Weierstraß-Institut

für Angewandte Analysis und Stochastik

Leibniz-Institut im Forschungsverbund Berlin e. V.

Gradient and Generic systems in the space of fluxes, applied to reacting particle systems

D.R. Michiel Renger[1]

submitted: July 4, 2018

[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: michiel.renger@wias-berlin.de

No. 2516

Berlin 2018

Gradient and Generic systems in the space of fluxes, applied to reacting particle systems

D.R. Michiel Renger

Abstract

In a previous work we devised a framework to derive generalised gradient systems for an evolution equation from the large deviations of an underlying microscopic system, in the spirit of the Onsager-Machlup relations. Of particular interest is the case where the microscopic system consists of random particles, and the macroscopic quantity is the empirical measure or concentration. In this work we take the particle flux as the macroscopic quantity, which is related to the concentration via a continuity equation. By a similar argument the large deviations can induce a generalised gradient or Generic system in the space of fluxes. In a general setting we study how flux gradient or generic systems are related to gradient systems of concentrations. The arguments are explained by the example of reacting particle systems, which is later expanded to include spatial diffusion as well.

1 Introduction

By the Boltzmann-Einstein relation, the free energy of a system is inherently related to the fluctuations of an underlying microscopic particle system. In this sense, two systems with the same macroscopic behaviour can be driven by completely different free energies, if their corresponding microscopic systems are different. Therefore, one of the main objectives of (equilibrium) statistical mechanics is to derive the 'physically correct' macroscopic free energy from fluctuations in microscopic systems. A similar principle can be applied to systems that evolve over time, where dynamic fluctuations may lead to a gradient flow, driven by the free energy. For stochastically reversible systems and close to equilibrium this is the classic Onsager-Machlup theory [Ons31, OM53]. Such relations are known to hold for many reversible dynamics, not necessarily close to equilibrium [BDSG$^+$04]. More recently, it was shown that microscopic reversibility always implies the emergence of a macroscopic gradient flow [MPR14], but in this generality one needs to allow for so-called generalised gradient flows. In brief, a generalised gradient structure (GGS) is defined by a possibly non-linear relation between velocities and affinities. Although there are exist some non-reversible models that lead to macroscopic gradient flows [Die15], these are considered non-typical; so in order to understand systems with non-reversible microscopic fluctuations, one needs to look for even further (thermodynamically consistent) generalisations of a gradient flow.

One such generalisation could be Generic [Ött05], which couples a gradient flow to a Hamiltonian system in such a way that the Hamiltonian energy is conserved and the free energy decays over time. Clearly, in order to derive Generic structures from dynamical large deviations in a general setting, one again needs to allow for non-linear relations, just like generalised gradient flows. One is thus lead to study generalised Generic structures (GGEN) [Mie11]. A recent work shows a set of necessary and sufficient conditions of which microscopic fluctuations induce such a generalised Generic structure [KLMP18a, KLMP18b].

Another generalisation worth considering is to study, in addition to the macroscopic state variables, the corresponding fluxes. These fluxes hold more information than the state variables due to the possible occurrence of 'divergence-free' fluxes that do not alter the states. The study of flux large deviations and corresponding thermodynamic properties is a key concept in Macroscopic Fluctuation Theory [BDSG$^+$15]. With the fluctuations on the fluxes one could for example extract a generalisation of a gradient structure: (what we call) a force structure [MN08, KJZ17, Mae17, Ren17], where the affinity/driving force may no longer be the gradient of some free energy.

In this paper we pursue another flux-based point of view, based on the observation that a GGS and GGEN can be interpreted as a free energy balance. If there would be work done that results in a divergence-free flux, than one might expect a gap the energy balance, so that the such systems can not induce a GGS or GGEN. Could one then still have GGS/GGEN structure in the space of fluxes? The main point of this paper is that this is generally impossible. We will see that if the fluctuations induce a GGS or GGEN in the space of fluxes, this is, up to some conditions, equivalent to the fluctuations inducing a GGS or GGEN in the state space.

Our leading example and main application will be that of an isothermal chemical reaction network, as studied in [MPPR17]. In Section 2 we recall the main arguments from that paper, applied to concentrations undergoing reactions. In Section 3 we expand these ideas and show how they can be applied to reaction fluxes. Based on this example, we then develop an abstract theory about induced GGSs and GGENs in flux space, in Section 4. To show the generality of these principles, we then show in Sections 5 how the theory applies to transport fluxes in a diffusing particles model, and in Section 6 we combine the arguments from Sections 3 and 6 to derive results for a transport and reaction fluxes in a simple reaction-diffusion model.

2   Leading example 1: chemical reactions

In this section we explain the main concepts, mostly by reiterating the arguments of [MPR14] and [MPPR17]. In particular, we argue that large deviations/fluctuations provide a 'physically correct' GGS for the evolution of concentrations undergoing chemical reactions.

Consider a network of isothermal chemical reactions, for example:

$$2\mathrm{Na} + \mathrm{Cl}_2 \to 2\mathrm{NaCl} \qquad \text{and} \qquad 2\mathrm{NaCl} \to 2\mathrm{Na} + \mathrm{Cl}_2.$$

We denote the set of species by $\mathcal{Y}$ (in this example $\{\mathrm{Na}, \mathrm{Cl}_2, \mathrm{NaCl}\}$) and the set of reactions by $\mathcal{R}$; for each reaction r we consider a forward and backward reaction (so that here $\mathcal{R}$ consists of one element). The stoichiometric coefficients are denoted by $(\alpha_{r,y})_{y \in \mathcal{Y}}$ for the forward reactants (here $(2,1,0)$) and $(\beta_{r,y})_{y \in \mathcal{Y}}$ for the forward products (here $(0,0,2)$), which yields the state change matrix $\Gamma := (\gamma_{r,y}) := (\beta_{r,y} - \alpha_{r,y})$. The evolution of the concentrations $\rho_t \in \mathbb{R}^{\mathcal{Y}}$ is then described by the *Reaction Rate Equation*,

$$\dot{\rho}_t = \sum_{r \in \mathcal{R}} \gamma_r \big(k_{r,\mathrm{fw}}(\rho_t) - k_{r,\mathrm{bw}}(\rho_t) \big) = \Gamma \bar{k}(\rho_t), \tag{1}$$

with concentration-dependent reaction rates $\bar{k}_r := k_{r,\mathrm{fw}} - k_{r,\mathrm{bw}}$.

Typically (and often in this paper), the reaction rates will be of the form

$$k_{r,\mathrm{fw}}(\rho) := \kappa_{r,\mathrm{fw}} \rho^{\alpha_r} \qquad \text{and} \qquad k_{r,\mathrm{bw}}(\rho) := \kappa_{r,\mathrm{bw}} \rho^{\beta_r} \tag{2}$$

for some constants $\kappa_{r,\mathrm{fw}}, \kappa_{r,\mathrm{bw}}$, using the notation $\rho^{\alpha_r} = \prod_{y \in \mathcal{Y}} \rho_y^{\alpha_{r,y}}$. In that case the network is said to be of *mass-action kinetics*. For more background on chemical reaction networks we refer to the survey [AK11]. More details about the fluctuations and induced GGSs for chemical reaction networks can be found in [MPPR17]; for completeness we shall recall these results in this section.

To notationally stress the similarity and differences between different concepts throughout this paper, we shall always net quantities by a bar (¯), and we distinguish functionals on state space from functionals on flux space by a hat (ˆ). In this section we study concentrations only, which we consider to be states.

2.1 Reacting particle system

A classical microscopic particle system underlying the evolution (1) is the following [Kur72]. Let V be a large, well-mixed volume that contains, at time t, a total number $N_t^{(V)}$ of particles of species $Y_{t,i}^{(V)}$, $i = 1, \ldots, N_t^{(V)}$. A reaction r occurs randomly with some propensities (jump rates) $\lambda_{r,\mathrm{fw}}^{(V)}(\rho), \lambda_{r,\mathrm{bw}}^{(V)}(\rho)$. Whenever a forward reaction r occurs, α_r particles are removed and β_r particles are created, and vice versa for a backward reaction. Hence each reaction requires a cumbersome relabelling of particles $Y_{t,i}$. It is therefore more practical to work directly with the (particles per volume) concentration $\rho_{t,y}^{(V)} := \frac{1}{V} \# \{Y_{t,i} = y, i = 1, \ldots N_t^{(V)}\}$. This quantity will also play the role of the macroscopic state variable. Whenever a forward or backward reaction r takes place, the concentration can now be simply updated by a jump $\rho_t^{(V)} = \rho_{t-}^{(V)} \pm \frac{1}{V} \gamma_r$. Then the $\mathcal{Y}$-dimensional vector $\rho_t^{(V)}$ is a Markov jump process, which satisfies the master equation

$$\dot{\hat{P}}_t^{(V)}(\rho) = \sum_{r \in \mathcal{R}} \lambda_{r,\mathrm{fw}}^{(V)}(\rho - \tfrac{1}{V}\gamma_r)\hat{P}_t^{(V)}(\rho - \tfrac{1}{V}\gamma_r) - \lambda_{r,\mathrm{fw}}^{(V)}(\rho)\hat{P}_t^{(V)}(\rho) +$$
$$\lambda_{r,\mathrm{bw}}^{(V)}(\rho + \tfrac{1}{V}\gamma_r)\hat{P}_t^{(V)}(\rho + \tfrac{1}{V}\gamma_r) - \lambda_{r,\mathrm{bw}}^{(V)}(\rho)\hat{P}_t^{(V)}(\rho).$$

It will be beneficial to work with the corresponding generator, which is the adjoint of the right-hand side of the master equation, with respect to the dual pairing $\langle \hat{P}_t^{(V)}, f \rangle = \sum_\rho f(\rho)\,\hat{P}_t^{(V)}(\rho)$ with an arbitrary test function:

$$(\hat{\mathcal{Q}}^{(V)}f)(\rho) := \sum_{r \in \mathcal{R}} \lambda_{r,\mathrm{fw}}^{(V)}(\rho)\big(f(\rho + \tfrac{1}{V}\gamma_r) - f(\rho)\big) + \lambda_{r,\mathrm{bw}}^{(V)}(\rho)\big(f(\rho - \tfrac{1}{V}\gamma_r) - f(\rho)\big). \tag{3}$$

The propensities that are usually used in the so-called chemical master equation are derived from combinatoric considerations, and yield the mass-action kinetics in the limit [Kur72, AK11, MPPR17]:

$$\tfrac{1}{V}\lambda_{r,\mathrm{fw}}^{(V)}(\rho) := \frac{1}{V} \cdot \frac{\kappa_{r,\mathrm{fw}}}{V^{\alpha_{r,\mathrm{tot}}-1}} \frac{(\rho V)!}{(\rho V - \alpha_r)!} \xrightarrow{V \to \infty} \kappa_{r,\mathrm{fw}}\rho^{\alpha_r} = k_{r,\mathrm{fw}}(\rho),$$

$$\tfrac{1}{V}\lambda_{r,\mathrm{bw}}^{(V)}(\rho) := \frac{1}{V} \cdot \frac{\kappa_{r,\mathrm{bw}}}{V^{\beta_{r,\mathrm{tot}}-1}} \frac{(\rho V)!}{(\rho V - \beta_r)!} \xrightarrow{V \to \infty} \kappa_{r,\mathrm{bw}}\rho^{\beta_r} = k_{r,\mathrm{bw}}(\rho), \tag{4}$$

using the notation $\alpha_{r,\mathrm{tot}} := \sum_{y \in \mathcal{Y}} \alpha_{r,y}$ and $\alpha_r! := \prod_{y \in \mathcal{Y}} \alpha_{r,y}!$.

2.2 Equilibrium: limit, large deviations and free energy

We will from now on (throughout this section) assume that the reaction network is of mass-action kinetics (2), and chemically detailed balanced, i.e. there exists a $\rho^* \in \mathbb{R}_+^{\mathcal{Y}}$ for which

$$\kappa_{r,\mathrm{fw}}\rho^{*\alpha_r} = \kappa_{r,\mathrm{bw}}\rho^{*\beta_r} \qquad \text{for all } r \in \mathcal{R}. \tag{5}$$

Naturally, ρ^* is an equilibrium under the deterministic evolution (1) [1]. Under this assumption, the invariant distribution of the stochastic model is known to be [AK11]:

$$\hat{P}^{(V)}_\infty(\rho) = \prod_{y\in\mathcal{Y}} \frac{(V\rho^*_y)^{V\rho_y}}{(V\rho_y)!} \mathrm{e}^{-V\rho^*_y}.$$

Letting $V\to\infty$, this invariant distribution concentrates on the equilibrium state:

$$\hat{P}^{(\infty)}_\infty(\rho) = \begin{cases} 1, & \rho = \rho^*, \\ 0, & \text{otherwise.} \end{cases}$$

One can then extract the free energy by considering the corresponding large deviations, i.e. the exponential rate which with $\hat{P}^{(V)}_\infty$ converges to zero. Indeed, by Stirling's formula,

$$-\frac{1}{V}\log\hat{P}^{(V)}_\infty(\rho) \xrightarrow{V\to\infty} \sum_{y\in\mathcal{Y}} \rho_y \log\frac{\rho_y}{\rho^*_y} - \rho_y + \rho^*_y =: h(\rho|\rho^*). \tag{6}$$

Such limit is known as a large-deviation principle; the function on the right characterises the stochastic cost of microscopic fluctuations. If the reaction rates are related to an internal energy via Arrhenius' law, than the expression $h(\rho|\rho^*)$ is really the Helmholtz free energy, apart from a normalisation term and a constant scaling, as explained in more detail in [MPPR17, Sec. 2.2&2.3].

2.3 Dynamics: limit and large deviations

Observe that in the microscopic model, the process speeds up as V increases with order V while the jump sizes are of size $1/V$. Therefore by (4), as $V\to\infty$ the generator $\hat{\mathcal{Q}}^{(V)}$ converges to the limit generator

$$(\hat{\mathcal{Q}}^{(\infty)}f)(\rho) := \sum_{r\in\mathcal{R}} \bar{k}_r(\rho)\,\nabla f(\rho)\cdot\gamma_r.$$

Since this generator depends on the test function f through $\nabla f(\rho)$ only, we can make the ansatz that the limit process is deterministic $\hat{P}^{(\infty)}_t(\tilde{\rho}) = \delta_{\rho_t}(\tilde{\rho})$, for some curve $\rho_{(\cdot)}$. Plugging this into the definition of the generator yields:

$$\nabla f(\rho_t)\cdot\dot{\rho}_t = \partial_t f(\rho_t) = \partial_t\langle\hat{P}^{(\infty)}_t, f\rangle = \langle\hat{P}^{(\infty)}_t, \mathcal{Q}^{(\infty)}f\rangle$$
$$= (\hat{\mathcal{Q}}^{(\infty)}f)(\rho_t) = \nabla f(\rho_t)\cdot\sum_{r\in\mathcal{R}}\bar{k}_r(\rho_t)\gamma_r.$$

As this relation holds for any test function f, we see that the ansatz was justified if the postulated curve ρ_t satisfies the Reaction Rate Equation (1). Hence the stochastic process $\rho^{(V)}_t$ converges (pathwise in probability) to the deterministic solution ρ_t of the Reaction Rate Equation.

