

PAPER • OPEN ACCESS

Generalising thermodynamic efficiency of interactions: inferential, information-geometric and computational perspectives

To cite this article: Qianyang Chen *et al* 2026 *J. Phys. Complex.* **7** 025002

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

You may also like

- [Information geometry and non-equilibrium thermodynamic relations in the over-damped stochastic processes](#)
Eun-jin Kim
- [Efficiency Bound of Learning with Coarse Graining](#)
Minghao Li, , Shihao Xia et al.
- [Thermodynamic efficiency of information and heat flow](#)
Armen E Allahverdyan, Dominik Janzing and Guenter Mahler



PAPER

OPEN ACCESS

RECEIVED
11 December 2025REVISED
12 March 2026ACCEPTED FOR PUBLICATION
31 March 2026PUBLISHED
10 April 2026

Original content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



Generalising thermodynamic efficiency of interactions: inferential, information-geometric and computational perspectives

Qianyang Chen^{1,*} , Nihat Ay^{2,3}  and Mikhail Prokopenko¹ ¹ Centre for Complex Systems, Faculty of Engineering, The University of Sydney, Sydney, NSW 2006, Australia² Institute for Data Science Foundations, Hamburg University of Technology, 21073 Hamburg, Germany³ Santa Fe Institute, Santa Fe, NM 87501, United States of America

* Author to whom any correspondence should be addressed.

E-mail: qianyang.chen@sydney.edu.au**Keywords:** self-organisation, collective behaviour, information geometry, statistical physics, statistical inference, thermodynamic efficiency**Abstract**

Self-organizing systems consume energy to generate internal order. The concept of thermodynamic efficiency, drawing from statistical physics and information theory, has previously been proposed to characterise a change in control parameter by relating the resulting predictability gain to the required amount of work. However, previous studies have taken a *system-centric* perspective and considered only single control parameters. Here, considering systems at or near equilibrium, we generalise thermodynamic efficiency to multiple control parameters and extend the definition of thermodynamic efficiency to protocols in arbitrary directions, by introducing *directional efficiency*. Taking an *observer-centric* perspective, we derive two novel formulations. The first, an *inferential form*, relates efficiency to fluctuations of macroscopic observables, interpreting thermodynamic efficiency in terms of how well the system parameters can be inferred from observable macroscopic behaviour. The second, an *information-geometric form*, expresses efficiency in terms of the Fisher information matrix, interpreting it with respect to how difficult it is to navigate the statistical manifold defined by the control protocol. This observer-centric perspective is contrasted with the existing system-centric view, where efficiency is considered an intrinsic property of the system.

1. Introduction

Collective systems, such as groups of interacting particles, flocks of birds, schools of fish, or active matter, create and maintain internal order by consuming energy and performing work [1–4]. Their large-scale complex patterns or coordinated movements arise from simple local interactions, a phenomenon known as self-organisation. For example, flocking behaviour emerges as each bird tries to follow its neighbours while avoiding collision. The efficiency of local interactions can be measured by *thermodynamic efficiency*: the ratio of entropy reduction or predictability gain to the generalised work required to vary a system's control parameter [5–9].

Thermodynamic efficiency has been studied for various complex dynamical systems, including canonical magnetisation models, such as the Curie–Weiss model [8] and the Ising model [9], self-propelled particles [6], urban dynamics [5], and contagion networks [7]. Each study examined the thermodynamic efficiency along a single protocol, focusing on a control parameter driving the system across distinct macroscopic states (i.e. phase transition). For example, in canonical magnetisation models, phase transitions were induced by changing parameters such as temperature, external magnetic field strength, and coupling strength [8, 9]. In modelling the motion of self-propelled particles, increasing the particle alignment strength was observed to drive collective motion from disorder to ordered states [6]. In urban systems, modifying the social disposition parameter, a factor that prioritises the attractiveness of suburbs relative to travel costs, shifts the city layouts between monocentric (single mega-suburb) and polycentric (multiple affluent suburbs) configurations [5]. In modelling an epidemic spread, the infection transmission rate controls the transition between non-epidemic and epidemic phases [7]. These studies

observed that thermodynamic efficiency peaks or diverges at a phase transition, and interpreted this in terms of the system's ability to increase its intrinsic cohesiveness relative to the work carried out near critical regimes. In other words, these approaches adopted a *system-centric view* on the efficiency of collective interactions during self-organisation.

Despite different contexts, these studies demonstrated that self-organizing systems are most energetically efficient in increasing internal order (i.e. predictability) when poised at the critical regimes. Building on these results, recent work [9] proposed the thermodynamic efficiency as an intrinsic utility for self-organisation and introduced the *principle of super-efficiency*. The principle postulates that collective systems self-organise to criticality because it is the regime where (a) given the amount of work available to change the control parameter, the gain in predictability is maximised, or (b) given the desired predictability gain, the required work is minimised. The study also proposed extending thermodynamic efficiency to systems with multiple control parameters; however, formulating this extension remained an open challenge.

Moreover, the intrinsic utility interpretation treated thermodynamic efficiency as the system's property—independent of how it is observed and measured. However, the emergence of order in collective systems is inherently linked to the specifically chosen observables. Our choice of order parameters, control parameters, or the coordinate representations used to encode the parameter space can all affect the computed value of thermodynamic efficiency. Yet, no formal framework exists to capture this *observer-centric* perspective, even for the single parameter case.

This paper aims to generalise thermodynamic efficiency to settings with multiple control parameters for systems at or near equilibrium. We derive two new equivalent formulations and offer different interpretations: one expresses efficiency in terms of the fluctuations of macroscopic observables, interpreted from an inferential perspective; the other in terms of Fisher information, interpreted from the information-geometric perspective. These two interpretations are observer-centric, and we contrast them with the previously developed computational perspective [6, 9], which is system-centric.

This paper is structured as follows. Section 2 reviews the definition of thermodynamic efficiency for a single control parameter and extends the definition for multi-parameter systems explored using arbitrary control protocols, by introducing *directional efficiency*. Sections 3 and 4 extend the concept to multi-parameter settings, introducing two new formulations: *inferential* and *information-geometric*. In addition, these sections offer observer-centric interpretations of these two forms. Section 5 illustrates the idea through simulations of the 2D Ising model. Section 6 concludes with a discussion of the key findings and their implications.

2. Defining thermodynamic efficiency

Thermodynamic efficiency [5–8] quantifies how efficiently a collective system self-organises. It is a function of the control parameter λ , which modulates the interactions among the individuals within the collective, and hence changes the system's overall predictability. For example, in the Ising model—a simple system of interacting binary spins on a lattice—we may consider the coupling strength between neighbouring spins as a control parameter, and increasing the coupling strength increases the spins' tendency to align, hence bringing order to the system. Thermodynamic efficiency $\eta(\lambda)$ measures how much work done by tuning λ translates to increased order. The thermodynamic efficiency of interaction connects deeply to information theory under Jaynes' generalised statistical mechanics framework [10].

For a collective system governed by a set of control parameters (or generalised forces) $\{\lambda_i\}$, each corresponding to a conjugate observable (or collective variable) X_i [11], if the system is subject to constraints given by the expected values of the observables $\langle X_i \rangle$, the probability distribution the system's microstate x is given by the maximum entropy distribution:

$$p(x; \lambda) = \frac{1}{Z} e^{-\sum_i \lambda_i X_i(x)} \quad (1)$$

where the sum $\sum_i \lambda_i X_i(x)$ can be viewed as generalised total energy.

For example, in the canonical square-lattice Ising model, the generalised forces $\{\lambda_i\} = \{-\beta h, -\beta J\}$, with $\beta = \frac{1}{k_B T}$ being the inverse temperature, h the external magnetic field, and J the coupling strength. The corresponding conjugate observables are $\{X_i(x)\} = \{\sum_n \sigma_n, \sum_{\langle m, n \rangle} \sigma_m \sigma_n\}$, where σ_n is the spin of the n th site, and the observables represent the total magnetisation and the total pairwise interactions, respectively.

The partition function Z is:

$$Z(\underline{\lambda}) = \sum_x e^{-\sum_i \lambda_i X_i(x)} \quad (2)$$

Under information theory, Shannon entropy S quantifies the uncertainty of a probability distribution $p(x)$ and is defined as [12]:

$$S = - \sum_x p(x) \log p(x) \quad (3)$$

In this generalised framework, log of the partition function, or the Massieu potential Ψ (analogous to the Gibbs potential in thermodynamics), is the difference between the Shannon entropy and the generalised total energy [11]:

$$\Psi = \log Z = - \sum_i \lambda_i \langle X_i \rangle + S \quad (4)$$

where $\langle X_i \rangle$ is the average of the observable X_i .

Thermodynamic efficiency η for a given control parameter λ_i is defined as the ratio of reduction in Shannon entropy S (or gain in predictability) to the generalised work $\langle \beta \mathbb{W} \rangle$ expended for tuning the parameter (\mathbb{W} refers to thermodynamic work). Mathematically, it can be expressed as [5–8]:

$$\eta(\lambda_j) = - \frac{\partial S(\lambda_j)}{\partial \lambda_j} / \frac{\partial \langle \beta \mathbb{W} \rangle(\lambda_j)}{\partial \lambda_j} \quad (5)$$

where partial derivatives are used to emphasise that the definition holds for settings with multiple control parameters.

In a quasi-static process, the infinitesimal work done by the system to change the control parameter is equal to the reduction in Gibbs free energy $d\langle \mathbb{W} \rangle = -d\mathbb{G}$ [13, section 15]. Therefore, the thermodynamic efficiency can be expressed as:

$$\eta(\lambda_j) = \frac{\partial S(\lambda_j)}{\partial \lambda_j} / \frac{\partial \beta \mathbb{G}(\lambda_j)}{\partial \lambda_j} \quad (6)$$

where \mathbb{G} is the Gibbs potential, so that [13, section 31]:

$$\mathbb{G} = - \frac{1}{\beta} \log Z \quad (7)$$

Similar to (5), we define generalised directional efficiency evaluated at $\underline{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_n)$ for the protocol moving in the direction $\underline{\nu} = (\nu_1, \dots, \nu_n)$ as:

$$\eta_{\underline{\nu}}(\underline{\lambda}) = \frac{-D_{\underline{\nu}} S(\underline{\lambda})}{D_{\underline{\nu}} \langle \beta \mathbb{W} \rangle(\underline{\lambda})} \quad (8)$$

where the directional derivative of a scalar function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ in the direction of $\underline{\nu}$ at the point $\underline{\lambda}$ is defined as [14, pp. 147–148]:

$$D_{\underline{\nu}} f(\underline{\lambda}) := \lim_{t \rightarrow 0} \frac{f(\underline{\lambda} + t\underline{\nu}) - f(\underline{\lambda})}{t} = \nabla f \cdot \underline{\nu}. \quad (9)$$

For a single-parameter case, the efficiency $\eta(\lambda_j)$ can be interpreted as the ratio of the amount of information gained about the system to the amount of generalised work done to change the control parameter by a small amount $\delta\lambda_j$. We provide interpretations for the multi-parameter case in more detail in sections 3.2 and 4.2.

