Maximum caliber inference and the stochastic Ising model

Carlo Cafaro[1] and Sean Alan Ali[2]
[1]*SUNY Polytechnic Institute, Albany, New York 12203, USA*
[2]*Albany College of Pharmacy and Health Sciences, Albany, New York 12208, USA*
(Received 3 August 2016; revised manuscript received 4 November 2016; published 28 November 2016)

We investigate the maximum caliber variational principle as an inference algorithm used to predict dynamical properties of complex nonequilibrium, stationary, statistical systems in the presence of incomplete information. Specifically, we maximize the path entropy over discrete time step trajectories subject to normalization, stationarity, and detailed balance constraints together with a path-dependent dynamical information constraint reflecting a given average global behavior of the complex system. A general expression for the transition probability values associated with the stationary random Markov processes describing the nonequilibrium stationary system is computed. By virtue of our analysis, we uncover that a convenient choice of the dynamical information constraint together with a perturbative asymptotic expansion with respect to its corresponding Lagrange multiplier of the general expression for the transition probability leads to a formal overlap with the well-known Glauber hyperbolic tangent rule for the transition probability for the stochastic Ising model in the limit of very high temperatures of the heat reservoir.

DOI: 10.1103/PhysRevE.94.052145

I. INTRODUCTION

A detailed knowledge of microscopic dynamics may be neither necessary nor sufficient to understand the macroscopic behavior of a complex system. For example, entropy is a crucially important feature of macrostates that cannot be determined from microstates. Following the statistical mechanical works of Gibbs [1] and inspired by Shannon's advances in information theory [2], Jaynes is the undisputed pioneer in the use of entropy for optimal information processing leading to reliable macroscopic predictions in the presence of incomplete information [3–7]. We point out that information processing is considered optimal when it takes into account all available knowledge of the microphysics as well as all the available macroscopic data. No arbitrary assumptions beyond that are introduced.

A. Jaynes's MaxEnt

Jaynes's MaxEnt is an inductive method of reasoning for making predictions about *equilibrium properties* of macroscopic phenomena in the presence of incomplete information [3]. The only type of initial information allowed is values of quantities which are observed to be constant in time. In synthesis, MaxEnt is a *state-based* variational method of information-theoretic nature aiming at inferring macroscopic (conserved) properties of complex systems at equilibrium in the presence of limited information about the microscopic nature of the phenomena being investigated. Macroscopic predictions generated by the MaxEnt inference algorithm are related to the experimental behavior of actual physical systems only when, and to the extent that, they lead to sharply peaked probability distributions. Furthermore, if it occurs that there is experimental evidence that a given MaxEnt prediction is incorrect, one should reasonably conclude that the enumeration of the possible microstates suggested by our knowledge of the laws of physics was not correctly given. The MaxEnt inference algorithm should be applied again by modifying the set of input information constraints corresponding to the newly updated enumeration of the microstates of the system. Indeed, failure of the MaxEnt algorithm could be regarded as more valuable than its success since this can lead to fundamental advances in science [7]. For instance, the failures of classical statistical mechanics that are rectified by quantum theory constitute valid examples of such a possibility. Despite a few weak technical aspects of Jaynes's approach [8], it is unjustified to state that physics becomes irrelevant in the MaxEnt formalism. As pointed out earlier, failures of the MaxEnt inference algorithm are ascribed to its physical rather than statistical aspects. The MaxEnt inference is merely an algorithm, a messenger [9]. The MaxEnt inference algorithm ensures the objectivity of its predictions but does not claim deductive certainty for them. Ultimately, conclusive successes and failures belong to physics. Having said that, we also have to point out that the application of MaxEnt is not limited to statistical mechanics. Practical applications of the MaxEnt algorithm include, but are not limited to, image reconstruction in radio and x-ray astronomy [10,11], image reconstruction in medical tomography [12], x-ray crystallography [13], molecular biology [14], nuclear magnetic resonance spectroscopy [15], and the collective dynamics of a population of neurons [16]. From a more theoretical perspective, several applications of the MaxEnt algorithm appear in the characterization of complexity of natural phenomena [17–20], energy levels statistics [21,22], quantum entanglement [23,24], and ferromagnetic materials [25].

