

PAPER

Multidimensional thermodynamic uncertainty relations

To cite this article: Andreas Dechant 2019 *J. Phys. A: Math. Theor.* **52** 035001

View the [article online](#) for updates and enhancements.

You may also like

- [Stronger reverse uncertainty relation for multiple incompatible observables](#)
Xiao Zheng, Ai-Ling Ji and Guo-Feng Zhang
- [Uncertainty, joint uncertainty, and the quantum uncertainty principle](#)
Varun Narasimhachar, Alireza Poostindouz and Gilad Gour
- [Uncertainty relation based on metric-adjusted skew information with quantum memory](#)
Ruonan Ren and Yongming Li

Multidimensional thermodynamic uncertainty relations

Andreas Dechant 

WPI-Advanced Institute of Materials Research (WPI-AIMR), Tohoku University,
Sendai 980-8577, Japan

E-mail: andreas.dechant@outlook.com

Received 27 September 2018, revised 16 November 2018

Accepted for publication 26 November 2018

Published 18 December 2018



CrossMark

Abstract

We extend a class of recently derived thermodynamic uncertainty relations to vector-valued observables. In contrast to the scalar-valued observables examined previously, this multidimensional thermodynamic uncertainty relation provides a natural way to study currents in high-dimensional systems and to obtain relations between different observables. Our proof is based on the generalized Crámer–Rao inequality, which we interpret as a relation between physical observables and the Fisher information. This allows us to develop high-dimensional versions of both the original, steady state uncertainty relation and the more recently obtained generalized uncertainty relation for time-periodic systems. We apply the multidimensional uncertainty relation to obtain a new constraint on the performance of steady-state heat engines, which is tighter than previous bounds and reveals the role of heat-work correlations. As a second application, we show that the uncertainty relation is connected to a bound on the differential mobility. As a result of this connection, we find that a necessary condition for equality in the uncertainty relation is that the system obeys the equilibrium fluctuation–dissipation relation.

Keywords: current fluctuations, entropy production, mobility, heat engines, information geometry

(Some figures may appear in colour only in the online journal)

A remarkable property of stochastic transport in steady-state systems is the so-called thermodynamic uncertainty relation (TUR), which was first conjectured by Barato *et al* [1] and subsequently proven by Gingrich *et al* [2]. This relation states that the square of the average current is bounded from above by the variance of the current times the entropy production. It thus provides a universal relation between a current, its fluctuations and the thermodynamic cost of driving the current. This relation was later generalized from the long-time limit to steady state systems at finite time [3–5], to time-periodic systems [6–8] and to specific cases

of stochastic time-dependent driving [9]. In all these formulations, the stochastic current is a scalar quantity. While many toy models involve one-dimensional settings, where considering scalar observables is sufficient, realistic physical systems are placed in three-dimensional space, where both the direction and the magnitude of a current may be important. Further, one may be interested in the behavior of several currents, which are generally correlated. Thus, it is desirable to obtain a bound in the spirit of the TUR, which explicitly takes into account the dimensionality of the system.

In this work, we extend the family of TURs to vector-valued observables. In the process, we will re-derive the TUR from a central result of information theory, the Crámer–Rao bound [10, 11]. This derivation shows that the TUR can be understood as a consequence of the information-geometric properties of the space of path probability densities [12]. The resulting multidimensional thermodynamic uncertainty relation (MTUR) encompasses both the finite-time result for steady states [3–5] and the more recently obtained result for time-periodic systems [6–8]. The derivation from the Crámer–Rao bound not only allows for a straightforward extension to multidimensional systems with vector-valued observables. It further shows that the result obtained in [7] is not restricted to time-periodic driving, but also holds for arbitrary time-dependent systems under the assumption of certain regularity conditions. Throughout this paper, we will focus on Langevin dynamics, however, we stress that all the results also hold for Markov jump dynamics, with the specific expressions given in the appendix.

1. Fisher information and the Crámer–Rao bound

The mathematical foundation for our results is the generalized Crámer–Rao bound. Suppose that we have a probability density $P(\omega, \theta) = P(\Omega = \omega | \theta)$ for some general random variable Ω and depending on a set of M parameters $\theta = \{\theta_1, \dots, \theta_M\}$. Note that Ω may present the instantaneous value of some stochastic process $\mathbf{x}(t)$, but may equally well be the path of the stochastic process $\{\mathbf{x}(t)\}_{t \in [0, \mathcal{T}]}$ during the time interval $[0, \mathcal{T}]$. We further consider a set of K observables $\mathbf{r}(\Omega) = \{r_1(\Omega), \dots, r_K(\Omega)\}$ with average $\langle \mathbf{r} \rangle_\theta = \int d\omega \mathbf{r}(\omega) P(\omega, \theta)$. Depending on the random variable Ω , the integral $d\omega$ can represent a sum over discrete states, an integral over a set of continuous variables or a path integral. The generalized Crámer–Rao bound is written as an operator inequality [13]

$$\mathbf{J}_r(\theta)^T \Xi_r(\theta)^{-1} \mathbf{J}_r(\theta) \leq \mathbf{I}(\theta), \quad (1)$$

where T denotes transposition and we defined the $K \times M$ Jacobian $\mathbf{J}_r(\theta)$ of $\langle \mathbf{r} \rangle_\theta$ with respect to θ ,

$$(\mathbf{J}_r(\theta))_{ij} = \partial_{\theta_j} \langle r_i \rangle_\theta, \quad (2)$$

the positive definite $K \times K$ covariance matrix $\Xi_r(\theta)$,

$$(\Xi_r(\theta))_{ij} = \langle r_i r_j \rangle_\theta - \langle r_i \rangle_\theta \langle r_j \rangle_\theta, \quad (3)$$

and the positive semidefinite $M \times M$ Fisher information matrix

$$(\mathbf{I}(\theta))_{ij} = \int d\omega \frac{\partial_{\theta_i} P(\omega, \theta) \partial_{\theta_j} P(\omega, \theta)}{P(\omega, \theta)}. \quad (4)$$

The operator inequality (1) is interpreted as $\mathbf{I} - \mathbf{J}_r^T \Xi_r^{-1} \mathbf{J}_r$ being a positive semidefinite matrix, i.e. for some arbitrary vector $\mathbf{v} \in \mathbb{R}^M$ we have $\mathbf{v}^T (\mathbf{I} - \mathbf{J}_r^T \Xi_r^{-1} \mathbf{J}_r) \mathbf{v} \geq 0$. The physical interpretation of the generalized Crámer–Rao bound is most conveniently made clear by focusing on

the case of a single observable r and parameter θ_i , in which case equation (1) simplifies to the inequality

$$\frac{(\partial_{\theta_i} \langle r \rangle_\theta)^2}{\langle \Delta r^2 \rangle_\theta} \leq (\mathbf{I}(\boldsymbol{\theta}))_{ii}. \quad (5)$$

The left-hand side is the change in the average of the observable r due to a change in the parameter θ_i , relative to the fluctuations of r . The left-hand thus side tells us how much information on the parameter θ_i the observable r contains: if the average of r is almost independent of θ_i (or the fluctuations of r are very large), then a measurement of r will not allow us to make any statement about the value of θ_i . On the other hand, if $\langle r \rangle_\theta$ changes substantially by varying θ_i , then the measurement of r can potentially provide us with a good estimate of θ_i . The Crámer–Rao bound states that this information on θ_i contained in any observable is always less than the corresponding Fisher information $\mathbf{I}_{ii}$. In other words, the probability density itself contains the maximum amount of information; measuring any observable can only yield less information.