3. Relating efficiency to fluctuation

The previous studies focus on a single control parameter and its corresponding observable. In the following section, we turn our focus on the inter-dependencies of multiple observables $\{X_i\}$ in response to changing a specific control parameter λ_j . We aim to reformulate the thermodynamic efficiency $\eta(\lambda_j)$ in terms of the covariance of observables $\{X_i\}$, which is a measure of how the observables fluctuate together, and provide a statistical perspective on the efficiency of self-organisation.

3.1. Reformulating efficiency in terms of fluctuations

Using well-known results from generalised statistical physics (see appendix A), the rate of entropy change with respect to the control parameter λ_j can be expressed as:

$$\frac{\partial S(\underline{\lambda})}{\partial \lambda_j} = - \sum_i \lambda_i \text{Cov}(X_i, X_j). \quad (10)$$

This key expression yields the numerator of the thermodynamic efficiency (6) in terms of the covariance across multiple observables. Turning our attention to the denominator, we use the following representation, which is typically derived by combining (7) and (A1) (see appendix A):

$$\frac{\partial \beta \mathbb{G}(\underline{\lambda})}{\partial \lambda_j} = - \frac{\partial \log Z(\underline{\lambda})}{\partial \lambda_j} = \langle X_j \rangle. \quad (11)$$

Substituting (10) and (11) into (6), we can reformulate the thermodynamic efficiency $\eta(\lambda_j)$ as:

$$\eta^{\text{inf}}(\lambda_j) = - \frac{\sum_i \lambda_i \text{Cov}(X_i, X_j)}{\langle X_j \rangle} \quad (12)$$

We refer to this as the inferential form of thermodynamic efficiency, denoted by η^{inf} . For a single control parameter λ and conjugate observable X , the efficiency can be simplified to:

$$\eta^{\text{inf}}(\lambda) = - \frac{\lambda \text{Var}(X)}{\langle X \rangle}. \quad (13)$$

Equations (12) and (13) offer a key insight: thermodynamic efficiency can be expressed in terms of fluctuations in the observables driven by changing a specific control parameter λ . Specifically, the efficiency in (12) is the total fluctuation, weighted by all the control parameters, and normalised by the mean value of the observable conjugate to the parameter of interest. This formulation links the efficiency of interactions to the statistical inference of the system's probability distribution, and the next section explores this, *observer-centric*, interpretation further.

3.2. The statistical perspective of efficiency

In 1957, Jaynes in his seminal work [10] posed a key question: What probability distribution best describes our state of knowledge about a physical system, given some average measurement outcomes of macroscopic observables? He concluded that by assigning probabilities which maximise the information entropy subject to relevant constraints, one can arrive at the same probability distribution given by canonical methods of statistical mechanics. The resulting probability distribution is given by $p(x; \underline{\lambda})$, where λ_i are the Lagrange multipliers associated with the constraints on the average values of observables $\{X_i\}$.

The same problem can be rephrased differently: Given average measurement values of the observables $\langle X_1(x) \rangle, \langle X_2(x) \rangle, \dots$, how can we infer the parameters of the probability distribution $p(x; \underline{\lambda})$, such that the distribution agrees with the average measurements and maximises the information entropy? This can be thought of as estimating $\underline{\lambda}$ based on sampling the macroscopic observables $X_1(x), X_2(x), \dots$. The error of estimating $\underline{\lambda}$ from the given data is measured by the covariance matrix of the estimator:

$$A_{ij} = \text{Cov}(\lambda_i, \lambda_j).$$

The covariance matrix of the estimators A has an inverse relationship with the covariance matrix of the observables $B_{ij} = \text{Cov}(X_i, X_j)$, which measures how much the macroscopic observables X_i and X_j change together [15, p. 190]:

$$A = B^{-1}.$$

This means that when the covariance (and variance) of the macroscopic observables $\text{Cov}(X_i, X_j)$ is large, the covariance (and variance) of the parameter estimators $\text{Cov}(\lambda_i, \lambda_j)$ is small, hence the observer can estimate $\underline{\lambda}$ with smaller errors.

Therefore, equations (12) and (13) show that the efficiency of self-organisation goes hand-in-hand with the *inferential efficiency*. A statistical estimator is considered efficient when it is unbiased and has minimum variance. When the system self-organises into a state where the observables tend to fluctuate together, it has created internal structure where an observer can recover its hidden parameters more accurately. In other words, when the self-organizing system is efficient, observers can make a better guess of the system's hidden parameters $\{\lambda\}$ by observing the macroscopic observables $\{X\}$. The denominator $\langle X_j \rangle$ normalises the efficiency so that η becomes comparable across the set of control parameters $\{\lambda\}$.

4. Relating efficiency to fisher information

4.1. Reformulating efficiency in terms of fisher information

Thermodynamic efficiency can also be reformulated, using information theory, in terms of the Fisher information [16, 17]. As shown in appendix B, $\mathcal{I}_{ij}(\lambda) = \text{Cov}(X_i, X_j)$. Substituting this to (10) gives the numerator of the efficiency $\eta(\lambda_j)$ in terms of Fisher information:

$$\frac{\partial S(\lambda)}{\partial \lambda_j} = - \sum_i \lambda_i \mathcal{I}_{ij}(\lambda). \quad (14)$$

Using (B3) (see appendix B), the denominator of the efficiency $\eta(\lambda_j)$ can be expressed via the integral of the Fisher information \mathcal{I}_{jj} over the control parameter λ_j (derivation for single parameter case is given in [6, 18]):

$$\begin{aligned} \frac{\partial \beta \mathbb{G}(\lambda)}{\partial \lambda_j} &= - \frac{\partial \log Z(\lambda)}{\partial \lambda_j} \\ &= - \int_{\lambda_j^*}^{\lambda_j} \mathcal{I}_{jj}(\lambda_1, \dots, \lambda_j', \dots, \lambda_n) d\lambda_j' \end{aligned} \quad (15)$$

where λ_j^* is the point where a change in the control parameter λ_j does not extract or perform any work, or the zero-response point. That is:

$$\left. \frac{\partial \langle \beta \mathbb{W} \rangle(\lambda)}{\partial \lambda_j} \right|_{\lambda_j = \lambda_j^*} = 0. \quad (16)$$

Under a quasi-static process, this leads to:

$$- \left. \frac{\partial \beta \mathbb{G}(\lambda)}{\partial \lambda_j} \right|_{\lambda_j = \lambda_j^*} = - \langle X_j \rangle = 0. \quad (17)$$

For example, in a system of interacting spins, such as a 2D Ising model, the zero-response point with respect to the coupling strength is the point where the nearest-neighbour correlations vanish, which is the point where $\sum \sigma_i \sigma_j = 0$.

Combining (14) and (15) yields:

$$\eta^\Delta(\lambda_j) = \frac{\sum_i \lambda_i \mathcal{I}_{ij}(\lambda)}{\int_{\lambda_j^*}^{\lambda_j} \mathcal{I}_{jj}(\lambda_1, \dots, \lambda_j', \dots, \lambda_n) d\lambda_j'}. \quad (18)$$

We refer to this as the information-geometric form of thermodynamic efficiency, denoted by η^Δ . For a single control parameter λ , the efficiency can be simplified to:

$$\eta^\Delta(\lambda) = \frac{\lambda \mathcal{I}(\lambda)}{\int_{\lambda^*}^{\lambda} \mathcal{I}(\lambda') d\lambda'}. \quad (19)$$

The Cramer–Rao inequality states that the covariance matrix of a set of unbiased estimators for the parameters $\{\lambda\}$ is bounded by the inverse of the Fisher information matrix [17]:

$$\text{Cov}(\hat{\lambda}) \geq \mathcal{I}^{-1}(\lambda). \quad (20)$$

This inequality implies that larger Fisher information leads to a smaller lower bound of the estimation variance, that is, the parameters can be estimated more precisely. In this context, higher thermodynamic efficiency implies the ability to more precisely estimate control parameters λ from the macroscopic observables X , relative to the total variance accumulated along the control protocol from λ_j to λ_j^* . In other words, continuing along the protocol towards the critical point, a potentially higher precision quantified by the numerator, is contrasted with a potentially higher total precision accumulated along the protocol, quantified by the denominator. This aligns with our earlier point in section 3.2 that thermodynamic efficiency is linked to inferential efficiency. However, fundamentally, the equations (18) and (19) can be interpreted information-geometrically, as discussed in the next section.

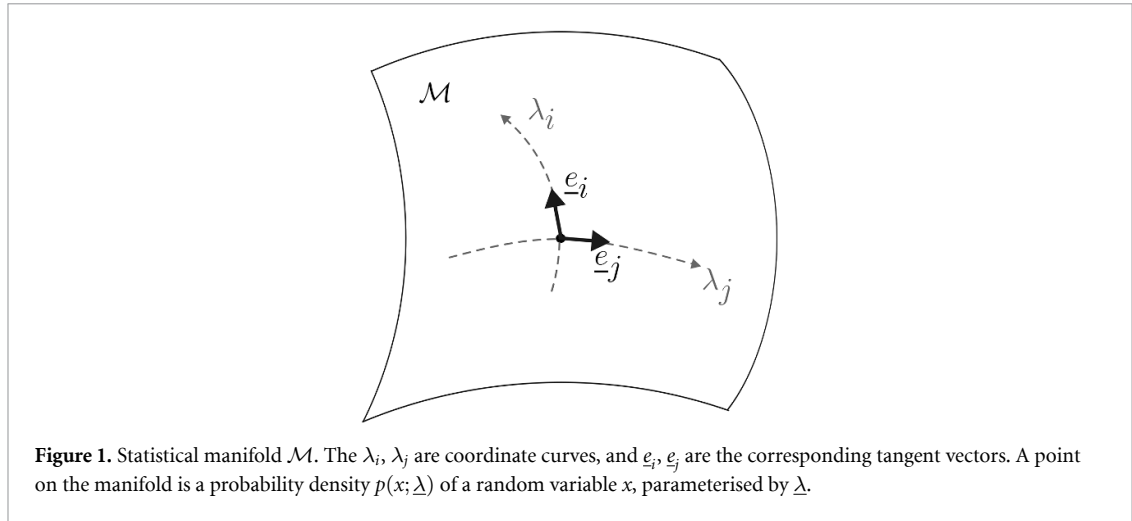


Figure 1. Statistical manifold \mathcal{M} . The λ_i, λ_j are coordinate curves, and e_i, e_j are the corresponding tangent vectors. A point on the manifold is a probability density $p(x; \underline{\lambda})$ of a random variable x , parameterised by $\underline{\lambda}$.