B. Jaynes's MaxCal

The extension of Jaynes's MaxEnt principle to nonequilibrium statistical mechanical phenomena is known as Jaynes's MaxCal principle [5,6]. It constitutes an inductive method of reasoning for making predictions of *nonequilibrium properties* of macroscopic phenomena in the presence of incomplete information. The type of initial information allowed is extended to values of quantities which are observed to be nonconstant in time. Unlike MaxEnt, MaxCal is a *path-based* variational

 ©2016 American Physical Society

method of information-theoretic nature aiming at inferring macroscopic (dynamical) properties of complex systems out of equilibrium in the presence of partial knowledge about the microscopic nature of the phenomena of interest. From an applied perspective, it is found that MaxCal is especially suitable for solving problems involving small systems that are out of equilibrium where fluctuations become important. Systems of this type appear naturally in biology and nanotechnology [26–28]. Specifically, in Refs. [26,27] MaxCal was applied to derive several flux distributions in diffusive systems: Fick's law of particle transport, Fourier's law of heat transport, and Newton's viscosity law of momentum transport. In Ref. [28], instead, MaxCal was experimentally tested in the study of a single colloidal particle transitioning between two energy wells. In recent years, MaxCal has also found application to the dynamics of collectively moving animal groups [29]. Recent years have also witnessed a significant number of more theoretically oriented applications of the MaxCal inference algorithm [9,30–33]. In particular, MaxCal has been employed to derive master equations [34,35] and Markov processes [36], it has been challenged to make inferences in the realm of non-Markovian dynamics [37], it has been used to recover classical Newtonian equations of motion [32] and the Fokker-Planck equation [33], and, finally, it has been proposed as a general variational principle for nonequilibrium statistical mechanics [38]. For an extensive review on the MaxCal formalism, we refer to Ref. [39]. We emphasize that when taking into consideration nonequilibrium scenarios, Jaynes was concerned with continuous paths satisfying deterministic Hamiltonian equations of motion [5,6] while Filyukov and Karpov focused their attention on systems having discrete dynamical states where trajectories are composed of discrete time steps [40–42]. Specifically, they assumed that the evolution of the system may be specified by a Markov chain with discrete times and a finite number of states. In particular, they noticed that, in addition to the state probabilities, path probabilities were required to study nonequilibrium processes since it was important to know the rates of transition from one state to another [40,41]. This discrete line of investigation is generating very interesting findings. In Ref. [43], the MaxCal inference procedure was used to infer the transition probabilities in a stationary Markov process given the knowledge of both the stationary-state populations and an average global dynamical quantity. Specifically, Dixit and Dill showed that their work yields the correct values of dynamical quantities (normalized occupancy autocorrelation) in an example of molecular dynamics simulations of a water solvation shell around a single water molecule (the average fluctuation of the number of molecules in the hydration shell was assumed known). In Ref. [44], the MaxCal inference procedure was used to infer both the stationary-state distributions and the transition probabilities in a stationary Markov process in the presence of both state- and path-dependent information constraints. Specifically, Dixit considered a particle diffusing on a two-dimensional square lattice where the external potential field was used as the average global dynamical quantity in the path entropy maximization procedure. In Ref. [45], MaxCal was applied to network science in order to infer transition rates between nodes of a network in the presence of partial knowledge in the form of both state- and path-dependent information. In particular, when the average global dynamical quantity known was represented by the mean jump rate, it was uncovered that the transition rates ω_{ij} exhibit a square-root dependence on the stationary state populations at nodes i and j.