Similar to the calculation of the fluctuations of the equilibrium (6), we can study the large deviations of the path probabilities $\hat{P}^{(V)}$; this is known as a dynamic large-deviation principle. These dynamical

---

[1] It should be stressed that, given an initial concentration ρ_0, both the deterministic evolution and the stochastic model is confined to the 'stoichiometric compatibility class' $\rho_0 + \mathrm{Ran}\,\Gamma = \{\rho_0 + \Gamma w : w\in\mathbb{R}^{\mathcal{R}}_+\}$. Therefore, it is not clear whether this equilibrium ρ^* lies within the compatibility class that corresponds to the initial concentration. However, if there exists a detailed balanced concentration ρ^*, then there exists a unique detailed balanced concentration within each such class, see [AK11] and the references therein. Without loss of generality, we can therefore implicitly assume that the detailed balanced equilibrium is unique, and lies within the correct compatibility class.

fluctuations can be formally calculated with the framework of [FK06]. To this aim we study the non-linear generator:

$$
\begin{aligned}
(\hat{\mathcal{H}}^{(V)} f)(\rho) &:= \tfrac{1}{V} \mathrm{e}^{-Vf(\rho)} (\hat{\mathcal{Q}}^{(V)} \mathrm{e}^{Vf})(\rho) \\
&= \sum_{r\in\mathcal{R}} \tfrac{1}{V} \lambda_{r,\mathrm{fw}}^{(V)}(\rho) \big(\mathrm{e}^{Vf(\rho+\frac{1}{V}\gamma_r)-Vf(\rho)} - 1 \big) \\
&\quad + \tfrac{1}{V} \lambda_{r,\mathrm{bw}}^{(V)}(\rho) \big(\mathrm{e}^{Vf(\rho-\frac{1}{V}\gamma_r)-Vf(\rho)} - 1 \big) \\
&\xrightarrow{V\to\infty} \sum_{r\in\mathcal{R}} k_{r,\mathrm{fw}}(\rho) \big(\mathrm{e}^{\nabla f(\rho)\cdot\gamma_r} - 1 \big) + k_{r,\mathrm{bw}}(\rho) \big(\mathrm{e}^{-\nabla f(\rho)\cdot\gamma_r} - 1 \big).
\end{aligned}
$$

As before, the limit depends on the test function through $\nabla f(\rho)$ only, which is consistent with the fact that the limit is deterministic. We then define, by a slight abuse of notation,

$$
\hat{\mathcal{H}}(\rho,\xi) := \sum_{r\in\mathcal{R}} k_{r,\mathrm{fw}}(\rho) \big(\mathrm{e}^{\xi\cdot\gamma_r} - 1 \big) + k_{r,\mathrm{bw}}(\rho) \big(\mathrm{e}^{-\xi\cdot\gamma_r} - 1 \big), \qquad \text{and} \qquad (7)
$$

$$
\hat{\mathcal{L}}(\rho,s) := \sup_{\xi\in\mathbb{R}^{\mathcal{Y}}} \xi \cdot s - \hat{\mathcal{H}}(\rho,\xi). \qquad (8)
$$

The dynamic large-deviation principle now states that [2]

$$
\mathrm{Prob}^{(V)}\big(\rho_{(\cdot)}^{(V)} \approx \rho_{(\cdot)}\big) \overset{V\to\infty}{\sim} \mathrm{e}^{-V\int_0^T \hat{\mathcal{L}}(\rho_t,\dot\rho_t)\,dt}. \qquad (9)
$$

Remark 2.1. The function (8) is implicitly defined as a supremum; although the supremum can be calculated explicitly, this leads to very cumbersome expressions. However, it does have a dual formulation in terms of a minimisation problem:

$$
\hat{\mathcal{L}}(\rho,s) = \inf_{\substack{j_{\mathrm{fw}},j_{\mathrm{bw}}\in\mathbb{R}_+^{\mathcal{R}}: \\ \Gamma(j_{\mathrm{fw}}-j_{\mathrm{bw}})=s}} h\big(j_{\mathrm{fw}}|k_{\mathrm{fw}}(\rho)\big) + h\big(j_{\mathrm{bw}}|k_{\mathrm{bw}}(\rho)\big),
$$

where $h(j|k) := \sum_{r\in\mathcal{R}} j_r \log(j_r/k_r) - j_r + k_r$, similar to (6). Although the relative entropy h appears in both expressions, they should not be confused: $h(\rho|\rho^*)$ is an equilibrium rate whereas $h(j|k(\rho))$ is a dynamic quantity. We shall see later on that the latter can be directly (without the infimum) be interpreted as a large-deviation rate, where the variable j is a reaction flux. $\qquad\square$

2.4 GGS, energy balance, and relation with fluctuations

The (naive) aim is to rewrite the macroscopic equation as a gradient flow of some free energy $\mathcal{F}$:

$$
\dot\rho_t = -\hat{K}(\rho_t)\,\nabla\hat{\mathcal{F}}(\rho_t) =: -\mathrm{grad}_{\rho_t}\hat{\mathcal{F}}(\rho_t), \qquad (10)
$$

where $\hat{K}(\rho)$ is some linear symmetric, positive definite (linear response) operator that maps thermodynamic forces to velocities. Mathematically, this operator can be interpreted as the inverse of the metric tensor of some manifold, so that the right-hand side is the gradient on this manifold.

---

[2]The rigorous definition of the large-deviation principle, the heuristics behind this method, and the rigorous proof of this statement is all beyond the scope of this paper. For the precise details we refer to [FK06], and for the rigorous proof for this particular system (by more classical methods) to [DRW16, ADE17, PR18]. For the sake of brevity, we assume that the randomness in the initial condition is sufficiently small (e.g. deterministic) so that we do not obtain initial fluctuations.

Clearly, (10) is equivalent to requiring

$$0 = \tfrac{1}{2}\|\dot\rho_t + \hat{K}(\rho_t)\,\nabla\hat{\mathcal{F}}(\rho_t)\|^2_{\hat{K}(\rho_t)^{-1}} = \tfrac{1}{2}\|\dot\rho_t\|^2_{\hat{K}(\rho_t)^{-1}} + \tfrac{1}{2}\|\nabla\hat{\mathcal{F}}(\rho_t)\|^2_{\hat{K}(\rho_t)} + \nabla\hat{\mathcal{F}}(\rho_t)\cdot\dot\rho_t,$$

if we set $\|\xi\|^2_{\hat{K}(\rho)} := \langle \xi, \hat{K}(\rho)\xi\rangle$ and $\|s\|^2_{\hat{K}(\rho)^{-1}} := \langle s, \hat{K}(\rho)^{-1}s\rangle$. Integrated over a time interval $(0, T)$, this reads

$$0 = \int_0^T \left(\tfrac{1}{2}\|\dot\rho_t\|^2_{\hat{K}(\rho_t)^{-1}} + \tfrac{1}{2}\|\nabla\hat{\mathcal{F}}(\rho_t)\|^2_{\hat{K}(\rho_t)} \right) dt + \hat{\mathcal{F}}(\rho_T) - \hat{\mathcal{F}}(\rho_0).$$

The last two terms describe the free energy loss (or entropy production), and the first two terms describe the dissipation[3]; as such this equation represents a free energy balance.

Observe that this expression is always non-negative, and 0 exactly on the gradient flow (10). Moreover, we see that this expression has the same dimension as $\hat{\mathcal{F}}$, the free energy; it is indeed the free energy cost to deviate from the macroscopic dynamics. This interpretation shows that this cost should be equal to the cost $\int_0^T \hat{\mathcal{L}}(\rho_t, \dot\rho_t)\,dt$ of microscopic fluctuations (9).

This is in many cases, and particularly in this case of chemical reactions, impossible. Since the large-deviation function (8) is non-quadratic, one should allow for non-quadratic dissipation terms. We therefore replace the two squared norms by a pair of dual *dissipation potentials*:

$$0 = \int_0^T \left(\hat{\Psi}(\rho_t, \dot\rho_t) + \hat{\Psi}^*\big(\rho_t, -\nabla\hat{\mathcal{F}}(\rho_t)\big) \right) dt + \hat{\mathcal{F}}(\rho_T) - \hat{\mathcal{F}}(\rho_0), \tag{11}$$

where, as in the quadratic case, the potentials are convex duals of each other, i.e. $\hat{\Psi}(\rho, s) = \sup_\xi \xi \cdot s - \hat{\Psi}^*(\rho, \xi)$. Moreover, we assume that $\hat{\Psi}$ and $\hat{\Psi}^*$ are both non-negative; from (11) we then see that the free energy $\mathcal{F}$ is non-increasing on the flow.

Due to the convex duality, the right-hand side of (11) is again always non-negative, and 0 exactly when

$$\dot\rho_t = \nabla_\xi \hat{\Psi}^*\big(\rho_t, -\nabla\hat{\mathcal{F}}(\rho_t)\big).$$

We call such equation a *generalised gradient flow*, and the underlying structure $(\mathbb{R}^{\mathcal{Y}}, \hat{\Psi}, \hat{\mathcal{F}})$ a *generalised gradient structure (GGS)*. Note that the generalisation with respect to (10) lies in the fact that we allow for a non-linear relation between forces and velocities. We moreover say that a gradient structure is *induced* by a cost function $\hat{\mathcal{L}}$ whenever $\hat{\mathcal{L}}(\rho, s) = \hat{\Psi}(\rho, s) + \hat{\Psi}^*\big(\rho, -\nabla\hat{\mathcal{F}}(\rho)\big) + \nabla\hat{\mathcal{F}}(\rho) \cdot s$.

In Section 4 we will recall the relation between fluctuation costs and GGSs, as described in [MPR14]. Applied to the current setting of chemical reactions, we reiterate the following result from [MPPR17]. If we again assume mass-action kinetics (2) and chemical detailed balance (5), then there exists a unique GGS $(\mathbb{R}^{\mathcal{Y}}, \hat{\Psi}, \hat{\mathcal{F}})$ induced by the large-deviation cost $\hat{\mathcal{L}}$ from (8), where

$$\hat{\mathcal{F}}(\rho) := \tfrac{1}{2}h(\rho|\rho^*), \qquad \text{and} \qquad \hat{\Psi}^*(\rho, \xi) := \sum_{r\in\mathcal{R}} \hat\sigma_r(\rho)\big(\cosh(\xi\cdot\gamma_r) - 1\big), \tag{12}$$

with $\hat\sigma_r(\rho) := 2\sqrt{k_{r,\mathrm{fw}}(\rho)k_{r,\mathrm{bw}}(\rho)}$. Since $\hat{\Psi}^*$ appears as a sum, the expression for Ψ becomes a so-called inf-convolution where all reactions are strongly intertwined. For more details on these inf-convolutions and the factor $1/2$ in front of the free energy, we again refer to [MPPR17, Sec. 3.4].

---

[3]For a linear gradient flow (10), $\tfrac{1}{2}\|\dot\rho_t\|^2_{\hat{K}(\rho_t)^{-1}} + \tfrac{1}{2}\|\nabla\hat{\mathcal{F}}(\rho_t)\|^2_{\hat{K}(\rho_t)} = \|\dot\rho_t\|^2_{\hat{K}(\rho)^{-1}}$; hence the dissipation can be seen as a kinetic energy. This is however no longer true for general GGSs.

3 Leading example 2: fast-slow reaction fluxes

As mentioned in the introduction, the motivation behind the current paper is to search for thermodynamically consistent structures for systems that are not detailed balanced. The idea is that we increase the space by taking fluxes into account. However, in order to see the connection between structures in flux space and structures in state space, we dedicate this section to an example that is 'almost' detailed balanced. More general systems will be dealt with when discussing the general theory in Section 4.

3.1 Reacting particle system and macroscopic equation

We consider a system of fast and slow chemical reactions, $\mathcal{R} = \mathcal{R}_{\text{slow}} \cup \mathcal{R}_{\text{fast}}$, where the slow reactions are assumed to be of mass-action kinetics and detailed balanced, as in the previous section. By contrast, we will not assume anything of the like for the fast reactions, neither shall we assume that each reaction consist of a forward and a backward reaction. However, in the microscopic model, we shall assume that the fast reactions happen on a faster time-scale, i.e.

$$\frac{1}{V^2}\lambda_r^{(V)}(\rho) \to \tilde{k}_r(\rho) \qquad \text{for } r \in \mathcal{R}_{\text{fast}}, \tag{13}$$

but their effect on the concentrations is smaller, i.e. $\rho^{(V)}(t) = \rho^{(V)}(t^-) + \frac{1}{V^2}\gamma_r$ whenever a reaction $r \in \mathcal{R}_{\text{fast}}$ occurs at time t. To notationally distinguish the two time scales we use a tilde ($\tilde{\ }$) to denote one-way fast quantities. Let us briefly mention that the additional fast reactions are not essential to see the relation between structures in flux and state space; they just lead to a richer example that is interesting in its own right.

The main idea is now to increase the state space by bookkeeping the events in the microscopic system, in this case, by counting the number of reactions that have occurred up to a given time. With the right scaling, this defines the following *integrated reaction fluxes*[4]:

$$\bar{W}_{t,r}^{(V)} := \frac{1}{V}\#\big\{\text{forward reactions } r \text{ occurred in } (0,t)\big\}$$
$$- \frac{1}{V}\#\big\{\text{backward reactions } r \text{ occurred in } (0,t)\big\}, \qquad r \in \mathcal{R}_{\text{slow}},$$
$$\tilde{W}_{t,r}^{(V)} := \frac{1}{V^2}\#\big\{\text{reactions } r \text{ occurred in } (0,t)\big\}, \qquad r \in \mathcal{R}_{\text{fast}},$$

and to shorten notation we sometimes write

$$W_t^{(V)} := \big((\bar{W}_{t,r}^{(V)})_{r\in\mathcal{R}_{\text{slow}}}, (\tilde{W}_{t,r}^{(V)})_{r\in\mathcal{R}_{\text{fast}}}\big). \tag{14}$$

We shall always assume that the initial condition is known (deterministically) a priori, so that the concentrations can be retrieved from the integrated fluxes via the *continuity equation*:

$$\rho_t^{(V)} = \phi^{(V)}[W_t^{(V)}] := \rho_0^{(V)} + \Gamma W_t^{(V)}.$$

In this sense the integrated fluxes encode more information than the concentrations.

---

[4]The term "integrated" signifies that these fluxes are cumulative over time. This simplifies the microscopic analysis since the corresponding process is Markovian; on a macroscopic scale only the time-derivatives will play a role. We also mention that we consider net rather than one-way slow fluxes, else the slow dynamics would not induce any force or gradient structure, see [Ren17, Sec. 4.6].