4.2. Information-geometric perspective of efficiency

Sections 3.1 and 4.1 extended thermodynamic efficiency to multi-parameter systems, but constrain the protocol to follow along the axis of one control parameter at a time. We now examine the thermodynamic efficiency of protocols in arbitrary directions—referred to as *directional efficiency* (equation (8)).

The numerator of equation (8), which is the directional derivative of entropy reduction in direction \underline{v} evaluated at $\underline{\lambda}$, is given by:

$$-D_{\underline{v}}S(\underline{\lambda}) = -\nabla S(\underline{\lambda}) \cdot \underline{v}. \tag{21}$$

Substituting each partial derivative $\frac{\partial S}{\partial \lambda_j}$ ($j = 1, \dots, n$) in equation (21) with expression (14) yields:

$$-D_{\underline{v}}S(\underline{\lambda}) = \sum_j v_j \sum_i \lambda_i \mathcal{I}_{ij}(\underline{\lambda}) \tag{22}$$

In information geometry, the Fisher information matrix defines a Riemannian metric on a statistical manifold, known as the Fisher–Rao metric. Consider a statistical model $\mathcal{M} = \{p(x; \underline{\lambda})\}$ parameterised by $\underline{\lambda} = (\lambda_1, \dots, \lambda_n)$. \mathcal{M} is an m -dimensional manifold with $\underline{\lambda}$ being a coordinate system (illustrated in figure 1). We define the map from the parameter space to the statistical manifold as $\varphi : \underline{\lambda} \mapsto p(\cdot; \underline{\lambda})$ and the tangent vector e_i of the i th coordinate curve on the manifold as:

$$e_i = \frac{\partial \varphi}{\partial \lambda_i}. \tag{23}$$

A component \mathcal{I}_{ij} of Fisher information matrix is the inner product of the tangent vectors e_i, e_j of \mathcal{M} in the directions of the coordinate curves λ_i, λ_j (illustrated in figure 1) [19–22]:

$$\mathcal{I}_{ij} = \langle e_i, e_j \rangle \tag{24}$$

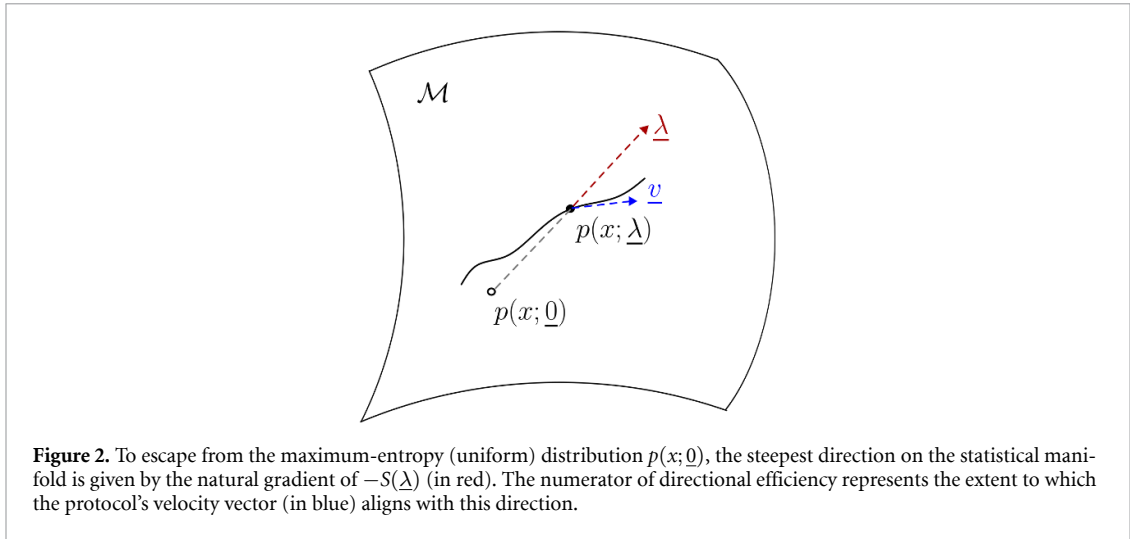
where $\langle \cdot, \cdot \rangle$ denotes the inner product of two vectors. $\mathcal{I}_{ii} = \langle e_i, e_i \rangle = \|e_i\|^2$ is therefore the squared norm of the tangent vector.

Comparing identity (24) to the numerator of $\eta_{\underline{v}}$ in equation (22), we derive that:

$$\begin{aligned} -D_{\underline{v}}S(\underline{\lambda}) &= \sum_j v_j \sum_i \lambda_i \mathcal{I}_{ij}(\underline{\lambda}) \\ &= \sum_j \sum_i v_j \lambda_i \langle e_i, e_j \rangle \\ &= \langle \underline{v}, \underline{\lambda} \rangle. \end{aligned} \tag{25}$$

Note that equation (14) written in vector form is:

$$\nabla S(\underline{\lambda}) = -\mathcal{I}(\underline{\lambda}) \underline{\lambda}. \tag{26}$$



Multiplying both sides of the equation by the inverse of the Fisher information matrix and negative one, we obtain:

$$-\mathcal{I}^{-1}(\underline{\lambda}) \nabla S(\underline{\lambda}) = \underline{\lambda}. \tag{27}$$

The LHS of the equation is known as the natural gradient of $-S(\underline{\lambda})$, which is the steepest descent direction of $S(\underline{\lambda})$ on the statistical manifold [23].

The vector $\underline{\lambda}$ thus represents both the position on the manifold and the direction of fastest entropy reduction. At the origin, $\underline{\lambda} = (0, \dots, 0)$, the statistical model corresponds to the maximum-entropy (uniform) distribution. The vector $\underline{\lambda}$, which points from the origin to the current position, provides the optimal direction to move away from maximum entropy.

In this context, the numerator of η_v in equation (25) can be considered as the inner product of the velocity \underline{v} and the natural gradient of descending entropy evaluated at the point $\underline{\lambda}$. It represents the extent to which the protocol's velocity aligns with the most efficient direction (illustrated in figure 2).

On the other hand, the denominator of η_v is the directional derivative of the generalised work. Assuming a quasi-static protocol, the denominator can be expressed as:

$$D_{\underline{v}} \langle \beta \mathbb{W} \rangle (\underline{\lambda}) = D_{\underline{v}} \log Z(\underline{\lambda}) = \nabla \log Z(\underline{\lambda}) \cdot \underline{v}. \tag{28}$$

Substituting each partial derivative $\frac{\partial \log Z}{\partial \lambda_j}$ in equation (28) with equation (15) yields:

$$D_{\underline{v}} \langle \beta \mathbb{W} \rangle (\underline{\lambda}) = \sum_j v_j \int_{\lambda_j^*}^{\lambda_j} \mathcal{I}_{jj}(\lambda_1, \dots, \lambda_j', \dots, \lambda_n) d\lambda_j'. \tag{29}$$

For each dimension $j = 0, \dots, n$, we compute the integral of the corresponding diagonal elements of the Fisher information matrix, starting from the zero-response value λ_j^* to the current value λ_j .

To interpret this equation information-geometrically, we consider the distance and energy of a path on the manifold. The geodesic metric distance for the statistical manifold between two probability distributions $p(x; \underline{\lambda})$ and $p(x; \underline{\lambda}')$, i.e. Fisher-Rao distance [24], is defined as:

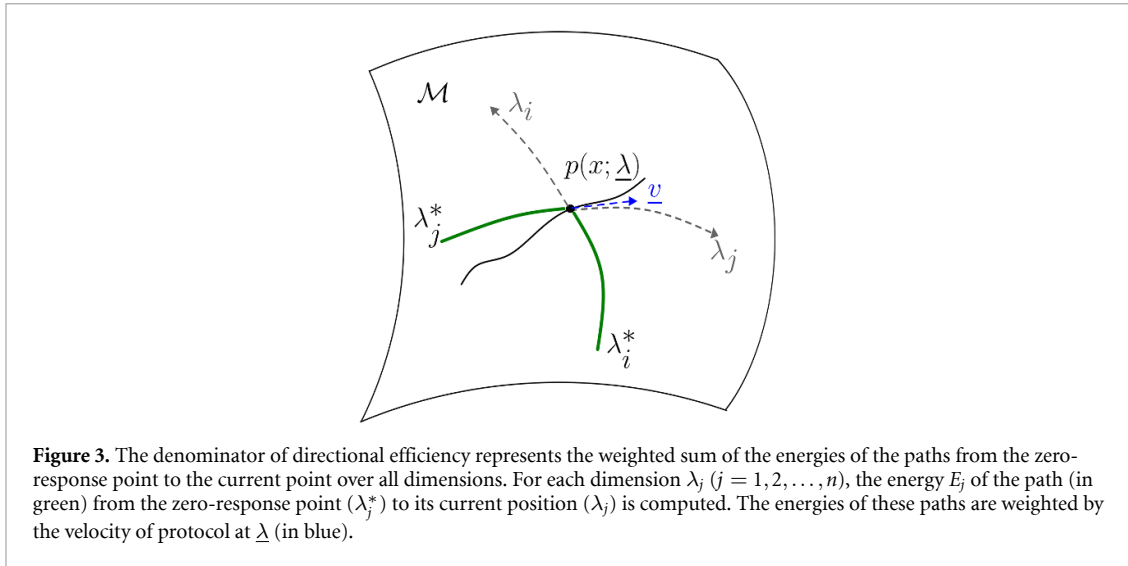
$$D_{FR}(p_{\underline{\lambda}}, p_{\underline{\lambda}'}) := \int_0^1 \sqrt{\dot{\gamma}(t)^\top \mathcal{I}_{\gamma(t)} \dot{\gamma}(t)} dt. \tag{30}$$

where γ denotes the geodesic passing through $\gamma(0) = \underline{\lambda}$ and $\gamma(1) = \underline{\lambda}'$. The dot notation represents the time derivative of the curve, i.e. the velocity.

