +\infty} \frac{1}{V} \log \mathbb{P}^V_{[T,0]}(\vec{x}^V(\cdot) \in \Gamma)
$$

$$
= \int_0^T \left(L(-\dot{\vec{x}}^*(t), \vec{x}^*(t)) - L(\dot{\vec{x}}^*(t), \vec{x}^*(t)) \right) \mathrm{d}t \tag{4.44}
$$

$$
= - \int_0^T \dot{\vec{x}}^* \cdot \nabla \psi^{ss}(\vec{x}^*) \, \mathrm{d}t = \psi^{ss}(\vec{x}_0) - \psi^{ss}(\vec{x}_T).
$$

Therefore, the symmetric Hamiltonian implies that in the large number limit, the fluctuating entropy production rate Z_v defined in (4.40) only depends on the given initial $\vec{x}_0$, end states $\vec{x}_T$ and its value is given by the path affinity $Act(\vec{x}^R(\cdot)) - Act(\vec{x}(\cdot)) = \psi^{ss}(\vec{x}_0) - \psi^{ss}(\vec{x}_T)$.
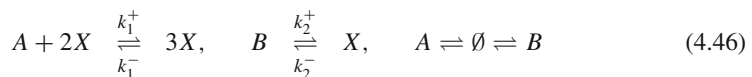
As a special example, when RRE (1.6) is detailed balanced, H satisfies the symmetry (1.26) with $\psi^{ss}(\vec{x}) = KL(\vec{x}||\vec{x}^s)$. Then the modified reversed momentum $\vec{p}^{MR}(t)$ is $\vec{p}^{MR}(t) = \log \frac{\vec{x}^R(t)}{\vec{x}^s} - \vec{p}^R(t)$. The corresponding minimum action can be calculated as

$$
Act(\vec{x}^R(\cdot)) = Act(\vec{x}(\cdot)) + \int_0^T \frac{\mathrm{d}}{\mathrm{d}t} KL(\vec{x}^R(t)||\vec{x}^s) \, \mathrm{d}t
$$

$$
= Act(\vec{x}(\cdot)) + KL(\vec{x}^R_T||\vec{x}^s) - KL(\vec{x}^R_0||\vec{x}^s). \tag{4.45}
$$

For a special case that $\vec{x}(t)$ being the solution to RRE (1.6) with $\vec{x}(0) = \vec{x}^B$ and $\vec{x}(T) = \vec{x}^A$ for some finite time T, then $\vec{x}(t)$ is a least action solution with zero action cost (the 'downhill' path). The time reversed curve $\vec{x}^R(t) = \vec{x}(T - t)$ is the most probable path (the 'uphill' path) from $\vec{x}^A$ to $\vec{x}^B$ with action cost $KL(\vec{x}^B||\vec{x}^s) - KL(\vec{x}^A||\vec{x}^s)$. As time evolves, the solution to the RRE and the reversed one stay at the same level set of the Hamiltonian $H \equiv 0$.

4.4 Non-equilibrium Example: A Bistable Schlögl Catalysis Model

In this section, we show the symmetric Hamiltonian, brought by the Markov chain detailed balance, does include a class of non-equilibrium enzyme reactions due to the flux grouping property (4.1). This type of non-equilibrium enzyme reactions plays important roles in a living cell, for instance in the phosphorylation-dephosphorylation with 2-autocatalysis described in Appendix D, the enzyme plays as an intermediate species but dramatically lower the energy barrier. The flux grouping within a same reaction vector leads to multiple steady states and nonzero steady state fluxes that maintains a ecosystem. We will illustrate the idea of constructing the optimally controlled time reserved solution in a well-known bistable reaction in an open system. Consider Schlögl catalysis model [68] with environment Ø, chemostats A, B and internal specie X

$$
A + 2X \underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} 3X, \qquad B \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} X, \qquad A \rightleftharpoons \text{Ø} \rightleftharpoons B \tag{4.46}
$$

where $k_1^+, k_1^-, k_2^+, k_2^- > 0$ are reaction rates. X plays a role of enzyme in biological system and is usually called an intermediate or an autocatalyst [8, 74]. Denote the concentration of X as x and the concentration of A, B as a, b. Here a, b are assumed to be constants that are sustained by the environment. Below, we will see there is a bifurcation ratio of a, b to classify the reaction rate system as a non-equilibrium reaction system, except for a special value of $\frac{a}{b}$. We will use the Schlögl model to describe non-equilibrium steady state behaviors, which have three typical features: (i) multiple steady states; (ii) nonzero steady state fluxes; (iii)

positive entropy production rates at non-equilibrium steady states. We first observe the flux grouping degeneracy in the same reaction vector, which usually exists in enzyme reactions such as the Michaelis–Menten kinetics; see Appendix D for a realistic phosphorylation-dephosphorylation model. This is the main reason leading to coexistence of multiple steady states. The RRE for the Schlögl model indeed can be formulated as a Ginzburg–Landau model with double well potential. Based on Proposition 4.6, we study the transition path between two stable steady states passing through an unstable state with an associated energy barrier. We will see the energy barrier is not computed from the Ginzburg–Landau double well potential but rather the energy landscape given by the HJE steady solution. That is to say, a simple diffusion approximation can not be used to compute transition path problem in the large deviation regime; c.f. [4].

In detail, given $a, b > 0$, the forward/backward fluxes of these two reactions are

$$\Phi_1^+(x) = k_1^+ a x^2, \quad \Phi_1^-(x) = k_1^- x^3, \quad \Phi_2^+(x) = k_2^+ b, \quad \Phi_2^-(x) = k_2^- x. \quad (4.47)$$

Since $\nu_{11} = 1, \nu_{21} = 1$ for internal species X, we have the macroscopic RRE for x

$$\dot{x} = k_1^+ a x^2 - k_1^- x^3 + k_2^+ b - k_2^- x =: f(x). \quad (4.48)$$

Given $a, b > 0$, the steady states of (4.48) is solved by $f(x_s) = 0$, while the RRE detailed balance condition reads

$$k_1^+ a x_s^2 = k_1^- x_s^3, \quad k_2^+ b = k_2^- x_s. \quad (4.49)$$

This means only when the ratio $\frac{a}{b} = \frac{k_1^- k_2^+}{k_1^+ k_2^-}$, there exists a unique detailed balanced equilibrium x_s. Indeed, in this case, $f(x) = \left(\frac{k_1^-}{k_2^-} x^2 + 1 \right) \left(k_2^+ b - k_2^- x \right)$. Especially, there is no external flux between chemostats A, B and the environment and thus at the detailed balanced equilibrium, the system can be regarded as a closed system. Notice in this simple example, the complex balance condition (2.6) is same as detailed balance condition (1.12), so Schlögl model (4.48) is a non-detailed/complex balanced RRE system. One can also check the deficiency of this model (as defined in Sect. (2.2.1)) is $\delta = 4 - 2 - 1 = 1$.

In general, assume $\frac{a}{b} \neq \frac{k_1^- k_2^+}{k_1^+ k_2^-}$ and $f(x)$ has three zero points. Two of them are stable steady states while one of them is an unstable steady state. We denote the corresponding antiderivative of $-f(x)$ as $\Upsilon(x)$. $\Upsilon(x)$ is a Ginzburg–Landau double well potential with two stable states, which determines the bifurcation and the first order phase transition. At a non-equilibrium steady states x^s, by elementary calculations, there is nonzero steady flux $f_1(x^s) := \Phi_1^+(x^s) - \Phi_1^-(x^s) < 0 < f_2(x^s) := \Phi_2^+(x^s) - \Phi_2^-(x^s)$. The nonzero steady flux maintains a source-production circulation $\emptyset \longrightarrow B \longrightarrow X \longrightarrow A \longrightarrow \emptyset$ in this open ecosystem at either one of the non-equilibrium steady states. This ecosystem continues exchanging both chemical energy and materials with its environment. Compared with equilibrium, the open system continues converting the chemical energy into heat at non-equilibrium steady states. The positive entropy production rate is

$$T\dot{S} = k_{\mathrm{B}} T (\Phi_1^+(x^s) - \Phi_1^-(x^s)) \log \frac{\Phi_1^+(x^s)}{\Phi_1^-(x^s)} + k_{\mathrm{B}} T (\Phi_2^+(x^s) - \Phi_2^-(x^s)) \log \frac{\Phi_2^+(x^s)}{\Phi_2^-(x^s)} > 0,$$

which characterizes the irreversible process.