The integrated fluxes are again a Markov process, with generator (cf. (3)):

$$
\begin{aligned}
(\mathcal{Q}^{(V)}f)(w) := \sum_{r\in\mathcal{R}_{\text{slow}}} &\lambda_{r,\text{fw}}^{(V)}\big(\phi^{(V)}[w]\big)\big(f(w+\tfrac{1}{V}\mathbb{1}_r)-f(w)\big) \\
&+ \lambda_{r,\text{bw}}^{(V)}\big(\phi^{(V)}[w]\big)\big(f(w-\tfrac{1}{V}\mathbb{1}_r)-f(w)\big) \\
+ \sum_{r\in\mathcal{R}_{\text{fast}}} &\lambda_r^{(V)}\big(\phi^{(V)}[w]\big)\big(f(w+\tfrac{1}{V^2}\mathbb{1}_r)-f(w)\big),
\end{aligned}
$$

3.2 Limit and large deviations

We now mimic the arguments of Subsection 2.3, but now in the space of fluxes. Let $\rho_0^{(V)}\to\rho_0$, and so $\phi^{(V)}[w]\to\phi[w]:=\rho_0+\Gamma w$. Then by the same argument as in Subsection 2.3, using the scalings (4) and (13), one finds that as $V\to\infty$, the random process $W_t^{(V)}$ converges (pathwise in probability) to the solution of the macroscopic equations:

$$
\dot{\bar{w}}_{t,r}=\bar{k}_r\big(\phi[w_t]\big),\quad r\in\mathcal{R}_{\text{slow}},\qquad\text{and}\qquad \dot{\tilde{w}}_{t,r}=\tilde{k}_r\big(\phi[w_t]\big),\quad r\in\mathcal{R}_{\text{fast}},\qquad(15)
$$

again using the notation (14). Indeed, combining these equations leads to the macroscopic equation $\dot\rho_t=\Gamma k(\rho_t)$ for the concentrations.

Similarly, we study the fluctuations through the non-linear generator:

$$
\begin{aligned}
(\mathcal{H}^{(V)}f)(w) :=& \tfrac{1}{V}\mathrm{e}^{-Vf(w)}(\mathcal{Q}^{(V)}\mathrm{e}^{Vf})(w) \\
=& \sum_{r\in\mathcal{R}_{\text{slow}}} \tfrac{1}{V}\lambda_{r,\text{fw}}^{(V)}\big(\phi^{(V)}[w]\big)\big(\mathrm{e}^{Vf(w+\frac{1}{V}\mathbb{1}_r)-Vf(w)}-1\big) \\
&\qquad + \tfrac{1}{V}\lambda_{r,\text{bw}}^{(V)}\big(\phi^{(V)}[w]\big)\big(\mathrm{e}^{Vf(w-\frac{1}{V}\mathbb{1}_r)-Vf(w)}-1\big) \\
&+ \sum_{r\in\mathcal{R}_{\text{fast}}} \tfrac{1}{V}\lambda_r^{(V)}\big(\phi^{(V)}[w]\big)\big(\mathrm{e}^{Vf(w+\frac{1}{V^2}\mathbb{1}_r)-Vf(w)}-1\big) \\
\xrightarrow{V\to\infty}& \sum_{r\in\mathcal{R}_{\text{slow}}} k_{r,\text{fw}}\big(\phi[w]\big)\big(\mathrm{e}^{\partial_{w_r}f(w)}-1\big)+k_{r,\text{bw}}\big(\phi[w]\big)\big(\mathrm{e}^{-\partial_{w_r}f(w)}-1\big) \\
&\qquad\qquad + \sum_{r\in\mathcal{R}_{\text{fast}}} \tilde{k}_r\big(\phi[w]\big)\partial_{w_r}f(w).\qquad(16)
\end{aligned}
$$

As in Section 2.3, this limit depends on the gradient $\nabla f(w)$ only, which is consistent with the deterministic limit (15). Again we set, by a slight abuse of notation,

$$
\begin{aligned}
\mathcal{H}(w,\zeta) :=& \sum_{r\in\mathcal{R}_{\text{slow}}} k_{r,\text{fw}}\big(\phi[w]\big)\big(\mathrm{e}^{\bar\zeta_r}-1\big)+k_{r,\text{bw}}\big(\phi[w]\big)\big(\mathrm{e}^{-\bar\zeta_r}-1\big) \\
&\qquad\qquad\qquad + \sum_{r\in\mathcal{R}_{\text{fast}}} \tilde{k}_r\big(\phi[w]\big)\tilde\zeta_r,\qquad(17)
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{L}(w,j) :=& \sup_{\zeta\in\mathbb{R}^{\mathcal{R}}}\zeta\cdot j-\mathcal{H}(w,\zeta) \\
=& \inf_{j_{\text{fw}}-j_{\text{bw}}=\bar{j}} h\big(j_{\text{fw}}|k_{\text{fw}}(\phi[w])\big)+h\big(j_{\text{bw}}|k_{\text{bw}}(\phi[w])\big)+\chi_{\{\tilde{j}=\tilde{k}(\phi[w])\}},
\end{aligned}
$$

using the notation $\chi_{\{\tilde{j}=\tilde{k}(\phi[w])\}}=0$ if $\tilde{j}=\tilde{k}(\phi[w])$ and ∞ otherwise. Then, the dynamic large-deviation principle for the integrated fluxes $W^{(V)}(t)$ state that (see [PR18] for a rigorous proof):

$$
\mathrm{Prob}\big(W_{(\cdot)}^{(V)}\approx w_{(\cdot)}\big)\overset{V\to\infty}{\sim}\mathrm{e}^{-V\int_0^T\mathcal{L}(w_t,\dot{w}_t)\,dt}.\qquad(18)
$$

Comparing this dynamic large-deviations principle with (9), we see – not coincidentally – strong similarities. Let us assume that the limit fast fluxes do not influence the concentration, i.e.

$$\Gamma \tilde{k}\big(\phi[w]\big) = 0 \qquad \text{for all } w. \tag{19}$$

Naturally, this macroscopic condition entails a sort of decoupling between the slow and fast dynamics, e.g. when the species involved in the slow dynamics act as a catalyst for the fast dynamics. In that case $\rho_t^{(V)} = \phi^{(V)}[W_t^{(V)}]$ is exactly the process with generator (3). The contraction principle of large deviations theory [DZ87, Th. 4.2.1] then states that the two large-deviation costs are related via:

$$\int_0^T \hat{\mathcal{L}}(\rho_t, \dot{\rho}_t)\, dt = \inf_{w_{(\cdot)}:\rho_t=\phi[w_t]} \int_0^T \mathcal{L}(w_t, \dot{w}_t)\, dt.$$

In this setting, this infimum is only with respect to the second variable, because $\mathcal{L}(w, j)$ depends on w through $\phi[w]$ only, which in turn arises naturally from the fact that the jump rates depend on the state and not on the integrated flux. Therefore the relation above simplifies further to:

$$\hat{\mathcal{L}}\big(\phi[w], s\big) = \inf_{j:s=\Gamma j} \mathcal{L}(w, j) \qquad \text{for all } w, s, \tag{20}$$

which is consistent with Remark 2.1. These relations will be the starting point of the general theory that we develop in Section 4. Let us only mention here that as a consequence we also have the relation $\hat{\mathcal{H}}\big(\phi[w], \xi\big) = \mathcal{H}(w, \Gamma^\mathsf{T} \xi)$ for all w, ξ, cf. (7) and (17).

3.3 Induced Generic structure

We now investigate whether the large deviations (18) induces some structure in the space of fluxes. It turns out that this is indeed the case. As was found in [Ren17, Cor. 4.8], under the detailed balance assumption (5) the slow fluxes induce the GGS $(\mathbb{R}_+^{\mathcal{R}_{\text{slow}}}, \Psi, \mathcal{F})$, where

$$\mathcal{F}(w) := \hat{\mathcal{F}}\big(\phi[w]\big) = \tfrac{1}{2} h(\rho_0 + \Gamma w | \rho^*), \tag{21}$$

$$\Psi^*(w, \bar{\zeta}) := \sum_{r \in \mathcal{R}_{\text{slow}}} \sigma_r(w)\big(\cosh(\bar{\zeta}_r) - 1\big) \qquad \text{and} \tag{22}$$

$$\Psi(w, \bar{\jmath}) := \sum_{r \in \mathcal{R}_{\text{slow}}} \sigma_r(w)\Big(\cosh^*\big(\tfrac{\bar{\jmath}_r}{\sigma_r(w)}\big) + 1 \Big), \tag{23}$$

with $\sigma_r(w) := 2\sqrt{k_{r,\text{fw}}\big(\phi[w]\big) k_{r,\text{bw}}\big(\phi[w]\big)}, r \in \mathcal{R}_{\text{slow}}$. Extending these dissipation potentials to the full flux space by setting $\Psi^*(w, \zeta) := \Psi^*(w, \bar{\zeta})$ and $\Psi(w, j) := \Psi(w, \bar{\jmath}) + \chi_{\{\tilde{\jmath}=0\}}$, we can decompose the large-deviation cost function as

$$\mathcal{L}(w, j) = \Psi\big(w, j - \tilde{k}(\phi[w])\big) + \Psi^*\big(w, -\nabla\mathcal{F}(w)\big) + \nabla\mathcal{F}(w) \cdot \big(j - \tilde{k}(\phi[w])\big),$$

and accordingly, the macroscopic evolution (15) as:

$$\dot{w}_t = \nabla_\zeta \Psi^*\big(w_t, -\nabla\mathcal{F}(w_t)\big) + \tilde{k}(\phi[w]).$$

Due to (21) and the decoupling condition (19), the fast fluxes are orthogonal to the driving force, i.e.

$$\nabla_w \mathcal{F}(w) \cdot \tilde{k}(\phi[w]) = \nabla_\rho \hat{\mathcal{F}}\big(\phi[w]\big) \cdot \Gamma\tilde{k}(\phi[w]) = 0. \tag{24}$$

The quadruple $(\mathbb{R}_+^{\mathcal{R}}, \Psi, \mathcal{F}, \tilde{k} \circ \phi)$ satisfying this condition falls within the class of what is recently coined *pre-(Generalised) Generic (pGGEN)* [KLMP18a, KLMP18b].

It was shown in those works that non-interaction condition (24) is a necessary and sufficient condition for the existence of an underlying *Generalised Generic (GGEN)* structure $(\mathbb{R}_+^{\mathcal{R}}, \Psi, \mathcal{F}, L, \mathcal{E})$[5]. This means that the fast flux term is Hamiltonian $\tilde{k}(\phi[w]) = L(w) \nabla \mathcal{E}(w)$ for some Poisson structure L satisfying the Jacobi identity (see Section 4), $\mathcal{E}$ is some Hamiltonian energy, and the following two non-interaction conditions are satisfied:

$$L(w) \nabla \mathcal{F}(w) = 0 \qquad \text{and} \qquad \Psi^*\big(w, \zeta + z \nabla \mathcal{E}(w)\big) = \Psi^*(w, \zeta) \qquad (25)$$

for all $w \in \mathbb{R}_+^{\mathcal{R}}, \zeta \in \mathbb{R}^{\mathcal{R}}$ and $z \in \mathbb{R}$. These two conditions guarantee that along solutions the free energy is non-increasing and the Hamiltonian energy is conserved:

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{F}(w_t) = \underbrace{\nabla \mathcal{F}(w) \cdot \nabla_\zeta \Psi^*\big(w_t, -\nabla \mathcal{F}(w)\big)}_{\leq 0 \text{ by convexity}} + \underbrace{\nabla \mathcal{F}(w) \cdot L(w) \nabla \mathcal{E}(w)}_{=0 \text{ by } (25)} \leq 0,$$

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}(w_t) = \underbrace{\nabla \mathcal{E}(w) \cdot \nabla_\zeta \Psi^*\big(w_t, -\nabla \mathcal{F}(w)\big)}_{=0 \text{ by } (25)} + \underbrace{\nabla \mathcal{E}(w) \cdot L(w) \nabla \mathcal{E}(w)}_{=0 \text{ by skewsymmetry}} = 0.$$

One main message of this section is that the flux large-deviation cost $\mathcal{L}$ induces a unique pGGEN system $(\mathbb{R}_+^{\mathcal{R}}, \Psi, \mathcal{F}, \tilde{k} \circ \phi)$. Although this implies the existence of a GGEN system induced by $\mathcal{L}$, one can not uniquely decide on the basis of $\mathcal{L}$ what the 'correct' Hamiltonian structure $(\mathbb{R}_+^{\mathcal{R}}, L, \mathcal{E})$ should be. Additional physical or mathematical arguments needed to uniquely fix the Hamiltonian structure are beyond the scope of this paper; for possible constructions of Poisson structures L and energies $\mathcal{E}$, we refer the reader to [KLMP18b, Sec. 4]).

It should be noted that the pGGEN system $(\mathbb{R}_+^{\mathcal{R}}, \Psi, \mathcal{F}, \tilde{k} \circ \phi)$ is rather special in that (24) holds for any $\hat{\mathcal{F}}$, since the drift $\tilde{k} \circ \phi$ is 'divergence-free'. We will see that such systems play a special role in connecting systems in flux and state space.

Another main message is that the GGS part (21) on flux space is strongly related to the GGS (12) on state space, just like both cost functions are related by (20). These observations will be the basis of the general theory of the next section.

4 General theory

Following the examples of Sections 2 and 3, we now develop a more abstract framework to study to the relation between energy-driven structures in flux and in state space.

4.1 Geometry and notation

Throughout this section we assume to be given:

- ◼ A differentiable manifold $\mathcal{W}$ ("flux space"), where tangents are denoted by $(w, j) \in T\mathcal{W}$ and cotangents by $(w, \zeta) \in T^*W$. Note that we distinguish between tangents and cotangents; we write the dual pairing between them as $_{T_w^*}\langle \zeta, j \rangle_{T_w}$ or simply $\langle \zeta, j \rangle$;

---

[5]One often needs to introduce an auxiliary energy (e.g. a heat bath) to force conservation of energy, which enlarges the degrees of freedom in the system. To keep notation accessible we ignore this issue.
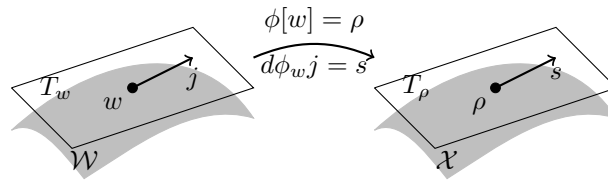
Figure 1: The 'continuity' mapping between the flux and state manifolds.

- A differentiable manifold $\mathcal{X}$ ("state space"), where tangents are denoted by $(\rho, s) \in T\mathcal{X}$ and cotangents by $(\rho, \xi) \in T^*\mathcal{X}$;

- A surjective differentiable operator $\phi : \mathcal{W} \to \mathcal{X}$ with bounded linear differential $d\phi_w : T_w \to T_{\Gamma(w)}$ and adjoint differential $d\phi_w^\mathsf{T} : T^*_{\phi[w]} \to T^*_w$. This defines an abstract continuity equation $\phi[w] = \rho$, or in differentiated form $d\phi_w j = s$, see Figure 1. Contrary to Sections 2 and 3, the second continuity equation may now also depend on w. In practice, the continuity mapping $\phi[w]$ depends on the initial state ρ_0, which we assume to be fixed.