The energy of the curve γ is given by [25, p. 20]:

$$E(\gamma) := \frac{1}{2} \int_0^1 \dot{\gamma}(t)^\top \mathcal{I}_{\gamma(t)} \dot{\gamma}(t) dt \tag{31}$$

$E(\gamma)$ is also known as 'action of the curve' [25, p.19], typically defined as the time integral of the difference between kinetic energy and potential energy along the trajectory γ .



When the system moves in the direction of a coordinate curve λ_j , the velocity vector $\dot{\gamma}(t)$ is one along axis λ_j and zero for all other axes. The energy along the path is therefore simplified to:

$$E(\gamma) = \frac{1}{2} \int_0^1 \mathcal{I}_{jj}(\gamma(t)) dt \tag{32}$$

Comparing equation (32) to equation (29), and letting $\gamma(0) = \lambda_j^*$ and $\gamma(1) = \lambda_j$, we note that the j th integral in equation (29) can be considered as the energy of the path going from the reference point λ_j^* to the point λ_j along the j th axis (subject to a constant scalar multiplier of 2). The denominator of $\eta_{\underline{v}}$ is therefore the sum of these individual path energies across all dimensions, weighted by the control protocol's velocity (figure 3).

Combining equations (22) and (29) yields:

$$\eta_{\underline{v}}(\underline{\lambda}) = \frac{\langle \underline{v}, \underline{\lambda} \rangle}{\sum_j v_j \int_{\lambda_j^*}^{\lambda_j} \mathcal{I}_{jj}(\lambda_1, \dots, \lambda_j', \dots, \lambda_n) d\lambda_j'} \tag{33}$$

When the protocol is constrained along one axis λ_j at the speed of one, equation (33) simplifies to equation (18) in section 4.1:

$$\begin{aligned} \eta_{(0, \dots, 1, \dots, 0)}(\underline{\lambda}) &= \frac{\sum_k v_k \sum_i \lambda_i \mathcal{I}_{ik}(\underline{\lambda})}{\sum_k v_k \int_{\lambda_k^*}^{\lambda_k} \mathcal{I}_{kk}(\lambda_1, \dots, \lambda_k', \dots, \lambda_n) d\lambda_k'} \\ &= \frac{\sum_i \lambda_i \mathcal{I}_{ij}(\underline{\lambda})}{\int_{\lambda_j^*}^{\lambda_j} \mathcal{I}_{jj}(\lambda_1, \dots, \lambda_j', \dots, \lambda_n) d\lambda_j'} \end{aligned} \tag{34}$$

since $v_k = 0$ when $k \neq j$ and $v_k = 1$ when $k = j$. Thus we recover $\eta^\Delta(\lambda_j)$ as defined in equation (18). In this special case, the numerator quantifies the deviation of the protocol's velocity—restricted along axis j —from the natural gradient, while the denominator represents the path energy from zero-response point to current position along the j th axis.

Under the information-geometric view, directional efficiency quantifies the efficiency of navigation on the statistical manifold following a control protocol. The directional efficiency is determined by both the system's position ($\underline{\lambda}$) and its velocity (\underline{v}) under the protocol. It measures how well the protocol aligns with the optimal direction, given the weighted sum of path energies across all coordinate directions. For a given position, one can compare various protocols by moving the system in different directions \underline{v} . Aligning \underline{v} with the natural gradient of the manifold maximises the numerator, allowing the system to gain predictability quickly. However, this may also increase the denominator if \underline{v} places a heavy weight on high-energy axis directions. An efficient protocol must strike a balance between the gain in predictability and the associated energy.

4.3. Efficiency and parameter space compression

Previous studies [26, 27] show that physical or biological systems with complex microscopic behaviour can often be effectively described by simple models with a few parameters. The reduction in parameter space dimensionality, or parameter space compression, is possible because the system's behaviour is more sensitive to certain parameter combinations than others. The Fisher information matrix $\mathcal{I}(\underline{\lambda})$ captures this sensitivity: its eigenvalues identify dominant directions in parameter space. A large eigenvalue corresponds to a 'stiff' direction, along which small perturbations to parameter λ_i produce large changes in the system's behaviour.

Thermodynamic efficiency in its Fisher information formulation (18) or (19) naturally connects to the concept of parameter space compression. Specifically, $\eta(\lambda_j)$ measures the gain in macroscopic order per unit energetic cost when changing a control parameter λ_j . If λ_j aligns with a stiff direction in the parameter space, then it will have high thermodynamic efficiency because the system responds sharply to it. Conversely, a 'sloppy' parameter direction will contribute very little to macroscopic changes, and so it will have low thermodynamic efficiency.

Therefore, thermodynamic efficiency relates to the compressibility of the parameter space. It is highest along parameter directions that are least compressible—those that dominate observable behaviours—and therefore reflects the *hierarchy* of parameter importance (captured by the corresponding eigenvalues of the Fisher information matrix) with respect to parameter space compression. A self-organizing system that tunes itself to maximise efficiency 'collapses' into its effective parameter space where a few dominating control parameters govern the macroscopic behaviour of the system.

4.4. Computational interpretation of efficiency

Considering equations (18) and (19), we can interpret thermodynamic efficiency not only in terms of Fisher information (i.e. in terms of estimation precision), but also in broad computational terms, connecting the efficiency of interactions to distributed computation within the system [6].

It has been suggested in previous studies [6, 9] that the thermodynamic efficiency can be computed in computational terms, for the single parameter case⁴:

$$\eta^\dagger(\lambda) = -\frac{dS(\lambda)}{d\lambda} \bigg/ \int_{\lambda^*}^{\lambda} \mathcal{I}(\lambda') d\lambda'. \quad (35)$$

Given our analysis, we can easily generalise this computational interpretation to multiple control parameters:

$$\eta^\dagger(\lambda_j) = -\frac{\partial S(\underline{\lambda})}{\partial \lambda_j} \bigg/ \int_{\lambda_j^*}^{\lambda_j} \mathcal{I}_{jj}(\lambda_1, \dots, \lambda'_j, \dots, \lambda_n) d\lambda'_j. \quad (36)$$

This computational, *system-centric*, interpretation quantifies the ratio of the predictability gain while computing the next state of the system to the total sensitivity accumulated on a computational path from the reference point (where no work is performed) to the current control parameter value.

5. An example

To compare the three different forms of thermodynamic efficiency, we consider the canonical 2D Ising model. The Hamiltonian of the system is given by:

$$\begin{aligned} \mathcal{H}(\underline{\sigma}; J, h) &= -J \sum_{\langle i, j \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i \\ &= -JX_1(\underline{\sigma}) - hX_2(\underline{\sigma}) \end{aligned} \quad (37)$$

where $\sigma_i \in \{-1, 1\}$ denotes the spin of the i th site and the macroscopic observables $\{X_1, X_2\}$ capture the nearest-neighbour interactions and total magnetisation respectively. The control parameters considered are the coupling strength J and the external field h . We define J and h as the dimensionless parameters that absorb the inverse temperature $\beta = \frac{1}{k_B T}$ (i.e. they correspond to $\frac{J}{k_B T}$ and $\frac{h}{k_B T}$ in the standard statistical physics notation, where $k_B \approx 1.38 \times 10^{-23}$ is the Boltzmann constant). We set $k_B T = 1$ for this analysis.

⁴ We added superscript \dagger to differentiate from other formulations.

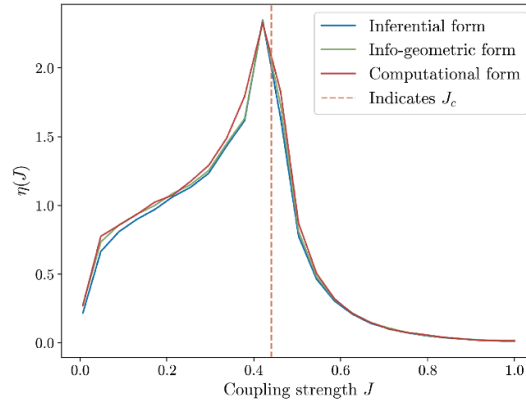


Figure 4. Thermodynamic efficiency computed for the inferential form $\eta^{\infty}(J)$, information-geometric form $\eta^{\Delta}(J)$, and computational form $\eta^{\dagger}(J)$, using 2D Ising model with zero external field. Lattice size: 50×50 .

The inferential form of thermodynamic efficiency is given by (12):

$$\eta^{\infty}(J) = -\frac{J\text{Var}(X_1) + h\text{Cov}(X_1, X_2)}{\langle X_1 \rangle} \quad (38a)$$

$$\eta^{\infty}(h) = -\frac{J\text{Cov}(X_1, X_2) + h\text{Var}(X_2)}{\langle X_2 \rangle} \quad (38b)$$

the information-geometric form is given by (18):

$$\eta^{\Delta}(J) = \frac{J\mathcal{I}_{JJ} + h\mathcal{I}_{Jh}}{\int_{J^*}^J \mathcal{I}_{JJ}(J', h) dJ'} \quad (39a)$$

$$\eta^{\Delta}(h) = \frac{J\mathcal{I}_{hJ} + h\mathcal{I}_{hh}}{\int_{h^*}^h \mathcal{I}_{hh}(J, h') dh'} \quad (39b)$$

and the computational form is given by (36):

$$\eta^{\dagger}(J) = -\frac{\partial S(J, h)}{\partial J} \bigg/ \int_{J^*}^J \mathcal{I}_{JJ}(J', h) dJ' \quad (40a)$$

$$\eta^{\dagger}(h) = -\frac{\partial S(J, h)}{\partial h} \bigg/ \int_{h^*}^h \mathcal{I}_{hh}(J, h') dh'. \quad (40b)$$

Zero-response points for J and h can be identified via the average of the corresponding conjugate observables. For a given external field h , the zero-response coupling strength J^* is the value of J that yields zero nearest-neighbour correlation, i.e. $\langle X_1 \rangle = -\langle \sum \sigma_i \sigma_j \rangle = 0$. For example, when $h = 0$, $J^* = 0$. The zero-response point for h is always $h^* = 0$ due to symmetry, as the average magnetisation $\langle X_2 \rangle = -\langle \sum \sigma_i \rangle = 0$ at zero field.