With this simple non-equilibrium reaction system, we illustrate how to find a transition path as a least action curve (in the sense of the large deviation theory in [2, Theorem 1.6])

connecting two non-equilibrium steady states. In general, $\Upsilon(x)$ is not symmetric but without loss of generality we simply assume the following symmetric form

$$\Upsilon(x) = \frac{k_1^-}{4}[(x - \theta)^2 - r^2]^2. \qquad (4.50)$$

This typical symmetric double well potential has two stable local minimums $\theta \pm r$ and an unstable critical point θ provided $(k_1^+)^2 a^2 - 3k_1^- k_2^- > 0$. Then $f(x) = -\partial_x \Upsilon(x)$ implies $\theta = \frac{a k_1^+}{2 k_1^-}, r = \frac{\sqrt{a^2 (k_1^+)^2 - 3k_1^- k_2^-}}{\sqrt{3} k_1^-}$. Here k_2^+ will be a slaver parameter to ensure symmetry. With the fluxes in (4.47), the Hamiltonian is

$$\begin{aligned} H(p, x) &= [\Phi_1^+(x) + \Phi_2^+(x)](e^p - 1) + [\Phi_1^-(x) + \Phi_2^-(x)](e^{-p} - 1) \\ &= (k_1^+ a x^2 + k_2^+ b)(e^p - 1) + (k_1^- x^3 + k_2^- x)(e^{-p} - 1), \end{aligned} \qquad (4.51)$$

which is strictly convex w.r.t $\vec{p}$. This degeneracy is due to the same reaction vector yields a flux grouping. Denote

$$\alpha(x) := \frac{\Phi_1^-(x) + \Phi_2^-(x)}{\Phi_1^+(x) + \Phi_2^+(x)} = \frac{k_1^- x^3 + k_2^- x}{k_1^+ a x^2 + k_2^+ b}.$$

Then we know $p = 0$ or $p = \log \alpha(x)$ are solutions to $H(p, x) = 0$. By elementary calculations, one can verify

$$\nabla_p H(p, x)\big|_{p=0} = f(x) = -\nabla_p H(p, x)\big|_{p=\log \alpha}, \qquad (4.52)$$

and the symmetry w.r.t $\frac{\log \alpha}{2}$

$$H(p, x) = H(\log \alpha - p, x), \quad \forall x, p. \qquad (4.53)$$

Assume the 'downhill' RRE starts from some initial state $\vec{x}(0)$ and then goes to one stable state $\vec{x}(T_\varepsilon) \to \theta - r$. Here $\varepsilon > 0$ is the ε-neighborhood of $\theta - r$ and $T_\varepsilon \to +\infty$ as $\varepsilon \to 0^+$. Then by Proposition 4.6, the modified time reversed solution $\vec{x}^R(t)$ and $\vec{p}^{MR}(t) = -\vec{p}^R(t) + \vec{q}(\vec{x}^R(t)) = \log \alpha$ is still a least action solution starting from $\vec{x}_0^R = \vec{x}(T_\varepsilon) \approx \theta - r$ and ending at $\vec{x}_{T_\varepsilon}^R = \vec{x}(0)$ and satisfy the same Hamiltonian dynamics.

In this 1D example, there always exists a potential function $\psi^{ss}(x)$ such that $\log \alpha(x) = \partial_x \psi^{ss}(x)$. This ψ^{ss} is the steady solution to the HJE, and is a Lyapunov function to RRE (4.48). From Proposition 4.6, the least action value is given by the path affinity

$$Act(\vec{x}^R(\cdot)) = \psi^{ss}(\vec{x}_T^R) - \psi^{ss}(\vec{x}_0^R) = \psi^{ss}(x(0)) - \psi^{ss}(\theta - r). \qquad (4.54)$$

However, we point out energy landscape $\psi^{ss}(x)$ computed from $\log \alpha = \partial_x \psi^{ss}$ is not same as the double well potential $\Upsilon(x)$ in RRE (4.48). They are two different Lyapunov functions but have same increasing/decreasing regimes. Indeed, Since $\partial_x \psi^{ss}(x) = \log \alpha$ is the steady solution to the HJE, so by Proposition 3.1, $\psi^{ss}(\vec{x})$ is a Lyapunov function satisfying

$$\frac{d\psi^{ss}}{dt} = \partial_x \psi^{ss} \dot{x} = -\partial_x \Upsilon(x) \partial_x \psi^{ss}(x) \leq 0. \qquad (4.55)$$

Therefore, $\partial_x \Upsilon$ and $\partial_x \psi^{ss}$ has the same monotonicity. Although ψ^{ss} is also a double well potential with the same stable/unstable points as Υ, the affinity of the path is given by the difference in terms of ψ^{ss} instead of Υ. At each basin of attraction of stable states, ψ^{ss} coincides with the so-called quasipotential, as explained in (4.38).

Below, we also study the effects of perturbations of chemostats in this sustained non-equilibrium system, specifically, the linear response of the energy landscape ψ^{ss} to a perturbation of the external flux represented by chemostats. Denote the concentration of chemostats as a generic parameter b with perturbation $\varepsilon\tilde{b}$, $\varepsilon \ll 1$. Then the energy landscape satisfies the steady HJE with parameter b

$$H(\nabla\psi^{ss}(\vec{x}),\vec{x},b) = 0 = H(\nabla(\psi^{ss}(\vec{x}) + \varepsilon\tilde{\psi}^{ss}(\vec{x})),\vec{x},b+\varepsilon\tilde{b}). \qquad (4.56)$$

Here $\psi^{ss}(\vec{x}) + \varepsilon\tilde{\psi}^{ss}(\vec{x})$ is the new energy landscape under perturbed chemostats. Then Taylor's expansion w.r.t. ε gives the leading order equation

$$\nabla_p H(\nabla\psi^{ss}(\vec{x}),\vec{x},b) \cdot \nabla\tilde{\psi}^{ss}(\vec{x}) + \nabla_b H(\nabla\psi^{ss}(\vec{x}),\vec{x},b) \cdot \tilde{b} = 0. \qquad (4.57)$$

If further assume the symmetry (1.26) for Hamiltonian, the response energy landscape perturbation is given by

$$\frac{\mathrm{d}\tilde{\psi}^{ss}(\vec{x}(t))}{\mathrm{d}t} = \nabla\tilde{\psi}^{ss}(\vec{x}(t)) \cdot \dot{\vec{x}} = \nabla_b H(\nabla\psi^{ss}(\vec{x}),\vec{x},b) \cdot \tilde{b}. \qquad (4.58)$$

The rigorous justification for this linear response relation can follow the method in Hairer and Majda [32].

5 Existence of ψ^{ss} and Diffusion Approximation for Transition Paths

We have shown in previous sections that the stationary solution $\psi^{ss}(\vec{x})$ to HJE (1.17) serves as the energy landscape of chemical reactions, facilitates the conservative–dissipative decomposition for RRE, and also determines both the energy barrier and thermodynamics of chemical reactions. For a detailed/complex balanced RRE, we simply have a convex stationary solution $\psi^{ss}(\vec{x}) = \mathrm{KL}(\vec{x}||\vec{x}^s)$; see Lemma 3.6. For general chemical reactions, the existence of ψ^{ss} and obtaining ψ^{ss} via optimal control problem will be discussed in this section. Based on the strong gradient formulation (4.17) under detailed balance assumption, a drift-diffusion approximation, which shares the same energy landscape and same symmetric Hamiltonian structure, gives a good quadratic approximation near not only the 'downhill' solution to the macroscopic RRE but also the 'uphill' least action curve; see Sect. 5.2.

If a positive steady state $\pi(\vec{x}_v)$ to the mesoscopic master equation (1.5) exists, then one way to obtain ψ^{ss} is from the WKB expansion $\psi^{ss}(\vec{x}) = \lim_{V\to+\infty}\frac{-\log\pi_V(\vec{x}_V)}{V}$. Rigorously, under the assumption that there exists a positive detailed balanced $\pi_{\vec{x}_v}$, an USC viscosity solution to the stationary HJE (1.17) in the Barron-Jensen's sense [9] was constructed in [24] by using this invariant measure $\pi(\vec{x}_v)$. For the general case without detailed balance, the existence of viscosity solutions to HJE can also be obtained using the dynamic programming method. In [72, Theorem 2.2, Theorem 2.41], viscosity solutions are constructed via the minimization of the action functional $\inf_{T,\vec{x}(\cdot)} Act_{x_0,T}$ defined in Sect. 2.4. So we use Hamiltonian $H(\vec{p},\vec{x})$ to reformulate the transition path problem as a dual problem of the Maupertuis's principle of the least action problem (4.39). That is to say, we regard $\vec{p}$ as a control variable, then the least action problem is a constrained optimal control problem in a undefined time horizon (a.k.a infinite time horizon with an optimal terminal time [20]); see Sect. 5.1. Then the energy landscape ψ^{ss} is then represented as the unique weak KAM solution to HJE satisfying given boundary data on the projected Aubry set since the projected Aubry set is a uniqueness set for weak KAM solutions [25].