- An L-function $\mathcal{L} : T\mathcal{W} \to \mathbb{R}_+$ on flux space (see below for the definition of L-functions). This function could be a dynamic large-deviation cost function corresponding to random fluxes in some microscopic system, but throughout this section it could also be a more general expression.

- An L-function $\hat{\mathcal{L}} : T\mathcal{X} \to \mathbb{R}_+$ on state space, related to the flux space L-function via

$$\hat{\mathcal{L}}(\rho, s) := \inf_{\substack{(w,j) \in T\mathcal{W}: \\ \phi[w]=\rho,\, d\phi[w]j=s}} \mathcal{L}(w, j). \tag{26}$$

  This relation is again inspired by large-deviation theory, where such relation holds due to the contraction principle [DZ87, Th. 4.2.1].

- Corresponding to the L-functions are their convex duals with respect to their second variable, i.e. $\mathcal{H} : T^*\mathcal{W} \to \mathbb{R}$ and $\hat{\mathcal{H}} : T^*\mathcal{X} \to \mathbb{R}$ with

$$\mathcal{H}(w, \zeta) := \sup_{j \in T_w} \langle \zeta, j \rangle - \mathcal{L}(w, j) \qquad \text{and} \qquad \hat{\mathcal{H}}(\rho, \xi) := \sup_{s \in T_\rho} \langle \xi, s \rangle - \mathcal{L}(\rho, \xi).$$

  We express assumptions in terms of these duals, since in practice they are often more explicitly given than their corresponding L-functions.

Recall that in the previous section we saw that $\mathcal{L}(w, j)$ depends on w through $\rho = \phi[w]$ only. This condition becomes slightly more complicated in non-flat spaces, for a number of reasons. Firstly, the flux j can not be kept fixed while changing w, unless one has a path-independent notion of parallel transport, i.e. the space is flat. Secondly, even if $\mathcal{L}(w, j)$ would not depend on w, this dependence could re-enter through the continuity equation in the infimum $\inf_{j: d\phi_w j=s} \mathcal{L}(w, j)$. Therefore, the condition that we need is that $\hat{\mathcal{L}}(\phi[w], s) = \inf_{\substack{j \in T_w: \\ d\phi_w j=s}} \mathcal{L}(w, j)$ for all $w \in \mathcal{W}, s \in T_{\phi[w]}$. It is easily seen that this condition is equivalent to the following flux invariance:

> for fixed $(\rho, \xi) \in T^*\mathcal{X}$, the function $\phi^{-1}[\rho] \ni w \mapsto \mathcal{H}(w, d\phi_w^\mathsf{T} \xi)$ does not depend on w. (27)

All manifolds and functionals are assumed to be sufficiently differentiable wherever needed. For a (differentiable) functional $F : \mathcal{W} \to \mathbb{R}$ (and similarly on flux space) we write $dF(w) \in T_w^*$ for the derivative, in the sense that on a curve $\frac{d}{dt} F(w_t) = \langle dF(w_t), \dot{w}_t \rangle$.

4.2 Definitions

We now define the notions of L-functions, dissipation potentials, GGS, pGGEN, Poisson operators and GGEN on flux space; the same concepts on state space are defined analogously. Naturally all notions are compatible with the exposition from Section 2 and 3.

Definition 4.1. *We say $\mathcal{L} : T\mathcal{W} \to \mathbb{R}_+$ is an L-function whenever for all $w \in \mathcal{W}$:*

  (i) $j \mapsto \mathcal{L}(w, j)$ is convex;

  (ii) $\mathcal{L}(w, j) = 0 \iff j = \mathcal{A}_\mathcal{L}(w)$ for some unique vector field $\mathcal{A}_\mathcal{L}$.

Due to the convexity, $\mathcal{L}$ is also the convex dual of $\mathcal{H}$. Central to GGS, pGGEN and GGEN is the notion of dissipation potentials:

Definition 4.2. *A function $\Psi : T\mathcal{W} \to \mathbb{R}_+$ is called a* dissipation potential *whenever for all $w \in \mathcal{W}$:*

  (i) $j \mapsto \Psi(w, j)$ is convex;

  (ii) $\Psi(w, 0) = 0$;

If these conditions hold, then the same conditions hold for the (pre-)dual dissipation potential

$$\Psi^*(w, \zeta) := \sup_{j \in T_w} \langle \zeta, j \rangle - \Psi(w, j). \tag{28}$$

We also say that (Ψ, Ψ^*) is a dissipation potential *pair whenever Ψ is a dissipation potential.*

Definition 4.3. *A generalised gradient system (GGS) is a triple $(\mathcal{W}, \Psi, \mathcal{F})$, where $\mathcal{W}$ is a differentiable manifold, $\mathcal{F} : \mathcal{W} \to \mathbb{R}$ and $\Psi : T\mathcal{W} \to \mathbb{R}_+$ is a dissipation potential. We say that an L-function $\mathcal{L}$ induces a GGS $(\mathcal{W}, \Psi, \mathcal{F})$ whenever for all $(w, j) \in T\mathcal{W}$:*

$$\mathcal{L}(w, j) = \Psi(w, j) + \Psi^*\big(w, -d\mathcal{F}(w)\big) + \langle d\mathcal{F}(w), j \rangle. \tag{29}$$

As explained in Subsection 3.3, by extending a GGS with an orthogonal drift one arrives at

Definition 4.4 ([KLMP18b])**.** *A* Generalised pre-Generic system *(pGGEN) is a quadruple $(\mathcal{W}, \Psi, \mathcal{F}, b)$, where $\mathcal{W}$ is a differentiable manifold, Ψ is a dissipation potential, $\mathcal{F} : \mathcal{W} \to \mathbb{R}$, and $b(w) \in T_w$ is a vector field such that:*

$$\langle d\mathcal{F}(w), b(w) \rangle = 0 \qquad \text{for all } w \in \mathcal{W}.$$

We say that an L-function induces a pGGEN $(\mathcal{W}, \Psi, \mathcal{F}, b)$ whenever for all $(w, j) \in T\mathcal{W}$:

$$\mathcal{L}(w, j) = \Psi\big(w, j - b(w)\big) + \Psi^*\big(w, -d\mathcal{F}(w)\big) + \langle d\mathcal{F}(w), j \rangle.$$

Finally, if the drift has the form of an Hamiltonian system that behaves more or less independently of the GGS part we arrive at a Generalised Generic system. In order to define this we first define:

Definition 4.5. *A linear operator $L : T^*\mathcal{W} \to T\mathcal{W}$ is called a Poisson structure if it satisfies the Jacobi identity*

$$\big\{\{\mathcal{F}_1, \mathcal{F}_2\}_L, \mathcal{F}_3\big\}_L + \big\{\{\mathcal{F}_2, \mathcal{F}_3\}_L, \mathcal{F}_1\big\}_L + \big\{\{\mathcal{F}_3, \mathcal{F}_1\}_L, \mathcal{F}_2\big\}_L = 0$$

for all $\mathcal{F}_{1,2,3} : \mathcal{W} \to \mathbb{R}$, where $\{\mathcal{F}_1, \mathcal{F}_2\}_L(w) := \langle d\mathcal{F}_1(w), L(w)d\mathcal{F}_2(w) \rangle$;

Jacobi's identity implies skew symmetry, i.e. $\langle \zeta_1, L(w)\zeta_2 \rangle = -\langle \zeta_2, L(w)\zeta_1 \rangle$ for all $w \in \mathcal{W}$, $\zeta_1, \zeta_2 \in T_w$; in particular one has $\langle \zeta, L(w)\zeta \rangle = 0$. Finally,

Definition 4.6 ([Mie11, Sect. 2.5]). *A generalised Generic system (GGEN) is defined as a quintuple* $(\mathcal{W}, \Psi, \mathcal{F}, L, \mathcal{E})$, *where* $\mathcal{W}$ *is a differentiable manifold,* Ψ *is a dissipation potential,* $\mathcal{E}, \mathcal{F} : \mathcal{W} \to \mathbb{R}$, $L : T^*\mathcal{W} \to T\mathcal{W}$ *is a Poisson structure, and the two non-interaction conditions are satisfied:*

$$L(w)d\mathcal{F}(w) = 0 \qquad \text{for all } w \in \mathcal{W}, \text{ and} \tag{30}$$

$$\Psi^*\big(w, \zeta + \lambda d\mathcal{E}(w)\big) = \Psi^*(w, \zeta) \qquad \text{for all } (w, \zeta) \in T^*\mathcal{W} \text{ and } \lambda \in \mathbb{R}. \tag{31}$$

We say that an L-function induces a GGEN $(\mathcal{W}, \Psi, \mathcal{F}, L, \mathcal{E})$ *whenever for all* $(w, j) \in T\mathcal{W}$:

$$\mathcal{L}(w, j) = \Psi\big(w, j - L(w)d\mathcal{E}(w)\big) + \Psi^*\big(w, -d\mathcal{F}(w)\big) + \langle d\mathcal{F}(w), j \rangle. \tag{32}$$

Remark 4.7. If a GGS, pGGEN or GGEN is given on a manifold $\mathcal{W}$, and the dissipation potentials are quadratic (as explained in Section 2.4), then one can use the positive definite operator $K(w)\xi := d_\zeta \Psi^*(w, \zeta)$ to define a new manifold, and redefine everything on this manifold. This allows to study the structures from a more geometric point of view. □

4.3 From L-functions to GGS, pGGEN and GGEN

We now recall some of the main results from [MPR14] and [KLMP18b], that give necessary and sufficient conditions for an L-function to induce a GGS or pGGEN. A similar result for GGEN does not exist, since an L-function does not uniquely determine a Poisson operator and Hamiltonian energy. However, from a pGGEN one can always construct a GGEN (in a non-unique way); for that result we refer the reader to [KLMP18b, Sec. 4].

Again, the following results are described, but not restricted to flux space.

Theorem 4.8 ([MPR14, Prop. 2.1 & Th. 2.1]). *Let* $\mathcal{L} : T\mathcal{W} \to \mathbb{R}_+$ *be an L-function with convex dual* $\mathcal{H}$, *and let* $\mathcal{F} : \mathcal{W} \to \mathbb{R}$ *be given. Then the following statements are equivalent:*

(i) $\mathcal{L}$ *induces a GGS* $(\mathcal{W}, \Psi, \mathcal{F})$ *for some dissipation potential* Ψ,

(ii) $\mathcal{H}(w, \zeta) = \Psi^*\big(w, \zeta - d_w\mathcal{F}(w)\big) - \Psi^*\big(w, -d\mathcal{F}(w)\big)$ *for some diss. pot.* Ψ^*,

(iii) $d_\zeta \mathcal{H}\big(w, d_w\mathcal{F}(w)\big) = 0.$ \hfill (33)

(iv) $d_j \mathcal{L}(w, 0) = d_w \mathcal{F}(w),$ \hfill (34)

In that case Ψ^* *(and indirectly* Ψ*) is uniquely determined by*

$$\Psi^*(w, \zeta) = \mathcal{H}\big(w, \zeta + d_w\mathcal{F}(w)\big) - \mathcal{H}\big(w, d_w\mathcal{F}(w)\big). \tag{35}$$

From condition (34) we see that $\mathcal{F}$, if it exists, is uniquely given up to constants. Note in particular that conditions (33) and (34) do not involve Ψ.

Theorem 4.9 ([KLMP18b, Th. 3.6]). *Let $\mathcal{L} : T\mathcal{W} \to \mathbb{R}_+$ be an L-function with convex dual $\mathcal{H}$, and let $\mathcal{F} : \mathcal{W} \to \mathbb{R}$ and a vector field $b(w) \in T_w$ be given for which $\langle d_w\mathcal{F}(w), b(w)\rangle = 0$. Then the following statements are equivalent:*

(i) $\mathcal{L}$ induces a pGGEN $(\mathcal{W}, \Psi, \mathcal{F}, b)$ for some dissipation potential Ψ,

(ii) $\mathcal{H}(w, \zeta) = \Psi^*\big(w, \zeta - d_w\mathcal{F}(w)\big) - \Psi^*\big(w, -d_w\mathcal{F}(w)\big) + \langle \zeta, b(w)\rangle$ *for some diss. pot. Ψ^*,*

$$\text{(36)}$$

(iii) $d_\zeta\mathcal{H}\big(w, d_w\mathcal{F}(w)\big) = b(w)$,

(iv) $d_j\mathcal{L}\big(w, b(w)\big) = d_w\mathcal{F}(w)$. $\hspace{6cm}$ (37)

In that case Ψ^* is uniquely determined by

$$\Psi^*(w, \zeta) = \mathcal{H}\big(w, \zeta + d_w\mathcal{F}(w)\big) - \mathcal{H}\big(w, d_w\mathcal{F}(w)\big) - \langle \zeta, b(w)\rangle. \tag{38}$$

From condition (36) we see that $\mathcal{H}$ must consist of a convex part and a linear part $\langle \zeta, b(w)\rangle$, so that the drift b is a priori and uniquely fixed by $\mathcal{H}$. Therefore $\mathcal{F}$ is again uniquely fixed by condition (37). This is different from the Generic setting; $Ld\mathcal{E}$ uniquely defines $d\mathcal{F}$ and vice versa, but the whole quintuple may not be unique. However, one can still state a Generic analogue of Theorems 4.8 and 4.9 as follows:

Proposition 4.10. *Let $\mathcal{L} : T\mathcal{W} \to \mathbb{R}_+$ be an L-function with convex dual $\mathcal{H}$, and let a Poisson structure $L : T^*\mathcal{W} \to T\mathcal{W}$ and energies $\mathcal{E}, \mathcal{F} : \mathcal{W} \to \mathbb{R}$ be given such that the non-interaction condition $Ld\mathcal{F} = 0$ holds. Then the following statements are equivalent:*

(i) $\mathcal{L}$ induces a GGEN $(\mathcal{W}, \Psi, \mathcal{F}, L, \mathcal{E})$ for some dissipation potential Ψ,

(ii) $d_s\mathcal{L}\big(w, L(w)d_w\mathcal{E}(w)\big) = d_w\mathcal{F}(w)$ *and*

$$\mathcal{L}(w, j) = \infty \qquad \text{for all } j \in T_w \text{ for which } \langle d\mathcal{E}, j\rangle \neq 0, \tag{39}$$

(iii) $d_\zeta\mathcal{H}\big(w, d_w\mathcal{F}(w)\big) = L(w)d_w\mathcal{E}(w)$ *and* $\hspace{4cm}$ (40)

$$\mathcal{H}\big(w, \zeta + \lambda d\mathcal{E}(w)\big) = \mathcal{H}(w, \zeta) \qquad \text{for all } (w, \zeta) \in T^*\mathcal{W} \text{ and } \lambda \in \mathbb{R}. \tag{41}$$

In that case Ψ^* is uniquely determined by

$$\Psi^*(w, \zeta) = \mathcal{H}\big(w, \zeta + d_w\mathcal{F}(w)\big) - \mathcal{H}\big(w, d_w\mathcal{F}(w)\big). \tag{42}$$

Proof. By Lemma 4.11 (see below), we can apply Theorem 4.8 to the shifted L-function $\tilde{\mathcal{L}}(w, j) := \mathcal{L}\big(w, j + L(w)d_w\mathcal{E}(w)\big)$ and back. This yields the three equivalences, apart from the other non-interaction condition (30). From the explicit formula (42) one finds that the missing non-interaction condition is equivalent to (39) and to (41). $\hspace{4cm}$ $\square$

The previous proof made use of the following lemma:

Lemma 4.11. *Let $L : T^*\mathcal{W} \to T\mathcal{W}$ be a Poisson operator, $\mathcal{E}, \mathcal{F} : \mathcal{W} \to \mathbb{R}$ be energies and Ψ be a dissipation potential such that the non-interaction conditions (30) and (31) hold. An L-function $\mathcal{L}$ induces an GGEN $(\mathcal{W}, \Psi, \mathcal{F}, L, \mathcal{E})$ if and only if the shifted L-function $\tilde{\mathcal{L}}(w, j) := \mathcal{L}\big(w, j + L(w)d_w\mathcal{E}(w)\big)$ induces the GGS $(\mathcal{W}, \Psi, \mathcal{F})$.*

Proof. Because of the non-interaction condition (31), the shift transforms relation (32) into (29), and analogously for the other direction. $\hspace{6cm}$ $\square$

4.4   Relation between structures in flux and state space

We now consider L-functions $\mathcal{L}$ and $\hat{\mathcal{L}}$ on flux and state space, and study how their induced structures are related.