2. Path Fisher information for Langevin dynamics

To make the connection between these information-theoretic ideas and a concrete physical situation, consider the N -dimensional diffusion process $\mathbf{x}(t) = \{x_1(t), \dots, x_N(t)\}$ described by the Itô–Langevin equation [14],

$$\dot{\mathbf{x}}(t) = \mathbf{a}(\mathbf{x}(t), t, \boldsymbol{\theta}) + \sqrt{2\mathbf{B}(\mathbf{x}(t), t)} \cdot \boldsymbol{\xi}(t), \quad (6)$$

with mutually independent Gaussian white noises $\xi_i(t)$. Equivalently, we have the Fokker–Planck equation for the probability density $P(\mathbf{x}, t, \boldsymbol{\theta})$ and current $\mathbf{j}(\mathbf{x}, t, \boldsymbol{\theta})$ [14],

$$\begin{aligned} \partial_t P(\mathbf{x}, t, \boldsymbol{\theta}) &= -\nabla \cdot \mathbf{j}(\mathbf{x}, t, \boldsymbol{\theta}) \\ \mathbf{j}(\mathbf{x}, t, \boldsymbol{\theta}) &= \left(\mathbf{a}(\mathbf{x}, t, \boldsymbol{\theta}) - \nabla \mathbf{B}(\mathbf{x}, t) \right) P(\mathbf{x}, t, \boldsymbol{\theta}) \end{aligned} \quad (7)$$

with N -dimensional drift vector (or generalized forces) $\mathbf{a}(\mathbf{x}, t, \boldsymbol{\theta})$ and symmetric, positive definite $N \times N$ diffusion matrix $\mathbf{B}(\mathbf{x}, t)$. We assume that the generalized forces $\mathbf{a}(\mathbf{x}, t, \boldsymbol{\theta})$ depend on a set of control parameters $\boldsymbol{\theta}$ as

$$\mathbf{a}(\mathbf{x}, t, \boldsymbol{\theta}) = \mathbf{a}_0(\mathbf{x}, t) + \sum_{i=1}^M \theta_i \mathbf{a}_i(\mathbf{x}, t). \quad (8)$$

Physically, we may take $\mathbf{a}_0$ to define a reference system and the $\mathbf{a}_i$ to be perturbations to this reference system. For small values of the parameters $\boldsymbol{\theta}$ the Fisher information matrix then describes the linear response behavior of the system. We define the random variable ω as the path $\{\mathbf{x}(t)\}_{t \in [0, T]}$ of the diffusion process corresponding to the above Langevin dynamics. The probability density of the path is given by the Onsager–Machlup functional [14]

$$\begin{aligned} \mathbb{P}(\{\mathbf{x}(t)\}_{t \in [0, T]}) &\propto \exp \left[-\mathcal{S}[\mathbf{x}(t), \boldsymbol{\theta}] \right] P_0(\mathbf{x}(0), \boldsymbol{\theta}) \\ \text{with} \quad \mathcal{S}[\mathbf{x}(t), \boldsymbol{\theta}] &= \frac{1}{4} \int_0^T dt \left(\dot{\mathbf{x}}(t) - \mathbf{a}(t) \right)^T \mathbf{B}(t)^{-1} \left(\dot{\mathbf{x}}(t) - \mathbf{a}(t) \right), \end{aligned} \quad (9)$$

and a prefactor that depends only on the diffusion matrix $\mathbf{B}$. Note that here and in the following we use the short-hand notation $f(t) = f(\mathbf{x}(t), t)$ to denote functions evaluated along the

trajectory. For now, we ignore a possible dependence of the initial state $P_0(\mathbf{x}, \boldsymbol{\theta})$ on the parameters $\boldsymbol{\theta}$. Then, the derivative of the path probability with respect to θ_i is given by

$$\frac{\partial_{\theta_i} \mathbb{P}(\{\mathbf{x}(t)\}_{t \in [0, T]})}{\mathbb{P}(\{\mathbf{x}(t)\}_{t \in [0, T]})} = \frac{1}{2} \int_0^T dt \mathbf{a}_i(t)^T \mathbf{B}(t)^{-1} (\dot{\mathbf{x}}(t) - \mathbf{a}(t)), \quad (10)$$

where we used the symmetry of $\mathbf{B}$. The Fisher information matrix of the path probability density is then given by the following path integral

$$\begin{aligned} (\mathbb{I}(\boldsymbol{\theta}))_{ij} = & \frac{1}{4} \int \mathcal{D}\mathbf{x}(t) \left(\int_0^T dt \int_0^T ds \mathbf{a}_i(t)^T \mathbf{B}(t)^{-1} (\dot{\mathbf{x}}(t) - \mathbf{a}(t)) \right. \\ & \left. \times \mathbf{a}_j(s)^T \mathbf{B}(s)^{-1} (\dot{\mathbf{x}}(s) - \mathbf{a}(s)) \right) \mathbb{P}(\{\mathbf{x}(t)\}_{t \in [0, T]}). \end{aligned} \quad (11)$$

Along any path of the diffusion process, we have $\dot{\mathbf{x}}(t) - \mathbf{a}(t) = \sqrt{2\mathbf{B}(t)} \cdot \boldsymbol{\xi}(t)$, and thus

$$\begin{aligned} (\mathbb{I}(\boldsymbol{\theta}))_{ij} = & \frac{1}{2} \int \mathcal{D}\mathbf{x}(t) \left(\int_0^T dt \int_0^T ds \mathbf{a}_i(t)^T \sqrt{\mathbf{B}(t)^{-1}} \boldsymbol{\xi}(t) \right. \\ & \left. \times \boldsymbol{\xi}(s)^T \sqrt{\mathbf{B}(s)^{-1}} \mathbf{a}_j(s) \right) \mathbb{P}(\{\mathbf{x}(t)\}_{t \in [0, T]}). \end{aligned} \quad (12)$$

Since the noises are white and uncorrelated, we can write

$$\boldsymbol{\xi}(t) \boldsymbol{\xi}(s)^T = \mathbf{1} \delta(t - s), \quad (13)$$

where $\mathbf{1}$ is the $N \times N$ identity matrix. Thus, the integral over s becomes trivial,

$$(\mathbb{I}(\boldsymbol{\theta}))_{ij} = \frac{1}{2} \int \mathcal{D}\mathbf{x}(t) \left(\int_0^T dt \mathbf{a}_i(t)^T \mathbf{B}(t)^{-1} \mathbf{a}_j(t) \right) \mathbb{P}(\{\mathbf{x}(t)\}_{t \in [0, T]}). \quad (14)$$

Since the quantity over which the path integral is performed only depends on the single time t , we can replace the path probability density with the one-time probability density and obtain,

$$(\mathbb{I}(\boldsymbol{\theta}))_{ij} = \frac{1}{2} \int_0^T dt \langle \mathbf{a}_i^T \mathbf{B}^{-1} \mathbf{a}_j \rangle_{t, \boldsymbol{\theta}} + (\mathbf{I}_0(\boldsymbol{\theta}))_{ij}, \quad (15)$$

where the average $\langle \dots \rangle_{t, \boldsymbol{\theta}}$ is taken with respect to the solution of equation (7) with parameter values $\boldsymbol{\theta}$ and $\mathbf{I}_0(\boldsymbol{\theta})$ is the Fisher information matrix of the initial state. For the case of a single parameter, this expression was recently obtained in [12]. The Fisher information matrix of the path probability density is thus explicitly expressed in terms of the perturbing generalized forces $\mathbf{a}_i$. By the Crámer–Rao bound (1), this quantity bounds the response of *any* path-dependent observable $\mathbf{r}[\mathbf{x}(t)]$ to a change in the parameters $\boldsymbol{\theta}$. We can write the Crámer–Rao bound in a more intuitive way by defining

$$d\langle \mathbf{r} \rangle_{\boldsymbol{\theta}} = \langle \mathbf{r} \rangle_{\boldsymbol{\theta} + d\boldsymbol{\theta}} - \langle \mathbf{r} \rangle_{\boldsymbol{\theta}} = \mathbf{J}_r(\boldsymbol{\theta}) d\boldsymbol{\theta}, \quad (16)$$

which results in

$$d\langle \mathbf{r} \rangle_{\boldsymbol{\theta}}^T \boldsymbol{\Xi}_r(\boldsymbol{\theta})^{-1} d\langle \mathbf{r} \rangle_{\boldsymbol{\theta}} \leq d\boldsymbol{\theta}^T \mathbb{I}(\boldsymbol{\theta}) d\boldsymbol{\theta}. \quad (17)$$