5.1 Existence of the Stationary Solution ψ^{ss} via Optimal Control and Viscosity Solution

Assume $\vec{x}^A$ and $\vec{x}^B$ are two steady states to RRE belonging to the same stoichiometric compatibility class such that $\vec{x}^A - \vec{x}^B \in G$. For the most probable path described by a least action problem with L defined in (2.29), recall the minimization problem (4.38) based on Maupertuis's principle. Without the explicit formula of L, using the Hamiltonian $H(\vec{p}, \vec{x})$, we first reformulate (4.38) as the following control problem. Regarding $\vec{p}$ as a control variable, we minimize the running cost described by the action functional in an undefined time horizon

$$v(\vec{y}; \vec{x}^A, c) = \inf_{T, \vec{p}} \int_0^T \left(\vec{p} \cdot \nabla_p H(\vec{p}, \vec{x}) - H(\vec{p}, \vec{x}) + c \right) dt,$$
$$\text{s.t. } \dot{\vec{x}} = \nabla_p H(\vec{p}, \vec{x}), \; t \in (0, T), \quad \vec{x}_0 = \vec{x}^A, \; \vec{x}_T = \vec{y}. \tag{5.1}$$

Here $v(\vec{y}; \vec{x}^A, c)$ is called the value function, $c \geq c_0$ is an energy level and c_0 is a critical minimum energy level such that

$$\inf_{T, \vec{x}(\cdot)} \int_0^T (L(\dot{\vec{x}}(t), \vec{x}(t)) + c_0) \, dt \geq 0 \tag{5.2}$$

and if $c < c_0$ this inf becomes $-\infty$. From the definition of the critical minimum energy level [11]

$$c_0 = \sup\{c \in \mathbb{R}; \, \exists \text{ closed curve } \vec{x}(\cdot) \text{ s.t. } \int_0^T (L(\dot{\vec{x}}(t), \vec{x}(t)) + c) \, dt < 0\}, \tag{5.3}$$

it is easy to see that for the Lagrangian L and Hamiltonian H in chemical reactions, the critical level $c_0 = 0$. Indeed, on the one hand, since $L \geq 0$ due to Lemma 2.5, so we know at least $c_0 \leq 0$. On the other hand, if $c_0 < 0$, then one can choose a standing curve $\vec{x}(t) \equiv y$ at a steady state y of RRE such that $\dot{\vec{x}} = \vec{R}(y) \equiv 0$. Then one have $L(\dot{\vec{x}}, \vec{x}) \equiv 0$ while $\int_0^T (L(\dot{\vec{x}}(t), \vec{x}(t)) + c_0) \, dt < 0$.

From [72, Theorem 2.39 and Theorem 2.47], we know for any $c \geq c_0$, the value function $v(y; \vec{x}^A, c)$ is a viscosity solution to the following static HJE

$$H(\nabla v(\vec{y}), \vec{y}) = c, \quad \forall \vec{y} \tag{5.4}$$

due to $\vec{x}^A$ is a steady state of RRE. In the chemical reactions, $c = 0 = c_0$, the above $v(y; \vec{x}^A) = v(y; \vec{x}^A, 0)$ is a viscosity solution to HJE $H(\nabla v(\vec{y}), \vec{y}) = 0$. However, these viscosity solutions are not unique.

Now we describe a selection principle via the weak KAM theory [34] and then the global energy landscape ψ^{ss} can be represented via the following weak KAM solution. Assume there are only finite steady solutions to RRE, denoted as the Aubry set $\mathcal{A} = \{\vec{x}_i^A\}_{i=1}^J$. Then

$$\psi^{ss}(\vec{x}) = \min_{\vec{x}^A \in \mathcal{A}} \left(\psi^{ss}(\vec{x}^A) + v(\vec{x}; \vec{x}^A) \right) \tag{5.5}$$

is the unique weak KAM solution to stationary HJE satisfying given boundary data on the projected Aubry set since the projected Aubry set is a uniqueness set for weak KAM solutions [25].

Remark 5.1 In [45], Lazarescu et al. used a biased Hamiltonian H with observations for the time-averaged flux and concentration to study the dynamic phase transitions in a long time limit. In an open system without detailed balance, with mixed boundary condition and

properly chosen bias for fluxes and concentrations, trajectories converging to a constructed global attractor was obtained in [45] while the optimality of the biased Hamiltonian dynamics in the optimal control context was still unclear. Beside the deterministic optimal control problem described above, one can also directly investigate the stochastic optimal control problem from the original large number process C^v with a fixed volume V. The transition path theory theory (TPT) was first proposed by E and Vanden-Eijnden in [76], particularly in [52] for Markov jumping process, to obtain transition paths and transition rates at a finite noise level by calculating the committor function, i.e., the stationary solution to the backward equation with two boundary conditions at two stable states A and B. In [26], an optimally controlled random walk is constructed based on the committor function, which realized Monte Carlo simulations for the transition path almost surely.

5.2 Construction of a Drift-Diffusion Process with the Same Energy Barrier for the Transition Path

We have shown the law of large numbers gives the macroscopic RRE however the transition path is in the large deviation regime. In this section, we construct a diffusion approximation, which can also be used to approximate the transition path. The most efficient way for constructing a diffusion approximation is through the Kramers–Moyal approximation for the master equation. We will show it is exactly equivalent to the quadratic approximation of the Hamiltonian near the solution to the macroscopic RRE. Then using the symmetric Hamiltonian, we give a new construction of diffusion approximation that shares the same energy barrier for transition paths.

Near the minimizer of $Act(\cdot)$, i.e., the curve solves RRE (1.6), we have the following quadratic approximation for the running cost. Denote $\vec{s}^* := \nabla_p H(\vec{p}, \vec{x})\big|_{\vec{p}=0} = \sum_{j=1}^{M} \vec{v}_j \left(\Phi_j^+(\vec{x}) - \Phi_j^-(\vec{x}) \right)$. Then we have

$$H(\vec{p}, \vec{x}) = \vec{s}^* \cdot \vec{p} + \frac{1}{2} \vec{p}^T \nabla_{pp}^2 H(0, \vec{x}) \vec{p} + o(|\vec{p}|^2) \tag{5.6}$$

and for $\vec{s} \in G$,

$$L(\vec{s}, \vec{x}) = \max_{\vec{p} \in G} (\vec{s} - \vec{s}^*) \cdot \vec{p} - \frac{1}{2} \vec{p}^T \nabla_{pp}^2 H(\vec{0}, \vec{x}) \vec{p} + o(|\vec{p}|^2). \tag{5.7}$$

Then approximately we have $\vec{s} - \vec{s}^* = \nabla_{pp}^2 H(0, \vec{x}) \vec{p}^*$ and

$$L(\vec{s}, \vec{x}) \approx \frac{1}{2} \vec{p}^{*T} \nabla_{pp}^2 H(\vec{0}, \vec{x}) \vec{p}^*. \tag{5.8}$$

One way of constructing a Langevin equation with the corresponding quadratic Hamiltonian (5.6) is

$$d\vec{x} = \nabla_p H(\vec{0}, \vec{x}) \, dt + \sqrt{\frac{1}{V} \nabla_{pp}^2 H(0, \vec{x})} \, dB. \tag{5.9}$$

Particularly, for our chemical reaction Hamiltonian, $\nabla_{pp}^2 H(0, \vec{x}) = \sum_j \left(\Phi_j^+(\vec{x}) + \Phi_j^-(\vec{x}) \right) \vec{v}_j \otimes \vec{v}_j$. The above equation is known as the chemical Langevin equation [23].

We now explain the above quadratic approximation exactly corresponds to the Kramers–Moyal approximation for the CME (1.5). The CME (1.5) can be regarded as a monotone

scheme for the RRE (1.6); see [24]. The leading Taylor expansion for (1.5) upto the second order yields a diffusion approximation

$$\partial_t p = -\nabla \cdot \left(p \sum_{j=1}^{M} \vec{v}_j \left(\Phi_j^+(\vec{x}) - \Phi_j^-(\vec{x}) \right) \right)$$
$$+ \frac{1}{2V} \sum_{j=1}^{M} \langle \nabla^2 \left(p(\Phi_j^+(\vec{x}) + \Phi_j^-(\vec{x})) \right) \vec{v}_j, \vec{v}_j \rangle + O\left(\frac{1}{V^2} \right). \quad (5.10)$$

This is known as the Kramers–Moyal expansion for the CME. This was also used as 'system size expansion' by van Kampen in [74] and in numerical analysis, it is also called a modified equation. The corresponding Hamiltonian of (5.10) via the WKB expansion is

$$H(\vec{p}, \vec{x}) = \sum_{j=1}^{M} \left(\Phi_j^+(\vec{x}) - \Phi_j^-(\vec{x}) \right) \vec{p} \cdot \vec{v}_j + \frac{1}{2} \sum_{j=1}^{M} \left(\Phi_j^+(\vec{x}) + \Phi_j^-(\vec{x}) \right) (\vec{p} \cdot \vec{v}_j)^2. \quad (5.11)$$

This is exactly the same as the quadratic approximation (5.6) of the original Hamiltonian at $\vec{p} = \vec{0}$.

However, as illustrated in the Schlögl catalysis model in Sect. 4.4, we point out the above quadratic approximation for the Hamiltonian works only for a region close to solutions to the 'downhill' macroscopic RRE. On the contrary, the 'uphill' transition path is apparently a rare transition path in the large deviation regime that is not closed to solutions to the RRE. [14] also quantified the failure of the simple diffusion approximation via the Kramers–Moyal expansion when studying the extinction problem for stochastic population model, which is also an exit problem in the large deviation regime.

Below, we follow the standard procedure for achieving the fluctuation–dissipation relation to construct a diffusion approximation such that (i) the diffusion model satisfies a fluctuation–dissipation relation and yields the same energy landscape as the original chemical reaction process; (ii) the corresponding quadratic Hamiltonian has the same symmetric property; (iii) the diffusion approximation valid near both the 'downhill' RRE solution and the 'uphill' most probable path.