Figure 4 shows simulation results for thermodynamic efficiency with respect to parameter J under zero magnetic field ($h = 0$). All three methods yield consistent results, with maximum thermodynamic efficiency achieved at the critical coupling strength J_c . Under our convention ($k_B T = 1$), this peak aligns with Onsager's exact solution [28, 29] $J_c/(k_B T) = \frac{1}{2} \log(1 + \sqrt{2}) \approx 0.4407$, corresponding to the critical temperature $k_B T_c/J = \frac{2}{\log(1 + \sqrt{2})} \approx 2.2691$. Further details regarding the simulation setup are provided in appendix C.

Thermodynamic efficiency for control parameter h exhibits discontinuity at $h = 0$, as shown in figure 5 using the inferential form $\eta^{\infty}(h)$ as an example. The thermodynamic efficiency of h is symmetric about $h = 0$, reflecting the spatial symmetry of the magnetic field orientation. Crucially, $\eta^{\infty}(h)$ is undefined at $h = 0$ because both its numerator and denominator become zero (see appendix C figure 10 for more details). This singularity corresponds to a first-order phase transition and spontaneous symmetry breaking as the external magnetic field crosses $h = 0$. Furthermore, $\eta^{\infty}(h)$ is higher for supercritical coupling strength (or equivalent to temperature below T_c), capturing the fact that spontaneous magnetisation is more significant in the order phase.

When we vary both the coupling strength J and the external field h , we can evaluate the thermodynamic efficiency across the full parameter space. Figure 6 shows, for each parameter combination,

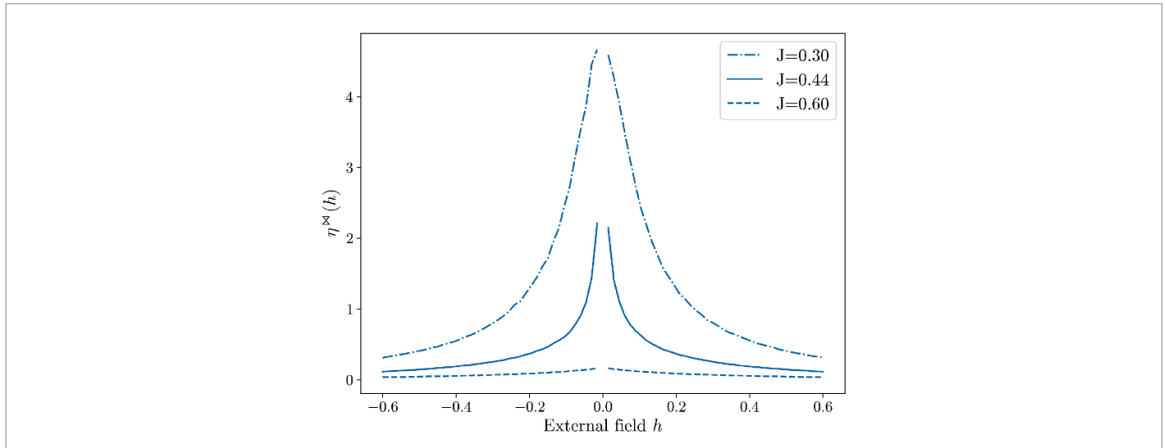


Figure 5. Thermodynamic efficiency of h computed for the inferential form $\eta^{\infty}(h)$, shown for different coupling strengths. $\eta^{\infty}(h)$ is undefined at $h = 0$ as its denominator becomes zero. The singularity demonstrates a discontinuous transition for the control parameter h , and the system undergoes spontaneous symmetry breaking at $h = 0$. Lattice size: 50×50 . Simulations initialised with random configurations.

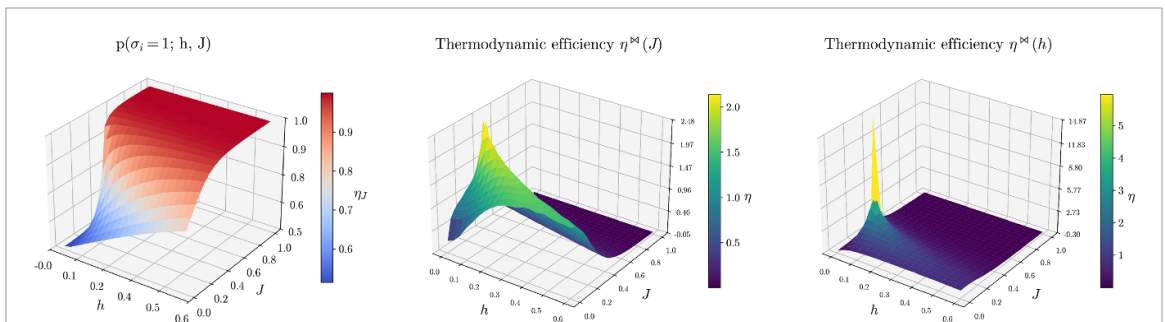


Figure 6. Probability landscape and thermodynamic efficiency η^{∞} (inferential form). **Left:** probability of positive spin at equilibrium for different combinations of external field strength h and coupling strength J . $P(\sigma_i = 1)$ closer to 0.5 (blue colour) indicates a disordered state of the system, where spins are equally likely to be up or down; $P(\sigma_i = 1)$ closer to 1 (red colour) indicates an ordered state, where spins are predominantly up. **Middle and right:** thermodynamic efficiency η^{∞} with respect to J and h (inferential form) for the same parameter combinations. Notice that the edge where the transition from disorder to order is steep corresponds to high thermodynamic efficiency. Lattice size: 50×50 .

the equilibrium probability of an up-spin, $P(\sigma_i = 1)$ (top) and the thermodynamic efficiency η^{∞} with respect to J (middle) and h (right), computed using the inferential form. Note that the probability distribution of spins is not the same as the distribution of microscopic configurations, as multiple configurations may yield the same probability distribution of up and down spins. Nonetheless, the top panel illustrates the regions of order and disorder equilibrium states in the parameter space. The edge with steep transitions from disorder to order equilibrium states aligns with the region of high thermodynamic efficiency. Thermodynamic efficiencies computed using the information-geometric and computational forms yield similar results (see appendix C).

6. Conclusion and discussion

In this paper, we have considered two different observer-centric interpretations of thermodynamic efficiency by taking the inferential and the information-geometric views, and contrasted them with the previously proposed system-centric, computational view of thermodynamic efficiency. In doing so, we derived two different forms of thermodynamic efficiency while generalising the framework to systems with multiple control parameters, governed by protocols in arbitrary directions.

We first derived the *inferential form*, which is expressed as a normalised, weighted sum of variance and covariances of the macroscopic observables. This formulation captures how changing a specific control parameter λ_j influences the joint fluctuations of all the observables $\{X_i\}$, emphasising their interdependencies. In this view, thermodynamic efficiency reflects the system’s ability to reveal hidden parameters through observable behaviour, thereby connecting the efficiency of self-organisation to that of

statistical inference. Furthermore, the inferential form provides a more robust and practical method to compute efficiency from simulations.

We also derived the second formulation, namely the *information-geometric form*, and extended thermodynamic efficiency to protocols in arbitrary directions by introducing *directional efficiency*. In this view, directional efficiency is interpreted through the lens of navigation on the statistical manifold following a control protocol trajectory. Given a protocol, directional efficiency evaluates how closely the path direction aligns with the natural gradient of entropy descent, relative to the weighted sum of path energies across all coordinate axes of the manifold. A protocol with high efficiency balances the rate of entropy reduction with the associated total energy. This perspective also connects to parameter space compression, where moving in a stiffer direction leads to a stronger effect on the system's response, corresponding to higher efficiency.

Both interpretations are observer-centric: they focus on the observer's ability to extract information about the system. In contrast, the computational form—as explored in previous studies [6, 9]—is system-centric, focusing on the system's ability to perform work to increase predictability in computing its next state. We do not suggest that the system-centric perspective on the thermodynamic efficiency offers an observer-independent measure of the quantity, because estimation of it necessarily requires a particular resolution. However, we argue that thermodynamic efficiency provides an intrinsic utility to the collective system that guides the system's behaviour independent of the observer. The two observer-centric views considered in our work explicitly capture the efficiency with respect to observable quantities, either in terms of accuracy of the statistical inference or in information geometric terms describing measurable perturbations.

Our study assumes the system remains at or near equilibrium, and that the control parameters are varied quasi-statically. This near-equilibrium dynamics leads to exponential family probability distributions of the system's states, which is important for the aforementioned formulations to hold. In addition to the well-known Boltzmann–Gibbs distribution, log-normal distributions are also used in studies of near-equilibrium dynamics, for example, in the context of weak dissipation or multiplicative phenomena [30]. Although the use of exponential family distributions is common in physical, biological and social systems [31–36], other heavier-tailed distributions such as power-law distributions are also used, especially for systems operating far from equilibrium such as financial markets [37, 38], city population [39] and various network systems [40–43]. The extension to non-equilibrium systems is possible, but requires a more careful treatment of the probability distribution and the work done by the system. When the distribution can be well approximated by an exponential family distribution, the generalised formulae can estimate the efficiency landscape of the high-dimensional parameter space. When no standard statistical distributions can be identified for the system states, thermodynamic efficiency or directional efficiency can still be computed from first principles: by perturbing the control parameter and measuring the change in system entropy and the associated work. This will be an important direction for future work, as many real-world systems operate far from equilibrium.

Our findings demonstrate that varying a control parameter affects the efficiency of a self-organizing system through the inter-dependencies of the macroscopic observables. Furthermore, the results bridge the observer-centric and system-centric views of self-organisation, and open up new interpretations grounded in statistical inference and information geometry. This unified view deepens our understanding of the fundamental principles governing self-organisation and provides a valuable framework for the study of guided self-organisation.

Acknowledgments

Q C is supported by the University of Sydney Postgraduate Award (UPA). We wish to acknowledge the support of the Sydney Informatics Hub at the University of Sydney and the National Computing Infrastructure (NCI) Australia for HPC resources that have contributed to the research results reported within this paper.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: [10.5281/zenodo.18946762](https://doi.org/10.5281/zenodo.18946762) [52].