The left-hand side is the response of the observable $\mathbf{r}$ relative to its fluctuations. The right-hand side can be related to another information-theoretic quantity, the Kullback–Leibler divergence between two probability densities P and Q

$$D_{\text{KL}}(P|Q) = \int d\omega P \ln \left(\frac{P}{Q} \right). \quad (18)$$

The Kullback–Leibler divergence is positive and vanishes only at $P = Q$, its global minimum. The curvature around the minimum is given by the Fisher information matrix,

$$D_{\text{KL}}(P(\boldsymbol{\theta} + d\boldsymbol{\theta})|P(\boldsymbol{\theta})) = \frac{1}{2} d\boldsymbol{\theta}^T \mathbb{I}(\boldsymbol{\theta}) d\boldsymbol{\theta} + O(d\boldsymbol{\theta}^3) \quad (19)$$

and we thus arrive at

$$d\langle \mathbf{r} \rangle_{\boldsymbol{\theta}}^T \Xi_r(\boldsymbol{\theta})^{-1} d\langle \mathbf{r} \rangle_{\boldsymbol{\theta}} \leq 2D_{\text{KL}}(\mathbb{P}(\boldsymbol{\theta} + d\boldsymbol{\theta})|\mathbb{P}(\boldsymbol{\theta})). \quad (20)$$

This is the extension of the fluctuation-response inequality (FRI) derived in [15] to vector-valued observables and more than one perturbation. In the multidimensional case, the variance of the observable $\mathbf{r}$ is replaced by its covariance matrix. This relation can be understood as the connection between the macroscopic and microscopic response of the system to a change of the parameters $\boldsymbol{\theta}$. The left-hand side is the response of some observable (i.e. a macroscopic, ensemble-averaged quantity) relative to its fluctuations. By contrast, the right hand side quantifies the change in the path probability (i.e. the microscopic dynamics) as a result of the perturbation. The FRI thus states that the response of any macroscopic observable is bounded by the change in the microscopic dynamics.

3. Multidimensional thermodynamic uncertainty relations

In [15] the steady state thermodynamic uncertainty relation (TUR) was derived from the FRI (20) by using a special choice for the perturbation. Here, we extend this derivation to the multidimensional case and explicitly time-dependent dynamics. The type of observables described by the TUR are time-integrated currents, defined by

$$\mathbf{r}[\mathbf{x}(t)] = \int_0^T dt \mathbf{Z}(\mathbf{x}(t), t) \circ \dot{\mathbf{x}}(t), \quad (21)$$

where $\mathbf{Z}$ is an arbitrary $K \times N$ matrix-valued function and $\circ$ denotes the Stratonovich-product. The average of such a time-integrated current is given in terms of the probability current $\mathbf{j}$,

$$\langle \mathbf{r} \rangle_{\boldsymbol{\theta}} = \int_0^T dt \int d\mathbf{x} \mathbf{Z}(\mathbf{x}, t) \mathbf{j}(\mathbf{x}, t, \boldsymbol{\theta}). \quad (22)$$

We now consider a single perturbation given by

$$\mathbf{a}_1(\mathbf{x}, t) = \frac{(\mathbf{a}_0(\mathbf{x}, t) - \nabla \mathbf{B}(\mathbf{x}, t)) \mathcal{P}(\mathbf{x})}{P(\mathbf{x}, t, 0)}, \quad (23)$$

with an arbitrary, time-independent probability density $\mathcal{P}(\mathbf{x}) > 0$, $\int d\mathbf{x} \mathcal{P}(\mathbf{x}) = 1$. For this choice, the Fokker–Planck equation (7) reads

$$\begin{aligned} \partial_t P(\mathbf{x}, t, \theta) = & -\nabla \left(\mathbf{a}_0(\mathbf{x}, t) + \theta \frac{(\mathbf{a}_0(\mathbf{x}, t) - \nabla \mathbf{B}(\mathbf{x}, t)) \mathcal{P}(\mathbf{x})}{P(\mathbf{x}, t, 0)} \right. \\ & \left. - \nabla \mathbf{B}(\mathbf{x}, t) \right) P(\mathbf{x}, t, \theta). \end{aligned} \quad (24)$$

Formally, this equation is solved by

$$P(\mathbf{x}, t, \theta) = P(\mathbf{x}, t, 0) + \theta \left(P(\mathbf{x}, t, 0) - \mathcal{P}(\mathbf{x}) \right). \quad (25)$$

However, while this solution is normalized, it is not positive for arbitrary $\mathcal{P}$ and θ . The reason is that, for an arbitrary choice of $\mathcal{P}$, the drift coefficient $\mathbf{a}_1$ can become very large at points where the probability density in the unperturbed system is small and thus the dynamics is no longer well-defined. Demanding that the solution should be a proper positive probability density restricts the possible choices of $\mathcal{P}$,

$$\mathcal{P}(\mathbf{x}) < \frac{1 + \theta}{\theta} P(\mathbf{x}, t, 0). \quad (26)$$

This typically can be satisfied by a normalized probability density only for sufficiently small $\theta \ll 1$. Supposing we have such a $\mathcal{P}$, the probability current is given by

$$\mathbf{j}(\mathbf{x}, t, \theta) = (1 + \theta) \mathbf{j}(\mathbf{x}, t, 0), \quad (27)$$

i.e. the additional drift vector leads to a rescaling of the probability currents. The same is then obviously true for the average of the time-integrated current equation (21),

$$\partial_\theta \langle \mathbf{r} \rangle_\theta = \langle \mathbf{r} \rangle_0. \quad (28)$$

From equations (1) and (15), we then immediately have the multidimensional version of the generalized thermodynamic uncertainty relation (GTUR),

$$\langle \mathbf{r} \rangle^T \Xi_r^{-1} \langle \mathbf{r} \rangle \leq \frac{1}{2} \Sigma, \quad (29)$$

where the quantity Σ has a structure similar to the entropy production,

$$\Sigma = \int_0^\mathcal{T} dt \int d\mathbf{x} \frac{\boldsymbol{\nu}^T(\mathbf{x}, t) \mathbf{B}^{-1}(\mathbf{x}, t) \boldsymbol{\nu}(\mathbf{x}, t)}{P(\mathbf{x}, t)} \\ \text{with } \boldsymbol{\nu}(\mathbf{x}, t) = \left(\mathbf{a}_0(\mathbf{x}, t) - \nabla \mathbf{B}(\mathbf{x}, t) \right) \mathcal{P}(\mathbf{x}). \quad (30)$$

For a scalar current and periodic driving, this GTUR has been derived for a jump process in [7]. The above derivation shows that the GTUR has a straightforward extension to vector-valued currents by replacing the variance of the current with the covariance matrix. Further, the result holds not only for periodic but for arbitrary time-dependent systems. We remark that for a time-periodic system with $P(\mathbf{x}, \mathcal{T}) = P(\mathbf{x}, 0)$ and a current without explicit time-dependence $\mathbf{Z}(\mathbf{x}, t) \equiv \mathbf{Z}(\mathbf{x})$, we obtain the bound derived in [6] for the choice

$$\mathbf{a}_1(\mathbf{x}, t) = \frac{\int_0^\mathcal{T} dt \mathbf{j}(\mathbf{x}, t)}{\mathcal{T} P(\mathbf{x}, t)}, \quad (31)$$

yielding the inequality

$$\langle \mathbf{r} \rangle^T \Xi_r^{-1} \langle \mathbf{r} \rangle \leq \frac{1}{2} \overline{\Sigma} \quad (32)$$

with the quantity $\bar{\Sigma}$ defined in terms of the time-averaged probability current

$$\bar{\Sigma} = \int_0^T dt \int d\mathbf{x} \frac{\bar{\mathbf{j}}^T(\mathbf{x}) \mathbf{B}^{-1}(\mathbf{x}, t) \bar{\mathbf{j}}(\mathbf{x})}{P(\mathbf{x}, t)} \quad (33)$$

with $\bar{\mathbf{j}}(\mathbf{x}) = \frac{1}{T} \int_0^T dt \mathbf{j}(\mathbf{x}, t)$.