Under symmetric Hamiltonian condition (1.26), recall the strong gradient flow in terms of energy landscape

$$\frac{d}{dt}\vec{x} = -K(\vec{x})\nabla \psi^{ss}(\vec{x}),$$

$K(\vec{x}) = \frac{1}{2} \int_0^1 \nabla_{pp}^2 H(\theta \nabla \psi^{ss}(\vec{x})) \, d\theta$. Following the standard technique for achieving the fluctuation–dissipation relation, we use the backward Ito's integral to construct a drift-diffusion process

$$d\vec{x} = K(\vec{x})\nabla \psi^{ss}(\vec{x}) \, dt + \sqrt{\frac{2}{V}K} \, \widehat{d}B, \quad (5.12)$$

where $\widehat{d}B$ means the multiplicative noise in the backward Ito's integral sense [42]. In the standard forward Ito's integral sense, this reads

$$d\vec{x} = -K(\vec{x})\nabla \psi^{ss}(\vec{x}) \, dt + \frac{1}{V}\nabla \cdot K \, dt + \sqrt{\frac{2}{V}K} \, dB. \quad (5.13)$$

Then the Fokker-Planck equation is

$$\frac{\partial \rho}{\partial t} = \frac{1}{V} \nabla \cdot \left(e^{-V\psi^{ss}} K \nabla \left(\rho e^{V\psi^{ss}} \right) \right).$$

(i) This equation has an invariant measure $\pi = e^{-V\psi^{ss}}$; (ii) The invariant measure satisfies the detailed balance condition, since the Fokker-Planck operator $\nabla \cdot \left(\pi K \nabla \left(\frac{\rho}{\pi} \right) \right)$ is self-adjoint in $L^2(\pi^{-1})$; (iii) For any convex function $\phi(x)$, the dissipation relation holds

$$\frac{\mathrm{d}}{\mathrm{d}t} \int \pi \phi \left(\frac{\rho}{\pi} \right) \mathrm{d}x = \langle \phi' \left(\frac{\rho}{\pi} \right), \partial_t \rho \rangle = -\frac{1}{V} \langle K \nabla \frac{\rho}{\pi}, \phi'' \left(\frac{\rho}{\pi} \right) \nabla \frac{\rho}{\pi} \rangle \leq 0. \quad (5.14)$$

Here the invariant measure yields exactly the same energy landscape as the original chemical large number process $\psi^{ss} = -\frac{\log \pi}{V}$. The corresponding quadratic Hamiltonian is symmetric w.r.t. $\nabla \psi^{ss}(\vec{x})$

$$H(\vec{p}, \vec{x}) = \left(\vec{p} - \nabla \psi^{ss}(\vec{x}) \right) \cdot K \vec{p} = H(\nabla \psi^{ss}(\vec{x}) - \vec{p}, \vec{x}). \quad (5.15)$$

We point out this diffusion approximation (5.12) has a covariance $\int_0^1 \nabla_{pp}^2 H(\theta \nabla \psi^{ss}(\vec{x}), \vec{x}) \mathrm{d}\theta$ but the diffusion approximation using chemical Langevin equation (5.9) has a different covariance $\nabla_{pp}^2 H(0, \vec{x})$ is different from the previous diffusion approximation (5.9) near RRE at the central limit regime. We remark the diffusion approximation (5.12) satisfying fluctuation–dissipation relation is also used in [22, 29, 56] to study the stochastic uncertainty relation for a general process.

6 Discussion

In this paper, we revisit the macroscopic dynamics for some non-equilibrium chemical reactions from a Hamiltonian viewpoint. The concentration of chemical species is modeled by the nonlinear RRE system, which is the thermodynamic limiting equation from the law of large numbers for the random time-changed Poisson representation of chemical reactions. The Hamiltonian defined from the WKB expansion determines a HJE, and the minimizer of the dynamic solution recovers the solution to the RRE. The stationary solution ψ^{ss} to HJE serves as the energy landscape for general non-equilibrium reactions. The existence of ψ^{ss} is represented as an optimal control problem in an undefined time horizon, which can be represented as a weak KAM solution to HJE. More importantly, we use ψ^{ss} to decompose RRE into a conservative part and dissipative part, which, together with the additional mass conservation law, gives raise a GENERIC formalism for RRE. Through ψ^{ss}, the thermodynamics for non-equilibrium reactions can also be decomposed as nonadiabatic and adiabatic parts, where the later one maintains a positive entropy production rate at NESS.

We then study the energy dissipation relation at both mesoscopic and macroscopic levels and prove the passage from the mesoscopic one to the other. A non-convex energy landscape ψ^{ss} emerges from the convex mesoscopic relative entropy functional $\mathrm{KL}(\rho_v || \pi_v)$ in the large number limit, which picks up the non-equilibrium features. This mean-field limit passage also applies to the symmetric property in a chemical reaction. Particularly, the mesoscopic Markov chain detailed balance leads to a symmetric Hamiltonian, while the markov chain detailed balance is not equivalent to the more constrained chemical version of detailed balance. The non-convexity of the macroscopic energy landscape ψ^{ss}, naturally brought by a grouped polynomial probability flux, enables us to study a class of non-equilibrium chemical reaction with multiple steady states, for instance the bistable Schlögl model. However, we point

out multiple steady states and non-convex energy are also common in other equilibrium statistical physics such as the Lagenvin dynamics with non-convex potential and Ising model of ferromagnetism.

We then focus on finding transition paths between coexistent stable steady states in some non-equilibrium biochemical reactions using a symmetric Hamiltonian w.r.t. the stationary solution $\nabla \psi^{ss}$. Under this symmetric condition, the transition path is explicitly given by piecewise least action curves, where the 'uphill' curve is a $\nabla \psi^{ss}$-modified time reversal of the 'downhill' least action curve, where ψ^{ss} also gives the energy barriers and path affinities.

The bistability and bifurcation in Schlögl's model appear in many general forms, such as the Stuart–Landau equation for general sustained nonlinear oscillating system with application for the Belousov–Zhabotinsky reaction. When including spatial variation in the reaction–diffusion equation for spontaneous spatial pattern formation, the double well bistability generates the Turing pattern while the Fisher-KPP bistability generates traveling waves. We also study a quadratic approximation for the Hamiltonian near the RRE solution, i.e., the mean path in the sense of the law of large numbers. However, we point out the transition path problem in chemical reaction is in the large deviation regime and the associated energy barrier can not be computed by a simple quadratic approximation. Instead, based on the strong form of gradient flow in terms of free energy ψ^{ss}, we construct anther drift–diffusion approximation which shares the same symmetric Hamiltonian and energy barrier for the most probable path connecting two non-equilibrium steady states.

Data Availibility All data generated or analysed during this study are included in this published article.

Appendix A: Master Equation and Generator

A.1 Master Equation Derivation

We will only compute the generator for the portion of the forward reactions with the forward Poisson process $Y_j = Y_j^+$ in (1.2), because the backward portion is exactly same. Consider

$$\vec{X}(t) = \vec{X}(0) + \sum_{j=1}^{M} \vec{v}_j \mathbb{1}_{\{\vec{X}(t_-)+\vec{v}_j \geq 0\}} Y_j^+ \left(\int_0^t \varphi_j^+(X(s)) \, ds \right) \tag{A.1}$$

For any test function $f \in C_b$, since $R_j^+(t) := Y_j^+ \left(\int_0^t \varphi_j^+(\vec{X}(s)) \, ds \right)$ is a counting process representing the j-th reaction, so

$$f(\vec{X}(t)) = f(\vec{X}(0)) + \sum_{j=1}^{M} \int_0^t \mathbb{1}_{\{\vec{X}(s_-)+\vec{v}_j \geq 0\}} \left(f(\vec{X}(s_-) + \vec{v}_j) - f(\vec{X}(s_-)) \right) \, dR_j^+(s).$$

$$\tag{A.2}$$

From [3, Thm 1.10], $M_j^+(t) := Y_j^+ \left(\int_0^t \varphi_j^+(\vec{X}(s)) \, ds \right) - \int_0^t \varphi_j^+(\vec{X}(s)) \, ds$ is a Martingale. Thus (A.2) becomes

$$
\begin{aligned}
f(\vec{X}(t)) = & f(\vec{X}(0)) + \sum_{j=1}^M \int_0^t \mathbb{1}_{\{\vec{X}(s)+\vec{v}_j \geq 0\}} \varphi_j^+(\vec{X}(s)) \left(f(\vec{X}(s)+\vec{v}_j) - f(\vec{X}(s)) \right) \, ds \\
& + \sum_{j=1}^M \int_0^t \mathbb{1}_{\{\vec{X}(s_-)+\vec{v}_j \geq 0\}} \left(f(\vec{X}(s_-)+\vec{v}_j) - f(\vec{X}(s_-)) \right) \, dM_j^+(s).
\end{aligned}
\tag{A.3}
$$

