Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

Website: http://AIMsciences.org

pp. **X–XX**

ORTHOGONALITY OF FLUXES IN GENERAL NONLINEAR REACTION NETWORKS

D R. MICHIEL RENGER AND JOHANNES ZIMMER

Dedicated, in gratitude, to Alex Mielke on the occasion of his 60th birthday

D R. Michiel Renger*

Weierstrass Institute for Applied Analysis and Stochastics
Mohrenstrasse 39
10117 Berlin, Germany

Johannes Zimmer

Department of Mathematical Sciences
University of Bath
Bath BA2 7AY, United Kingdom

(Communicated by the associate editor name)

Abstract. We consider the chemical reaction networks and study currents in these systems. Reviewing recent decomposition of rate functionals from large deviation theory for Markov processes, we adapt these results for reaction networks. In particular, we state a suitable generalisation of orthogonality of forces in these systems, and derive an inequality that bounds the free energy loss and Fisher information by the rate functional.

1. **Introduction.** It is now becoming widely accepted that fluxes hold a key to understanding many phenomena in non-equilibrium thermodynamics. Typically, non-equilibrium systems are forced out of equilibrium by external forces. Since such forces may cause 'divergence-free' fluxes, the forces can not be balanced by changes in mass densities unless one takes the full fluxes into account. In fact, a lot of thermodynamic information can be extracted from microscopic fluctuations of fluxes on the large-deviations scale; this idea is the basis of Macroscopic Fluctuation Theory (MFT) [3].

Although the setting of our paper will be slightly different, we briefly describe the typical setting considered in MFT. One considers a random particle system with particle density $C^{(n)}(x,t)$ (we write C instead of ρ as usual for a density to emphasise the analogy for what follows) and particle flux $J^{(n)}(x,t)$ connected via the continuity equation $\dot{C}^{(n)}(t) = -\operatorname{div} J^{(n)}(t)$, where div is a divergence operator. The parameter n controls the number of elements in the system. When this number is sent to infinity, this can lead — in a suitable scaling of space and time – to a

2000 *Mathematics Subject Classification.* Primary: 82C35; Secondary: 82C31, 60J27.
Key words and phrases. Chemical reactions, large deviations, orthogonality.

macroscopic limit:

$$\dot{c}(t) = \Gamma j(t), \text{ with} \tag{1.1}$$
$$j(t) = \kappa(c(t)), \tag{1.2}$$

for some model-specific operator κ; here $\Gamma\xi = -\operatorname{div}\xi$ (we will consider similar structures on the level of chemical networks below, see (2.2) below). The microscopic fluctuations around this macroscopic limit are often characterised by a large-deviation principle. In the context considered here, this means that the probability of observing trajectories which deviate from the macroscopic limit converges exponentially:

$$\operatorname{Prob}\left((c^{(n)}, j^{(n)}) \approx (c, j)\right) \overset{n\to\infty}{\sim} \exp\left(-n\left[\mathcal{I}_0(c(0)) + \int_0^T \|j(t) - \kappa(c(t))\|^2_{L^2(f(c(t))}\, dt\right]\right), \tag{1.3}$$

for some model-specific function f; $\mathcal{I}_0$ is a functional depending on the initial data $(c^{(n)}(0))_n$. For example, for independent Brownian particles, $f(c) = c^{-1}$, and for the simple symmetric exclusion process $f(c) = c^{-1}(1-c)^{-1}$.

The corresponding inner product is a very powerful tool that can be used to define orthogonal decompositions of fluxes into "reversible" and "irreversible" parts. With this decomposition the large-deviation rate function also decomposes,

$$\|j(t) - \kappa(c(t))\|^2_{L^2(f(c(t))} = \|j(t)^{\mathrm{rev}} - \kappa^{\mathrm{rev}}(c(t))\|^2_{L^2(f(c(t))} + \|j^{\mathrm{irr}}(t) - \kappa^{\mathrm{irr}}(c(t))\|^2_{L^2(f(c(t))}. \tag{1.4}$$

Moreover, by duality one also obtains an inner product on (generalised) forces, allowing to orthogonally decompose forces into reversible and irreversible parts.

In this work we consider a similar setting, however for a class of systems with a different noise so that the large deviations are entropic rather than quadratic,

$$\operatorname{Prob}\left((c^{(n)}, j^{(n)}) \approx (c, j)\right) \overset{n\to\infty}{\sim} \exp\left(-n\left[\mathcal{I}_0(c(0)) + \int_0^T \mathcal{S}\big(j(t) \mid \kappa(c(t))\big)\, dt\right]\right), \tag{1.5}$$

see (2.8) below for the precise definition. To facilitate the intuitive picture, we interpret these systems as models for chemical reactions, although the range of applicability is much wider. In particular, these models are more natural than white-noise driven models when the underlying state space remains discrete. Their non-quadratic fluctuation costs do not suggest a natural inner product, and so it is not evident how to meaningfully decompose the cost as in (1.4) for the quadratic case.

Recently, a new concept of generalised orthogonality was introduced for entropic cost functions corresponding to independent particles [5]. In that case, the operator κ is linear and the continuity operator will be the negative discrete divergence $\Gamma = -\operatorname{div}$. In this work we extend this orthogonality concept to the more general case of randomly inter(re-)acting particles. In Sections 2 and 3, we collect results from the literature, adapting and extending them where needed to the context of chemical reaction networks. We therefore omit the proofs and sometimes the precise technical assumptions, which can be found in the quoted literature. Sections 4 and 5 give orthogonal decomposition results and bounds on the entropy and a term resembling the Fisher information in the classical setting.

2. **Microscopic model, macroscopic limit and large deviations.** As mentioned we study a model for chemical reactions, although the mathematical structure is applicable to any Markov jump process for which the large-deviation principle of the form (1.5) holds [5, 12]. Let $\mathcal{Y}$ be the set of reactants involved, for example $\mathcal{Y} := \{\mathrm{Na}, \mathrm{CL}_2, \mathrm{NaCL}\}$. We denote by $\mathcal{R}$ the set of reactions, in the example

$$\mathcal{R} := \{2\mathrm{Na} + \mathrm{CL}_2 \to 2\mathrm{NaCL},\ 2\mathrm{NaCL} \to 2\mathrm{Na} + \mathrm{CL}_2\}\,.$$

The stoichiometric information of a reaction r is captured in the state change vector $\gamma^{(r)} \in \mathbb{R}^{\mathcal{Y}}$, which describes with negative entries how many reactants of each species are consumed and with positive entries the creation of products, for example $\gamma^{(r)} := (-2, -1, 2)$ for the reaction $r\colon 2\mathrm{Na} + \mathrm{CL}_2 \to 2\mathrm{NaCL}$. We shall assume that $\mathcal{R}$ is the disjoint union of the set of forward reactions $\mathcal{R}_{\mathrm{fw}}$, in this example $\mathcal{R}_{\mathrm{fw}} := \{2\mathrm{Na} + \mathrm{CL}_2 \to 2\mathrm{NaCL}\}$, and the set of backward reactions $\mathcal{R}_{\mathrm{bw}}$, such that for each forward reaction $r \in \mathcal{R}_{\mathrm{fw}}$ there is exactly one backward reaction $\mathrm{bw}(r) \in \mathcal{R}_{\mathrm{bw}}$ with:

$$\gamma^{\mathrm{bw}(r)} = -\gamma^{(r)}. \tag{2.1}$$

This assumption is no loss of generality, as we can always introduce phantom reactions with zero reaction rates, i.e., reactions that exist on paper but do not take place. We sometimes also write $\mathrm{bw}(r) \in \mathcal{R}_{\mathrm{fw}}$ for $r \in \mathcal{R}_{\mathrm{bw}}$. The state change vectors are collected in the matrix $\Gamma := [\gamma^{(r)}]_{r\in\mathcal{R}} \in \mathbb{R}^{\mathcal{Y}\times\mathcal{R}}$. We deliberately use the same notation as in (1.1), to emphasise the analogy of that equation to (2.2) below.