Author contributions

Qianyang Chen  0000-0001-5382-1396

Conceptualization (equal), Formal analysis (lead), Software (lead), Writing – original draft (lead), Writing – review & editing (equal)

Nihat Ay  0000-0002-8527-2579

Conceptualization (equal), Formal analysis (supporting), Writing – review & editing (equal)

Mikhail Prokopenko  0000-0002-4215-0344

Conceptualization (equal), Formal analysis (supporting), Supervision (lead), Writing – review & editing (equal)

Appendix A. Preliminaries and core relations in generalised statistical physics

The first derivative of log partition function Z with respect to the control parameter λ_j gives the expected value of the observable X_j [11, 15, 44]:

$$\frac{\partial \log Z(\underline{\lambda})}{\partial \lambda_j} = -\langle X_j \rangle \tag{A1}$$

while its second derivative gives the covariance of the observables (X_i, X_j) [11, 15, 44]:

$$\frac{\partial^2 \log Z}{\partial \lambda_i \partial \lambda_j} = -\frac{\partial \langle X_i \rangle}{\partial \lambda_j} = -\frac{\partial \langle X_j \rangle}{\partial \lambda_i} = \text{Cov}(X_i, X_j). \tag{A2}$$

It is known that the derivative of entropy can be expressed as follows [15, p. 191], cf [45, p. 37]:

$$\begin{aligned} \frac{\partial S(\underline{\lambda})}{\partial \lambda_j} &= \frac{\partial \log Z}{\partial \lambda_j} + \sum_i \lambda_i \frac{\partial \langle X_i \rangle}{\partial \lambda_j} + \sum_i \frac{\partial \lambda_i}{\partial \lambda_j} \langle X_i \rangle \\ &= \sum_i \lambda_i \frac{\partial \langle X_i \rangle}{\partial \lambda_j}. \end{aligned} \tag{A3}$$

Hence, a combination of (A2) and (A3) yields

$$\frac{\partial S(\underline{\lambda})}{\partial \lambda_j} = -\sum_i \lambda_i \text{Cov}(X_i, X_j). \tag{A4}$$

Table (1) summarises and compares relevant information-theoretic and statistical physics quantities.

Table 1. Comparing relevant information-theoretic and statistical physics quantities.

	Information theory	Statistical physics
Microstates, observables	$x, X_i(x)$	$x, X_i(x)$
(generalised) control parameter/ lagrange multiplier ^a	λ_i	$\beta \hat{\lambda}_i$
(generalised) energy ^b	Weighted cost: $\sum_i \lambda_i X_i(x)$	Hamiltonian $\mathbb{H}(x)$: $\mathbb{H}(x) = \sum_i \hat{\lambda}_i X_i(x)$
Equilibrium distribution	$p(x) = \frac{1}{Z} e^{-\sum_i \lambda_i X_i(x)}$	$p(x) = \frac{1}{Z} e^{-\beta \sum_i \hat{\lambda}_i X_i(x)}$
Entropy	Shannon entropy S : $S = -\sum_x p(x) \log p(x)$	Thermodynamic entropy \mathbb{S} : $\mathbb{S} = -k_B \sum_x p(x) \log p(x)$
Potential	Massieu potential Ψ : $\Psi = \log Z = -\sum_i \lambda_i \langle X_i \rangle + S$	Gibbs potential \mathbb{G} : $\mathbb{G} = -\frac{1}{\beta} \log Z = \langle \mathbb{H} \rangle - T\mathbb{S}$

^a λ_i is the generalised control parameter, being the product of the control parameter $\hat{\lambda}_i$ and the corresponding Lagrange multiplier, e.g. the inverse temperature $\beta = 1/(k_B T)$.

^b In information-theoretic studies, the generalised energy is often interpreted as a linear combination of specific cost functions, weighted by their conjugate multipliers [46–48].

^c An extra scalar of $\log_2(e)$ is required if log of base 2 is used in Shannon entropy and the natural log is used in thermodynamic entropy.

Appendix B. Fisher information and its numerical approximation

Fisher information measures how sensitive the probability distribution is to changes in the control parameter λ . Let $p(x; \lambda)$ denote the probability density of the random variable X , governed by the parameter λ . The Fisher information $\mathcal{I}(\lambda)$ of the probability distribution $p(x|\lambda)$ is defined as [17]:

$$\mathcal{I}(\lambda) = \sum_x p(x; \lambda) \left(\frac{d \log p(x; \lambda)}{d\lambda} \right)^2 \tag{B1}$$

and for multiple control parameters, the Fisher information matrix $\mathcal{I}_{ij}(\lambda)$ is used [17]:

$$\mathcal{I}_{ij}(\lambda) = \sum_x p(x; \lambda) \frac{\partial \log p(x; \lambda)}{\partial \lambda_i} \frac{\partial \log p(x; \lambda)}{\partial \lambda_j}. \tag{B2}$$

The Fisher information matrix has been shown to be the same as the covariance matrix of observables X_i [11, 49]:

$$\mathcal{I}_{ij}(\lambda) = \frac{\partial^2 \log Z}{\partial \lambda_i \partial \lambda_j} = \text{Cov}(X_i, X_j). \tag{B3}$$

Numerically, the Fisher information matrix is computed using the covariance matrix of the observables instead of using first principles, as the covariance matrix is computationally more stable.

Appendix C. Simulation of ising model

We consider the canonical Ising model to compare the numerical results of the two different forms of thermodynamic efficiency [52].

The simulations are performed on a toroidal lattice of size 50×50 using Metropolis algorithm [50]. We also explored lattice sizes 25×25 and 75×75 for comparison. The coupling strength J is varied from 0 to 1 in increments of 0.05, and the external field strength h ranges from 0 to 0.6 in increments of 0.03. For each value of J and h combination, 500 independent simulations are run, and the average result is used to produce the plot. Each simulation runs for a total of 40.2 million time steps (equivalent to 16 080 lattice sweeps for a 50×50 lattice) to ensure the system reaches equilibrium, with the first 40 million steps discarded as transients.

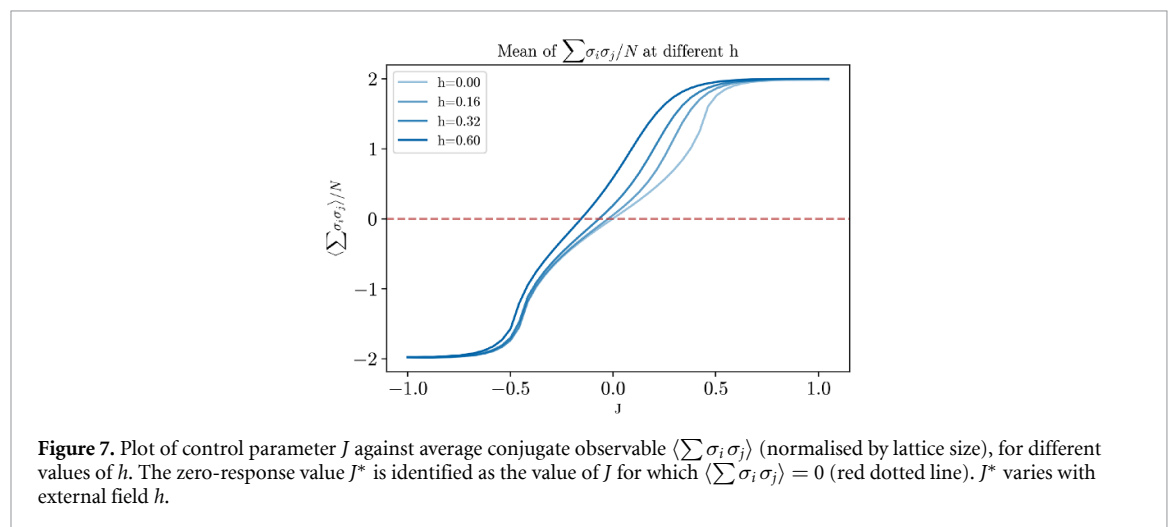
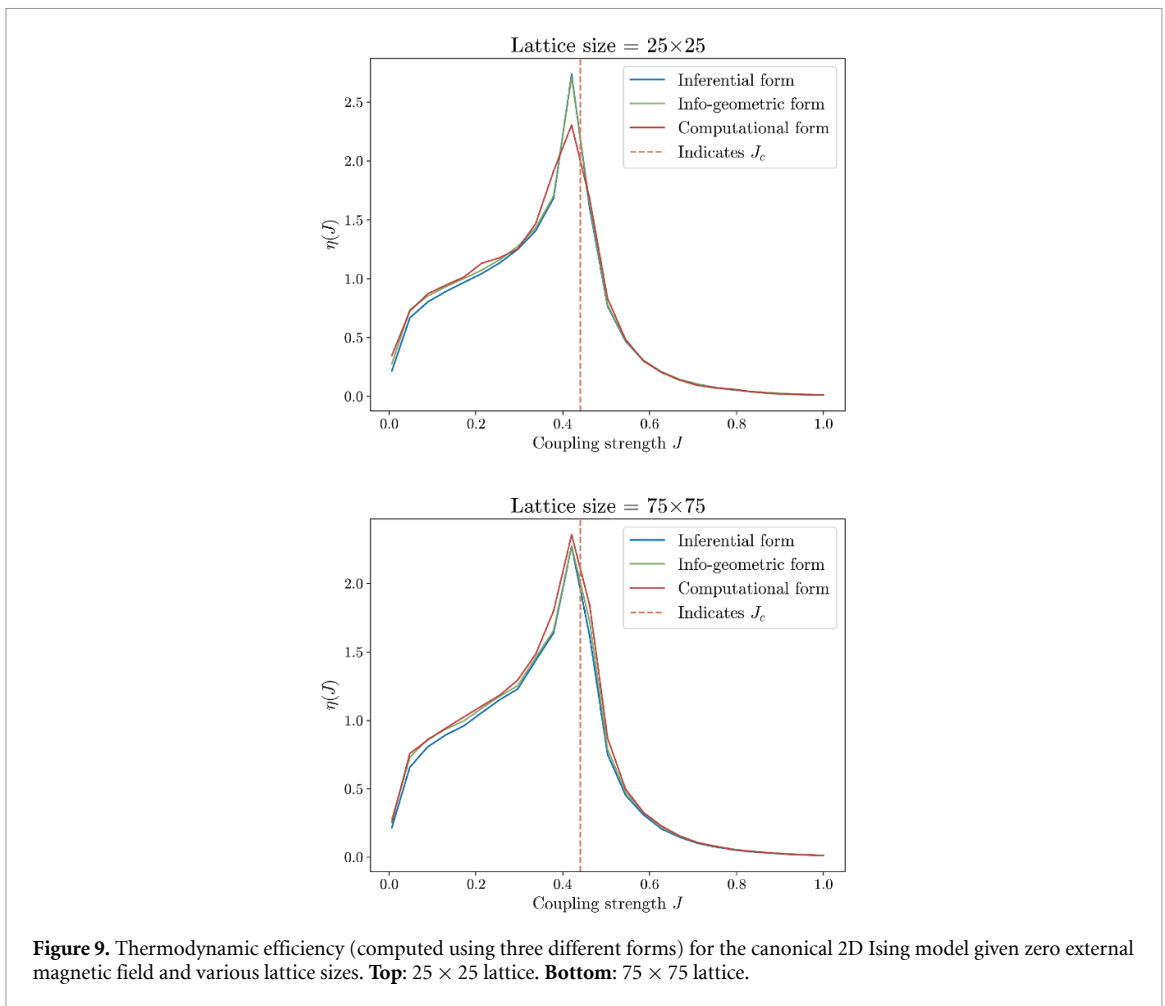
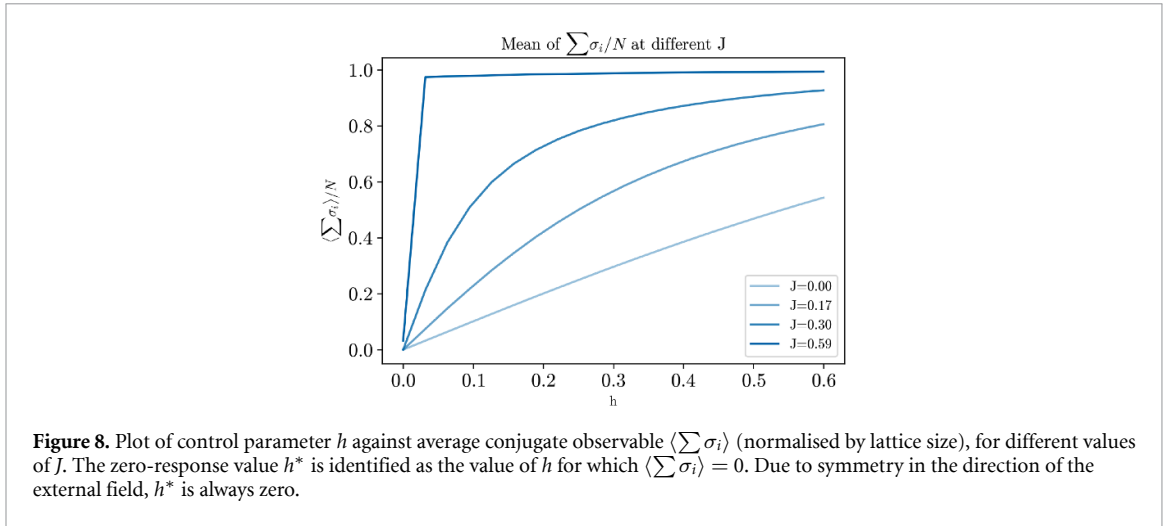