For steady-state systems, we can choose $\mathcal{P}(\mathbf{x}) = P^{\text{st}}(\mathbf{x})$. For this choice, we have $\nu(\mathbf{x}) = \mathbf{j}^{\text{st}}(\mathbf{x})$ and the quantity Σ is precisely the entropy production ΔS during the time interval $[0, T]$, thus yielding the steady-state multidimensional thermodynamic uncertainty relation (MTUR)

$$\langle \mathbf{r} \rangle^T \bar{\Sigma}_r^{-1} \langle \mathbf{r} \rangle \leq \frac{1}{2} \Delta S. \quad (34)$$

For scalar currents, this inequality has been extensively discussed in the literature [1–5]. The present generalization to vector-valued currents has several advantages. First, it allows discussing currents in systems with more than one spatial dimension in a natural way. Second, the bound provided by equation (34) is tighter than the bound on any scalar current formed by a linear combination of the individual currents, see below. Finally, the MTUR explicitly takes into account correlations, providing additional insight into the relation between different currents. If the observables r_i are independent of each other, then their covariance matrix is diagonal, $\langle \Delta r_i \Delta r_j \rangle = \delta_{ij} \langle \Delta r_i^2 \rangle$, and we obtain

$$\sum_i \frac{\langle r_i \rangle^2}{\langle \Delta r_i^2 \rangle} \leq \frac{1}{2} \Delta S. \quad (35)$$

Since all the terms on the left-hand side are positive, it is obvious that each of the observables obeys the uncertainty relation on its own. However, the bound on the sum is obviously tighter. For two observables r_1 and r_2 , we can write the bound explicitly,

$$\begin{aligned} \langle \Delta r_2^2 \rangle \langle r_1 \rangle^2 - 2 \langle \Delta r_1 \Delta r_2 \rangle \langle r_1 \rangle \langle r_2 \rangle + \langle \Delta r_1^2 \rangle \langle r_2 \rangle^2 \\ \leq \frac{1}{2} \left(\langle \Delta r_1^2 \rangle \langle \Delta r_2^2 \rangle - \langle \Delta r_1 \Delta r_2 \rangle^2 \right) \Delta S. \end{aligned} \quad (36)$$

This bound involves the variances of the individual currents as well as their correlation. We will discuss some consequence of the explicit dependence on the correlations in the next section. We remark that we may also obtain a joint bound on $\langle r_1 \rangle$ and $\langle r_2 \rangle$ by considering the scalar observable $\rho(\mathbf{x}) = \cos(\varphi) r_1(\mathbf{x}) + \sin(\varphi) r_2(\mathbf{x})$, i.e. the projection of $\mathbf{r}$ onto an arbitrary unit vector. The corresponding scalar uncertainty relation $\langle \rho \rangle^2 / \langle \Delta \rho^2 \rangle \leq \Delta S / 2$ then yields a bound that is generally less tight than equation (36) and tends to equation (36) upon maximization with respect to the angle φ . Thus the bound on vector-valued observables equation (34) is always tighter than the bound on any scalar formed by a linear combination of the entries of $\mathbf{r}$.

4. Consequences of the multidimensional TUR

4.1. Power-efficiency tradeoff relations

One of the consequences of the TUR is a tradeoff relation between power and efficiency for steady-state heat engines [5, 7, 16–18]. For an engine operating between two heat baths at temperatures T_c and $T_h > T_c$, the steady-state entropy production rate $\sigma^{\text{st}} = \Delta S / T$ can be written as

$$\sigma^{\text{st}} = -\frac{1}{T_c} \langle \dot{q}_c \rangle + \frac{1}{T_h} \langle \dot{q}_h \rangle, \quad (37)$$

where q_c and q_h are the amounts of heat dissipated into the cold, respectively absorbed from the hot heat bath, both of which are time-integrated currents of the type equation (21). In terms of the power output of the engine $\langle \dot{w} \rangle = \langle \dot{q}_h \rangle - \langle \dot{q}_c \rangle$ and the Carnot efficiency $\eta_C = 1 - T_c/T_h$, this can be written as

$$\sigma^{\text{st}} = -\frac{1}{T_c} (\langle \dot{w} \rangle - \eta_C \langle \dot{q}_h \rangle) = \frac{\langle \dot{q}_h \rangle}{T_c} (\eta_C - \eta), \quad (38)$$

where we introduced the efficiency $\eta = \langle \dot{w} \rangle / \langle \dot{q}_h \rangle$. Obviously, the condition that $\sigma^{\text{st}} \geq 0$ implies $\eta \leq \eta_C$ if the engine is supposed to perform work at a positive rate. From the one-dimensional TUR for the heat current $\dot{q}_h$ and work current $\dot{w}$ respectively, we have

$$\begin{aligned} \langle \dot{q}_h \rangle^2 &\leq D_{qq} \sigma^{\text{st}} = \frac{D_{qq}}{T_c} \langle \dot{q}_h \rangle (\eta_C - \eta) \\ \langle \dot{w} \rangle^2 &\leq D_{ww} \sigma^{\text{st}} = \frac{D_{ww}}{T_c} \frac{\langle \dot{w} \rangle}{\eta} (\eta_C - \eta), \end{aligned} \quad (39)$$

where $D_{qq} = \lim_{\mathcal{T} \rightarrow \infty} \langle \Delta q_h^2 \rangle / (2\mathcal{T})$ and $D_{ww} = \lim_{\mathcal{T} \rightarrow \infty} \langle \Delta w^2 \rangle / (2\mathcal{T})$ characterize the fluctuations of the input heat and output work. These bounds imply the tradeoff-relation [16, 18]

$$\langle \dot{w} \rangle \leq \min \left(\frac{D_{qq}}{T_c} \eta (\eta_C - \eta), \frac{D_{ww}}{T_c} \frac{\eta_C - \eta}{\eta} \right), \quad (40)$$

which states that work current has to vanish as the efficiency approaches the Carnot efficiency, assuming the fluctuations of heat and work remain finite. By contrast, the MTUR (36) yields the bound

$$\begin{aligned} \langle \dot{q}_h \rangle^2 D_{ww} - 2 \langle \dot{q}_h \rangle \langle \dot{w} \rangle D_{qw} + \langle \dot{w} \rangle^2 D_{qq} \\ \leq \frac{1}{T_c} (D_{qq} D_{ww} - D_{qw}^2) (\eta_C - \eta) \langle \dot{q}_h \rangle. \end{aligned} \quad (41)$$

Replacing $\langle \dot{q}_h \rangle = \langle \dot{w} \rangle / \eta$ this can be written as

$$\langle \dot{w} \rangle \leq \frac{1}{T_c} \frac{D_{qq} D_{ww} - D_{qw}^2}{\eta^2 D_{qq} - 2\eta D_{qw} + D_{ww}} \eta (\eta_C - \eta). \quad (42)$$

This bound is tighter than the both bounds in equation (40) for any $\mathbf{D}$ and η , which are included as limiting cases for $D_{ww} \gg D_{qq}$ and $D_{ww} \ll D_{qq}$, respectively. Further, it reveals that the correlations $D_{qw} = \lim_{\mathcal{T} \rightarrow \infty} \langle \Delta q_h \Delta w \rangle / (2\mathcal{T})$ between heat and work play an important role in determining the maximal output power of the engine. At first sight, it seems that, if input heat and work are strongly correlated, $D_{qw} / \sqrt{D_{qq} D_{ww}} \approx 1$, the output power of the engine has to vanish. However, the condition $D_{qw} / \sqrt{D_{qq} D_{ww}} \approx 1$ also implies that heat and work are linearly dependent, i.e. $w \approx \eta q_h$ holds not only for the averages but also for the fluctuating heat and work. We expect this to hold in the long-time limit if the operation of the engine is ergodic in the sense that work and heat reproduce their ensemble-averaged values in the long-time limit. In this limit, all three bounds, equations (40) and (42), are identical. Thus, since the bound equation (42) is generally tighter than equation (40), the maximum bound on the output power results for a heat engine whose operation is ergodic. In general, we find that, if the work fluctuations are larger than the heat fluctuations, $D_{ww} > D_{qq}$, the bound on the output power is small and maximized at $\eta = \eta_C/2$ and for weak correlations between heat

and work $D_{qw}^2 \ll D_{qq}D_{ww}$ (see the top-left panel of figure 1). Increasing the heat fluctuations such that $D_{qq} > D_{ww}$, the attainable power increases and the maximum of the bound equation (42) shifts towards lower efficiencies and stronger heat-work correlations (top-right panel of figure 1). If the heat fluctuations are much larger than those of the work $D_{qq} \gg D_{ww}$, the bound develops a pronounced maximum at $\eta = \sqrt{D_{ww}/D_{qq}}$, which implies that the fluctuations of heat and work scale in the same manner as the averages, $\langle \dot{w} \rangle / \langle \dot{q} \rangle = \sqrt{D_{ww}/D_{qq}}$, and thus strong correlations between the fluctuating heat and work.