Now we derive the master equation for $\vec{X}(t) \in \mathbb{N}^N$. Denote the (time marginal) law of $\vec{X}(t)$ as

$$
p(\vec{n}, t) = \mathbb{E} \left(\mathbb{1}_{\vec{X}(t)}(\vec{n}) \right),
\tag{A.4}
$$

where $\mathbb{1}$ is the indicator function. For any $f : \mathbb{Z}^N \to \mathbb{R}$, $f(\vec{X}) = \sum_{\vec{n}} f(\vec{n}) \mathbb{1}_{\vec{X}}(\vec{n})$, and

$$
\mathbb{E}(f(\vec{X})) = \sum_{\vec{n}} f(\vec{n}) \mathbb{E}(\mathbb{1}_{\vec{X}}(\vec{n})) = \sum_{\vec{n}} f(\vec{n}) p(\vec{n}, t).
\tag{A.5}
$$

Taking expectation for (A.3), we have the Dynkin's formula

$$
\mathbb{E} f(\vec{X}(t)) = \mathbb{E} f(\vec{X}(0)) + \sum_{j=1}^M \int_0^t \mathbb{E} \left(\mathbb{1}_{\{\vec{X}(s)+\vec{v}_j \geq 0\}} \varphi_j^+(\vec{X}(s)) \left(f(\vec{X}(s)+\vec{v}_j) - f(\vec{X}(s)) \right) \right) \, ds.
\tag{A.6}
$$

Taking derivative yields

$$
\begin{aligned}
\frac{d}{dt} \sum_{\vec{n}} f(\vec{n}) p(\vec{n}, t) = & \sum_{j=1}^M \sum_{\vec{n} \geq 0, \, \vec{n}+\vec{v}_j \geq 0} \varphi_j^+(\vec{n}) \left(f(\vec{n}+\vec{v}_j) - f(\vec{n}) \right) p(\vec{n}, t) \\
= & \sum_{j=1}^M \sum_{\vec{n} \geq 0, \, \vec{n}+\vec{v}_j \geq 0} \varphi_j^+(\vec{n}) f(\vec{n}+\vec{v}_j) p(\vec{n}, t) \\
& - \sum_{j=1}^M \sum_{\vec{n} \geq 0, \, \vec{n}+\vec{v}_j \geq 0} \varphi_j^+(\vec{n}) f(\vec{n}) p(\vec{n}, t) \\
= & \sum_{j=1}^M \sum_{\vec{n} \geq 0, \, \vec{n}-\vec{v}_j \geq 0} \varphi_j^+(\vec{n}-\vec{v}_j) f(\vec{n}) p(\vec{n}-\vec{v}_j, t) \\
& - \sum_{j=1}^M \sum_{\vec{n} \geq 0, \, \vec{n}+\vec{v}_j \geq 0} \varphi_j^+(\vec{n}) f(\vec{n}) p(\vec{n}, t) \\
= & \sum_{\vec{n} \geq 0} f(\vec{n}) \left(\sum_{j=1, \, \vec{n}-\vec{v}_j \geq 0}^M \varphi_j^+(\vec{n}-\vec{v}_j) p(\vec{n}-\vec{v}_j, t) \right. \\
& \left. - \sum_{j=1, \, \vec{n}+\vec{v}_j \geq 0}^M \varphi_j^+(\vec{n}) p(\vec{n}, t) \right).
\end{aligned}
\tag{A.7}
$$

Then the master equation for $p(\vec{n}, t)$ is

$$\frac{d}{dt} p(\vec{n}, t) = \sum_{j=1, \vec{n}-\vec{v}_j \geq 0}^{M} \varphi_j^+(\vec{n}-\vec{v}_j) p(\vec{n}-\vec{v}_j, t) - \sum_{j=1, \vec{n}+\vec{v}_j \geq 0}^{M} \varphi_j^+(\vec{n}) p(\vec{n}, t). \quad \text{(A.8)}$$

After including the backward reactions with Y^-,

$$\frac{d}{dt} p(\vec{n}, t) = \sum_{j=1, \vec{n}-\vec{v}_j \geq 0}^{M} \varphi_j^+(\vec{n}-\vec{v}_j) p(\vec{n}-\vec{v}_j, t) - \sum_{j=1, \vec{n}+\vec{v}_j \geq 0}^{M} \varphi_j^+(\vec{n}) p(\vec{n}, t)$$
$$+ \sum_{j=1, \vec{n}+\vec{v}_j \geq 0}^{M} \varphi_j^-(\vec{n}+\vec{v}_j) p(\vec{n}+\vec{v}_j, t) - \sum_{j=1, \vec{n}-\vec{v}_j \geq 0}^{M} \varphi_j^+(\vec{n}) p(\vec{n}, t) \quad \text{(A.9)}$$

Therefore, for the chemical reaction described by (1.2), the master equation is

$$\frac{d}{dt} p(\vec{n}, t) = \sum_{j=1, \vec{n}-\vec{v}_j \geq 0}^{M} \left(\varphi_j^+(\vec{n}-\vec{v}_j) p(\vec{n}-\vec{v}_j, t) - \varphi_j^-(\vec{n}) p(\vec{n}, t) \right)$$
$$+ \sum_{j=1, \vec{n}+\vec{v}_j \geq 0}^{M} \left(\varphi_j^-(\vec{n}+\vec{v}_j) p(\vec{n}+\vec{v}_j, t) - \varphi_j^+(\vec{n}) p(\vec{n}, t) \right). \quad \text{(A.10)}$$

Similarly, one can derive the master equation for the rescaled large number jumping process $C^v(t)$.

We only compute the generator for the portion of the forward reactions. Notice $R_j^+(t) = Y_j \left(V \int_0^t \tilde{\Phi}_j^+(C^v(s)) \, ds \right)$ is a counting process and $M_j^+(t) = \frac{1}{V} Y_j^+ \left(V \int_0^t \tilde{\Phi}_j^+(C^v(s)) \, ds \right) - \int_0^t \tilde{\Phi}_j^+(C^v(s)) \, ds$ is a martingale. Similar to (A.3), we obtain for any $f \in C_b$,

$$f(C^v(t)) = f(C^v(0)) + \sum_{j=1}^{M} \int_0^t \mathbb{1}_{\{C^v(s_-)+\frac{\vec{v}_j}{V} \geq 0\}} \left(f(C^v(s_-) + \frac{\vec{v}_j}{V}) - f(C^v(s_-)) \right) dR_j^+(s)$$
$$= f(C^v(0)) + \sum_{j=1}^{M} \int_0^t V \mathbb{1}_{\{C^v(s)+\frac{\vec{v}_j}{V} \geq 0\}} \tilde{\Phi}_j(C^v(s)) \left(f(C^v(s) + \frac{\vec{v}_j}{V}) - f(C^v(s)) \right) ds$$
$$+ \sum_{j=1}^{M} \int_0^t V \mathbb{1}_{\{C^v(s_-)+\frac{\vec{v}_j}{V} \geq 0\}} \left(f(C^v(s_-) + \frac{\vec{v}_j}{V}) - f(C^v(s_-)) \right) dM_j^+(s). \quad \text{(A.11)}$$

Then using $\mathbb{E} f(C^v(t)) = \frac{1}{V} \sum_{\vec{x}_v} f(\vec{x}_v) p(\vec{x}_v, t)$, we obtain the generator Q_v for the large number process $C^v(t)$ for fixed V

$$\frac{d}{dt}\sum_{\vec{x}_v} f(\vec{x}_v)p(\vec{x}_v,t) = V\sum_{\vec{x}_v \geq 0}\left[\sum_{j=1,\,\vec{x}_v+\frac{\vec{v}_j}{V} \geq 0}^{M} \tilde{\Phi}_j^+(\vec{x}_v)\left(f(\vec{x}_v+\frac{\vec{v}_j}{V})-f(\vec{x}_v)\right)p(\vec{x}_v,t)\right.$$

$$+\sum_{j=1,\,\vec{x}_v-\frac{\vec{v}_j}{V} \geq 0}^{M} \left.\tilde{\Phi}_j^-(\vec{x}_v)\left(f(\vec{x}_v-\frac{\vec{v}_j}{V})-f(\vec{x}_v)\right)p(\vec{x}_v,t)\right] \tag{A.12}$$

$$=:\sum_{\vec{x}_v}(Q_v f)(\vec{x}_v)p(\vec{x}_v,t).$$

Here in the definition of generator, one can define a zero extension for the region outside $\vec{x}_v \geq 0$.

Appendix B: Mean-Field Limit RRE for C^v

Since the original proof for the mean filed equation of chemical reaction C^v in [3, 43] omitted the 'no reaction' constraint outside nonnegative region, so we provide a pedagogical proof after including the constraint $\vec{x}_v \pm \frac{\vec{v}_j}{V} \geq 0$.

Assume there exists a solution $\vec{x}_v(\cdot) \in C^1([0,T];\mathbb{R}_+^N)$ to RRE (1.6) and $x_i(t) > 0$ for all $t \in [0,T]$ and each component i. Recall $\vec{R}(\vec{x}) = \sum_{j=1}^{M}\vec{v}_j\left(\Phi_j^+(\vec{x})-\Phi_j^-(\vec{x})\right)$ defined in (2.2).