We now further specify the microscopic particle system that models such reaction network, and introduce the time-reversed process. Next, we describe the macroscopic limit and the corresponding large deviations. We then briefly summarise in the remainder of this section some results from [13] which will be needed in the current work: time-reversal symmetries and the relations between forward and backward reaction rates that can be derived.

2.1. **Microscopic model.** The microscopic model will be a Markov jump process consisting of randomly reacting particles in a large volume of size V, which now controls number of particles in the system. We assume that there is no spatial dependence of the reactions. The *empirical concentration measure* is

$$C_y^{(V)}(t) := \frac{1}{V}\#\big\{\text{particles of species } y \text{ present at time } t\big\}.$$

A reaction $r \in \mathcal{R}$ occurs randomly with concentration-dependent intensity $k_r^{(V)} = k_r^{(V)}(C^{(V)}(t))$, upon which the concentration is updated with $C^{(V)}(t) := C^{(V)}(t^-) + V^{-1}\gamma^{(r)}$, where $C^{(V)}(t^-)$ is the limit from the left. These intensities $k_r^{(V)}$ are also called the *jump rates* or *propensities*. In addition, we would like to introduce a reaction flux $J_r^{(V)}(t)$ that measures the amount of reactions taking place at time t. However, since the process has jumps, it is mathematically easier to introduce the *(time-)integrated flux*,

$$W_r^{(V)}(t) := \frac{1}{V}\#\left\{\text{reactions } r \text{ occurred in time } (0,t]\right\}.$$

The fluxes and concentrations are then related by the continuity equation $C^{(V)}(t) = C^{(V)}(0) + \Gamma W^{(V)}(t)$, or formally

$$\dot{C}^{(V)}(t) = \Gamma \dot{W}^{(V)}(t) = \Gamma J^{(V)}(t). \tag{2.2}$$

We point out that this equation is the equivalent for chemical reactions to (1.1), for the MFT setting sketched in the introduction. The pair $(C^{(V)}(t), W^{(V)}(t))$ is now a

Markov process in $\mathbb{R}^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}}$; its generator is given by

$$(\mathcal{Q}^{(V)} f)(c,w) := \sum_{r\in\mathcal{R}} k_r^{(V)}(c)\left[f(c+\tfrac{1}{V}\gamma^{(r)}, w+\tfrac{1}{V}\mathbb{1}_r) - f(c,w)\right].$$

We remark that these are the sole reasons for considering integrated fluxes: although the concentration $C^{(V)}(t)$ is a Markov process, $C^{(V)}(t)$ paired with the non-integrated flux $J^{(V)}(t)$ is not, and moreover, (2.2) only holds in a weak, measure-valued sense. However in the macroscopic regime $V\to\infty$, these issues do not play a role anymore, so that later we can focus on the non-integrated fluxes.

We collect the needed technical assumptions on the Markov chain as follows.

Assumption 2.1. *Assume that for all V the process $C^{(V)}(t)$ is non-explosive on $(0,T)$ (that is, for almost all starting points, almost all trajectories do not exhibit infinitely many jumps in $(0,T)$) and remains almost surely within a compact set, typically a simplex that describes mass conservation. Furthermore, the process has a unique, coordinate-wise positive invariant measure $0<\pi^{(V)}\in\mathcal{P}(\frac{1}{V}\mathbb{N}_0^I)$. Integrated fluxes are assumed to satisfy $W^{(V)}(0)=0$ almost surely in V.*

Example 2.2. *The typical example that is used to model microscopic chemical reactions is $k_r^{(V)}(c) = \omega^{(r)} V \prod_{y\in\mathcal{Y}} (1/V)^{\alpha_y^{(r)}} \alpha_y^{(r)}! \binom{c_y V}{\alpha_y^{(r)}}$, where $\omega^{(r)}$ is a reaction-specific constant and $\alpha_y^{(r)}$ are the number of y-reactants consumed in reaction r. In that case the master equation for the Markov process $C^{(V)}(t)$ is called the "Chemical Master Equation", and in many cases the invariant measure $\pi^{(V)}$ is explicitly known and positive [2]. In general, explosion may occur for such models, but it can be ruled out by imposing an additional mass conservation assumption [12].*

2.2. **The time-reversed process.** Here we introduce the time-reversed process which will be crucial throughout the paper. For a path $(C^{(V)}(t), W^{(V)}(t))$ and the given final time T, we define the time-reversed path as

$$\left(\overleftarrow{C}^{(V)}(t), \overleftarrow{W}^{(V)}(t)\right) := \left(C^{(V)}(T-t), {W^{(V)}}^{\mathsf{T}}(T) - {W^{(V)}}^{\mathsf{T}}(T-t)\right),$$

where ${W^{(V)}}^{\mathsf{T}}_{\mathrm{bw}(r)}(t) := W_r^{(V)}(t)$; by construction, these time-reversed integrated fluxes $\overleftarrow{W}^{(V)}(t)$ remain non-negative, non-decreasing, and initially calibrated at 0, as the original integrated fluxes $\overleftarrow{W}^{(V)}(t)$.

Theorem 2.3 ([13, Proposition 4.1], time-reversal)**.** *Let $\mathbb{P}^{(V)}_{\mathcal{Q}^{(V)},\pi^{(V)}}$ be the path measure of the Markov process $(C^{(V)}(t), W^{(V)}(t))$ with generator $\mathcal{Q}^{(V)}$ and initial distribution $\pi^{(V)}\times\delta_0$. Then*

$$\mathbb{P}^{(V)}_{\mathcal{Q}^{(V)},\pi^{(V)}}\big((\overleftarrow{C}^{(V)}(t), \overleftarrow{W}^{(V)}(t)) \in \mathcal{A}\big) = \mathbb{P}^{(V)}_{\overleftarrow{\mathcal{Q}}^{(V)},\pi^{(V)}}\big((C^{(V)}(t), W^{(V)}(t)) \in \mathcal{A}\big), \tag{2.3}$$

where the reversed generator is

$$(\overleftarrow{\mathcal{Q}} f)(w) = \sum_{r\in\mathcal{R}} \overleftarrow{k}^{(V)}_{\mathrm{bw}(r)}(c)\big[f(c+\tfrac{1}{V}\gamma^{\mathrm{bw}(r)}, w+\tfrac{1}{V}\mathbb{1}_{\mathrm{bw}(r)}) - f(c,w)\big], \tag{2.4}$$

and the reversed rates are related to the forward rates through

$$\overleftarrow{k}^{(V)}_{\mathrm{bw}(r)}(c) := \frac{\pi^{(V)}(c+\frac{1}{V}\gamma^{\mathrm{bw}(r)})}{\pi^{(V)}(c)} k_r^{(V)}(c+\tfrac{1}{V}\gamma^{\mathrm{bw}(r)}). \tag{2.5}$$

Remark 2.4. A special role is played for systems that satisfy "microscopic detailed balance", i.e., for $V > 0$:

$$\pi^{(V)}(c)k_r^{(V)}(c) = \pi^{(V)}(c + \tfrac{1}{V}\gamma^{(r)})k_r^{(V)}(c + \tfrac{1}{V}\gamma^{(r)}) \qquad \text{for all } c \in \mathbb{R}^{\mathcal{Y}} \text{ and } r \in \mathcal{R}.$$

Indeed, by (2.5), this is equivalent to $\overleftarrow{k}_r^{(V)}(c) = k_r^{(V)}(c)$, and hence to the reversibility of the Markov process, i.e., $\overleftarrow{\mathcal{Q}}^{(V)} = \mathcal{Q}^{(V)}$ in (2.3) (not to be confused with thermodynamic reversibility). As we will see, on a macroscopic level this condition corresponds to conservative forces, or the absence of external forces. Since we are mainly interesting in the behaviour of systems undergoing external forces, we shall not assume this condition, but merely use it as a validity check at some places. □

2.3. **Macroscopic limit and large deviations.** We now study the macroscopic behaviour of the system in the limit $V \to \infty$. To this aim we need to make the following assumption.