4.2. TUR and mobility

In the derivation and discussion of the TUR, we so far used only a single perturbation proportional to the probability current in equation (1). However, the general formulation of the Crámer–Rao bound (1) allows us to consider additional perturbations and as a consequence, to connect the TUR to other inequalities. Let us assume that we are dealing with a steady-state system, so that, similar to the derivation of the TUR, the choice $\mathbf{a}_1(\mathbf{x}) = \mathbf{j}^{\text{st}}(\mathbf{x})/P^{\text{st}}(\mathbf{x})$ leaves the probability density unchanged. We further explicitly consider a system of overdamped particles in contact with a heat bath at temperature T with a position-independent (positive definite and symmetric) mobility matrix $\mathbf{M}$, i.e. $\mathbf{B} = T\mathbf{M}$,

$$\dot{\mathbf{x}}(t) = \mathbf{M}\mathbf{F}(\mathbf{x}(t)) + \sqrt{2T\mathbf{M}}\boldsymbol{\xi}(t), \quad (43)$$

where $\mathbf{F}(\mathbf{x})$ is an arbitrary force that may include interactions between the particles and conservative as well as non-conservative external forces. We now introduce a second perturbation $\mathbf{a}_2 = \mathbf{M}\mathbf{f}$ with a constant force $\mathbf{f}$. As the observable, we take the time-integrated displacement vector $\mathbf{r}[\mathbf{x}(t)] = \int_0^T dt \dot{\mathbf{x}}(t)$ with average $\langle \mathbf{r} \rangle_\theta = T\mathbf{v}_\theta$, where $\mathbf{v}_\theta$ is the steady-state drift velocity. Then, the bound (1), evaluated at $\theta_1 = \theta_2 = 0$ (i.e. in the linear response limit) reads,

$$\begin{aligned} & \left(\begin{array}{c} \partial_{\theta_1} \mathbf{v}_\theta^T \\ \partial_{\theta_2} \mathbf{v}_\theta^T \end{array} \right) \mathbf{D}^{-1} (\partial_{\theta_1} \mathbf{v}_\theta, \partial_{\theta_2} \mathbf{v}_\theta) \Big|_{\theta=0} \\ & \leq \left(\begin{array}{cc} \sigma^{\text{st}} & \frac{1}{T} \int d\mathbf{x} \mathbf{f}^T \mathbf{M} \mathbf{M}^{-1} \mathbf{j}^{\text{st}}(\mathbf{x}) \\ \frac{1}{T} \int d\mathbf{x} \mathbf{f}^T \mathbf{M} \mathbf{M}^{-1} \mathbf{j}^{\text{st}}(\mathbf{x}) & \frac{1}{T} \int d\mathbf{x} \mathbf{f}^T \mathbf{M} \mathbf{M}^{-1} \mathbf{M} \mathbf{f} P^{\text{st}}(\mathbf{x}) \end{array} \right) \\ & = \left(\begin{array}{cc} \sigma^{\text{st}} & \frac{1}{T} \mathbf{f}^T \mathbf{v}_0 \\ \frac{1}{T} \mathbf{f}^T \mathbf{v}_0 & \frac{1}{T} \mathbf{f}^T \mathbf{M} \mathbf{f} \end{array} \right), \end{aligned} \quad (44)$$

where $(\mathbf{D})_{ij} = \lim_{T \rightarrow \infty} \langle \Delta r_i \Delta r_j \rangle / (2T)$ is the matrix of diffusivities. We also define the differential mobility matrix $\mathcal{M}$ via

$$\partial_{\theta_2} \mathbf{v}_\theta \Big|_{\theta=0} = \mathcal{M} \mathbf{f}. \quad (45)$$

The elements $(\mathcal{M})_{ij}$ of this matrix encode how much the drift velocity in direction x_i changes by applying a small constant force in direction j . Since the perturbation corresponding to θ_1 , proportional to the probability current, rescales all currents and thus also the drift velocity to $\mathbf{v}_\theta \simeq (1 + \theta_1) \mathbf{v}_0$, we have $\partial_{\theta_1} \mathbf{v}_\theta \Big|_{\theta=0} = \mathbf{v}_0$. The bound equation (44) is then equivalent to the set of inequalities

$$\mathbf{v}_0^T \mathbf{D}^{-1} \mathbf{v}_0 \leq \sigma^{\text{st}} \quad (46)$$

$$\mathcal{M}^T \mathbf{D}^{-1} \mathcal{M} \leq \frac{\mathbf{M}}{T} \quad (47)$$

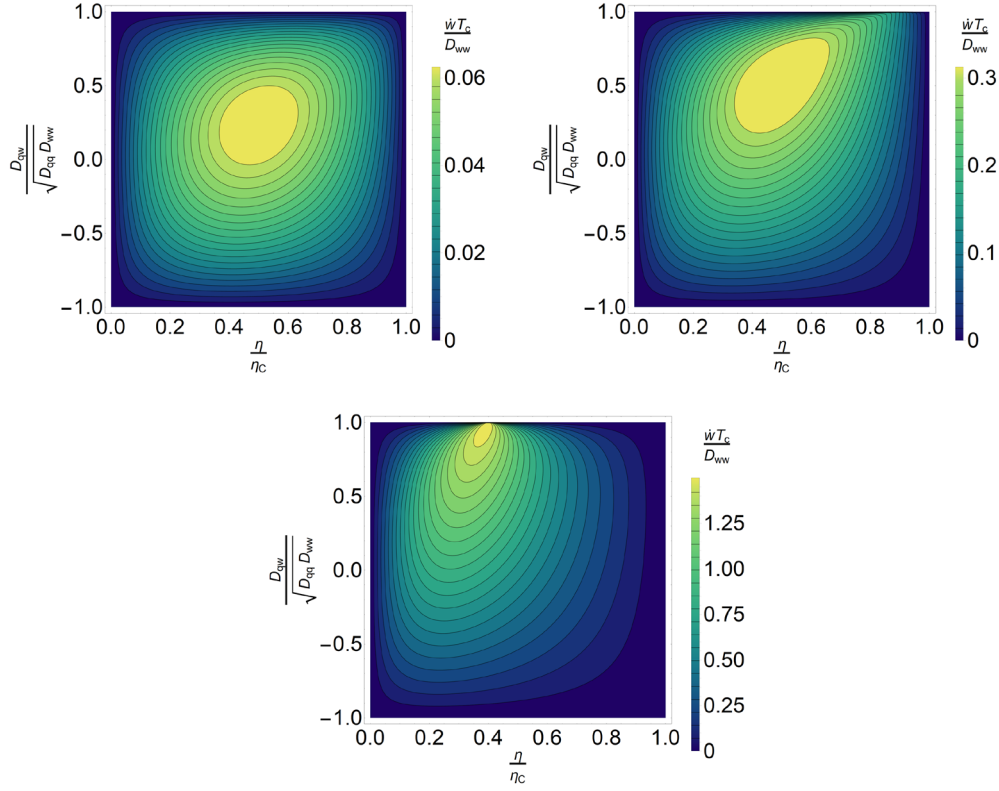


Figure 1. The bound equation (42) on the output power of a steady-state heat engine, as a function of the efficiency relative to the Carnot efficiency $\eta_C = 0.5$ and the relative size of the correlations between heat and work. In the top left panel, the fluctuations of the work are larger than those of the heat $D_{qq}/D_{ww} = 1/2$. In the top right and bottom panel, we increase the heat fluctuations to $D_{qq}/D_{ww} = 5$ and $D_{qq}/D_{ww} = 25$, respectively. The remaining parameters are $D_{ww} = 1$ and $T_c = 1$.