Figure 7. Plot of control parameter J against average conjugate observable $\langle \sum \sigma_i \sigma_j \rangle$ (normalised by lattice size), for different values of h . The zero-response value J^* is identified as the value of J for which $\langle \sum \sigma_i \sigma_j \rangle = 0$ (red dotted line). J^* varies with external field h .

To compute thermodynamic efficiency, configuration distributions are sampled from the last 200 000 steps at intervals corresponding to one lattice sweep. The average distribution from these samples is used to compute the Fisher information. Configuration entropy is computed by averaging over 500 simulations, each calculated using the final lattice configuration from the simulation using Kikuchi approximation [51]:

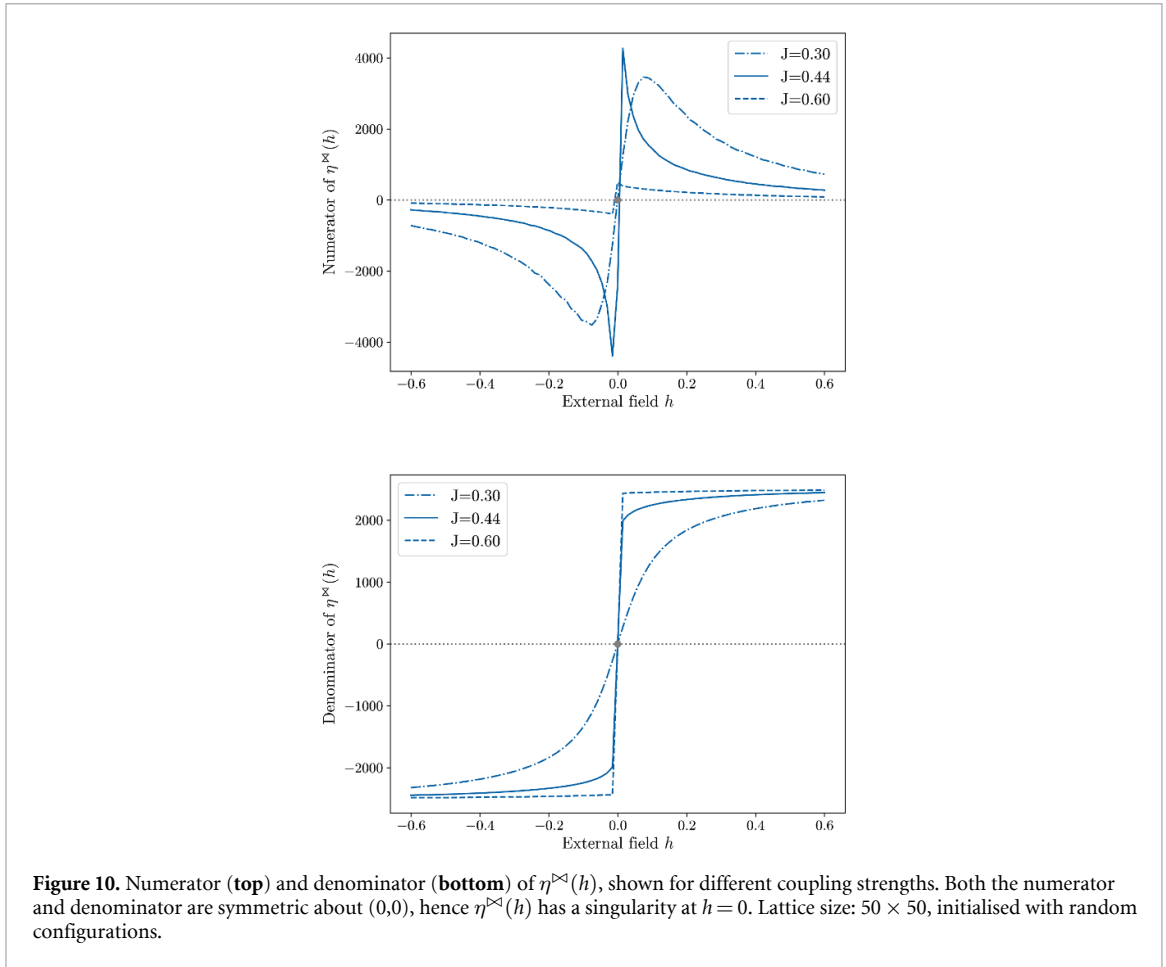
$$S = S_4 - 2S_2 + S_1 \tag{C1}$$



where S_k is the entropy of the size- k sublattices.

The lattice is initialised in a fully ordered state. At high coupling strength, single spin-flip algorithms (such as Metropolis or Glauber dynamics) risk trapping the system in a local minimum. Initialising at a fully ordered state allows the system to reach equilibrium faster at high coupling strength, without affecting outcomes at low coupling strength. Since only equilibrium states are analysed and transients are excluded, this method gives the same result as using random initialisation but significantly accelerates convergence.

To compute the integral of Fisher information in computational form η^\dagger and information-geometric form η^Δ , one must first identify the zero-response point for each combination of control parameters. The zero-response point for a given control parameter is the value at which the ensemble average of its



conjugate observable is zero. For coupling strength J , this occurs when the nearest-neighbour correlation is zero (figure 7); for the external field h , it corresponds to zero average magnetisation (figure 8).

Figure 9 provides a comparison of thermodynamic efficiency across different lattice sizes, simulated with zero external magnetic field. Notice that the information-geometric form, with Fisher information in both the numerator and denominator, is more sensitive to the change of lattice size than the other two forms, giving a much higher value in magnitude. However, all three forms are robust in terms of where the peak occurs (at the proximity of the critical regime).

To provide further insights into the singularity of $\eta(h)$ observed in figure 5, we here examine the behaviour of the numerator and denominator of thermodynamic efficiency. Figure 10 shows the two components separately for inferential form $\eta^{\infty}(h)$ with $h \in [-0.6, +0.6]$. The numerator and denominator are symmetric about the origin due to the spatial symmetry of the orientation of the magnetic field. At $h=0$, both quantities are zero, leading to a singularity of $\eta(h)$ at $h=0$.

Results for different combinations of external field strengths h and coupling strength J are shown in figure 11. While the methods have different sensitivities to finite-size effects and simulation resolutions, they both indicate the same region of high efficiency, aligning with the sharp transition between ordered and disordered equilibrium states.