In principle, the Liouville equation is the conventional starting point for the description of time-dependent statistical systems [46]. However, in most spin systems, the lack of complete knowledge leads to the absence of an explicit form of the interaction Hamiltonian between the spins and the heat reservoir [47]. This fact remains true even in the case of the linear Ising chain, the simplest Ising model [48]. For this reason, the Liouville equation is not useful whereas the method of the master equation allows one to overcome this technical difficulty. Indeed, within this latter method, it is only necessary to know the Hamiltonian of the spin system while the interaction that causes the time transition is assumed to be stochastic. This line of reasoning leads to the so-called Glauber stochastic Ising model [49]. As pointed out by Glauber himself in Ref. [49], the functional form of the transition probabilities proposed by him was motivated by simplicity requirements rather than generality conditions.

In this article, motivated by the considerations presented in Ref. [49] and especially inspired by the works presented in Refs. [43–45], we maximize the path entropy over discrete time step trajectories subject to normalization, stationarity, and detailed balance constraints together with a path-dependent dynamical information constraint reflecting a given average global behavior of the complex spin system being considered. We compute a general expression for the transition probability values associated with the stationary random Markov processes describing the nonequilibrium stationary system. Finally, we uncover that a convenient choice of the dynamical information constraint together with a perturbative expansion with respect to the corresponding Lagrange multiplier of the general expression for the transition probability leads to a formal overlap with the well-known Glauber hyperbolic tangent rule for the transition probability for the stochastic Ising model [50] in the limit of very high temperatures of the heat reservoir.

The layout of this article is as follows. In Sec. II, we briefly describe the stochastic Ising model as proposed by Glauber. In particular, we recast Glauber's hyperbolic tangent rule in a form more convenient for our analysis. In Sec. III, we present a synthetic picture of the philosophy underlying the MaxCal formalism. We focus our attention on the statistical inference of nonequilibrium properties of complex systems characterized by discrete dynamical states where paths are composed of discrete time steps. Section IV is divided into two parts. In the first part, we show a detailed computation of the path entropy maximization procedure in the presence of various information constraints, including a path ensemble average of the product of two neighboring spin values. In the second part, we formally compare the outcome of our computation with Glauber's transition probability expression in two limiting cases. Finally, our conclusions are reported in Sec. V.

II. THE STOCHASTIC ISING MODEL

As pointed out in the introduction, the Liouville equation would be the standard starting point for the description of

time-dependent statistical systems. However, in most cases, an explicit form of the interaction Hamiltonian between the spins and the heat reservoir is not fully known. For this reason, the Liouville equation is not useful in such a case and the method of the master equation allows one to overcome this technical difficulty. Indeed, within this method, it is only necessary to know the Hamiltonian of the spin system while the interaction that causes the time transition is assumed to be stochastic. The master equation is given by [51]

$$\frac{\partial p_i(t)}{\partial t} = -\sum_{i\neq j}[p_i(t)\omega_{i\to j} - p_j(t)\omega_{j\to i}], \tag{1}$$

where $p_i(t)$ is the probability of the system being in the state i at time t, while $\omega_{i\to j}$ is the transition rate for $i \to j$. At equilibrium,

$$\frac{\partial p_i(t)}{\partial t} = 0, \tag{2}$$

that is,

$$\frac{p_i(t)}{p_j(t)} = \frac{\omega_{j\to i}}{\omega_{i\to j}}. \tag{3}$$

Equation (3) is known as the detailed balance condition. Observe that, in explicit analogy with equilibrium statistical mechanics, the probability of the ith state occurring in a classical system is given by

$$p_i(t) = \frac{1}{\mathcal{Z}} e^{-\frac{E_i}{k_B T}}, \tag{4}$$

where $\mathcal{Z}$ is the partition function and E_i is the energy of the spin in the ith state. Usually, this probability is only approximately known due to the denominator [51].