$$\left(\mathbf{f}^T \left(\frac{\mathbf{v}_0}{T} - \mathcal{M}^T \mathbf{D}^{-1} \mathbf{v}_0 \right) \right)^2 \leq \left(\sigma^{\text{st}} - \mathbf{v}_0^T \mathbf{D}^{-1} \mathbf{v}_0 \right) \mathbf{f}^T \left(\frac{\mathbf{M}}{T} - \mathcal{M}^T \mathbf{D}^{-1} \mathcal{M} \right) \mathbf{f}. \quad (48)$$

The first inequality (46) is exactly the long-time version of the steady-state TUR equation (34). The second inequality (47) is the bound on the mobility matrix derived in [15]. The differential mobility $\mathcal{M}$ describes the response of the particles to a small force $\mathbf{f}$ in the presence of the force $\mathbf{F}$, whereas the bare mobility $\mathbf{M}$ describes both response and diffusivity in the absence of $\mathbf{F}$, i.e. for free diffusion. Compared to free diffusion, the force $\mathbf{F}$ can decrease or increase both the mobility and the diffusivity compared to their bare values. The inequality (47) states that these changes are not independent of each other but have to follow certain rules: increased mobility always is accompanied by enhanced diffusivity; whereas decreased diffusivity can only be achieved at the cost of reduced mobility [15]. We remark that this relation was first observed for a particle in a one-dimensional periodic potential and conjectured to hold in more general cases in [19]. We stress that this is valid for arbitrarily strong forces $\mathbf{F}$ and arbitrarily far from equilibrium.

Finally, the third inequality (48) shows that the uncertainty relation and the bound on mobility are not independent of each other. In particular, we can use it to find a condition for a system to saturate the TUR, i.e. to have equality in equation (46). This necessarily requires that the left-hand side of equation (48) vanishes for arbitrary $\mathbf{f}$, and thus

$$T\mathcal{M}^T D^{-1} \mathbf{v}_0 = \mathbf{v}_0. \quad (49)$$

This is generically satisfied only if $T\mathcal{M}^T D^{-1}$ is the identity matrix, i.e. for

$$D = T\mathcal{M}, \quad (50)$$

which is precisely the equilibrium fluctuation–dissipation relation [20, 21]. We thus arrive at the following statement: The TUR for the drift velocity $\mathbf{v}$ can be an equality only in systems that satisfy the equilibrium fluctuation–dissipation relation. For a typical non-equilibrium situation, in which the equilibrium fluctuation–dissipation relation is violated, the TUR thus presents a strict inequality.

5. Discussion

The multidimensional generalization of the GTUR and TUR developed here allows to investigate the properties of stochastic currents in realistic, high-dimensional transport situations and the interrelations between different stochastic current observables in a natural manner. On the application side, we have used the MTUR to establish new tradeoff relations for the performance of steady-state heat engines, involving the heat-work correlations. Such tradeoff relations in various flavors have recently been an active topic in the discussion of stochastic heat engines [5, 16–18, 22, 23] and the connection between work fluctuations and output power is essential for understanding if, and under what conditions, finite power at Carnot efficiency is realizable [24–26]. The MTUR shows that not just the fluctuations of work and heat, but also how they are correlated has a strong impact on the output power of an engine.

The derivation of the MTUR from the Crámer–Rao bound reveals that the family of uncertainty relations is actually a consequence of information-theoretic bounds [12, 15]. No observable can contain more information than the underlying probability distribution and thus the effect of a parameter change on an observable is bounded by the Fisher information. Choosing a suitable perturbation that turns the Fisher information into a physical observable (in case of the TUR the entropy production), the information-theoretic inequality translates into a relation between different physical observables [15].

We anticipate that generalizations of other uncertainty relations [27–29] to vector-valued observables can be derived in a similar manner. An open issue, on which some progress has been made recently [30], is the extension of TURs to non-Markovian systems. We speculate that the understanding of the TUR in terms of information theoretic bounds may provide a useful guideline to derive TURs also in non-Markovian systems.

Acknowledgments

This work was supported by the World Premier International Research Center Initiative (WPI), MEXT, Japan and by JSPS Grant-in-Aid for Scientific Research on Innovative Areas ‘Discrete Geometric Analysis for Materials Design’: Grant Number 17H06460. The author wishes to thank S-I Sasa and S Ito for stimulating discussions.

Appendix. Markov jump processes

We consider a Markov jump process on a finite state space of N states. In this case, the time-evolution of the occupation probabilities $p_k(t)$ with $k = 1, \dots, N$ is governed by the Master equation

$$\partial_t p_k(t) = \sum_l \left(W_{kl}(t) p_l(t) - W_{lk}(t) p_k(t) \right), \quad (\text{A.1})$$

where $W_{kl}(t) \geq 0$ are the transition rates from state l to state k , which we assume to satisfy the local detailed balance condition $W_{kl}(t) = 0 \Leftrightarrow W_{lk}(t) = 0$. We take the rates to be governed by a set of parameters

$$W_{kl}(t) = W_{kl}^0 \exp \left[\sum_{\mu=1}^M \theta_{\mu} \Omega_{kl}^{\mu}(t) \right]. \quad (\text{A.2})$$

We choose this exponential form with $W_{kl}^0(t) \geq 0$ to ensure that positive rates remain positive for all values of the parameters θ . For short times τ , the transition probability from state l to state k is to leading order given by

$$p(k, t + \tau | l, t) = \delta_{kl} + \tau \left(W_{kl}(t) - \delta_{kl} \sum_m W_{lm}(t) \right) + O(\tau^2). \quad (\text{A.3})$$

We can use this to define a path probability on the discretized time interval $[0, \mathcal{T}] = \cup_{n=1}^N [n\tau, (n+1)\tau]$ with $N\tau = \mathcal{T}$

$$\mathbb{P} = \prod_{n=1}^N p(k_n, t_n | k_{n-1}, t_{n-1}) p_{k_0}(0) \quad (\text{A.4})$$

whose derivative follows by applying the product rule

$$\partial_{\theta_{\mu}} \mathbb{P} = \left(\sum_{n=1}^N \frac{\partial_{\theta_{\mu}} p(k_n, t_n | k_{n-1}, t_{n-1})}{p(k_n, t_n | k_{n-1}, t_{n-1})} + \frac{\partial_{\theta_{\mu}} p_{k_0}(0)}{p_{k_0}(0)} \right) \mathbb{P}. \quad (\text{A.5})$$

The Fisher information of the path probability is then given by

$$\begin{aligned} (\mathbb{I}(\theta))_{\mu\nu} &= \sum_{k_N} \sum_{k_{N-1}} \dots \sum_{k_0} \frac{\partial_{\theta_{\mu}} \mathbb{P} \partial_{\theta_{\nu}} \mathbb{P}}{\mathbb{P}} \\ &= \sum_{k_N} \sum_{k_{N-1}} \dots \sum_{k_0} \left(\sum_{n=1}^N \frac{\partial_{\theta_{\mu}} p(k_n, t_n | k_{n-1}, t_{n-1})}{p(k_n, t_n | k_{n-1}, t_{n-1})} + \frac{\partial_{\theta_{\mu}} p_{k_0}(0)}{p_{k_0}(0)} \right) \\ &\quad \times \left(\sum_{m=1}^N \frac{\partial_{\theta_{\nu}} p(k_m, t_m | k_{m-1}, t_{m-1})}{p(k_m, t_m | k_{m-1}, t_{m-1})} + \frac{\partial_{\theta_{\nu}} p_{k_0}(0)}{p_{k_0}(0)} \right) \mathbb{P}. \end{aligned} \quad (\text{A.6})$$