Assumption 2.5. *Let the jump rates converge in average to reaction rates, that is,* $\sup_c |V^{-1}k_r^{(V)}(c) - \kappa_r(c)| \to 0$ *as* $V \to \infty$. *Furthermore, we assume that the rates* κ_r *satisfy the assumptions of* [12] *allowing for a large-deviation principle; in particular, a rate* $\kappa_r(c)$ *has to vanish if the associated reaction would lead to negative concentrations, and the mapping* $c \mapsto \kappa_r(c)$ *must be non-decreasing, Lipschitz continuous, superhomogeneous* [12, Assumption 2.2(vi)] *and bounded on the compact set of concentrations of Assumption 2.1. We require the invariant distributions* $\pi^{(V)} \in \mathcal{P}(\frac{1}{V}\mathbb{N}_0^{\mathcal{Y}})$ *to satisfy a large-deviation principle*

$$\pi^{(V)}(C^{(V)} \approx c) \overset{V\to\infty}{\sim} \exp\big(-V\mathcal{I}_0(c)\big), \tag{2.6}$$

for some $\mathcal{I}_0 \colon \mathbb{R}_+^{\mathcal{Y}} \to [0,\infty]$. *Finally, we assume that* $\mathcal{I}_0$ *is almost everywhere differentiable, as it holds for example for convex functionals.*

By a straightforward extension of Kurtz' Theorem [6], in the limit $V \to \infty$, the random path $\big(C^{(V)}(t), W^{(V)}(t)\big)$ converges (in measure) to the deterministic solution of the coupled equations

$$\dot{c}(t) = \Gamma\dot{w}(t), \text{ and } \dot{w}(t) = \kappa\big(c(t)\big), \tag{2.7}$$

which is the chemical reaction equivalent of (1.1) and (1.2).

Since the noise is essentially a re-scaled Poissonian, a large-deviations principle holds, which we now recall (see [12] for the precise statement). In the following theorem and throughout the paper we adapt the notation of the relative entropy:

$$\mathcal{S}(j \mid \kappa(c)) := \sum_{r\in\mathcal{R}} j_r \log\frac{j_r}{\kappa_r(c)} - j_r + \kappa_r(c). \tag{2.8}$$

Theorem 2.6 ([12])**.** *As in Theorem 2.3, let* $\mathbb{P}^{(V)}_{\mathcal{Q}^{(V)},\pi^{(V)}}$ *be the path measure of the Markov process* $(C^{(V)}(t), W^{(V)}(t))$ *with generator* $\mathcal{Q}^{(V)}$ *and initial distribution* $\pi^{(V)} \times \delta_0$. *Then for any* $(c,w) \in W^{1,1}(0,T;\mathbb{R}^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}})$

$$\mathbb{P}^{(V)}_{\mathcal{Q}^{(V)},\pi^{(V)}}((C^{(V)}, W^{(V)}) \approx (c,w)) \overset{V\to\infty}{\sim} \exp\left(-V\left[\mathcal{I}_0(c(0)) + \int_0^T \mathcal{S}\big(\dot{w}(t) \mid \kappa(c(t))\big)\,dt\right]\right).$$

The non-negative cost functional $\int_0^T \mathcal{S}\big(\dot w(t) \mid \kappa(c(t))\big)\,dt$ can be interpreted as the free energy required to force the system to follow a given path, (c,w) rather than the path that solves the macroscopic equation (2.7), whereas $\mathcal{I}_0$ measures the cost of deviating from the given initial data. This interpretation goes back to [11]; see also [10] for a more detailed account.

Example 2.7. *For the Chemical Master Equation discussed in Example 2.2, the corresponding reaction rates are* $\kappa^{(r)}(c) = \omega^{(r)} \prod_{y\in\mathcal{Y}} c_y^{\alpha_y^{(r)}}$. *This specific form of the reaction rates is called* mass-action kinetics [6, 2]. *Moreover, in that case* $\mathcal{I}_0(c) := \mathcal{S}(c \mid c^{\mathrm{eq}})$, *where* c^{eq} *is the equilibrium concentration* $0 = \Gamma\kappa(c^{\mathrm{eq}})$ [10].

2.4. **Time reversal symmetries and implications.** The large-deviation result Theorem 2.6 also applies to the time-reversed process $(\overleftarrow{C}^{(V)}(t), \overleftarrow{W}^{(V)}(t))$. Combining this with the time reversal result of Theorem 2.3 yields the following time-reversal symmetry:

Corollary 2.8. *For any path* $(c,w) \in W^{1,1}(0,T;\mathbb{R}^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}})$, *it holds*

$$\mathcal{I}_0(c(0)) + \int_0^T \mathcal{S}\big(\dot w(t) \mid \kappa(c(t))\big)\,dt = \mathcal{I}_0(c(T)) + \int_0^T \mathcal{S}\big(\dot w^{\mathsf{T}}(t) \mid \overleftarrow{\kappa}(c(t))\big)\,dt, \tag{2.9}$$

or (using $\dot c = \Gamma \dot w$*), for a dense set of* $(c,j) = (c,\dot w) \in \mathbb{R}^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}}$ *for which the chain rule holds,*

$$\mathcal{S}(j \mid \kappa(c))) - \mathcal{S}(j^{\mathsf{T}} \mid \overleftarrow{\kappa}(c)) = \nabla \mathcal{I}_0(c) \cdot \Gamma j. \tag{2.10}$$

Although the reversed propensities $\overleftarrow{k}_r^{(V)}(c)$ are explicitly known by (2.5) *if* the invariant measure $\pi^{(V)}$ is known, the reversed reaction rates $\overleftarrow{\kappa}_r(c)$ may generally not be explicit. However, using the fact that (2.10) holds for all j, one finds the relations [13, Section 4.2], if the quantities are defined,

$$\nabla \mathcal{I}_0(c) \cdot \gamma^{(r)} = \log \frac{\overleftarrow{\kappa}_{\mathrm{bw}(r)}(c)}{\kappa_r(c)}, \tag{2.11}$$

$$\sum_{r\in\mathcal{R}} \kappa_r(c) = \sum_{r\in\mathcal{R}} \overleftarrow{\kappa}_r(c), \tag{2.12}$$

$$\kappa_r(c)\kappa_{\mathrm{bw}(r)}(c) = \overleftarrow{\kappa}_r(c)\overleftarrow{\kappa}_{\mathrm{bw}(r)}(c). \tag{2.13}$$

Note in particular that the ratio $\overleftarrow{\kappa}_{\mathrm{bw}(r)}(c)/\kappa_r(c)$ is well-defined whenever $\nabla\mathcal{I}_0(c)$ is well-defined, i.e., on a dense subset. We mention that (2.13) was not stated in [13] but follows directly from (2.11) and the antisymmetry (2.1).

Remark 2.9. Building further upon Remark 2.4, if microscopic detailed balance holds, then this is also true for the reaction rates: $\overleftarrow{\kappa}(c) = \kappa(c)$, which is often called *chemical detailed balance*, or in case of mass-action kinetics, the *Wegscheider condition*. □

3. **Force structures.** In this section, we introduce force structures; these can be seen as a non-equilibrium generalisation of gradient flows. Such structure does in general not exist unless we consider net fluxes. Therefore, we first introduce net fluxes, then force structures, and we finally discuss the force structure that corresponds to the reversed dynamics.