Proposition 4.12. *Assume that an L-function* $\mathcal{L} : T\mathcal{W} \to \mathbb{R}_+$ *induces a pGGEN* $(\mathcal{W}, \Psi, \mathcal{F}, b)$ *where* $d\phi_w b(w) = 0$ *and*

$$\hat{\mathcal{F}}\big(\phi[w]\big) = \mathcal{F}(w), \qquad \text{(up to constants)}$$

for some $\hat{\mathcal{F}} : \mathcal{X} \to \mathbb{R}$. *Then the L-function* $\hat{\mathcal{L}}$ *given by* (26) *induces a GGS* $(\mathcal{X}, \hat{\Psi}, \hat{\mathcal{F}})$ *for some dissipation potential* $\hat{\Psi}$.

Proof. Since $d_w \hat{\mathcal{F}}(\phi[w]) = d\phi_w^\mathsf{T} d_\rho \hat{\mathcal{F}}(\phi[w])$ and $d\phi_w b(w) = 0$, we can rewrite

$$\hat{\mathcal{L}}(\rho, s) = \inf_{w \in \mathcal{W}:\phi[w]=\rho} \left(\inf_{j \in T_w:d\phi_w j=s} \Psi(w, j) \right) + \Psi^*\big(w, -d\phi_w^\mathsf{T} d_\rho \hat{\mathcal{F}}(\rho)\big) + \langle d_\rho \hat{\mathcal{F}}(\rho), s \rangle,$$

and because Ψ, Ψ^* is a dissipation potential pair, clearly

$$\hat{\mathcal{L}}(\rho, s) - \langle d_\rho \hat{\mathcal{F}}(\rho), s \rangle \geq \inf_{w \in \mathcal{W}:\phi[w]=\rho} \Psi^*\big(w, -d\phi_w^\mathsf{T} d_\rho \hat{\mathcal{F}}(\rho)\big) = \hat{\mathcal{L}}(\rho, 0).$$

This is equivalent to $d_\rho \hat{\mathcal{F}}(\rho) \in \partial_s \hat{\mathcal{L}}(\rho, 0) = \{d_s \hat{\mathcal{L}}(\rho, 0)\}$, which by Theorem 4.8 implies that $\hat{\mathcal{L}}$ induces a GGS $(\mathcal{X}, \hat{\Psi}, \hat{\mathcal{F}})$ for some $\hat{\Psi}$. $\qquad\square$

In the above proposition, $\mathcal{F}$ also on w through $\phi[w]$ only, which is a very physical assumption. It does imply however, that the equilibria of the flux gradient system can only be unique up to the kernel of ϕ; this kernel can be interpreted as a generalisation of divergence-free vector fields.

A natural question is now whether we can turn the statement of Proposition 4.12 around. Indeed, if the invariance condition (27) holds and we restrict to pGGEN with 'divergence-free drifts', then the statement becomes an equivalence, and we have an explicit relation between the flux and state dissipation potentials. This is a stronger version of the statement in [Ren17, Prop. 4.7], where we related GGS to so-called 'force structures'.

Theorem 4.13. *Assume that an L-function* $\mathcal{L} : T\mathcal{W} \to \mathbb{R}_+$ *with corresponding dual* $\mathcal{H}$ *satisfies the invariance condition* (27)*, and let the L-function* $\hat{\mathcal{L}} : \mathcal{X} \to \mathbb{R}_+$ *be given by* (26)*. Then* $\mathcal{L}$ *induces a pGGEN* $(\mathcal{W}, \Psi, \mathcal{F}, b)$ *with* $d\phi_w b(w) = 0$ *and* $\hat{\mathcal{F}} \circ \phi$ *for some* $\hat{\mathcal{F}} : \mathcal{X} \to \mathbb{R}$ *if and only if* $\hat{\mathcal{L}}$ *induces a GGS* $(\mathcal{X}, \hat{\Psi}^*, \hat{\mathcal{F}})$. *In that case the dissipation potentials* $\hat{\Psi}$ *and* $\hat{\Psi}^*$ *are related to* Ψ *and* Ψ^* *through*

$$\hat{\Psi}\big(\phi[w], s\big) = \inf_{\substack{j \in T_w: \\ d\phi_w j=s}} \Psi(w, j) \tag{43a}$$

$$\hat{\Psi}^*\big(\phi[w], \xi\big) = \Psi^*(w, d\phi_w^\mathsf{T} \xi). \tag{43b}$$

Proof. Assume that $\mathcal{L}$ induces a pGGEN $(\mathcal{W}, \Psi, \mathcal{F}, b)$ with $d\phi_w b(w) = 0$ and $\hat{\mathcal{F}} \circ \phi$. Since by assumption $\mathcal{H}(w, d\phi_w^\mathsf{T} \xi)$ does not depend on $w \in \phi^{-1}[\rho]$, by (35) the expression $\Psi^*(w, d\phi_w^\mathsf{T} \xi)$ is

also invariant under this choice. Therefore we can define $\hat{\Psi}^*(\rho, \xi)$ by (43b); it is easily checked that it its convex dual is given by (43a). We can write:

$$
\begin{aligned}
\hat{\mathcal{L}}(\rho, s) &= \inf_{d\phi_w j = s} \left\{ \Psi\big(w, j - b(w)\big) + \Psi^*\big(w, -d_w \mathcal{F}(w)\big) + \langle d_w \mathcal{F}(w), j \rangle \right\} \\
&= \inf_{d\phi_w j = s} \Psi(w, j) + \Psi^*\big(w, -d\phi_w^\mathsf{T} d_\rho \hat{\mathcal{F}}(\rho)\big) + \langle d_\rho \hat{\mathcal{F}}(\rho), s \rangle \\
&= \hat{\Psi}(\rho, s) + \hat{\Psi}^*\big(\rho, -d_\rho \hat{\mathcal{F}}(\rho)\big) + \langle d_\rho \hat{\mathcal{F}}(\rho), s \rangle,
\end{aligned}
$$

and hence $\hat{\mathcal{L}}$ induces the GGS $(\mathcal{X}, \hat{\Psi}, \hat{\mathcal{F}})$, which is unique by Theorem (4.8).

For the other direction, assume that $\hat{\mathcal{L}}$ induces a GGS $(\mathcal{X}, \hat{\Psi}^*, \hat{\mathcal{F}})$. Define $b(w) := d_\zeta \mathcal{H}\big(w, d_w \mathcal{F}(w)\big)$. Then by the invariance condition $\mathcal{H}(w, d\phi_w^\mathsf{T} \xi) = \hat{\mathcal{H}}(\rho, \xi)$ and by (33):

$$
\begin{aligned}
d\phi_w b(w) = d\phi_w d_\zeta \mathcal{H}\big(w, d_w \mathcal{F}(w)\big) = d\phi_w d_\zeta \mathcal{H}\big(w, d\phi_w d_\rho \hat{\mathcal{F}}(\rho)\big) \\
= d_\xi \mathcal{H}\big(\rho, d_\rho \hat{\mathcal{F}}(\rho)\big) = 0.
\end{aligned}
$$

Now define Ψ^* by (38). In particular $\Psi^*\big(w, -d\mathcal{F}(w)\big) = -\mathcal{H}\big(w, d_w \mathcal{F}(w)\big)$ since $\langle d_w \mathcal{F}(w), b(w) \rangle = 0$ and, by the definition of L-functions, $\mathcal{H}(w, 0) = 0$. Then (36) holds and hence by Theorem 4.9 the flux L-function $\mathcal{L}$ induces the pGGEN $(\mathcal{W}, \Psi, \mathcal{F}, b)$.

$\square$

Due to the non-uniqueness of induced GGEN systems, there is no similar 'if and only if' statement for the Generic setting. Nevertheless, in one direction, the GGEN analogue of Theorem 4.13 is:

Proposition 4.14. *Assume that an L-function $\mathcal{L} : T\mathcal{W} \to \mathbb{R}$ induces a GGEN $(\mathcal{W}, \Psi, \mathcal{F}, L, \mathcal{E})$ where*

$$
\hat{\mathcal{F}}\big(\phi[w]\big) = \mathcal{F}(w) \qquad \text{and} \qquad \hat{\mathcal{E}}\big(\phi[w]\big) = \mathcal{E}(w), \qquad \text{(up to constants)} \tag{44}
$$

for some $\hat{\mathcal{F}}, \hat{\mathcal{E}} : \mathcal{X} \to \mathbb{R}$, and that

$$
d\phi_w L(w) d\phi_w^\mathsf{T} =: \hat{L}(\rho) \text{ depends on } w \text{ through } \rho = \phi[w] \text{ only.} \tag{45}
$$

Then the L-function $\hat{\mathcal{L}}(\rho, s)$ given by (26) induces a GGEN $(\mathcal{X}, \hat{\Psi}, \hat{\mathcal{F}}, \hat{L}, \hat{\mathcal{E}})$ for some dissipation potential $\hat{\Psi}$.

If in addition, $\mathcal{H}$ satisfies the invariance principle (27), then $\hat{\Psi}$ and $\hat{\Psi}^*$ are related to Ψ and Ψ^* through (43a) and (43b).

Proof. We again apply Lemma 4.11 to transform the problem into a problem of GGSs. Indeed, the L-function $\tilde{\mathcal{L}}(w, j) := \mathcal{L}\big(w, j + L(w) d_w \mathcal{E}(w)\big)$ induces the GGS $(\mathcal{W}, \Psi, \mathcal{F})$. Hence by Proposition 4.12, a GGS $(\mathcal{X}, \hat{\Psi}, \hat{\mathcal{F}})$ (for some $\hat{\Psi}$) is induced by the L-function

$$
\begin{aligned}
\hat{\tilde{\mathcal{L}}}(\rho, s) &:= \inf_{\phi[w] = \rho} \inf_{d\phi_w j = s} \tilde{\mathcal{L}}(w, j) \\
&= \inf_{\phi[w] = \rho} \inf_{d\phi_w(j - L(w) d_w \mathcal{E}(w)) = s} \mathcal{L}(w, j) \\
&= \hat{\mathcal{L}}\big(\rho, s + \hat{L}(\rho) d_\rho \mathcal{E}(\rho)\big).
\end{aligned}
$$

If we can now validate that $\hat{L}$ is a Poisson structure, and that the non-interaction conditions are satisfied for $(\mathcal{X}, \hat{\Psi}, \hat{\mathcal{F}}, \hat{L}, \hat{\mathcal{E}})$, then Lemma 4.11 concludes the proof.

For the Poisson structure, note that, for any smooth $\hat{\mathcal{F}}_1, \hat{\mathcal{F}}_2 : \mathcal{X} \rightarrow \mathbb{R}$, the Lie bracket remains unaltered:

$$\{\hat{\mathcal{F}}_1, \hat{\mathcal{F}}_2\}_{\hat{L}}\big(\phi[w]\big) = \{\mathcal{F}_1 \circ \phi, \mathcal{F}_2 \circ \phi\}_L(w).$$

The non-interaction condition (31) is clearly satisfied as for any $w \in \mathcal{W}$ we have

$$\hat{L}\big(\phi[w]\big)d_\rho\hat{\mathcal{F}}\big(\phi[w]\big) = d\phi_w L(w)d\phi_w^\mathsf{T} d_\rho\hat{\mathcal{F}}\big(\phi[w]\big) = d\phi_w L(w)d_w\mathcal{F}(w) = 0.$$

To check the other non-interaction condition (30) we use the equivalent formulation (41). Indeed, for any $(\rho, \xi) \in T^*\mathcal{X}$ and $\lambda \in \mathbb{R}$,

$$\hat{\mathcal{H}}\big(\rho, \xi + \lambda d_\rho d\hat{\mathcal{E}}(\rho)\big) = \sup_{\phi[w]=\rho} \mathcal{H}\Big(w, d\phi_w^\mathsf{T}(\xi + \lambda d_\rho\hat{\mathcal{E}}(\phi[w]))\Big)$$

$$= \sup_{\phi[w]=\rho} \mathcal{H}\big(w, d\phi_w^\mathsf{T}\xi + \lambda d_w\mathcal{E}(w)\big) = \sup_{\phi[w]=\rho} \mathcal{H}\big(w, d\phi_w^\mathsf{T}\xi\big) = \hat{\mathcal{H}}\big(\rho, \xi\big).$$

Finally, if $\mathcal{H}$ satisfies the invariance property (27), then Proposition 4.12 yields relation (43a) and (43b). $\qquad\square$

Condition (45) is in a sense a natural one as the following result shows:

Proposition 4.15. *Assume that L-functions $\mathcal{L}$ and $\hat{\mathcal{L}}$ induce two GGENs $(\mathcal{W}, \Psi, \mathcal{F}, L, \mathcal{E})$, and respectively $(\mathcal{X}, \hat{\Psi}, \hat{\mathcal{F}}, \hat{L}, \hat{\mathcal{E}})$, where $\mathcal{F}, \mathcal{E}$ are related to $\hat{\mathcal{F}}, \hat{\mathcal{E}}$ by (44), and $\mathcal{H}$ satisfies the invariance property (27). Then*
$$\hat{L}\big(\phi[w]\big)d_\rho\hat{\mathcal{E}}\big(\phi[w]\big) = d\phi_w L(w)d\phi_w^\mathsf{T} d_\rho\hat{\mathcal{E}}\big(\phi[w]\big).$$

Proof. For any $w \in \mathcal{W}$ and $\rho = \phi[w]$, we may write $\hat{\mathcal{H}}(\rho, \xi) = \mathcal{H}(w, d\phi_w^\mathsf{T}\xi)$, and so by (40),

$$\hat{L}(\rho)d_\rho\hat{\mathcal{E}}(\rho) = d_\xi\hat{\mathcal{H}}\big(\rho, -d_\rho\hat{\mathcal{F}}(\rho)\big) = d\phi_w d_\zeta\mathcal{H}\big(w, -d\phi_w d_\rho\hat{\mathcal{F}}(\phi[w])\big)$$

$$= d\phi_w L(w)d\phi_w d_\rho\hat{\mathcal{E}}(\rho).$$

$\qquad\square$

5 Diffusion

In this section we apply the ideas of the previous section to a model for diffusion. The flux structure related to diffusion is interesting in its own right, and as far as the author is aware, previously unknown. In the next section we show how this model can be coupled with the results of Section 3 to obtain flux and state GGSs/pGGENs for reaction-diffusion systems.