We now show that only the diagonal terms in the double sum over m and n contribute. Consider

$$\begin{aligned} &\sum_{k_N} \sum_{k_{N-1}} \dots \sum_{k_0} \frac{\partial_{\theta_{\mu}} p(k_n, t_n | k_{n-1}, t_{n-1})}{p(k_n, t_n | k_{n-1}, t_{n-1})} \frac{\partial_{\theta_{\nu}} p(k_m, t_m | k_{m-1}, t_{m-1})}{p(k_m, t_m | k_{m-1}, t_{m-1})} \mathbb{P} \\ &= \sum_{k_n} \sum_{k_{n-1}} \dots \sum_{k_0} \partial_{\theta_{\mu}} p(k_n, t_n | k_{n-1}, t_{n-1}) \frac{\partial_{\theta_{\nu}} p(k_m, t_m | k_{m-1}, t_{m-1})}{p(k_m, t_m | k_{m-1}, t_{m-1})} \mathbb{P}^{n-1}, \end{aligned} \quad (\text{A.7})$$

where we assumed, without loss of generality, $n \geq m$ and $\mathbb{P}^{n-1}$ is the path probability up to step $n-1$. This expression is zero for $n > m$ since we have

$$\sum_{k_n} \partial_{\theta_\mu} p(k_n, t_n | k_{n-1}, t_{n-1}) = \partial_{\theta_\mu} \sum_{k_n} p(k_n, t_n | k_{n-1}, t_{n-1}) = \partial_{\theta_\mu} 1 = 0 \quad (\text{A.8})$$

and the only contribution comes from $n = m$, where the expression simplifies to

$$\begin{aligned} & \sum_{k_N} \sum_{k_{N-1}} \cdots \sum_{k_0} \frac{\partial_{\theta_\mu} p(k_n, t_n | k_{n-1}, t_{n-1}) \partial_{\theta_\nu} p(k_n, t_n | k_{n-1}, t_{n-1})}{p(k_n, t_n | k_{n-1}, t_{n-1})^2} \mathbb{P} \\ &= \sum_{k_n} \sum_{k_{n-1}} \frac{\partial_{\theta_\mu} p(k_n, t_n | k_{n-1}, t_{n-1}) \partial_{\theta_\nu} p(k_n, t_n | k_{n-1}, t_{n-1})}{p(k_n, t_n | k_{n-1}, t_{n-1})} p_{k_{n-1}}(t_{n-1}). \end{aligned} \quad (\text{A.9})$$

The Fisher information is thus additive in the individual steps and can be written as

$$\begin{aligned} (\mathbb{I}(\theta))_{\mu\nu} &= \sum_{n=1}^N \left(\sum_{k_n} \sum_{k_{n-1}} \frac{\partial_{\theta_\mu} p(k_n, t_n | k_{n-1}, t_{n-1}) \partial_{\theta_\nu} p(k_n, t_n | k_{n-1}, t_{n-1})}{p(k_n, t_n | k_{n-1}, t_{n-1})} \right. \\ &\quad \left. \times p_{k_{n-1}}(t_{n-1}) \right) + \sum_{k_0} \frac{(\partial_{\theta_\mu} p_{k_0}(0) \partial_{\theta_\nu} p_{k_0}(0))}{p_{k_0}(0)}. \end{aligned} \quad (\text{A.10})$$

The derivative of the transition probability with respect to a parameter θ_μ is given by

$$\begin{aligned} \partial_{\theta_\mu} p(k, t + \tau | l, t) &= \tau \left(\Omega_{kl}^\mu(t) W_{kl}(t) - \delta_{kl} \sum_m \Omega_{lm}^\mu(t) W_{lm}(t) \right) \\ &= \tau \frac{Z_{kl}^\mu(t) W_{kl}(t) - \delta_{kl} \sum_m \Omega_{lm}^\mu(t) W_{lm}(t)}{\delta_{kl} + \tau (W_{kl}(t) - \delta_{kl} \sum_m W_{lm}(t))} p(k, t + \tau | l, t). \end{aligned} \quad (\text{A.11})$$

We can use this to compute the Fisher information for a single step $n-1 \rightarrow n$,

$$\begin{aligned} (\mathbb{I}(\theta))_{\mu\nu}^{n-1 \rightarrow n} &= \tau^2 \sum_{k_n} \sum_{k_{n-1}} \left(\Omega_{k_n k_{n-1}}^\mu W_{k_n k_{n-1}} - \delta_{k_n k_{n-1}} \sum_m \Omega_{k_{n-1} m}^\mu W_{k_{n-1} m} \right) \\ &\quad \times \left(\Omega_{k_n k_{n-1}}^\nu W_{k_n k_{n-1}} - \delta_{k_n k_{n-1}} \sum_m \Omega_{k_{n-1} m}^\nu W_{k_{n-1} m} \right) \\ &\quad \times \frac{p_{k_{n-1}}}{\delta_{k_n k_{n-1}} + \tau (W_{k_n k_{n-1}} - \delta_{k_n k_{n-1}} \sum_m W_{k_{n-1} m})}, \end{aligned} \quad (\text{A.12})$$

where all time-dependent quantities are evaluated at t_{n-1} . Now, we distinguish the case $k_{n-1} = k_n$ and $k_{n-1} \neq k_n$, writing

$$\begin{aligned} (\mathbb{I}(\theta))_{\mu\nu}^{n-1 \rightarrow n} &= \tau \sum_{k_n \neq k_{n-1}} \sum_{k_{n-1}} \Omega_{k_n k_{n-1}}^\mu \Omega_{k_n k_{n-1}}^\nu W_{k_n k_{n-1}} p_{k_{n-1}} \\ &\quad + \tau^2 \sum_{k_n} \left(\Omega_{k_n k_n}^\mu W_{k_n k_n} - \sum_m \Omega_{k_n m}^\mu W_{k_n m} \right) \\ &\quad \times \left(\Omega_{k_n k_n}^\nu W_{k_n k_n} - \sum_m \Omega_{k_n m}^\nu W_{k_n m} \right) \\ &\quad \times \frac{p_{k_{n-1}}}{1 + \tau (W_{k_n k_{n-1}} - \sum_m W_{k_{n-1} m})} \\ &= \tau \sum_{k_n \neq k_{n-1}} \sum_{k_{n-1}} \Omega_{k_n k_{n-1}}^\mu \Omega_{k_n k_{n-1}}^\nu W_{k_n k_{n-1}} p_{k_{n-1}} + O(\tau^2). \end{aligned} \quad (\text{A.14})$$

We can then write the Fisher information along the path as

$$\begin{aligned} (\mathbb{I}(\theta))_{\mu\nu} = & \tau \sum_{n=1}^N \left(\sum_{k_n \neq k_{n-1}} \sum_{k_{n-1}} \Omega_{k_n, k_{n-1}}^\mu \Omega_{k_n, k_{n-1}}^\nu W_{k_n, k_{n-1}} p_{k_{n-1}} \right) \\ & + \sum_{k_0} \frac{(\partial_{\theta_\mu} p_{k_0}(0) \partial_{\theta_\nu} p_{k_0}(0))}{p_{k_0}(0)}. \end{aligned} \quad (\text{A.15})$$

Taking the continuous-time limit $\tau \rightarrow 0$ with $N\tau = \mathcal{T}$ fixed, the sum can be written as an integral, and we finally obtain,

$$(\mathbb{I}(\theta))_{\mu\nu} = \int_0^{\mathcal{T}} dt \sum_{k \neq l} \Omega_{kl}^\mu(t) \Omega_{kl}^\nu(t) W_{kl}(t) p_l(t) + \sum_k \frac{(\partial_{\theta_\mu} p_k(0) \partial_{\theta_\nu} p_k(0))}{p_k(0)}. \quad (\text{A.16})$$

This is the path Fisher information corresponding to the expression equation (15) for Langevin dynamics.