Fix $a_0 > 0$ such that the RRE solution tube $\Omega_{a_0} := \{\vec{y};\ \max_{t\in[0,T]}|\vec{y}-\vec{x}(t)| < a_0\} \subset \mathbb{R}_+^N$. For any $0 < a < a_0$, since $\vec{R}$ is locally Lipschitz, there exists K_a such that $|\vec{R}(\vec{x})-\vec{R}(\vec{y})| \leq K_a|\vec{x}-\vec{y}|$ for $\vec{x},\vec{y} \in \Omega_a$. Then we define a stopping time

$$\tau_{v,a} = \inf\{t;\ |C^v(t)-\vec{x}(t)| > a\}. \tag{B.1}$$

Then for $t \leq \tau_{v,a}$, $C^v \pm \frac{\vec{v}_j}{V} \subset \Omega_{a_0} \subset \mathbb{R}_+^N$ for V large enough, so the 'no reaction' constraint in process (1.4) does not turn on before $\tau_{v,a}$. Thus from the martingale decomposition (A.11), by Doob's continuous time optional stopping lemma,

$$C^v(t \wedge \tau_{v,a}) = C^v(0)-\vec{x}(0)+M_v(t \wedge \tau_{v,a})+\int_0^{t\wedge\tau_{v,a}}\vec{R}(C^v(s))\,ds, \tag{B.2}$$

where

$$M_v(t) = \sum_j \frac{\vec{v}_j}{V}\left(Y_j^+\left(V\int_0^t \Phi_j^+(C^v(s))\,ds\right)+Y_j^-\left(V\int_0^t \Phi_j^-(C^v(s))\,ds\right)\right)$$

$$-\vec{v}_j\int_0^t \Phi_j^+(C^v(s))+\Phi_j^-(C^v(s))\,ds \tag{B.3}$$

is a martingale. Here for simplicity, we assume the mesoscopic and macroscopic LMA are same $\tilde{\Phi}_j^\pm(\vec{x}) = \Phi_j^\pm(\vec{x})$ and then drop tilde. Compare the trajectory of SDE (1.4) with solution to RRE (1.6)

$$C^v(t \wedge \tau_{v,a})-\vec{x}(t \wedge \tau_{v,a}) = C^v(0)-\vec{x}(0)+M_v(t \wedge \tau_{v,a})$$

$$+\int_0^{t\wedge\tau_{v,a}}[\vec{R}(C^v(s))-\vec{R}(\vec{x}(s))]\,ds. \tag{B.4}$$

Thus

$$
\begin{aligned}
|C^{\mathrm{v}}(t \wedge \tau_{\mathrm{v},a}) - \vec{x}(t \wedge \tau_{\mathrm{v},a})| \leq &|C^{\mathrm{v}}(0) - \vec{x}(0)| \\
&+ |M_{\mathrm{v}}(t \wedge \tau_{\mathrm{v},a})| + K_a \int_0^{t \wedge \tau_{\mathrm{v},a}} |C^{\mathrm{v}}(s) - \vec{x}(s)| \, \mathrm{d}s|.
\end{aligned} \tag{B.5}
$$

Then by Gronwall's inequality, we have

$$
|C^{\mathrm{v}}(t \wedge \tau_{\mathrm{v},a}) - \vec{x}(t \wedge \tau_{\mathrm{v},a})| \leq \left(|C^{\mathrm{v}}(0) - \vec{x}(0) + \sup_{1 \leq s \leq t \wedge \tau_{\mathrm{v},a}} |M_{\mathrm{v}}(s)| \right) e^{K_a t \wedge \tau_{\mathrm{v},a}}. \tag{B.6}
$$

Notice that the process C^{v} is right continuous, then by definition of $\tau_{\mathrm{v},a}$, we have

$$
\{ \max_{0 \leq s \leq t} |C^{\mathrm{v}}(s) - \vec{x}(s)| > a \} \subset \{ |C^{\mathrm{v}}(t \wedge \tau_{\mathrm{v},a}) - \vec{x}(t \wedge \tau_{\mathrm{v},a})| \geq a \}. \tag{B.7}
$$

Then by (B.6),

$$
\begin{aligned}
\{ \max_{0 \leq s \leq t} |C^{\mathrm{v}}(s) - \vec{x}(s)| > a \} &\subset \{ |C^{\mathrm{v}}(0) - \vec{x}(0) + \sup_{1 \leq s \leq t \wedge \tau_{\mathrm{v},a}} |M_{\mathrm{v}}(s)| \geq a e^{-K_a t} \} \\
&\subset \{ |C^{\mathrm{v}}(0) - \vec{x}(0)| \geq \frac{a}{2} e^{-K_a t} \} \cup \{ \sup_{1 \leq s \leq t \wedge \tau_{\mathrm{v},a}} |M_{\mathrm{v}}(s)|^2 \geq \frac{a^2}{4} e^{-2K_a t} \}.
\end{aligned} \tag{B.8}
$$

Then by Doob's maximal inequality for submartingales, we know

$$
\mathbb{P}\{ \sup_{1 \leq s \leq t \wedge \tau_{\mathrm{v},a}} |M_{\mathrm{v}}(s)|^2 \geq \frac{a^2}{4} e^{-2K_a t} \} \leq \frac{4 e^{2K_a t}}{a^2} \mathbb{E}(|M_{\mathrm{v}}(t \wedge \tau_{\mathrm{v},a})|^2) \tag{B.9}
$$

Using the estimate of martingale M_{v}

$$
\mathbb{E}(|M_{\mathrm{v}}(t \wedge \tau_{\mathrm{v},a})|^2) = \sum_{j=1}^M \frac{|\vec{v}_j|^2}{V^2} \mathbb{E}\left(V \int_0^{t \wedge \tau_{\mathrm{v},a}} [\Phi_j^+(C^{\mathrm{v}}(s)) + \Phi_j^-(C^{\mathrm{v}}(s))] \, \mathrm{d}s \right),
$$

we know

$$
\begin{aligned}
\mathbb{P}\{ \max_{0 \leq s \leq t} |C^{\mathrm{v}}(s) - \vec{x}(s)| > a \} &\leq \mathbb{P}\{ |C^{\mathrm{v}}(0) - \vec{x}(0)| \geq \frac{a}{2} e^{-K_a t} \} + \frac{4 e^{2K_a t}}{a^2} \mathbb{E}(|M_{\mathrm{v}}(t \wedge \tau_{\mathrm{v},a})|^2) \\
&\leq \mathbb{P}\{ |C^{\mathrm{v}}(0) - \vec{x}(0)| \geq \frac{a}{2} e^{-K_a t} \} + C_a \frac{1}{V}.
\end{aligned} \tag{B.10}
$$

Then for arbitrarily small a, we conclude that if $C^{\mathrm{v}}(0) \to \vec{x}(0)$, (1.8) holds, i.e.,

$$
\lim_{V \to +\infty} \mathbb{P}\{ \max_{0 \leq s \leq t} |C^{\mathrm{v}}(s) - \vec{x}(s)| > a \} = 0. \tag{B.11}
$$

Appendix C: Proof of Lemma 3.6

We give the proof of Lemma 3.6 by some elementary computations and collecting existing results.

Proof of Lemma 3.6 Step 1 We prove the equivalence between (i) and (ii).

Plugging identity (2.12), we obtain the identity

$$
\begin{aligned}
H(\log \tfrac{\vec{x}}{\vec{x}^s}, \vec{x}) &= \sum_j \left(\Phi_j^-(\vec{x}) \left(\frac{\Phi_j^+(\vec{x}^s)}{\Phi_j^-(\vec{x}^s)} - 1 \right) + \Phi_j^+(\vec{x}) \left(\frac{\Phi_j^-(\vec{x}^s)}{\Phi_j^+(\vec{x}^s)} - 1 \right) \right) \\
&= \sum_j \left(\frac{\Phi_j^-(\vec{x})}{\Phi_j^-(\vec{x}^s)} \left(\Phi_j^+(\vec{x}^s) - \Phi_j^-(\vec{x}^s) \right) + \frac{\Phi_j^+(\vec{x})}{\Phi_j^+(\vec{x}^s)} \left(\Phi_j^-(\vec{x}^s) - \Phi_j^+(\vec{x}^s) \right) \right) \quad \text{(C.1)} \\
&= \sum_j \left(\frac{\vec{x}}{\vec{x}^s} \right)^{\vec{v}_j^-} \left(\Phi_j^+(\vec{x}^s) - \Phi_j^-(\vec{x}^s) \right) + \sum_j \left(\frac{\vec{x}}{\vec{x}^s} \right)^{\vec{v}_j^+} \left(\Phi_j^-(\vec{x}^s) - \Phi_j^+(\vec{x}^s) \right).
\end{aligned}
$$

Similar to (2.10), rearranging according to the reactant complex $\vec{\eta} \in \mathcal{C}$, we have

$$
H(\log \tfrac{\vec{x}}{\vec{x}^s}, \vec{x}) = \sum_{\vec{\eta} \in \mathcal{C}} \left(\frac{\vec{x}}{\vec{x}^s} \right)^{\vec{\eta}} \left(\sum_{j:\vec{v}_j^- = \vec{\eta}} \left(\Phi_j^+(\vec{x}^s) - \Phi_j^-(\vec{x}^s) \right) + \sum_{j:\vec{v}_j^+ = \vec{\eta}} \left(\Phi_j^-(\vec{x}^s) - \Phi_j^+(\vec{x}^s) \right) \right) \text{(C.2)}
$$

Then complex balance (2.6) is equivalent to $H(\log \tfrac{\vec{x}}{\vec{x}^s}, \vec{x}) = 0$.