3.1. **Net fluxes: macroscopic limit and large deviations.** Since reactions are ordered in forward-backward pairs with opposite state change vectors (2.1), one can also introduce the net integrated fluxes as $\bar{W}^{(V)}(t) := W^{(V)}(t) - W^{(V)}(t)^{\mathsf{T}}$, i.e.,

$$\bar{W}_r^{(V)}(t) := W_r^{(V)}(t) - W_{\mathrm{bw}(r)}^{(V)}(t), \qquad \text{for } r \in \mathcal{R}_{\mathrm{fw}}. \tag{3.1}$$

Keeping in mind that these net fluxes are only defined for $r \in \mathcal{R}_{\mathrm{fw}}$, we can use the same notation for the continuity equation, cf. (2.2):

$$\begin{aligned}\dot{C}^{(V)}(t) = \Gamma \dot{W}^{(V)}(t) &= \sum_{r\in\mathcal{R}_{\mathrm{fw}}} \gamma^{(r)} \dot{W}_r^{(V)}(t) + \gamma^{\mathrm{bw}(r)} \dot{W}_{\mathrm{bw}(r)}^{(V)}(t) \\ &= \sum_{r\in\mathcal{R}_{\mathrm{fw}}} \gamma^{(r)} \dot{\bar{W}}_r^{(V)}(t) = \Gamma \dot{\bar{W}}^{(V)}(t).\end{aligned}$$

Again, by Kurtz' Theorem [6], the pair $(C^{(V)}(t), \bar{W}^{(V)}(t))$ converges in the macroscopic limit as $V \to \infty$ to

$$\dot{c}(t) = \Gamma \dot{\bar{w}}(t), \text{ and } \dot{\bar{w}}_r(t) = \kappa_r\big(c(t)\big) - \kappa_{\mathrm{bw}(r)}\big(c(t)\big), \tag{3.2}$$

and, by a contraction principle also satisfies the following large-deviation principle.

Corollary 3.1.

$$\mathbb{P}^{(V)}_{\mathcal{Q}^{(V)},\pi^{(V)}}\big((C^{(V)}(t), \bar{W}^{(V)}(t)) \approx (c, \bar{w})\big) \overset{V\to\infty}{\sim} \exp\left(-V\left[\mathcal{I}_0(c) + \int_0^T \mathcal{L}\big(c(t), \dot{\bar{w}}(t)\big)\, dt\right]\right),$$

where for $(c, \bar{\jmath}) \in \mathbb{R}^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}_{\mathrm{fw}}}$*:*

$$\begin{aligned}\mathcal{L}(c, \bar{\jmath}) &:= \inf_{\substack{j:(0,T)\to\mathbb{R}^{\mathcal{R}}: \\ \bar{\jmath}=j-j^{\mathsf{T}}}} \mathcal{S}(j \mid \kappa(c)), \qquad (3.3)\\ &= \sum_{r\in\mathcal{R}_{\mathrm{fw}}} j_r \log\frac{j_r}{\kappa_r(c)} - j_r + \kappa_r(c) + (j_r - \bar{\jmath}_r)\log\frac{j_r - \bar{\jmath}_r}{\kappa_{\mathrm{bw}(r)}} - (j_r - \bar{\jmath}_r) + \kappa_{\mathrm{bw}(r)},\end{aligned}$$

$$\text{where } j_r := \tfrac{1}{2}\bar{\jmath}_r + \sqrt{\tfrac{1}{4}\bar{\jmath}_r^2 + \kappa_r(c)\kappa_{\mathrm{bw}(r)}(c)}.$$

3.2. **Force structure of the cost function.** Following [15, 7, 8, 5, 13], we now rewrite the cost function $\mathcal{L}$ in terms of forces, affinities and energies. As will be clear from the formulas, these forces are only defined for concentrations that satisfy the following condition:

$$\kappa(c) \lessgtr \kappa_{\mathrm{bw}}(c) :\Longleftrightarrow \big(\kappa_r(c) = 0 \iff \kappa_{\mathrm{bw}(r)}(c) = 0 \qquad \text{for all } r \in \mathcal{R}\big). \tag{3.4}$$

Systems for which this condition holds for *any* concentration c are sometimes said to be in *weak detailed balance*.

For any c for which (3.4) holds, the thermodynamic forces are defined as (setting $0/0 = 1$)

$$F_r(c) := \frac{1}{2}\log\frac{\kappa_r(c)}{\kappa_{\mathrm{bw}(r)}(c)}, \qquad \text{for } c \text{ satisfying (3.4) and } r \in \mathcal{R}_{\mathrm{fw}}; \tag{3.5}$$

these are the affinities of Schnakenberg [15]. We also define the dual and primal dissipation potentials (which are Legendre duals of each other):

$$\Phi^*(c,\zeta) := 2\sum_{r\in\mathcal{R}_{\mathrm{fw}}} \sqrt{\kappa_r(c)\kappa_{\mathrm{bw}(r)}(c)}\big(\cosh(\zeta_r)-1\big), \text{ and} \tag{3.6}$$

$$\begin{aligned}\Phi(c,\bar\jmath) &:= \sup_{\zeta\in\mathbb{R}^{\mathcal{R}_{\mathrm{fw}}}} \zeta\cdot\bar\jmath - \Phi^*(c,\zeta)\\ &= \sum_{r\in\mathcal{R}_{\mathrm{fw}}} 2\sqrt{\kappa_r(c)\kappa_{\mathrm{bw}(r)}(c)}\left(\cosh^*\left(\tfrac{\bar\jmath_r}{2\sqrt{\kappa_r(c)\kappa_{\mathrm{bw}(r)}(c)}}\right)+1\right).\end{aligned}$$

We can then decompose the cost function $\mathcal{L}$ as follows.

Theorem 3.2 ([7, 8, 5, 13])**.** *For any $c\in\mathbb{R}^{\mathcal{V}}_+$ for which $\kappa(c)\lessgtr\kappa_{\mathrm{bw}}(c)$ and any $\bar\jmath\in\mathbb{R}^{\mathcal{R}_{\mathrm{fw}}}$,*

$$\mathcal{L}(c,\bar\jmath) = \Phi(c,\bar\jmath) + \Phi^*\big(c,F(c)\big) - F(c)\cdot\bar\jmath, \tag{3.7}$$

and this choice of Φ^*,Φ and F is unique (assuming $\Phi^*(c,0)=0=\Phi(c,0)$ and $\Phi(c,\cdot),\Phi^*(c,\cdot)$ are Legendre duals of each other [9, Proposition 2.1]).

This decomposition has the following physical interpretation. The solution to (3.2) is the zero-cost flow $\mathcal{L}(c(t),\bar\jmath(t))=0$, which is equivalent to

$$\bar\jmath(t) = \nabla_\zeta\,\Phi^*\big(c(t),F(c(t))\big). \tag{3.8}$$

Since Φ^* is not quadratic, this represents a non-linear response between forces and fluxes. Indeed, this also means that the force F must be scaled to be dimensionless; see [10] for the correct scaling with physical constants. Of particular interest is the *Fisher information*:

$$\Phi^*\big(c,F(c)\big) = \sum_{r\in\mathcal{R}_{\mathrm{fw}}} \big(\sqrt{k_r(c)}-\sqrt{k_{\mathrm{bw}(r)}(c)}\big)^2.$$

We interpret this as a Fisher information since for the case of independent Brownian particles $\Phi^*(c,F(c)) = \int_{\mathbb{R}^d}|\nabla\sqrt{c(x)}|^2\,dx$, see also [1, 9]; the expression above has the same form, but with a discrete gradient and non-linear rates. The Fisher information will be discussed in more detail in Section 5.