Next, we define a stochastic current as

$$\mathbf{r}[k(t)] = \int_0^{\mathcal{T}} \sum_{k,l} \left(dn_{kl}(t) - dn_{lk}(t) \right) \mathbf{z}_{kl}(t), \quad (\text{A.17})$$

where $dn_{kl}(t)$ denotes the number of jumps from state l to state k during the time interval $[t, t + \tau]$ and $\mathbf{z}_{kl}(t)$ is an arbitrary K -vector, where K is the number of observables. The average of such a current is given by

$$\langle \mathbf{r} \rangle_\theta = \int_0^{\mathcal{T}} dt \sum_{k,l} \left(W_{kl}(t) p_l(t) - W_{lk}(t) p_k(t) \right) \mathbf{z}_{kl}(t). \quad (\text{A.18})$$

In order to derive the GTUR, we again need a proper choice of the perturbation Ω_{kl}^1 . Following [7], we take

$$\Omega_{kl}^1 = \frac{W_{kl}(t) \pi_l - W_{lk}(t) \pi_k}{W_{kl}(t) p_l^0(t) + W_{lk}(t) p_k^0(t)}, \quad (\text{A.19})$$

where $\pi_k > 0$ is an arbitrary set of positive parameters with $\sum_k \pi_k = 1$ and we denote by $p_k^0(t)$ the solution of the master equation (A.1) for $\theta = 0$. To linear order in θ , the master equation for the modified occupation probabilities is

$$\begin{aligned} \partial_t p_k(t) = & \sum_l \left(W_{kl}(t) p_l(t) - W_{lk}(t) p_k(t) \right) \\ & + \theta \left(\frac{W_{kl}(t) \pi_l - W_{lk}(t) \pi_k}{W_{kl}(t) p_l^0(t) + W_{lk}(t) p_k^0(t)} W_{kl}(t) p_l(t) \right. \\ & \left. + \frac{W_{kl}(t) \pi_l - W_{lk}(t) \pi_k}{W_{kl}(t) p_l^0(t) + W_{lk}(t) p_k^0(t)} W_{lk}(t) p_k(t) \right) + O(\theta^2). \end{aligned} \quad (\text{A.20})$$

It is easily verified by direct computation that this is solved by

$$p_k(t) = p_k^0(t) + \theta (p_k^0(t) - \pi_k) \quad (\text{A.21})$$

and that the average of the current changes to

$$\langle \mathbf{r} \rangle_\theta = (1 + \theta) \langle \mathbf{r} \rangle_0. \quad (\text{A.22})$$

On the other hand, the Fisher information corresponding to this transformation is

$$\mathbb{I}(0) = \frac{1}{2} \int_0^\tau dt \sum_{k \neq l} \frac{(W_{kl}(t)\pi_l - W_{lk}(t)\pi_k)^2}{W_{kl}(t)p_l^0(t) + W_{lk}(t)p_k^0(t)}. \quad (\text{A.23})$$

Using the inequality

$$\frac{(a-b)^2}{a+b} \leq \frac{1}{2}(a-b) \ln \left(\frac{a}{b} \right), \quad (\text{A.24})$$

which holds for arbitrary positive numbers a and b , we can bound this from above by

$$\mathbb{I}(0) \leq \frac{1}{2} \int_0^\tau dt \sum_{k \neq l} \left(\frac{W_{kl}(t)\pi_l - W_{lk}(t)\pi_k}{W_{kl}(t)p_l^0(t) - W_{lk}(t)p_k^0(t)} \right)^2 \sigma_{kl}(t), \quad (\text{A.25})$$

with the contribution to the entropy production from the transition from l to k ,

$$\sigma_{kl}(t) = \frac{1}{2} (W_{kl}(t)p_l(t) - W_{lk}(t)p_k(t)) \ln \left(\frac{W_{kl}(t)p_l(t)}{W_{lk}(t)p_k(t)} \right). \quad (\text{A.26})$$

The upper bound on $I(0)$ is precisely what is termed effective entropy production in [7]. Using the Crámer–Rao bound (1), we thus obtain the multidimensional GTUR for jump processes,

$$\begin{aligned} \langle \mathbf{r} \rangle^T \Xi_r^{-1} \langle \mathbf{r} \rangle &\leq \frac{1}{2} \Sigma \\ \text{with } \Sigma &= \int_0^\tau dt \sum_{k \neq l} \left(\frac{W_{kl}(t)\pi_l - W_{lk}(t)\pi_k}{W_{kl}(t)p_l^0(t) - W_{lk}(t)p_k^0(t)} \right)^2 \sigma_{kl}(t). \end{aligned} \quad (\text{A.27})$$

For a steady state process, we can choose $\pi_k = p_k^{\text{st}}$ and Σ reduces to the entropy production $\Delta S = \int_0^\tau dt \sum_{kl} \sigma_{kl}(t)$, recovering the MTUR for jump processes,

$$\langle \mathbf{r} \rangle^T \Xi_r^{-1} \langle \mathbf{r} \rangle \leq \frac{1}{2} \Delta S. \quad (\text{A.28})$$

ORCID iDs

Andreas Dechant  <https://orcid.org/0000-0002-5782-5893>

References

- [1] Barato A C and Seifert U 2015 *Phys. Rev. Lett.* **114** 158101
- [2] Gingrich T R, Horowitz J M, Perunov N and England J L 2016 *Phys. Rev. Lett.* **116** 120601
- [3] Pietzonka P, Ritort F and Seifert U 2017 *Phys. Rev. E* **96** 012101
- [4] Horowitz J M and Gingrich T R 2017 *Phys. Rev. E* **96** 020103
- [5] Dechant A and Sasa S I 2018 *J. Stat. Mech.* **063209**
- [6] Barato A C, Chetrite R, Faggionato A and Gabrielli D 2018 *New J. Phys.* **20** 103023
- [7] Koyuk T, Seifert U and Pietzonka P 2019 *J. Phys. A: Math. Theor.* **52** 02LT02

- [8] Barato A C, Chetrite R, Faggionato A and Gabrielli D 2018 (arXiv:[1810.11894](#))
- [9] Manikandan S K and Krishnamurthy S 2018 *J. Phys. A: Math. Theor.* **51** 11LT01
- [10] Radhakrishna Rao C 1945 *Bull. Calcutta Math. Soc.* **37** 81–91
- [11] Cramér H 2016 *Mathematical Methods of Statistics* vol 9 (Princeton, NJ: Princeton University Press)
- [12] Hasegawa Y and Van Vu T 2018 (arXiv:[1809.03292](#))
- [13] Kay S M 1993 *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory* (Englewood Cliffs, NJ: Prentice-Hall)
- [14] Risken H 1986 *The Fokker–Planck Equation* (Berlin: Springer)
- [15] Dechant A and Sasa S I 2018 (arXiv:[1804.08250](#))
- [16] Pietzonka P, Barato A C and Seifert U 2016 *J. Stat. Mech.* **124004**
- [17] Pietzonka P and Seifert U 2018 *Phys. Rev. Lett.* **120** 190602
- [18] Vroylandt H, Lacoste D and Verley G 2018 *J. Stat. Mech.* **023205**
- [19] Hayashi K and Sasa S I 2005 *Phys. Rev. E* **71** 020102
- [20] Kubo R 1957 *J. Phys. Soc. Japan* **12** 570–86
- [21] Hänggi P and Thomas H 1982 *Phys. Rep.* **88** 207–319
- [22] Shiraishi N, Saito K and Tasaki H 2016 *Phys. Rev. Lett.* **117** 190601
- [23] Dechant A and Sasa S I 2018 *Phys. Rev. E* **97** 062101
- [24] Campisi M and Fazio R 2016 *Nat. Commun.* **7** 11895
- [25] Holubec V and Ryabov A 2017 *Phys. Rev. E* **96** 062107
- [26] Holubec V and Ryabov A 2018 *Phys. Rev. Lett.* **121** 120601
- [27] Proesmans K and den Broeck C V 2017 *Europhys. Lett.* **119** 20001
- [28] Chiuchiu D and Pigolotti S 2018 *Phys. Rev. E* **97** 032109
- [29] Brandner K, Hanazato T and Saito K 2018 *Phys. Rev. Lett.* **120** 090601
- [30] Van Vu T and Hasegawa Y 2018 (arXiv:[1809.06610](#))