Step 2, assume (i), i.e., $\vec{x}^s$ satisfies complex balance condition (2.6), then one can construct a stationary distribution π_V via the product of Poisson distributions with intensity $V\vec{x}^s$ [3, Theorem 3.7] (see also [1])

$$
\log \pi_V(\vec{x}_v) = \sum_{i=1}^N \left(n_i \log(V x_i^s) - \log(n_i!) - V x_i^s \right), \quad \vec{n} := V\vec{x}_v \qquad \text{(C.3)}
$$

for the chemical master equation (1.5) with $\tilde{\Phi} = \varphi/V$ for a fixed volume V. Thus (i) implies (iii).

Step 3, assume (iii), for any $\vec{x} \in \mathbb{R}_+^N$, let $\vec{x}_V = \frac{\vec{n}}{V} \to \vec{x}$ as $V \to +\infty$, then the limit in WKB approximation for π_V exists

$$
\lim_{V \to +\infty} \frac{-\log \pi_V(\vec{x}_V)}{V} = \sum_{i=1}^N \left(x_i \log x_i - x_i \log x_i^s + x_i^s - x_i \right) = \mathrm{KL}(\vec{x}||\vec{x}^s) = \psi^{ss}(\vec{x}).
$$
$$\text{(C.4)}$$

Indeed, changing to variable $\vec{x}_V = \frac{\vec{n}}{V}$ and using the Stirling's formula, we have

$$
\begin{aligned}
\frac{\log \pi_V(\vec{x}_V)}{V} &= \frac{1}{V} \sum_{i=1}^N \left(n_i \log(V x_i^s) - n_i \log(n_i) + n_i - V x_i^s + O(\log n_i) \right) \\
&= \sum_{i=1}^N \left(x_i \log x_i^s - x_i \log x_i + x_i - x_i^s \right) + \frac{O(\sum_i^N \log n_i)}{V}.
\end{aligned}
\qquad \text{(C.5)}
$$
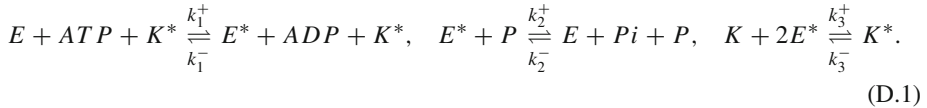
Then for any fixed $\vec{x} \in \mathbb{R}^N$, $\vec{x}_V = \frac{\vec{n}}{V} \to \vec{x}$ implies $\frac{O(\sum_i^N \log n_i)}{V} \to 0$ as $V \to +\infty$. Thus (ii) follows. □

Appendix D: Phosphorylation–Dephosphorylation with 2-Autocatalysis

The Schlögl model can be regarded as a simple but representative example which keeps the main features of non-equilibrium enzyme reactions. As the one of the most important enzyme

reaction in a single living cell, the phosphorylation–dephosphorylation reaction system (c.f. [63]) also fits into the symmetric Hamiltonian framework.

Here we briefly revisit the phosphorylation-dephosphorylation with 2-autocatalysis proposed by Fischer-Krebs in 1950s.

$$E + ATP + K^* \underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} E^* + ADP + K^*, \quad E^* + P \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} E + Pi + P, \quad K + 2E^* \underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} K^*.$$

$$(D.1)$$

Here the concentration of a protein in its open state (the phosphorylated E) is denoted as $x(t) = [E^*]$ while the concentration of a protein in its close state is $y(t) = [E]$. The third reaction equation representing the reversible binding is rapid and thus is assumed to be quasi-static. Under this quasi-static assumption, the active kinase K^* in the first reaction equation has a positive feedback from $2E^*$, which is known as 2-autocatalysis. We also regard the concentrations of the inactive kinase K, phosphatase P, adenosine triphosphate ATP, adenosine diphosphate ADP and phosphate group Pi as constant that sustained by environment.

From $\dot{y} = -\dot{x}$, we know the conservation of total mass of two proteins and thus $y(t) = [E_{tot}] - x(t)$. The RRE is given by

$$\dot{x} = (a_1^+[K]x^2y + a_2^- y) - (a_1^-[K]x^3 + a_2^+ x) =: [\Phi_1^+(x) + \Phi_2^+(x)] - [\Phi_1^-(x) + \Phi_2^-(x)]$$

$$(D.2)$$

where we lumped chemostats into rates and used quasi-static relation:

$$a_1^+ = k_1^+[ATP]\frac{k_3^+}{k_3^-}, \quad a_1^- = k_1^-[ADP]\frac{k_3^+}{k_3^-}, \quad a_2^+ = k^- + 2[P], \quad a_2^- = k_2^-[Pi][P].$$

We take [K] as a bifurcation parameter for the first order phase transition. The right-hand-side of (D.2) is a double well potential raising from the flux grouping property. This 2-autocatalysis model is basically same as the Schlögl model after effectively eliminating the quasi-static third reaction equation, so the mathematical analysis are same.

References

1. Anderson, D.F., Craciun, G., Kurtz, T.G.: Product-form stationary distributions for deficiency zero chemical reaction networks. Bull. Math. Biol. **72**(8), 1947–1970 (2010)
2. Agazzi, A., Dembo, A., Eckmann, J.-P.: Large deviations theory for Markov jump models of chemical reaction networks. Ann. Appl. Probab. **28**(3), 1821–1855 (2018)
3. Anderson, D.F., Kurtz, T.G.: Stochastic Analysis of Biochemical Systems. Springer, New York (2015)
4. Assaf, M., Meerson, B.: Wkb theory of large deviations in stochastic populations. J. Phys. A **50**(26), 263001 (2017)
5. Anderson, D.F.: Global asymptotic stability for a class of nonlinear chemical equations. SIAM J. Appl. Math. **68**(5), 1464–1476 (2008)
6. Bertini, L., De Sole, A., Gabrielli, D., Jona-Lasinio, G., Landim, C.: Macroscopic fluctuation theory for stationary non-equilibrium states. J. Stat. Phys. **107**(3), 635–675 (2002)
7. Bertini, L., De Sole, A., Gabrielli, D., Jona-Lasinio, G., Landim, C.: Macroscopic fluctuation theory. Rev. Mod. Phys. **87**(2), 593–636 (2015)
8. Bierman, A.: Studies on the effects of structure on the behavior of enzyme systems. Bull. Math. Biophys. **16**(3), 203–257 (1954)
9. Barron, E.N., Jensen, R.: Semicontinuous viscosity solutions for Hamilton-Jacobi equations with convex hamiltonians. Commun. Partial Differ. Equ. **15**(12), 293–309 (1990)
10. Bryc, W.: Large Deviations by the Asymptotic Value Method, pp. 447–472. Birkhäuser, Boston (1990)

11. Contreras, G., Iturriaga R.: Global minimizers of autonomous lagrangians. IMPA Rio de Janeiro, 22nd Brazilian Mathematics Colloquium (1999)

12. Dykman, M.I., Mori, E., Ross, J., Hunt, P.M.: Large fluctuations and optimal paths in chemical kinetics. J. Chem. Phys. **100**(8), 5735–5750 (1994)

13. Doi, M.: Second quantization representation for classical many-particle system. J. Phys. A **9**(9), 1465–1477 (1976)

14. Doering, C.R., Sargsyan, K.V., Sander, L.M.: Extinction times for birth-death processes: exact results, continuum asymptotics, and the failure of the Fokker-Planck approximation. Multiscale Model. Simul. **3**(2), 283–299 (2005)

15. Evans, D.J., Searles, D.J.: The fluctuation theorem. Adv. Phys. **51**(7), 1529–1585 (2002)

16. Feinberg, M.: On chemical kinetics of a certain class. Arch. Ration. Mech. Anal. **46**(1), 1–41 (1972)

17. Feinberg, M.: Foundations of Chemical Reaction Network Theory of Applied Mathematical Sciences, vol. 202. Springer, New York (2019)

18. Feng, J., Kurtz, T.G.: Large Deviations for Stochastic Processes, vol. 131. American Mathematical Soc., Providence (2006)

19. Fleming, W.H.: Optimal control of Markov processes. Technical report, Brown Univ Providence RI Lefschetz Center for Dynamical Systems (1983)

20. Fleming, W.H., Soner, H.M.: Controlled Markov Processes and Viscosity Solutions, 2nd edn. Springer, New York (2006)

21. Freidlin, M.I., Wentzell, A.D.: Random Perturbations of Dynamical Systems. Grundlehren der mathematischen Wissenschaften, vol. 260. Springer, Berlin (2012)

22. Gingrich, T.R., Horowitz, J.M., Perunov, N., England, J.L.: Dissipation bounds all steady-state current fluctuations. Phys. Rev. Lett. **116**(12), 120601 (2016)

23. Gillespie, D.T.: The chemical Langevin equation. J. Chem. Phys. **113**(1), 297–306 (2000)

24. Gao, Y., Liu, J.-G.: Thermodynamic limit of chemical master equation via nonlinear semigroup. arXiv:2205.09313 (2022)

25. Gao, Y., Liu, J.-G.: A selection principle for weak KAM solutions via Freidlin-Wentzell large deviation principle of invariant measures. arXiv:2208.11860 (2022)