Remark 3.3. If chemical detailed balance $\overleftarrow{\kappa}=\kappa$ holds, see Remark 2.9, then from (3.5) and (2.11) we find that the last term in (3.7) reads:

$$-F(c)\cdot\bar\jmath = \tfrac12\Gamma^{\mathsf{T}}\nabla\mathcal{I}_0(c)\cdot\bar\jmath = \tfrac12\nabla\mathcal{I}_0(c)\cdot\dot c,$$

so that the force is indeed conservative as anticipated in Remark 2.4. In this case, (3.8) describes a nonlinear gradient flow, either in the space of concentrations [10] or in the space of integrated net fluxes [14]. □

3.3. **Force structure for the reversed dynamics.** We can also give a similar decomposition applied to the reversed dynamics, explained in Sections 2.2 and 2.4. First, we apply Corollary 3.1 to the reversed dynamics, using (2.3):

Corollary 3.4. *Set $\overleftarrow{\bar W}^{(V)}(t) := -\bar W^{(V)}(T) + \bar W^{(V)}(T-t)$, cf. (3.1). Then:*

$$\begin{aligned}&\mathbb{P}^{(V)}_{\mathcal{Q}^{(V)},\pi^{(V)}}\Big((\overleftarrow{C}^{(V)}(t),\overleftarrow{\bar W}^{(V)}(t))\approx(c,\bar w)\Big)\\ &\qquad\qquad \overset{V\to\infty}{\sim} \exp\left(-V\left[\mathcal{I}_0(c)+\int_0^T \overleftarrow{\mathcal{L}}\big(c(t),\dot{\bar w}(t)\big)\,dt\right]\right),\end{aligned}$$

where for $(c,\bar{\jmath}) \in \mathbb{R}^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}_{\rm fw}}$*:*

$$\overleftarrow{\mathcal{L}}(c,\bar{\jmath}) := \inf_{\substack{j:(0,T)\to\mathbb{R}^{\mathcal{R}}:\\ \bar{\jmath}=j-j^{\mathsf{T}}}} \mathcal{S}(j \mid \overleftarrow{\kappa}(c)). \tag{3.9}$$

We may then also decompose the reverse cost function $\overleftarrow{\mathcal{L}}$ as a force structure.

Corollary 3.5 ([13])**.** *For any* $c \in \mathbb{R}^{\mathcal{Y}}_+$ *for which* $\overleftarrow{\kappa}(c) \lessgtr \overleftarrow{\kappa}_{\rm bw}(c)$ *and any* $\bar{\jmath} \in \mathbb{R}^{\mathcal{R}_{\rm fw}}$*,*

$$\overleftarrow{\mathcal{L}}(c,\bar{\jmath}) = \Phi(c,\bar{\jmath}) + \Phi^*\big(c, \overleftarrow{F}(c)\big) - \overleftarrow{F}(c)\cdot\bar{\jmath},$$

where Φ, Φ^* *are given by* (3.6)*, and*

$$\overleftarrow{F}_r(c) = \frac{1}{2}\log\frac{\overleftarrow{\kappa}_r(c)}{\overleftarrow{\kappa}_{{\rm bw}(r)}(c)},$$

and this choice of Φ^*, Φ *and* $\overleftarrow{F}$ *is unique.*

Observe that the potentials Φ, Φ^* are the same for the reversed dynamics due to (2.13). In fact also for the Fisher information $\Phi^*\big(c, F(c)\big) = \Phi^*\big(c, \overleftarrow{F}(c)\big)$ due to (2.12) and (2.13).

4. **Orthogonal decomposition.** We now mimic the arguments from [5] for the case of general nonlinear reactions. Throughout this section, we shall write forces $F(c)$ and gradients $\nabla\mathcal{I}_0(c)$ under the standing assumption that the quantities are well-defined and sufficiently regular. These assumptions will be relaxed in Section 5. First we explain how the dual dissipation potentials can be decomposed; next we discuss how the forces can be decomposed accordingly.

4.1. **Orthogonal decomposition of the dual dissipation potentials.** The idea of this short subsection is to decompose $\Phi^*(c,\xi+\zeta)$ in a similar fashion as one would do in a Hilbert space: $\frac{1}{2}\|\xi+\zeta\|^2 = \frac{1}{2}\|\xi\|^2 + \langle\xi,\zeta\rangle + \frac{1}{2}\|\zeta\|^2$, see for example [3, Section IIC]. However, since Ψ^* is not quadratic, the construction is a bit more involved: the pairing between forces becomes *nonlinear*, and one of the dual dissipation potentials must be modified:

$$\theta_c(\xi,\zeta) := 4\sum_{r\in\mathcal{R}_{\rm fw}} \sqrt{\kappa_r(c)\kappa_{{\rm bw}(r)}(c)}\,\sinh(\xi_r)\sinh(\zeta_r),$$

$$\Phi^*_\zeta(c,\xi) := 2\sum_{r\in\mathcal{R}_{\rm fw}} \sqrt{\kappa_r(c)\kappa_{{\rm bw}(r)}(c)}\,\cosh(\zeta_r)\big(\cosh(\xi_r)-1\big).$$

We can now write the decomposition as follows.

Proposition 4.1. *For any* $c \in \mathbb{R}^{\mathcal{Y}}$ *and* $\xi, \zeta \in \mathbb{R}^{\mathcal{R}}$*,*

$$\begin{aligned}\Phi^*(c,\xi+\zeta) &= \Phi^*_\zeta(c,\xi) + \theta_c(\xi,\zeta) + \Phi^*(c,\zeta)\\ &= \Phi^*(c,\xi) + \theta_c(\xi,\zeta) + \Phi^*_\xi(c,\zeta), \ \textit{and}\\ \theta_c(\xi,\zeta) &= \Phi^*(c,\xi+\zeta) - \Phi^*(c,\xi-\zeta).\end{aligned}$$

Proof. Both statements follow from the identity

$$\cosh(\xi+\zeta) = \cosh(\xi)\cosh(\zeta) + \sinh(\xi)\sinh(\zeta).$$

□

We stress that there is the choice which one of the two dissipation potentials on the right-hand side is modified. By a slight abuse of notation, we shall also write $\Phi^*_F(c,\xi) = \Phi^*_{F(c)}(c,\xi)$ for a force field F.

4.2. **Orthogonal decomposition of the forces.** The force and its reversed counterpart are connected via the relations (2.11), (2.12), (2.13), see [13, Section 4.6]:

$$F_r(c) + \overleftarrow{F}_r(c) = -\big(\Gamma^{\mathsf{T}} \nabla \mathcal{I}_0(c)\big)_r, \tag{4.1}$$

which is twice the force that one would have under chemical detailed balance, see Remark 3.3. This motivates the splitting of general forces into a symmetric and an antisymmetric part $F(c) = F^{\mathrm{sym}}(c) + F^{\mathrm{asym}}(c)$:

$$F_r^{\mathrm{sym}}(c) := \tfrac{1}{2}\big(F_r(c) + \overleftarrow{F}_r(c)\big) = \tfrac{1}{2}\log\frac{\kappa_r(c)}{\overleftarrow{\kappa}_{\mathrm{bw}(r)}(c)} = -\tfrac{1}{2}(\Gamma^{\mathsf{T}}\nabla\mathcal{I}_0(c))_r \qquad \text{and}$$
$$F_r^{\mathrm{asym}}(c) := \tfrac{1}{2}\big(F_r(c) - \overleftarrow{F}_r(c)\big) = \tfrac{1}{2}\log\frac{\overleftarrow{\kappa}_{\mathrm{bw}(r)}(c)}{\kappa_{\mathrm{bw}(r)}(c)},$$

where the latter indeed vanishes if chemical detailed balance $\kappa = \overleftarrow{\kappa}$ holds. More generally, the antisymmetric force measures how far the system is from chemical detailed balance. We briefly note that in case of linear reactions, the antisymmetric force is a constant, independent of the concentration [5], which is no longer true for nonlinear reactions.