Typical microscopic models of diffusion consist of Brownian particles, or discretised versions thereof, like random walkers or an exclusion process. Since empirical fluxes are a bit easier to define on a lattice, we focus on independent random walkers[6].

---

[6]With the scaling that we use, the system of independent random walkers is 'exponentially equivalent' to a system of Brownian motions, meaning they share the same hydrodynamic limit and large deviations.

5.1 Diffusing particle system

This microscopic actually has two scaling parameters: the number of particles, which we denote by V for consistency with the rest of this paper, and the lattice spacing ϵ_V. The speed with which $\epsilon_V \to 0$ as $V \to \infty$ is irrelevant. For fixed V, let $\left(X_{t,i}\right)_{i=1}^{V}$ be independent random walkers on the lattice $(\epsilon_V \mathbb{Z})^d$ with jump rate ϵ^{-2}. Define the random concentration and (integrated, net) flux by:

$$\rho_t^{(V)}(dx) := \frac{1}{V}\#\{i = 1, \ldots, V : X_{t,i} \in dx\}, \qquad \text{and}$$

$$\bar{W}_{t,l}^{(V)}(dx) := \frac{\epsilon_V}{V}\#\big\{\text{jumps } \tilde{x} \text{ to } \tilde{x} + \epsilon_V \mathbb{1}_l \text{ occurred in } (0, t) : \tilde{x} + \tfrac{1}{2}\epsilon_V \mathbb{1}_l \in dx\big\}$$
$$- \frac{\epsilon_V}{V}\#\big\{\text{jumps } \tilde{x} + \epsilon_V \mathbb{1}_l \text{ to } \tilde{x} \text{ occurred in } (0, t) : \tilde{x} + \tfrac{1}{2}\epsilon_V \mathbb{1}_l \in dx\big\}.$$

As usual dx denotes a spatial area, possibly a small box surrounding one lattice site, and $\rho_t^{(V)}$ and $\bar{W}_{t,l}^{(V)}$ are measures.

Now $\rho_t^{(V)}(dx)$ measures the number of particles present in (lattice points in) an area dx, while $\bar{W}_{t,l}^{(V)}(dx)$ measures the *net* number of particles that have jumped through all midpoints in dx, in direction $\mathbb{1}_l$, for $l = 1, \ldots, d$, see Figure 2. Note that both $\rho^{(V)}$ and $\bar{W}^{(V)}$ are defined as measures on the lattice with shrinking distance ϵ_V between lattice points; this measure-valued formulation is needed to pass to a continuum limit later on.
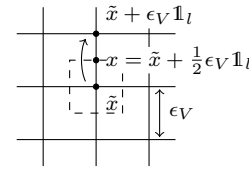

Figure 2: A jump through midpoint x in the direction with unit vector $\mathbb{1}_l$.

The concentrations and fluxes are related by the V-dependent continuity equation:

$$\rho_t^{(V)}(dx) = \phi^{(V)}[\bar{W}_t^{(V)}](dx) := \big(\rho_0^{(V)} - \operatorname{div}^{(\epsilon_V)} \bar{W}_t^{(V)}\big)(dx)$$

$$:= \rho_0^{(V)}(dx) - \frac{1}{\epsilon_V}\sum_{l=1}^{d}\big[\bar{W}_{t,l}(dx + \tfrac{1}{2}\epsilon + V\mathbb{1}_l) - \bar{W}_{t,l}(dx - \tfrac{1}{2}\epsilon_V \mathbb{1}_l)\big]. \tag{46}$$

Using that $\big(\tfrac{1}{2}\epsilon_V + \epsilon_V \mathbb{Z}\big)^d \subset \mathbb{R}^d$, the integrated flux $\bar{W}_t^{(V)}$ is a Markov process in $\mathcal{M}(\mathbb{R}^d)$ with generator

$$(\mathcal{Q}^{(V)}f)(\bar{w}) := \tag{47}$$

$$\frac{V}{\epsilon_V^2}\int \phi^{(V)}[\bar{w}](dx)\sum_{l=1}^{d}\big(f(\bar{w} - \tfrac{\epsilon_V}{V}\delta_{x-(\epsilon_V/2)\mathbb{1}_l}) - 2f(\bar{w}) + f(\bar{w} + \tfrac{\epsilon_V}{V}\delta_{x+(\epsilon_V/2)\mathbb{1}_l})\big).$$

Here, the factor V/ϵ_V^2 comes from the time scaling ϵ_V^2, together with the fact that we have $V\rho(dx)$ independent particles to choose from.

5.2 Limit and large deviations

For a test function $f \in C_b^1\big(\mathcal{M}(\mathbb{R}^d)\big)$, we set:

$$df(\bar{w})_l(x) := \lim_{\tau \to 0}\frac{f(\bar{w} + \tau\delta_x e_l) - f(\bar{w})}{\tau}.$$

The continuity equation (46) converges to the limit continuity equation (with the usual divergence operator):

$$\rho_t(dx) = \phi[\bar{w}_t](dx) := \rho_0(dx) - \mathrm{div}_x\, \bar{w}_t(dx). \tag{48}$$

With this notation, as $V \to \infty$ the generator (47) converges to (if $\phi[\bar{w}](dx) = \phi[\bar{w}](x)\, dx$):

$$(\mathcal{Q}^{(\infty)} f)(\bar{w}) := \int \mathrm{div}_x\, df(\bar{w})(x)\, \phi[\bar{w}](dx) = -\int df(\bar{w})(x) \cdot \nabla_x\, \phi[\bar{w}](x)\, dx.$$

As in Subsection 2.3, the limit generator depends on derivatives of the test function only, and so the process $\bar{W}_t^{(V)}$ converges (pathwise in probability) to the deterministic path satisfying Fick's Law:

$$\dot{\bar{w}}_t = -\nabla_x\, \phi[\bar{w}_t].$$

Naturally, combining this equation with the continuity equation (48) yields the diffusion equation for the empirical measure:

$$\dot{\rho}_t = \tfrac{d}{dt}\phi[\bar{w}_t] = d\phi_{\bar{w}_t}\dot{\bar{w}}_t = \Delta_x\, \phi[\bar{w}_t] = \Delta_x\, \rho_t.$$

Similarly, we derive the dynamic large deviations by studying the non-linear generator:

$$
\begin{aligned}
(\mathcal{H}^{(V)} f)(\bar{w}) &:= \frac{1}{V}\mathrm{e}^{-Vf(\bar{w})}\big(\mathcal{Q}^{(V)}\mathrm{e}^{Vf}\big)(\bar{w}) \\
&= \frac{1}{\epsilon_V^2}\int \phi^{(V)}[\bar{w}](dx) \times \\
&\quad \sum_{l=1}^{d} \mathrm{e}^{Vf(\bar{w}-\frac{\epsilon_V}{V}\delta_{x-(\epsilon_V/2)\mathbb{1}_l})-Vf(\bar{w})} - 2 + \mathrm{e}^{Vf(\bar{w}+\frac{\epsilon_V}{V}\delta_{x+(\epsilon_V/2)\mathbb{1}_l})-Vf(w)} \\
&\xrightarrow{V\to\infty} \int\big(\mathrm{div}_x\, df(\bar{w})(x) + \big|df(\bar{w})(x)\big|^2\big)\, \phi[\bar{w}](dx),
\end{aligned}
\tag{49}
$$

which follows from expanding the exponentials (with order-ϵ_V exponents) up to second order. Then the following large-deviation principle on flux space holds:

$$\mathrm{Prob}^{(V)}\big(\bar{W}_{(\cdot)}^{(V)} \approx \bar{w}_{(\cdot)}\big) \overset{V\to\infty}{\sim} \mathrm{e}^{-V\int_0^T \mathcal{L}(\bar{w}_t,\dot{\bar{w}}_t)\, dt},$$

with

$$\mathcal{H}(\bar{w}, \bar{\zeta}) := \int\big(\mathrm{div}_x\, \bar{\zeta}(x) + |\bar{\zeta}(x)|^2\big)\phi[\bar{w}](dx) = \|\bar{\zeta}\|_{L^2(\phi[\bar{w}])}^2 - \langle\bar{\zeta}, \nabla_x\, \phi[\bar{w}]\rangle, \quad \text{and} \tag{50}$$

$$\mathcal{L}(\bar{w}, \bar{j}) := \sup_{\bar{\zeta}}\, \langle\bar{\zeta}, \bar{j}\rangle - \mathcal{H}(\bar{w}, \bar{\zeta}) = \tfrac{1}{4}\|\bar{j} + \nabla_x\, \phi[\bar{w}]\|_{L^2(1/\phi[\bar{w}])}^2.$$

Note that $d\phi_{\bar{w}} = -\mathrm{div}_x$ is independent of $\bar{w}$, and by (50), the invariance condition (27) is satisfied. Hence by the contraction principle [DZ87, Th. 4.2.1], one obtains the large-deviation principle corresponding to the states (empirical measures):

$$\mathrm{Prob}^{(V)}\big(\rho_{(\cdot)}^{(V)} \approx \rho_{(\cdot)}\big) \overset{V\to\infty}{\sim} \mathrm{e}^{-V\int_0^T \hat{\mathcal{L}}(\rho_t,\dot{\rho}_t)\, dt},$$

with

$$\hat{\mathcal{L}}(\phi[\bar{w}], s) := \inf_{\bar{j}\in T_{\bar{w}}:\, -\mathrm{div}_x\, \bar{j}=s} \mathcal{L}(\bar{w}, \bar{j}) = \tfrac{1}{4}\|s - \Delta_x\, \phi[\bar{w}]\|_{\dot{H}^{-1}(\phi[\bar{w}])}^2, \tag{51}$$

using the notation $\|s\|_{\dot{H}^{-1}(\rho)}^2 := \sup_\xi 2s \cdot \xi - \|\xi\|_{\dot{H}^1(\rho)}^2 := \sup_\xi 2s \cdot \xi - \|\nabla_x\, \xi\|_{L^2(\rho)}^2.$

5.3 Induced GGSs in flux and state space

We can now apply Theorem 4.8 to extract a GGS from the L-function $\mathcal{L}$. We first choose the 'naive' flat manifold of non-negative vector measures $\mathcal{W} := \mathcal{M}_+(\mathbb{R}^d; \mathbb{R}^d)$ (equipped with the flat total variation metric). It is easily checked that condition (33) holds for the free energy given by [7]:

$$\mathcal{F}(\bar{w}) := \frac{1}{2} \int \phi[\bar{w}](dx) \log \phi[\bar{w}](x) - \phi[\bar{w}](dx), \tag{52}$$

where we identify $\phi[\bar{w}](dx) = \phi[\bar{w}](x)\,dx$, and we implicitly set $\mathcal{F}(w) = \infty$ whenever the measure is not absolutely continuous. The dissipation potentials are obtained from (35) and (28), which yields

$$\Psi^*(\bar{w}, \bar{\zeta}) := \|\bar{\zeta}\|^2_{L^2(\phi[\bar{w}])} \qquad \text{and} \qquad \Psi(\bar{w}, \bar{\jmath}) := \frac{1}{4} \|\bar{\jmath}\|^2_{L^2(1/\phi[\bar{w}])}. \tag{53}$$

Theorem 4.8 states that $\mathcal{L}$ induces the GGS $(\mathcal{M}_+(\mathbb{R}^d), \Psi, \mathcal{F})$ on flux space.

For the state space, it is well-known that the state L-function (51) induces the entropy-Wasserstein gradient flow of the entropy functional [ADPZ11, DLR13, EMR15, MPR14]. By the theory developed in Section 4, we can now see how this gradient structure is related to the flux gradient structure. Indeed, the flux free energy $\mathcal{F}$ depends on state only, i.e. $\mathcal{F}(\bar{w}) = \hat{\mathcal{F}}(\phi[\bar{w}]))$, where $\hat{\mathcal{F}}(\rho) = \frac{1}{2} \int \rho(dx) \log \rho(x) - \rho(dx)$, and so by Proposition 4.12 the state L-function $\hat{\mathcal{L}}$ induces a GGS driven by $\hat{\mathcal{F}}$. Moreover, since the invariance condition (27) holds, the dissipation potentials are related by (43):

$$\hat{\Psi}^*(\rho, \xi) = \|\nabla_x \xi\|^2_{L^2(\rho)} =: \|\xi\|^2_{\mathring{H}^1(\rho)},$$

$$\hat{\Psi}(\rho, s) = \inf_{-\operatorname{div}_x \bar{\jmath} = s} \frac{1}{4} \|\bar{\jmath}\|^2_{L^2(1/\rho)} =: \frac{1}{4} \|s\|^2_{\mathring{H}^{-1}(\rho)}.$$

5.4 A new geometry

The form of the dissipation potential Ψ and $\hat{\Psi}$ suggests that it is more natural to use different manifolds in the spirit of Remark 4.7. For the state space this points to the space $\mathcal{W} := \mathcal{P}_2(\mathbb{R}^d) = \{\rho \in \mathcal{P}(\mathbb{R}^d) : \int x^2 \rho(dx) < \infty\}$ of probability measures of finite second moment space, equipped with the Monge-Kantorovich-Wasserstein metric:

$$d_{\mathcal{P}_2}(\rho_0, \rho_1)^2 := \inf_{\substack{\gamma \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d): \\ \gamma(\cdot \times \mathbb{R}^d) = \rho_0(\cdot) \\ \gamma(\mathbb{R}^d \times \cdot) = \rho_1(\cdot)}} \iint_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \, \gamma(dx\,dy).$$

with tangent and cotangent space $T_\rho = \mathring{H}^{-1}(\rho)$ and $T_\rho^* = \mathring{H}^1(\rho)$. For this setting, the inverse metric tensor $K_{\mathcal{P}_2}(\rho) : T_\rho^* \to T_\rho$ is known by the Benamou-Brenier formula [Vil09, Th. 8.1] to be $K_{\mathcal{P}_2}(\rho)\xi := -2 \operatorname{div}_x \rho \nabla_x \xi$, so that the GGS is indeed the entropy-Wasserstein gradient flow [JKO98]:

$$\Delta_x \rho_t = \dot{\rho}_t = d_\xi \hat{\Psi}^*\big(\rho_t, -d\hat{\mathcal{F}}(\rho_t)\big) = -2K_{\mathcal{P}_2}(\rho_t)d\hat{\mathcal{F}}(\rho_t) =: -2 \operatorname{grad}_{\rho_t} \hat{\mathcal{F}}(\rho_t).$$

---

[7]This expression can again be seen as a relative entropy, cf. (21), but now with respect to the Lebesgue measure, where the measure of the whole space – in this case infinity – is omitted. See also [MPR14, Prop. 3.2] for a general result in locally finite measure spaces.