26. Gao, Y., Li, T., Li, X., Liu, J.-G.: Transition path theory for langevin dynamics on manifold: optimal control and data-driven solver, to appear in Multiscale Model. Simul. (2022)

27. Ge, H., Qian, H.: Mathematical formalism of nonequilibrium thermodynamics for nonlinear chemical reaction systems with general rate law. J. Stat. Phys. **166**(1), 190–209 (2017)

28. Gauckler, L., Yserentant, H.: Regularity and approximability of the solutions to the chemical master equation. ESAIM: Math. Modell. Numer. Anal. Modélisation Mathématique et Analyse Numérique **48**(6), 1757–1775 (2014)

29. Horowitz, J.M., Gingrich, T.R.: Proof of the finite-time thermodynamic uncertainty relation for steady-state currents. Phys. Rev. E **96**(2), 020103 (2017)

30. Hanggi, P., Grabert, H., Talkner, P., Thomas, H.: Bistable systems: master equation versus Fokker-Planck modeling. Phys. Rev. A **29**(1), 371–378 (1984)

31. Horn, F., Jackson, R.: General mass action kinetics. Arch. Ration. Mech. Anal. **47**, 81–116 (1972)

32. Hairer, M., Majda, A.J.: A simple framework to justify linear response theory. Nonlinearity **23**(4), 909–922 (2010). arXiv: 0909.4313

33. Hu, G.: Stationary solution of master equations in the large-system-size limit. Phys. Rev. A **36**(12), 5782–5790 (1987)

34. Ishii, H., Siconolfi, A.: The vanishing discount problem for Hamilton-Jacobi equations in the Euclidean space. Commun. Partial Differ. Equ. **45**(6), 525–560 (2020)

35. Joshi, B.: A detailed balanced reaction network is sufficient but not necessary for its Markov chain to be detailed balanced. Discret. Contin. Dyn. Syst. B **20**(4), 1077–1105 (2015)

36. Kirkwood, J.G.: Statistical mechanics of fluid mixtures. J. Chem. Phys. **3**(5), 300–313 (1935)

37. Kraaij, R.C., Lazarescu, A., Maes, C., Peletier, M.: Fluctuation symmetry leads to generic equations with non-quadratic dissipation. Stoch. Process. Appl. **130**(1), 139–170 (2020)

38. Kubo, R., Matsuo, K., Kitahara, K.: Fluctuation and relaxation of macrovariables. J. Stat. Phys. **9**(1), 51–96 (1973)

39. Kondepudi, D., Prigogine, I.: Modern Thermodynamics: From Heat Engines to Dissipative Structures. Wiley, New York (2014)

40. Kraaij, R.: Large deviations for Markov jump processes with mean-field interaction via the comparison principle for an associated Hamilton-Jacobi equation. J. Stat. Phys. **164**(2), 321–345 (2016)

41. Kraaij, R.C.: The exponential resolvent of a Markov process and large deviations for Markov processes via Hamilton-Jacobi equations. Electron. J. Probab. **25**, 1–39 (2020)

42. Kunitha, H.: On backward stochastic differential equations. Stochastics **6**(3–4), 293–313 (1982)

43. Kurtz, T.G.: Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. J. Appl. Probab. **8**(2), 344–356 (1971)
44. Kurtz, T.G.: Representations of markov processes as multiparameter time changes. Ann. Probab. **8**(4), 682–715 (1980)
45. Lazarescu, A., Cossetto, T., Falasco, G., Esposito, M.: Large deviations and dynamical phase transitions in stochastic chemical networks. J. Chem. Phys. **151**(6), 064117 (2019)
46. Lebowitz, J.L., Spohn, H.: A Gallavotti-Cohen-type symmetry in the large deviation functional for stochastic dynamics. J. Stat. Phys. **95**(1), 333–365 (1999)
47. Mielke, A.: On evolutionary gamma-convergence for gradient systems. In: Macroscopic and large scale phenomena: coarse graining, mean field limits and ergodicity, pp. 187–249. Springer (2016)
48. Mielke, A., Peletier, M.A., Liero, M., Michiel Renger, D.R.: On microscopic origins of generalized gradient structures. Discret. Contin. Dyn. Syst. S **10**(1), 1–35 (2017)
49. Maas, J., Mielke, A.: Modeling of chemical reaction systems with detailed balance using gradient structures. J. Stat. Phys. **181**(6), 2257–2303 (2020)
50. Mielke, A., Renger, D.R.M., Peletier, M.A.: On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion. Potential Anal. **41**(4), 1293–1327 (2014)
51. McQuarrie, D.A., Simon, J.D.: Physical Chemistry: A Molecular Approach, vol. 1. University Science Books, Sausalito (1997)
52. Metzner, P., Schütte, C., Vanden-Eijnden, E.: Transition path theory for Markov jump processes. Multiscale Model. Simul. **7**(3), 1192–1219 (2009)
53. Morpurgo, G., Touschek, B.F., Radicati, L.A.: On time reversal. Il Nuovo Cimento (1943–1954) **12**(5), 677–698 (1954)
54. Onsager, L.: Reciprocal relations in irreversible processes. I. Phys. Rev. **37**, 405–426 (1931)
55. Peliti, L.: Path integral approach to birth-death processes on a lattice. J. Phys. **46**(9), 1469–1483 (1985)
56. Polettini, M., Lazarescu, A., Esposito, M.: Tightening the uncertainty principle for stochastic currents. Phys. Rev. E **94**(5), 052104 (2016)
57. Patterson, R.I.A., Renger, D.R.M.: Large deviations of jump process fluxes. Math. Phys. Anal. Geom. **22**(3), 21 (2019)
58. Prigogine, I.: Introduction to Thermodynamics of Irreversible Processes. Interscience, New York (1967)
59. Patterson, R.I.A., Michiel Renger, D.R., Sharma, U.: Variational structures beyond gradient flows: a macroscopic fluctuation-theory perspective. arXiv:2103.14384 [math-ph] (2021)
60. Peletier, M.A., Redig, F., Vafayi, K.: Large deviations in stochastic heat-conduction processes provide a gradient-flow structure for heat conduction. J. Math. Phys. **55**(9), 093301 (2014)
61. Qian, H., Ge, H.: Stochastic Chemical Reaction Systems in Biology. Lecture Notes on Mathematical Modelling in the Life Sciences. Springer International Publishing (2021)
62. Qian, H.: Open-system nonequilibrium steady state: statistical thermodynamics, fluctuations, and chemical oscillations. J. Phys. Chem. B **110**(31), 15063–15074 (2006)
63. Qian, H.: Phosphorylation energy hypothesis: open chemical systems and their biological functions. Annu. Rev. Phys. Chem. **58**, 113–142 (2007)
64. Rao, R., Esposito, M.: Nonequilibrium thermodynamics of chemical reaction networks: wisdom from stochastic thermodynamics. Phys. Rev. X **6**(4), 041064 (2016)
65. Renger, D.R.M.: Flux large deviations of independent and reacting particle systems, with implications for macroscopic fluctuation theory. J. Stat. Phys. **172**, 1291–1326 (2018)
66. Roma, D.M., O'Flanagan, R.A., Ruckenstein, A.E., Sengupta, A.M., Mukhopadhyay, R.: Optimal path to epigenetic switching. Phys. Rev. E **71**(1), 011902 (2005)
67. Ruelle, D.P.: Extending the definition of entropy to nonequilibrium steady states. Proc. Natl. Acad. Sci. **100**(6), 3054–3058 (2003)
68. Schlogl, F.: Chemical reaction models for non-equilibrium phase transitions. Zeitschrift Phys. **253**(2), 147–161 (1972)
69. Sontag, E.D.: Structure and stability of certain chemical networks and applications to the kinetic proof-reading model of t-cell receptor signal transduction. IEEE Trans. Autom. Control **46**(7), 1028–1047 (2001)
70. Sandier, E., Serfaty, S.: Gamma-convergence of gradient flows with applications to Ginzburg-Landau. Commun. Pure Appl. Math. **57**(12), 1627–1672 (2004)
71. Shwartz, A., Weiss, A.: Large Deviations for Performance Analysis: Queues, Communication and Computing, vol. 5. CRC Press, Boca Raton (1995)
72. Tran, H.V.: Hamilton-Jacobi Equations: Theory and Applications, vol. 213. American Mathematical Soc, Providence (2021)
73. Varadhan, S.R.S.: Asymptotic probabilities and differential equations. Commun. Pure Appl. Math. **19**(3), 261–286 (1966)

74. Van Kampen, N.G.: Stochastic Processes in Physics and Chemistry, 3rd edn. Elsevier, Amsterdam (2007)
75. Weinan, E., Ren, W., Vanden-Eijnden, E.: String method for the study of rare events. Phys. Rev. B **66**(5), 052301 (2002)
76. Weinan, E., Eric, V.-E.: Towards a theory of transition paths. J. Stat. Phys. **123**(3), 503 (2006)
77. Whittle, P.: Systems in Stochastic Equilibrium. Wiley, New York (1986)
78. Yang, Y.-J., Qian, H.: Unified formalism for entropy production and fluctuation relations. Phys. Rev. E **101**(2), 022129 (2020)