The two force fields $F^{\mathrm{sym}}, F^{\mathrm{asym}}$ are indeed 'orthogonal' to each other if we use the nonlinear pairing $\theta_\cdot(c,\cdot)$ introduced above, as the next statement shows.

Proposition 4.2 (Generalisation of [5, Lemma 1])**.** *For any $c \in \mathbb{R}^{\mathcal{Y}}_+$ for which $F^{\mathrm{sym}}(c), F^{\mathrm{asym}}(c)$ are well-defined,*

$$\theta_c(F^{\mathrm{sym}}(c), F^{\mathrm{asym}}(c)) = 0.$$

Proof. This follows from (2.12) and (2.13). □

As a consequence of this and Proposition 4.1, the Fisher information can be decomposed as follows:

Corollary 4.3 (Generalisation of [5, Lemma 2 and Corollary 4])**.** *For any $c \in \mathbb{R}^{\mathcal{Y}}_+$ for which $F^{\mathrm{sym}}(c), F^{\mathrm{asym}}(c)$ are well-defined,*

$$\begin{aligned}\Phi^*\big(c, F^{\mathrm{sym}}(c) + F^{\mathrm{asym}}(c)\big) &= \Phi^*_{F^{\mathrm{asym}}}\big(c, F^{\mathrm{sym}}(c)\big) + \Phi^*\big(c, F^{\mathrm{asym}}(c)\big)\\ &= \Phi^*\big(c, F^{\mathrm{sym}}(c)\big) + \Phi^*_{F^{\mathrm{sym}}}\big(c, F^{\mathrm{asym}}(c)\big),\end{aligned}$$

and hence for any $c \in \mathbb{R}^{\mathcal{Y}}_+$ for which $\kappa_r(c) \lessgtr \kappa_{\mathrm{bw}(r)}(c)$ and any $\bar{\jmath} \in \mathbb{R}^{\mathcal{R}_{\mathrm{fw}}}$,

$$\begin{aligned}\mathcal{L}(c,\bar{\jmath}) &= \Phi(c,\bar{\jmath}) + \Phi^*\big(c, F^{\mathrm{asym}}(c)\big) - F^{\mathrm{asym}}(c)\cdot\bar{\jmath}\\ &\qquad + \Phi^*_{F^{\mathrm{asym}}}\big(c, F^{\mathrm{sym}}(c)\big) + \tfrac{1}{2}\nabla\mathcal{I}_0(c)\cdot\Gamma\bar{\jmath} \qquad (4.2)\\ &= \Phi(c,\bar{\jmath}) + \Phi^*\big(c, F^{\mathrm{sym}}(c)\big) + \tfrac{1}{2}\nabla\mathcal{I}_0(c)\cdot\Gamma\bar{\jmath}\\ &\qquad + \Phi^*_{F^{\mathrm{sym}}}\big(c, F^{\mathrm{asym}}(c)\big) - F^{\mathrm{asym}}(c)\cdot\bar{\jmath}. \qquad (4.3)\end{aligned}$$

5. **Fisher and entropy bounds.** Observe that the forces $F^{\mathrm{sym}}(c), F^{\mathrm{asym}}(c)$ are well-defined precisely if $\kappa(c) \lessgtr \overleftarrow{\kappa}_{\mathrm{bw}}(c) \lessgtr \kappa_{\mathrm{bw}}(c)$. It turns out that many terms in (4.2) and (4.3) remain well-defined even when this condition is violated, see Example 5.2. More precisely, we may introduce the following notation, where the

equalities on the right are true if $\kappa_r(c) \lessgtr \kappa_{\mathrm{bw}(r)}(c)$:

$$\mathrm{Fi}_{\mathrm{a}}^{\mathrm{s}}(c) := \frac{1}{2} \sum_{r \in \mathcal{R}_{\mathrm{fw}} \cup \mathcal{R}_{\mathrm{bw}}} \left(\sqrt{\kappa_r(c)} - \sqrt{\overleftarrow{\kappa}_{\mathrm{bw}(r)}(c)} \right)^2 = \Phi^*_{F^{\mathrm{asym}}}\big(c, F^{\mathrm{sym}}(c)\big), \text{ and}$$

$$\mathrm{Fi}_{\mathrm{s}}^{\mathrm{a}}(c) := \frac{1}{2} \sum_{r \in \mathcal{R}_{\mathrm{fw}} \cup \mathcal{R}_{\mathrm{bw}}} \left(\sqrt{\kappa_r(c)} - \sqrt{\overleftarrow{\kappa}_r(c)} \right)^2 = \Phi^*_{F^{\mathrm{sym}}}\big(c, F^{\mathrm{asym}}(c)\big).$$

Remark 5.1. If we interpret the differences in as abstract gradients, then both quantities are of the form $\frac{1}{2}\sum_r (\nabla \sqrt{\kappa(c)})^2$, which coincides with the usual continuous-space Fisher information $\frac{1}{2}\int |\nabla \sqrt{c(x)}|^2 \, dx = \frac{1}{2}\int \frac{|\nabla c(x)|^2}{c(x)} \, dx$. Recalling Remark 2.9, the second quantity $\mathrm{Fi}_{\mathrm{s}}^{\mathrm{a}}(c)$ measures how far the system is from being in detailed balance. □

Example 5.2. *Consider a two-state linear network with* $\kappa_{xy}(c) := a_{xy} c_x / c_0^{\mathrm{eq}}$ *and equilibrium concentration* $c^{\mathrm{eq}} > 0$*, see also* [5]*. The time-reversed rates are then given by* $\overleftarrow{\kappa}_{xy}(c) := a_{yx} c_x / c_x^{\mathrm{eq}}$*. Then the force is* $F_{xy}(c) = \frac{1}{2} \log \frac{a_{xy} c_x c_y^{\mathrm{eq}}}{a_{yx} c_y c_x^{\mathrm{eq}}}$*, which decomposes into* $F_{xy}^{\mathrm{sym}}(c) = \frac{1}{2} \log \frac{c_x c_y^{\mathrm{eq}}}{c_y c_x^{\mathrm{eq}}}$ *and* $F_{xy}^{\mathrm{asym}}(c) = \frac{1}{2} \log \frac{a_{xy}}{a_{yx}}$*, and the two Fisher informations are* $\mathrm{Fi}_{\mathrm{a}}^{\mathrm{s}}(c) = \frac{1}{2}(a_{12}+a_{21})(\sqrt{c_1/c_1^{\mathrm{eq}}} - \sqrt{c_2/c_2^{\mathrm{eq}}})^2$ *and* $\mathrm{Fi}_{\mathrm{s}}^{\mathrm{a}}(c) = \frac{1}{2}(c_1/c_1^{\mathrm{eq}} + c_2/c_2^{\mathrm{eq}})(\sqrt{a_{12}} - \sqrt{a_{21}})^2$*. We thus see that the forces* $F(c), F^{\mathrm{sym}}(c)$ *are not well-defined on the boundary* $c = (1,0), (0,1)$*, but the Fisher informations are.*

The following proposition then generalises Corollary 4.3 to all concentrations, which possibly violate $\kappa(c) \lessgtr \overleftarrow{\kappa}_{\mathrm{bw}}(c) \lessgtr \kappa_{\mathrm{bw}}(c)$. The resulting inequality, bounding Fisher information and entropy by the rate functional, is known as a FIR inequality [4]. The inequality we prove here corresponds to the inequality in that paper when choosing their parameter $\lambda = 1/2$. Apart from the different proof strategy, our method shows that the gap is precisely quantified by (4.2), at least for paths that stay away from the boundary,