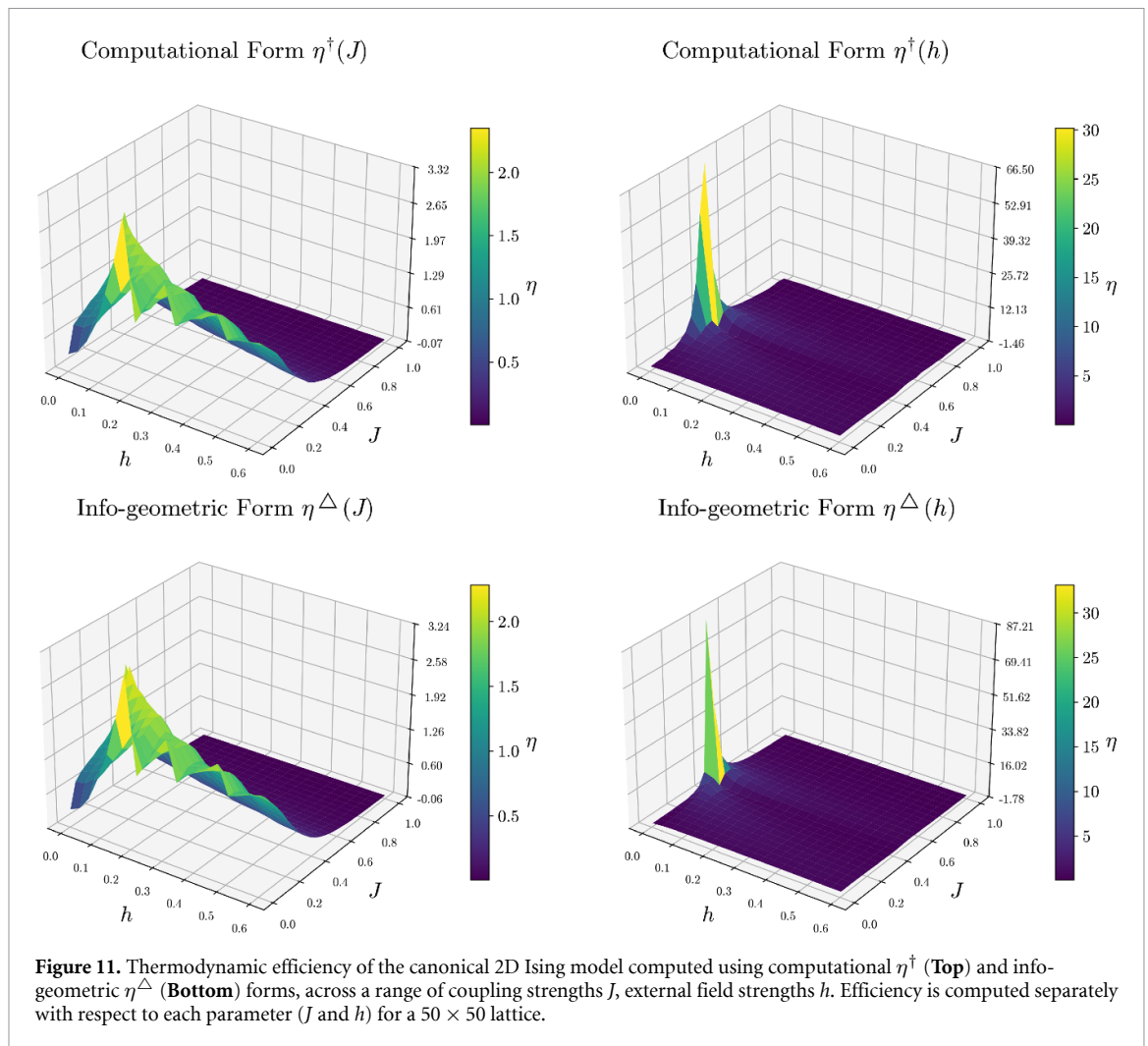


Figure 11. Thermodynamic efficiency of the canonical 2D Ising model computed using computational η^\dagger (Top) and info-geometric η^Δ (Bottom) forms, across a range of coupling strengths J , external field strengths h . Efficiency is computed separately with respect to each parameter (J and h) for a 50×50 lattice.

References

- [1] Ramaswamy S 2010 The mechanics and statistics of active matter *Annu. Rev. Condens. Matter Phys.* **1** 323
- [2] Marchetti M C, Joanny J F, Ramaswamy S, Liverpool T B, Prost J, Rao M and Simha R A 2013 Hydrodynamics of soft active matter *Rev. Mod. Phys.* **85** 1143
- [3] Vicsek T and Zafeiris A 2012 Collective motion *Phys. Rep.* **517** 71
- [4] Crosato E, Prokopenko M and Spinney R E 2019 Irreversibility and emergent structure in active matter *Phys. Rev. E* **100** 042613
- [5] Crosato E, Nigmatullin R and Prokopenko M 2018 On critical dynamics and thermodynamic efficiency of urban transformations *R. Soc. Open Sci.* **5** 180863
- [6] Crosato E, Spinney R E, Nigmatullin R, Lizier J T and Prokopenko M 2018 Thermodynamics and computation during collective motion near criticality *Phys. Rev. E* **97** 1
- [7] Harding N, Nigmatullin R and Prokopenko M 2018 Thermodynamic efficiency of contagions: a statistical mechanical analysis of the SIS epidemic model *Interface Focus* **8** 20180036
- [8] Nigmatullin R and Prokopenko M 2021 Thermodynamic efficiency of interactions in self-organizing systems *Entropy* **23** 757
- [9] Chen Q and Prokopenko M 2025 Why collective behaviours self-organize to criticality: A primer on information-theoretic and thermodynamic utility measures *R. Soc. Open Sci.* **12** 241655
- [10] Jaynes E T 1957 Information theory and statistical mechanics *Phys. Rev.* **106** 620
- [11] Crooks G E 2007 Measuring thermodynamic length *Phys. Rev. Lett.* **99** 100602
- [12] Shannon C E 1948 A mathematical theory of communication *Bell Syst. Tech. J.* **27** 623
- [13] Landau L D and Lifšic E M 2011 *Statistical Physics (Course of Theoretical Physics vol 5)* 3rd edn (Elsevier Butterworth Heinemann)
- [14] Loomis L H and Sternberg S 1990 *Advanced Calculus* rev. edn (Jones and Bartlett)
- [15] Jaynes E 1963 Information theory and statistical mechanics *Lectures in Theoretical Physics (Statistical Physics vol 3)* ed K Ford (Brandeis University Summer Institute, W. A. Benjamin, Inc.) pp 181–218
- [16] Fisher R A 1922 On the mathematical foundations of theoretical statistics *Phil. Trans. R. Soc. A* **222** 309
- [17] Cover T M and Thomas J A 2005 *Elements of Information Theory* 2nd edn (Wiley)
- [18] Prokopenko M, Lizier J T, Obst O and Wang X R 2011 Relating Fisher information to order parameters *Phys. Rev. E* **84** 041116
- [19] Amari S-I and Nagaoka H 2007 *Methods of Information Geometry (Translations of Mathematical Monographs)* (American Mathematical Society) p 191
- [20] Amari S-I 2016 *Information Geometry and Its Applications (Applied Mathematical Sciences vol 194)* (Springer)
- [21] Ay N, Jost J, Le H V and Schwachhöfer L 2017 *Information Geometry (Ergebnisse der Mathematik und Ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics vol 64)* (Springer)

- [22] Amari S-I 2021 Information geometry *Int. Stat. Rev.* **89** 250
- [23] Amari S-I 1998 Natural gradient works efficiently in learning *Neural Comput.* **10** 251
- [24] Nielsen F 2020 An elementary introduction to information geometry *Entropy* **22** 1100
- [25] Jost J 2017 *Riemannian Geometry and Geometric Analysis* (Springer)
- [26] Machta B B, Chachra R, Transtrum M K and Sethna J P 2013 Parameter space compression underlies emergent theories and predictive models *Science* **342** 604
- [27] Transtrum M K, Machta B B, Brown K S, Daniels B C, Myers C R and Sethna J P 2015 Perspective: sloppiness and emergent theories in physics, biology and beyond *J. Chem. Phys.* **143** 010901
- [28] Onsager L 1944 Crystal Statistics. I. a two-dimensional model with an order-disorder transition *Phys. Rev.* **65** 117
- [29] Walker J S 2023 *A Student's Guide to the Ising Model (Student's Guides)* (Cambridge University Press) ch 5, pp 125–49
- [30] Kleidon A, Gozzi C, Bucciati A and Sauro Graziano R 2024 Type of probability distribution reflects how close mixing dynamics in river chemistry are to thermodynamic equilibrium *Sci. Total Environ.* **941** 173409
- [31] Beck C 2007 Statistics of three-dimensional Lagrangian turbulence *Phys. Rev. Lett.* **98** 064502
- [32] Gustavsson K and Mehlig B 2016 Statistical models for spatial patterns of heavy particles in turbulence *Adv. Phys.* **65** 1
- [33] Bialek W, Cavagna A, Giardina I, Mora T, Silvestri E, Viale M and Walczak A M 2012 Statistical mechanics for natural flocks of birds *Proc. Natl Acad. Sci.* **109** 4786
- [34] Nghiem T-A, Telenczuk B, Marre O, Destexhe A and Ferrari U 2018 Maximum-entropy models reveal the excitatory and inhibitory correlation structures in cortical neuronal activity *Phys. Rev. E* **98** 012402
- [35] Zanoci C, Dehghani N and Tegmark M 2019 Ensemble inhibition and excitation in the human cortex: an ising-model analysis with uncertainties *Phys. Rev. E* **99** 032408
- [36] Hunter D R 2007 Curved exponential family models for social networks *Soc. Netw.* **29** 216
- [37] Mantegna R N and Stanley H E 1995 Scaling behaviour in the dynamics of an economic index *Nature* **376** 46
- [38] Jiang Z-Q and Zhou W-X 2010 Complex stock trading network among investors *Phys. A* **389** 4929
- [39] Marsili M and Zhang Y-C 1998 Interacting individuals leading to Zipf's Law *Phys. Rev. Lett.* **80** 2741
- [40] Barabási A-L and Albert R 1999 Emergence of scaling in random networks *Science* **286** 509
- [41] Newman M E J 2001 Clustering and preferential attachment in growing networks *Phys. Rev. E* **64** 025102
- [42] Barabási A-L 2005 The origin of bursts and heavy tails in human dynamics *Nature* **435** 207
- [43] Muchnik L, Pei S, Parra L C, Reis S D S, Andrade Jr J S, Havlin S and Makse H A 2013 Origins of power-law degree distribution in the heterogeneity of human activity in social networks *Sci. Rep.* **3** 1783
- [44] Niven R K and Andresen B 2010 Jaynes' maximum entropy principle, Riemannian metrics and generalised least action bound *Complex Physical, Biophysical and Econophysical Systems* (World Scientific) pp 283–317
- [45] de Groot S R and Mazur P 2013 *Non-Equilibrium Thermodynamics (Dover Books on Physics)* (Dover Publications)
- [46] Blahut R 2003 Computation of channel capacity and rate-distortion functions *IEEE Trans. Inf. Theory* **18** 460
- [47] Wilson A 2008 Boltzmann, Lotka and Volterra and spatial structural evolution: an integrated methodology for some dynamical systems *J. R. Soc. Interface* **5** 865
- [48] Salge C, Ay N, Polani D and Prokopenko M 2015 Zipf's law: balancing signal usage cost and communication efficiency *PLoS One* **10** e0139475
- [49] Brody D and Rivier N 1995 Geometrical aspects of statistical mechanics *Phys. Rev. E* **51** 1006
- [50] Metropolis N, Rosenbluth A W, Rosenbluth M N, Teller A H and Teller E 1953 equation of state calculations by fast computing machines *J. Chem. Phys.* **21** 1087
- [51] Kikuchi R 1951 A theory of cooperative phenomena *Phys. Rev.* **81** 988
- [52] Chen Q Y 2025 Ising model for thermodynamic efficiency (<https://doi.org/10.5281/zenodo.18946762>)