Motivated by this observation we can take for the flux manifold the space of signed vector measures of finite *first* moment $\mathcal{M}_1(\mathbb{R}^d; \mathbb{R}^d) := \{\bar{w} \in \mathcal{M}(\mathbb{R}^d; \mathbb{R}^d) : \int |x| \, |\bar{w}|(dx) < \infty\}$. This choice guarantees that the corresponding states have finite second moment (once $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$):

$$\int |x|^2 \, \rho(dx) = \int |x|^2 \, \rho_0(dx) + \int x \cdot \bar{w}(dx) - \textit{non-neg. bnd. term.} < \infty.$$

Moreover, we can now use the dissipation potential to construct a natural metric on $\mathcal{Y}$:

$$d_{\mathcal{M}_1}(\bar{w}_0, \bar{w}_1)^2 := \inf_{\substack{\tilde{w}:(0,1) \to \mathcal{W}: \\ \tilde{w}_0 = \bar{w}_0, \, \tilde{w}_1 = \bar{w}_1}} \int_0^1 \|\dot{\tilde{w}}_t\|^2_{L^2(1/(\rho_0 - \mathrm{div}_x \tilde{w}_t))} \, dt, \tag{54}$$

where the infimum runs over paths of fluxes for which $\rho_0 - \mathrm{div} \, \tilde{w}_t$ remains non-negative. The corresponding tangent and cotangent spaces (in the interior of the domain) are simply $T_{\bar{w}} = L^2(1/(\rho_0 - \mathrm{div}_x \bar{w}))$ and $T_{\bar{w}}^* = L^2(\rho_0 - \mathrm{div}_x \bar{w})$ and the inverse metric tensor $K_{\mathcal{M}_1}(\bar{w}) : T_{\bar{w}}^* \to T_{\bar{w}}$ is $K_{\mathcal{M}_1}(\bar{w})\bar{\zeta} = \bar{\zeta}/(\rho_0 - \mathrm{div}_x \bar{w})$. This yields an interesting geometry in flux space, which, as far as the author is aware, is still unknown in the literature.

6 A simple reaction-diffusion model

We now combine the models from Section 3 and 5 to study reaction-diffusion models in flux and state space. The stochastic particle system will now consist of 'reacting random walkers'. It is known that, if the reaction networks include reactions of different orders (unimolecular, bimolecular, etc.) and we only allow particles to react if the required number of particles are present within the same site/compartment, then the model may not converge to the expected reaction-diffusion equation [PP15]. The reason behind this is that for a multimolecular reaction, it becomes very unlikely that the required amount of reactants are all within one site/compartment; different order reactions would require different scalings. This is beyond the scope of the current paper. However we can already illustrate the combination of reaction and transport fluxes for a simple system of unimolecular equations of the type:

$$\mathsf{A} \xrightarrow{\kappa_{\mathrm{fw}}} \mathsf{B} \qquad\qquad \text{and} \qquad\qquad \mathsf{B} \xrightarrow{\kappa_{\mathrm{bw}}} \mathsf{A}.$$

In this section we consider GGSs only, hence we shall always consider net rather than one-way fluxes.

6.1 Reacting and diffusing particle system

Since we consider unimolecular reactions only, we can take *independent* reacting random walkers on the scaled lattice $(\epsilon_V \mathbb{Z})^d$, where each reaction occurs locally at each lattice site with rate κ_{fw} or κ_{bw} respectively, so that

$$\frac{1}{V} \lambda_{\mathrm{fw}}^{(V)}(\rho(dx)) \equiv \kappa_{\mathrm{fw}} \rho_{\mathsf{A}}(dx) \qquad\qquad \text{and} \qquad\qquad \frac{1}{V} \lambda_{\mathrm{bw}}^{(V)}(\rho(dx)) \equiv \kappa_{\mathrm{bw}} \rho_{\mathsf{B}}(dx).$$

For the transport mechanism, we assume that the two species A, B hop to neighbouring lattice sites with rates D_{A} and D_{B} respectively.

As before we consider the random concentrations, as well as the integrated (net) fluxes, where we now distinguish between <u>t</u>ransport fluxes and <u>r</u>eaction fluxes. If $X_{t,i} \in (\epsilon_V \mathbb{Z}^d) \subset \mathbb{R}^d$ is the position

and $Y_{t,i} \in \mathcal{Y} = \{A, B\}$ is the species of the i-th particle, then

$$\rho_{t,y}^{(V)}(dx) := \frac{1}{V} \#\{i = 1, \ldots, V : X_{t,i} \in dx \text{ and } Y_{t,i} = y\},$$

$$\bar{W}_{t,\text{tr},y,l}^{(V)}(dx) :=$$

$$\frac{\epsilon_V}{V} \#\big\{\text{jumps } \tilde{x} \text{ to } \tilde{x} + \epsilon_V \mathbb{1}_l \text{ of species } y \text{ occurred in } (0, t) : \tilde{x} + \tfrac{1}{2}\epsilon_V \mathbb{1}_l \in dx\big\}$$

$$- \frac{\epsilon_V}{V} \#\big\{\text{jumps } \tilde{x} + \epsilon_V \mathbb{1}_l \text{ to } \tilde{x} \text{ of species } y \text{ occurred in } (0, t) : \tilde{x} + \tfrac{1}{2}\epsilon_V \mathbb{1}_l \in dx\big\},$$

$$\bar{W}_{t,\text{re}}^{(V)}(dx) := \frac{1}{V} \#\big\{\text{forward reactions } r \text{ occurred in } (0, t) \text{ and in area } dx\big\},$$

$$- \frac{1}{V} \#\big\{\text{backward reactions } r \text{ occurred in } (0, t) \text{ and in area } dx\big\}.$$

The concentrations and fluxes are again related by a continuity equation:

$$\rho_{t,y}^{(V)}(dx) = \phi^{(V)}[\bar{W}_t^{(V)}]_y(dx) := \big(\rho_0^{(V)} - \text{div}^{(\epsilon_V)} \bar{W}_{t,\text{tr}}^{(V)} + \Gamma \bar{W}_{t,\text{re}}^{(V)}\big)_y(dx), \tag{55}$$

where the discrete divergence is as in (46), and $\Gamma = (-1, 1)$ is the matrix consisting of one state change vector corresponding to a forward reaction.

The pair $\bar{W}_t^{(V)} := (\bar{W}_{t,\text{tr}}^{(V)}, \bar{W}_{t,\text{re}}^{(V)})$ is then a Markov process with generator

$$(\mathcal{Q}^{(V)}f)(\bar{w}_\text{tr}, \bar{w}_\text{re}) := (\mathcal{Q}_\text{tr}^{(V)}f)(\bar{w}_\text{tr}, \bar{w}_\text{re}) + (\mathcal{Q}_\text{re}^{(V)}f)(\bar{w}_\text{tr}, \bar{w}_\text{re}),$$

where

$$(\mathcal{Q}_\text{tr}^{(V)}f)(\bar{w}_\text{tr}, \bar{w}_\text{re})$$

$$:= \frac{V}{\epsilon_V^2} \sum_{y=A,B} D_y \int \phi^{(V)}[\bar{w}]_y(dx) \sum_{l=1}^{d} \big(f(\bar{w}_\text{tr} - \tfrac{\epsilon_V}{V}\delta_{x-(\epsilon_V/2)\mathbb{1}_l}\mathbb{1}_y, \bar{w}_\text{re})$$

$$- 2f(\bar{w}_\text{tr}, \bar{w}_\text{re}) + f(\bar{w}_\text{tr} + \tfrac{\epsilon_V}{V}\delta_{x+(\epsilon_V/2)\mathbb{1}_l}\mathbb{1}_y, \bar{w}_\text{re})\big),$$

$$(\mathcal{Q}_\text{tr}^{(V)}f)(\bar{w}_\text{tr}, \bar{w}_\text{re})$$

$$:= V\kappa_\text{fw} \int \phi^{(V)}[\bar{w}]_A(dx)\big(f(\bar{w}_\text{tr}, \bar{w}_\text{re} + \tfrac{1}{V}\delta_x) - f(\bar{w}_\text{tr}, \bar{w}_\text{re})\big)$$

$$+ V\kappa_\text{bw} \int \phi^{(V)}[\bar{w}]_B(dx)\big(f(\bar{w}_\text{tr}, \bar{w}_\text{re} - \tfrac{1}{V}\delta_x) - f(\bar{w}_\text{tr}, \bar{w}_\text{re})\big).$$

6.2 Limit and large deviations

By the same procedure as in Sections 2.3 and 5.2, one finds that as $V \to \infty$ and $\epsilon_V \to 0$, the continuity operator (55) converges to (assuming $\phi[\bar{w}]_y(dx) = \phi[\bar{w}]_y(x)\,dx$):

$$\phi[\bar{w}]_y(x) := \rho_{0,y}(x) - \text{div}_x \bar{w}_{\text{tr},y}(x) + (\Gamma \bar{w}_\text{re})_y(x), \tag{56}$$

and the process converges (pathwise in probability) to the solution of the system:

$$\dot{\bar{w}}_{t,\text{tr},y}(x) = -D_y \nabla_x \phi[\bar{w}_t]_y(x),$$

$$\dot{\bar{w}}_{t,\text{re}}(x) = \kappa_\text{fw}\phi[\bar{w}_t]_A(x) - \kappa_\text{bw}\phi[\bar{w}_t]_B(x).$$

Indeed, putting these together yields the reaction-diffusion equation for the limit concentrations:

$$\dot{\rho}_{t,A}(x) = D_A \Delta_x \rho_{t,A}(x) - \kappa_\text{fw}\rho_{t,A}(x) + \kappa_\text{bw}\rho_{t,B}(x),$$

$$\dot{\rho}_{t,B}(x) = D_B \Delta_x \rho_{t,B}(x) + \kappa_\text{fw}\rho_{t,A}(x) - \kappa_\text{bw}\rho_{t,B}(x).$$

To find the corresponding large deviations, we combine (16) and (49) to calculate the non-linear generator:

$$
(\mathcal{H}^{(V)}f)(\bar{w}_{\mathrm{tr}}, \bar{w}_{\mathrm{re}}) := \frac{1}{V} \mathrm{e}^{-Vf(\bar{w}_{\mathrm{tr}}, \bar{w}_{\mathrm{re}})} \big(\mathcal{Q}^{(V)} \mathrm{e}^{Vf} \big)(\bar{w}_{\mathrm{tr}}, \bar{w}_{\mathrm{re}})
$$

$$
\xrightarrow{V \to \infty} \sum_{y=\mathsf{A},\mathsf{B}} D_y \int \big(\mathrm{div}_x \, \partial_{\bar{w}_{\mathrm{tr},y}} f(\bar{w}_{\mathrm{tr}}, \bar{w}_{\mathrm{re}})(x) + \big| \partial_{\bar{w}_{\mathrm{tr},y}} f(\bar{w}_{\mathrm{tr}}, \bar{w}_{\mathrm{re}})(x) \big|^2 \big) \phi[\bar{w}]_y(dx)
$$

$$
+ \kappa_{\mathrm{fw}} \int \phi[\bar{w}]_{\mathsf{A}}(dx) \big(\mathrm{e}^{\partial_{\bar{w}_{\mathrm{re}}} f(\bar{w}_{\mathrm{tr}}, \bar{w}_{\mathrm{re}})} - 1 \big)
$$

$$
+ \kappa_{\mathrm{bw}} \int \phi[\bar{w}]_{\mathsf{B}}(dx) \big(\mathrm{e}^{-\partial_{\bar{w}_{\mathrm{re}}} f(\bar{w}_{\mathrm{tr}}, \bar{w}_{\mathrm{re}})} - 1 \big).
$$

Let us again abbreviate $\bar{\zeta} = (\bar{\zeta}_{\mathrm{tr},y,l}(x), \bar{\zeta}_{\mathrm{re}}(x))$ and $\bar{\jmath} = (\bar{\jmath}_{\mathrm{tr},y,l}(x), \bar{\jmath}_{\mathrm{re}}(x))$. The limiting non-linear generator can now be split into

$$
\mathcal{H}(\bar{w}, \bar{\zeta}) := \mathcal{H}_{\mathrm{tr}}(\bar{w}, \bar{\zeta}_{\mathrm{tr}}) + \mathcal{H}_{\mathrm{re}}(\bar{w}, \bar{\zeta}_{\mathrm{re}}), \tag{57}
$$

$$
\mathcal{H}_{\mathrm{tr}}(\bar{w}, \bar{\zeta}_{\mathrm{tr}}) := \sum_{y=\mathsf{A},\mathsf{B}} D_y \big(\| \bar{\zeta}_{\mathrm{tr},y} \|^2_{L^2(\phi[\bar{w}]_y)} - \langle \bar{\zeta}_{\mathrm{tr},y}, \nabla_x \, \phi[\bar{w}]_y \rangle \big),
$$

$$
\mathcal{H}_{\mathrm{re}}(\bar{w}, \bar{\zeta}_{\mathrm{re}}) := \kappa_{\mathrm{fw}} \int \phi[\bar{w}]_{\mathsf{A}}(dx) \big(\mathrm{e}^{\bar{\zeta}_{\mathrm{re}}(x)} - 1 \big) + \kappa_{\mathrm{bw}} \int \phi[\bar{w}]_{\mathsf{B}}(dx) \big(\mathrm{e}^{-\bar{\zeta}_{\mathrm{re}}(x)} - 1 \big).
$$

Since each mechanism corresponds to a separate flux, the corresponding L-function also splits into two parts:

$$
\mathcal{L}(\bar{w}, \bar{\jmath}) := \sup_{\bar{\zeta}_{\mathrm{tr}}, \bar{\zeta}_{\mathrm{re}}} \langle \bar{\zeta}_{\mathrm{tr}}, \bar{\jmath}_{\mathrm{tr}} \rangle + \langle \bar{\zeta}_{\mathrm{re}}, \bar{\jmath}_{\mathrm{re}} \rangle - \mathcal{H}(\bar{w}, \bar{\zeta}) := \mathcal{L}_{\mathrm{tr}}(\bar{w}, \bar{\jmath}_{\mathrm{tr}}) + \mathcal{L}_{\mathrm{re}}(\bar{w}, \bar{\jmath}_{\mathrm{re}}), \tag{58}
$$

$$
\mathcal{L}_{\mathrm{tr}}(\bar{w}, \bar{\jmath}_{\mathrm{tr}}) := \sum_{y=\mathsf{A},\mathsf{B}} \frac{1}{4D_y} \| \bar{\jmath}_{\mathrm{tr},y} + D_y \nabla_x \, \phi[\bar{w}]_y \|^2_{L^2(1/\phi[\bar{w}]_y)},
$$

$$
\mathcal{L}_{\mathrm{re}}(\bar{w}, \bar{\jmath}_{\mathrm{re}}) := \inf_{j_{\mathrm{fw}} - j_{\mathrm{bw}} = \bar{\jmath}_{\mathrm{re}}} h\big(j_{\mathrm{fw}} | \kappa_{\mathrm{fw}} \phi[\bar{w}]_{\mathsf{A}}\big) + h\big(j_{\mathrm{bw}} | \kappa_{\mathrm{bw}} \phi[\bar{w}]_{\mathsf{B}}\big),
$$

using the usual the relative entropy between two measures, i.e: $h(j|k) := \int j(dx) \log(dj/dk(x)) - j(dx) + k(dx)$ if $j \ll k$, else $h(j|k) := \infty$.