Proposition 5.3. *Assume that* κ *satisfies the conditions of* [12] *needed for the large-deviations Theorem 2.6 to hold. Let* $\mathcal{I}_0$ *be lower semicontinuous. Then for any path* $(c, \bar{\jmath}) \in W^{1,1}(0,T;\mathbb{R}_+^{\mathcal{Y}}) \times L^1(0,T;\mathbb{R}^{\mathcal{R}_{\mathrm{fw}}})$ *such that* $\mathcal{I}_0$ *is continuous in a neighbourhood of* $c(0)$*,*

$$\int_0^T \mathcal{L}\big(c(t), \bar{\jmath}(t)\big) \, dt \geq \int_0^T \mathrm{Fi}_{\mathrm{a}}^{\mathrm{s}}\big(c(t)\big) \, dt + \tfrac{1}{2}\mathcal{I}_0\big(c(T)\big) - \tfrac{1}{2}\mathcal{I}_0\big(c(0)\big). \tag{5.1}$$

Proof. We use two approximation steps. In the first step we consider $(c, \bar{\jmath}) \in \mathbb{R}_+^{\mathcal{Y}} \times \mathbb{R}^{\mathcal{R}}$ for which only $\kappa(c) \lessgtr \overleftarrow{\kappa}_{\mathrm{bw}}(c)$ holds. While keeping this pair fixed, we define a new reaction network with propensities and corresponding reaction rates:

$$k_r^{(V,\epsilon)}(c) := k_r^{(V)}(c) + \epsilon \overleftarrow{k}_r^{(V)}(c), \qquad \text{and} \qquad \kappa_r^{(\epsilon)}(c) := \kappa_r(c) + \epsilon \overleftarrow{\kappa}_r(c).$$

Then $\kappa^{(\epsilon)}(c) \lessgtr \overleftarrow{\kappa}_{\mathrm{bw}}^{(\epsilon)}(c) \lessgtr \kappa_{\mathrm{bw}}^{(\epsilon)}(c)$, so (4.2) holds. By convex duality $\Phi(c, \bar{\jmath}) + \Phi^*\big(c, F^{\mathrm{asym}}(c)\big) - F^{\mathrm{asym}}(c) \cdot \bar{\jmath} \geq 0$, which yields

$$\mathcal{L}^{(\epsilon)}(c, \bar{\jmath}) \geq \mathrm{Fi}_{\mathrm{a}}^{\mathrm{s}\,(\epsilon)}(c) + \tfrac{1}{2} \nabla \mathcal{I}_0^{(\epsilon)} \cdot \Gamma \bar{\jmath},$$

where the superscript ϵ denotes the functions corresponding to the modified rates $\kappa^{(\epsilon)}$. Recall from Theorem 2.3 that the network with rates $\overleftarrow{k}^{(V)}$ has the same invariant measure $\pi^{(V)}$ as the network with rates $k^{(V)}$; by linearity the same is true for the network with rates $k^{(V,\epsilon)}$. It follows that the functional $\mathcal{I}_0^{(\epsilon)} = \mathcal{I}_0$

remains unaltered. The Fisher information $\mathrm{Fi}_{\mathrm{a}}^{\mathrm{s}\,(\epsilon)}(c)$ is continuous in $\kappa^{(\epsilon)}(c)$ and hence converges to $\mathrm{Fi}_{\mathrm{a}}^{\mathrm{s}}(c)$. Finally for the cost function it follows by continuity that

$$\begin{aligned}\mathrm{Fi}_{\mathrm{a}}^{\mathrm{s}}(c) + \tfrac{1}{2}\nabla\mathcal{I}_0(c)\cdot\Gamma\bar{\jmath} &= \lim_{\epsilon\to 0} \mathrm{Fi}_{\mathrm{a}}^{\mathrm{s}\,(\epsilon)}(c) + \tfrac{1}{2}\nabla\mathcal{I}_0^{(\epsilon)}(c)\cdot\Gamma\bar{\jmath}\\ &\le \inf_{j\in\mathbb{R}_+^{\mathcal{R}}:\, j_r - j_{\mathrm{bw}(r)} = \bar{\jmath}_r} \limsup_{\epsilon\to 0} \mathcal{S}\big(j \mid \kappa^{(\epsilon)}(c)\big)\\ &= \inf_{j\in\mathbb{R}_+^{\mathcal{R}}:\, j_r - j_{\mathrm{bw}(r)} = \bar{\jmath}_r} \mathcal{S}\big(j \mid \kappa(c)\big) = \mathcal{L}(c,\bar{\jmath}). \qquad (5.2)\end{aligned}$$

Hence this inequality holds for any $(c,\bar{\jmath})$ for which $\kappa(c) \lessgtr \overleftarrow{\kappa}_{\mathrm{bw}}(c)$.

In the second step, take an arbitrary path $(c,\bar{\jmath}) \in W^{1,1}(0,T;\mathbb{R}_+^{\mathcal{Y}}) \times L^1(0,T;\mathbb{R}^{\mathcal{R}_{\mathrm{fw}}})$ such that $\mathcal{I}_0$ is continuous in a neighbourhood of $c(0)$; without loss of generality we may assume that $\int_0^T \mathcal{L}\big(c(t),\bar{\jmath}(t)\big)\,dt < \infty$. Recall from (2.11) that $\kappa(c) \lessgtr \overleftarrow{\kappa}_{\mathrm{bw}}(c)$ for almost every c. We distinguish between two cases. In the first case the path $c(t) \equiv c$ is constant in time (and hence $\Gamma\bar{\jmath} = 0$). Then we may approximate $c^{(\epsilon)} \to c$ such that $\kappa(c^{(\epsilon)}) \lessgtr \overleftarrow{\kappa}_{\mathrm{bw}}(c^{(\epsilon)})$, hence (5.2) holds for all t, and also in time-integrated form. Then $\mathcal{I}_0(c^{(\epsilon)}(T)) - \mathcal{I}_0(c^{(\epsilon)}(0)) = 0 = \mathcal{I}_0(c(T)) - \mathcal{I}_0(c(0))$, and clearly the Fisher information converges. Then

$$\begin{aligned}&\int_0^T \mathrm{Fi}_{\mathrm{a}}^{\mathrm{s}}\big(c(t)\big) + \tfrac{1}{2}\mathcal{I}_0\big(c(T)\big) - \tfrac{1}{2}\mathcal{I}_0\big(c(0)\big)\\ &\quad= \lim_{\epsilon\to 0}\int_0^T \mathrm{Fi}_{\mathrm{a}}^{\mathrm{s}}\big(c^{(\epsilon)}(t)\big) + \tfrac{1}{2}\mathcal{I}_0\big(c^{(\epsilon)}(T)\big) - \tfrac{1}{2}\mathcal{I}_0\big(c^{(\epsilon)}(0)\big)\\ &\quad\le \limsup_{\epsilon\to 0}\int_0^T \mathcal{L}\big(c^{(\epsilon)},\bar{\jmath}(t)\big)\,dt \le \inf_{\substack{j\in L^1(0,T;\mathbb{R}_+^{\mathcal{R}})\\ j_r - j_{\mathrm{bw}(r)} = \bar{\jmath}_r}} \limsup_{\epsilon\to 0}\int_0^T \mathcal{S}\big(j(t) \mid k(c^{(\epsilon)})\big)\,dt\\ &\quad= \inf_{\substack{j\in L^1(0,T;\mathbb{R}_+^{\mathcal{R}})\\ j_r - j_{\mathrm{bw}(r)} = \bar{\jmath}_r}} \limsup_{\epsilon\to 0}\int_0^T \Big[\mathcal{S}\big(j(t) \mid k(c)\big) + \sum_{r\in\mathcal{R}} j_r(t)\log\frac{\kappa(c)}{\kappa(c^{(\epsilon)})}\Big]\,dt\\ &\quad= \int_0^T \mathcal{L}\big(c,\bar{\jmath}(t)\big)\,dt.\end{aligned}$$

In the second case the path $c(t)$ is not constant. Take any $j \in L^1(0,T;\mathbb{R}_+^{\mathcal{R}})$ for which $j_r - j_{\mathrm{bw}(r)} = \bar{\jmath}_r$. Following [12, Lemmas 3.8 & 3.9] we add a little mass and convolute with a heat kernel so that $c^{(\epsilon)}, j^{(\epsilon)} \to c, j$ strongly in L^1-norm and $\int_0^T \mathcal{S}\big(j^{(\epsilon)}(t) \mid \kappa(c^{(\epsilon)}(t))\big)\,dt \to \int_0^T \mathcal{S}\big(j(t) \mid \kappa(c(t))\big)\,dt$. Since $c(t)$ is not constant, the convolved path $c^{(\epsilon)}(t)$ only passes through points $c(t)$ that violate $\kappa(c^{(\epsilon)}(t)) \lessgtr \overleftarrow{\kappa}_{\mathrm{bw}}(c^{(\epsilon)}(t))$ on a t-null set. Hence inequality (5.2) holds for this pair $(c^{(\epsilon)},\bar{\jmath}^{(\epsilon)})$ and almost all t, and hence also in integrated form. The convergence of the Fisher information follows by dominated convergence, where we recall from Assumption 2.5 that the rates κ are bounded, and the reversed rates $\overleftarrow{\kappa}$ are bounded by (2.12). The convergence of $\mathcal{I}_0(c^{(\epsilon)}(T))$ is by lower semicontinuity and the convergence of $\mathcal{I}_0(c^{(\epsilon)}(0))$ by assumption. We conclude that the inequality (5.1) holds for any path as claimed. □

We can now also derive an inequality that bounds the other Fisher information together with the irreversible work.

Proposition 5.4. *Assume that κ satisfies the conditions of [12] needed for the large-deviations Theorem 2.6 to hold. Then for any path $(c,\bar{\jmath}) \in W^{1,1}(0,T;\mathbb{R}_+^{\mathcal{Y}}) \times$*

$L^1(0,T;\mathbb{R}^{\mathcal{R}})$ *such that* $\kappa_r(c(t)) \lessgtr \overleftarrow{\kappa}_r(c(t))$ *for almost all* $t \in (0,T)$ *and* $r \in \mathcal{R}$,

$$\int_0^T \mathcal{L}\big(c(t), \bar{\jmath}(t)\big)\, dt \geq \int_0^T \mathrm{Fi}_{\mathrm{s}}^{\mathrm{a}}\big(c(t)\big)\, dt - \int_0^T F^{\mathrm{asym}}(c(t)) \cdot \bar{\jmath}(t)\, dt.$$

Proof. This is the same argument as Proposition (5.3), but starting from (4.3) rather then (4.2). Note that the condition $\kappa_r(c(t)) \lessgtr \overleftarrow{\kappa}_r(c(t))$ is now needed for $F^{\mathrm{asym}}(c)$ to be well-defined. □

We finally remark that the irreversible work $\int_0^T F^{\mathrm{asym}}(c(t)) \cdot \bar{\jmath}(t)\, dt$ does not necessarily have a sign, but in the other direction we have by convex duality

$$\int_0^T F^{\mathrm{asym}}(c(t)) \cdot \bar{\jmath}(t)\, dt \leq \int_0^T \Phi\big(c(t), \bar{\jmath}(t)\big)\, dt + \int_0^T \Phi^*\big(c(t), F^{\mathrm{asym}}(c(t))\big)\, dt.$$

Acknowledgements. DRMR was funded by Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1114 "Scaling Cascades in Complex Systems", Project C08. JZ received funding through a Royal Society Wolfson Research Merit Award.

REFERENCES

[1] S. Adams, N. Dirr, M. Peletier and J. Zimmer, Large deviations and gradient flows, *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, **371** (2013), 20120341, 17, URL http://dx.doi.org/10.1098/rsta.2012.0341.

[2] D. F. Anderson and T. G. Kurtz, Continuous time Markov chain models for chemical reaction networks, in *Design and Analysis of Biomolecular Circuits* (eds. H. Koeppl, G. Setti, M. di Bernardo and D. Densmore), Springer, 2011, 3–42.

[3] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio and C. Landim, Macroscopic fluctuation theory, *Rev. Modern Phys.*, **87** (2015), 593–636, URL http://dx.doi.org/10.1103/RevModPhys.87.593.

[4] B. Hilder, M. A. Peletier, U. Sharma and O. Tse, An inequality connecting entropy distance, Fisher information and large deviations, *Stochastic Process. Appl.*, to appear, 2019, URL https://www.sciencedirect.com/science/article/pii/S0304414919300687.

[5] M. Kaiser, R. L. Jack and J. Zimmer, Canonical structure and orthogonality of forces and currents in irreversible Markov chains, *J. Stat. Phys.*, **170** (2018), 1019–1050, URL https://doi.org/10.1007/s10955-018-1986-0.

[6] T. G. Kurtz, Solutions of ordinary differential equations as limits of pure jump Markov processes, *J. Appl. Probability*, **7** (1970), 49–58, URL https://doi.org/10.2307/3212147.

[7] C. Maes and K. Netočný, Canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states, *Europhys. Lett. EPL*, **82** (2008), Art. 30003, 6, URL http://dx.doi.org/10.1209/0295-5075/82/30003.

[8] C. Maes, Frenetic bounds on the entropy production, *Phys. Rev. Lett.*, **119** (2017), 160601, URL https://link.aps.org/doi/10.1103/PhysRevLett.119.160601.

[9] 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 Anal.*, **41** (2014), 1293–1327, URL http://dx.doi.org/10.1007/s11118-014-9418-5.

[10] A. Mielke, R. I. A. Patterson, M. A. Peletier and D. R. M. Renger, Non-equilibrium thermodynamical principles for chemical reactions with mass-action kinetics, *SIAM J. Appl. Math.*, **77** (2017), 1562–1585, URL https://doi.org/10.1137/16M1102240.

[11] L. Onsager and S. Machlup, Fluctuations and irreversible processes., *Physical Rev. (2)*, **91** (1953), 1505–1512, URL https://journals.aps.org/pr/abstract/10.1103/PhysRev.91.1505.

[12] R. I. A. Patterson and D. R. M. Renger, Large deviations of jump process fluxes, *Math. Phys. Anal. Geom.*, **22** (2019), Paper No. 21, 32, URL https://doi.org/10.1007/s11040-019-9318-4.

[13] D. R. M. Renger, Flux Large Deviations of Independent and Reacting Particle Systems, with Implications for Macroscopic Fluctuation Theory, *J. Stat. Phys.*, **172** (2018), 1291–1326, URL https://doi.org/10.1007/s10955-018-2083-0.

[14] D. R. M. Renger, Gradient and GENERIC systems in the space of fluxes, applied to reacting particle systems, *Entropy*, **20**, URL http://www.mdpi.com/1099-4300/20/8/596.

[15] J. Schnakenberg, Network theory of microscopic and macroscopic behavior of master equation systems, *Rev. Modern Phys.*, **48** (1976), 571–585, URL http://dx.doi.org/10.1103/RevModPhys.48.571.

E-mail address: renger@wias-berlin.de

E-mail address: J.Zimmer@bath.ac.uk