As before, the calculation above formally shows that the flux large-deviation principle holds (see for example [BL12] for a similar but rigorous result):

$$
\mathrm{Prob}^{(V)} \big(\bar{W}^{(V)}_{(\cdot)} \approx \bar{w}_{(\cdot)} \big) \overset{V \to \infty}{\sim} \mathrm{e}^{-V \int_0^T \mathcal{L}(\bar{w}_t, \dot{\bar{w}}_t) \, dt}.
$$

The L-function (58) splits into two parts because the only interaction between the two mechanisms occurs through the state $\phi[\bar{w}]$. By contrast, the corresponding state space large-deviation are much more complicated. Observe that the continuity equation (56) is an affine function of $\bar{w}$, and so $d\phi_{\bar{w}} \bar{\jmath}$ is independent of $\bar{w}$, and the invariance condition (27) holds. As explained in the beginning of Section 4, this means that one can apply a straightforward contraction principle on the tangents to yield the large deviation cost function for the states/concentrations:

$$
\hat{\mathcal{L}}(\phi[\bar{w}], s) := \inf_{\substack{\bar{\jmath} = (\bar{\jmath}_{\mathrm{tr}}, \bar{\jmath}_{\mathrm{re}}): \\ s = -\mathrm{div} \, \bar{\jmath}_{\mathrm{tr}} + \Gamma \bar{\jmath}_{\mathrm{re}}}} \mathcal{L}(\bar{w}, \bar{\jmath}). \tag{59}
$$

This infimum reintroduces a strong interrelation between the two driving mechanisms. Indeed, for a given tangent (ρ, s), the fluxes in this infimum correspond to an optimal splitting between the two mechanisms, which can be seen as an inf-convolution. Similar interactions also arise when considering multiple reaction pairs, see [MPPR17, Sect. 3.4].

6.3 GGSs in flux and state space

We now apply Theorem 4.8 to the reaction-diffusion setting. The symmetry condition (33) holds for the function (57) if we choose the free energy functional

$$\mathcal{F}(\bar{w}) := \frac{1}{2} \int \phi[\bar{w}]_{\mathsf{A}}(dx) \log \kappa_{\mathrm{fw}} \phi[\bar{w}]_{\mathsf{A}}(x) - \phi[\bar{w}]_{\mathsf{A}}(dx)$$
$$+ \frac{1}{2} \int \phi[\bar{w}]_{\mathsf{B}}(dx) \log \kappa_{\mathrm{bw}} \phi[\bar{w}]_{\mathsf{B}}(x) - \phi[\bar{w}]_{\mathsf{B}}(dx).$$

Naturally, this functional can be seen as a combination of (21) and (52), where just like (52), it has the form of a relative entropy with respect to a locally finite invariant measure, namely $(\pi_y(dx))_{y=\mathsf{A},\mathsf{B}} \equiv (1/\kappa_{\mathrm{fw}}, 1/\kappa_{\mathrm{bw}})\, dx$. We find the corresponding dissipation potentials from (42), which is again a combination of the non-quadratic potentials (22),(23) and the quadratic potential (53):

$$\Psi^*(\bar{w}, \bar{\zeta}) := \Psi^*_{\mathrm{tr}}(\bar{w}, \bar{\zeta}_{\mathrm{tr}}) + \Psi^*_{\mathrm{re}}(\bar{w}, \bar{\zeta}_{\mathrm{re}}),$$
$$\Psi^*_{\mathrm{tr}}(\bar{w}, \bar{\zeta}_{\mathrm{tr}}) := D_{\mathsf{A}} \|\bar{\zeta}_{\mathrm{tr},\mathsf{A}}\|^2_{L^2(\phi[\bar{w}]_{\mathsf{A}})} + D_{\mathsf{B}} \|\bar{\zeta}_{\mathrm{tr},\mathsf{B}}\|^2_{L^2(\phi[\bar{w}]_{\mathsf{B}})},$$
$$\Psi^*_{\mathrm{re}}(\bar{w}, \bar{\zeta}_{\mathrm{re}}) := \int \sigma(\bar{w})(x) \big(\cosh(\bar{\zeta}_{\mathrm{re}}(x)) - 1 \big)\, dx, \qquad \text{and}$$
$$\Psi(\bar{w}, \bar{\jmath}) := \Psi_{\mathrm{tr}}(\bar{w}, \bar{\jmath}_{\mathrm{tr}}) + \Psi_{\mathrm{re}}(\bar{w}, \bar{\jmath}_{\mathrm{re}}),$$
$$\Psi_{\mathrm{tr}}(\bar{w}, \bar{\jmath}_{\mathrm{tr}}) := \frac{1}{4D_{\mathsf{A}}} \|\bar{\jmath}_{\mathrm{tr},\mathsf{A}}\|^2_{L^2(1/\phi[\bar{w}]_{\mathsf{A}})} + \frac{1}{4D_{\mathsf{B}}} \|\bar{\jmath}_{\mathrm{tr},\mathsf{B}}\|^2_{L^2(1/\phi[\bar{w}]_{\mathsf{B}})},$$
$$\Psi_{\mathrm{re}}(\bar{w}, \bar{\jmath}_{\mathrm{re}}) := \int \sigma(\bar{w})(x) \big(\cosh^*(\frac{\bar{\jmath}_{\mathrm{re}}(x)}{\sigma(\bar{w})(x)}) + 1 \big)\, dx,$$

with $\sigma(\bar{w})(x) := 2\sqrt{\kappa_{\mathrm{fw}}\kappa_{\mathrm{bw}}\phi[\bar{w}]_{\mathsf{A}}(x)\phi[\bar{w}]_{\mathsf{B}}(x)}$. Let the flux space be given by $\mathcal{W} = \mathcal{M}_1(\mathbb{R}^d; \mathbb{R}^d) \times \mathcal{M}_1(\mathbb{R}^d; \mathbb{R}^d) \times L^1(\mathbb{R}^d)$, where the first two spaces, corresponding to the transport fluxes, are equipped with the metric (54) introduced in the previous section. By Theorem 4.8 the flux cost function $\mathcal{L}$ induces the GGS $(\mathcal{W}, \Psi, \mathcal{F})$. We stress that the dissipation potential Ψ splits into two potentials for the transport and reaction mechanisms respectively, but the free energy is one and the same for both mechanisms.

Since a GGS is a special case of a pGGEN, by Theorem 4.13, the state cost function $\hat{\mathcal{L}}$ also induces a GGS $(\mathcal{X}, \hat{\Psi}, \hat{\mathcal{F}})$, in this case in the space $\mathcal{X} = \mathcal{P}_2(\mathbb{R}^d \times \{\mathsf{A}, \mathsf{B}\})$. The same result yields

$$\hat{\mathcal{F}}(\rho) := \frac{1}{2} \int \rho_{\mathsf{A}}(dx) \log \kappa_{\mathrm{fw}} \rho_{\mathsf{A}}(x) - \rho_{\mathsf{A}}(dx) + \frac{1}{2} \int \rho_{\mathsf{B}}(dx) \log \kappa_{\mathrm{bw}} \rho_{\mathsf{B}}(x) - \rho_{\mathsf{B}}(dx),$$

and, using $d\phi_{\bar{w}}^{\mathsf{T}} = \begin{pmatrix} \nabla_x & 0 \\ 0 & \nabla_x \\ -1 & 1 \end{pmatrix}$:

$$\hat{\Psi}^*(\rho, \xi) \overset{(43b)}{=} \Psi^*(\bar{w}, d\phi_{\bar{w}}^{\mathsf{T}}\xi) = \Psi^*_{\mathrm{tr}}(\bar{w}, \nabla_x \xi) + \Psi^*_{\mathrm{re}}(\bar{w}, \xi_{\mathsf{B}} - \xi_{\mathsf{A}})$$
$$= D_{\mathsf{A}} \|\nabla_x \xi_{\mathsf{A}}\|^2_{L^2(\rho_{\mathsf{A}})} + D_{\mathsf{B}} \|\nabla_x \xi_{\mathsf{B}}\|^2_{L^2(\rho_{\mathsf{B}})}$$
$$+ 2 \int \sqrt{\kappa_{\mathrm{fw}}\kappa_{\mathrm{bw}}\rho_{\mathsf{A}}(x)\rho_{\mathsf{B}}(x)} \big(\cosh(\xi_{\mathsf{B}}(x) - \xi_{\mathsf{A}}(x)) - 1 \big)\, dx,$$
$$\hat{\Psi}(\rho, s) \overset{(43a),(56)}{=} \inf_{s = -\operatorname{div}_x \bar{\jmath}_{\mathrm{tr}} + \Gamma \bar{\jmath}_{\mathrm{re}}} \Psi_{\mathrm{tr}}(\bar{w}, \bar{\jmath}_{\mathrm{tr}}) + \Psi_{\mathrm{re}}(\bar{w}, \bar{\jmath}_{\mathrm{re}}).$$

We stress that, analogous to the L-function (59), the dissipation potential $\hat{\Psi}$ on state space no longer splits into two parts.

7 Discussion

We studied gradient and (pre-) Generic structures induced by flux large deviations, and the relationship between structures induced by state large deviations. At a first glance, the resulting gradient or generic structures in flux may appear non-physical. However, in practice many induced flux structures have a free energy and dissipation potential that only depends on the integrated flux through the state of the system. Hence the main difference with energy-driven structures in state space is that the fluxes rather than velocities are being driven, which seems a very physical assumption.

It turns out that if the fluxes follow a gradient or generic flow, then so do the corresponding states. The same principle also holds in the other direction, but the flow in flux space could have an additional Hamiltonian term that is not observed when considering states only.

On a mathematical level, it can also be worthwhile to work with fluxes rather than states. As we saw in the examples, and Section 6 in particular, if a microscopic system consists of multiple driving mechanisms, then the corresponding Markov generator as well as the non-linear generator is a sum over these mechanisms. By considering separate fluxes for each of these mechanisms, the large-deviation L-function and its induced dissipation potential also splits into different terms for each mechanism. This decomposition can be beneficial for analytic and numerical purposes, e.g. using operator splitting techniques.

References

[ADE17]   A. Agazzi, A. Dembo, and J.-P. Eckmann. Large deviations theory for markov jump models of chemical reaction networks. arxiv.org/abs/1701.02126, 2017.

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

[AK11]   D.F. Anderson and T.G. Kurtz. Continuous time Markov chain models for chemical reaction networks. In Heinz Koeppl, Gianluca Setti, Mario di Bernardo, and Douglas Densmore, editors, *Design and Analysis of Biomolecular Circuits*, pages 3–42. Springer, New York, N.Y. U.S.A, 2011.

[BDSG⁺04] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim. Minimum dissipation principle in stationary non-equilibrium states. *Journal of Statistical Physics*, 116(1):831–841, 2004.

[BDSG⁺15] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim. Macroscopic fluctuation theory. *Reviews of Modern Physics*, 87(2), 2015.

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

[Die15]   H. Dietert. Characterisation of gradient flows on finite state Markov chains. *Electronic Communications in Probability*, 20(29):1–8, 2015.

[DLR13]   M. H. Duong, V. Laschos, and M. Renger. Wasserstein gradient flows from large deviations of many-particle limits. *ESAIM: Control, Optimisation and Calculus of Variations*, 19(4):1166–1188, 2013.

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

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

[EMR15]   M. Erbar, J. Maas, and D.R.M. Renger. From large deviations to Wasserstein gradient flows in multiple dimensions. *Electronic Communications in Probability*, 20(89):1–12, 2015.

[FK06]   J. Feng and T.G. Kurtz. *Large deviations for stochastic processes*, volume 131. American Mathematical Society, Providence, RI, USA, 2006.

[JKO98]   R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM Journal on Mathematical Analysis*, 29(1):1–17, 1998.

[KJZ17]   M. Kaiser, R.L. Jack, and J. Zimmer. Canonical structure and orthogonality of forces and currents in irreversible markov chains. arxiv.org/abs/1708.01453, 2017.

[KLMP18a] R. Kraaij, A. Lazarescu, C. Maes, and M.A. Peletier. Deriving GENERIC from a generalized fluctuation symmetry. *Journal of Statistical Physics*, 170(3):492–508, 2018.

[KLMP18b] R. Kraaij, A. Lazarescu, C. Maes, and M.A. Peletier. Fluctuation symmetry leads to GENERIC equations with non-quadratic dissipation. arxiv.org/abs/1712.10217v1, 2018.

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

[Mae17]   C. Maes. Frenetic bounds on the entropy production. *ArXiv Preprint*, 1705.07412, 2017.

[Mie11]   A. Mielke. Formulation of thermoelastic dissipative material behavior using GENERIC. *Continuum Mechanics and Thermodynamics*, 23(3):233–256, 2011.

[MN08]   C. Maes and K. Netočný. Canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states. *EPL*, 82(3):30003, 2008.

[MPPR17]   A. Mielke, I.A. Patterson, M.A. Peletier, and D.R.M. Renger. Non-equilibrium thermodynamical principles for chemical reactions with mass-action kinetics. *SIAM Journal on Applied Mathematics)*, 77(4), 2017.

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

[OM53]   L. Onsager and S. Machlup. Fluctuations and irreversible processes. *Phys. Rev.*, 91(6):1505–1512, Sep 1953.

[Ons31]   L. Onsager. Reciprocal relations in irreversible processes I. *Phys. Rev.*, 37(4):405–426, Feb 1931.

[Ött05]   H.C. Öttinger. *Beyond equilibrium thermodynamics*. Wiley-Interscience, Hoboken, NJ, USA, 2005.

[PP15]   P. Pfaffelhuber and L. Popovic. Scaling limits of spatial compartment models for chemical reaction networks. *The Annals of Applied Probability*, 25(6):3162–3208, 2015.

[PR18]   R.I.A. Patterson and D.R.M. Renger.   Large deviations of reaction fluxes. arxiv.org/abs/1802.02512, 2018.

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

[Vil09]   C. Villani. *Optimal transport: old and new*. Springer, Berlin-Heidelberg, Germany, 2009.