In the absence of any external magnetic field and in the working hypothesis that each spin σ_i is coupled through the transition probabilities $\omega_i(\sigma_i \to -\sigma_i)$ to only its nearest-neighbor spins σ_{i-1} and σ_{i+1}, and motivated by both simplicity and symmetry arguments, Glauber considered a stochastic Ising model described by a master equation in which the transition probabilities were given by [49]

$$\omega_i(\sigma_i \to -\sigma_i) = \frac{\Gamma}{2}\left\{1 - \frac{1}{2}\gamma\sigma_i(\sigma_{i-1} + \sigma_{i+1})\right\}. \tag{5}$$

In Eq. (5), Γ is a parameter that characterizes the time scale on which all transitions take place, $\Gamma/2$ denotes the rate per unit time at which the particle makes transitions from either state to the opposite, and γ is a parameter that describes the tendency of spins toward alignment. The explicit expression for γ in Eq. (5) can be obtained by imposing the detailed balancing condition at equilibrium at temperature T for the Ising model,

$$\frac{p_i(-\sigma_i)}{p_i(\sigma_i)} = \frac{\omega_i(\sigma_i \to -\sigma_i)}{\omega_i(-\sigma_i \to \sigma_i)}. \tag{6}$$

The quantity $p_i(\sigma_i)$ in Eq. (6) denotes the probability that the ith spin will assume the value σ_i and is proportional to the Maxwell-Boltzmann factor,

$$p_i(\sigma_i) \propto \exp\left[-\frac{J}{k_B T}\sigma_i(\sigma_{i-1} + \sigma_{i+1})\right], \tag{7}$$

where J is the exchange coupling constant and k_B is the Boltzmann constant. Substituting Eqs. (5) and (7) into Eq. (6), we obtain

$$\gamma = \gamma(k_B, J, T) = \tanh\left(\frac{2J}{k_B T}\right). \tag{8}$$

Substituting Eq. (8) into Eq. (5) and using the fact that $\sigma_i = \pm 1$ together with the point symmetry of the hyperbolic tangent function, we find

$$\omega_i(\sigma_i \to -\sigma_i) = \frac{\Gamma}{2}[1 - \sigma_i \tanh(\beta h_i)], \tag{9}$$

where $\beta \stackrel{\text{def}}{=} \frac{1}{k_B T}$ and h_i denotes a local magnetic field defined as

$$h_i \stackrel{\text{def}}{=} J(\sigma_{i-1} + \sigma_{i+1}). \tag{10}$$

Equation (9) is the so-called Glauber hyperbolic tangent rule. For future use, we recast this equation in a more convenient form. Using again the fact that $\sigma_i = \pm 1$ and exploiting the point symmetry of the hyperbolic tangent function, it follows that

$$\sigma_i \tanh(\beta h_i) = \tanh(\beta \sigma_i h_i). \tag{11}$$

Observe that the energy difference ΔE between a proposed new microstate j and an old microstate i is given by [52]

$$\Delta E \stackrel{\text{def}}{=} E_j - E_i = 2\sigma_i h_i. \tag{12}$$

Finally, inserting Eqs. (11) and (12) into Eq. (9), after some algebra we obtain

$$\omega_{ij} = \frac{\Gamma}{2}\left[1 - \tanh\left(\frac{\beta\Delta E}{2}\right)\right], \tag{13}$$

that is,

$$[\omega_{ij}]_{\text{Glauber}} = \Gamma \frac{e^{-\frac{\beta\Delta E}{2}}}{e^{\frac{\beta\Delta E}{2}} + e^{-\frac{\beta\Delta E}{2}}}. \tag{14}$$

For the sake of clarity, we remark that Glauber pointed out in Ref. [49] that in the working hypothesis of nearest-neighbor coupling among spins, the functional form of the transition probability ω_{ij} that leads to the same equilibrium state as the Ising model is not unique. The specific functional form of ω_{ij} in Eq. (5) as employed by Glauber in the master equation formalism was selected for simplifying the equations to handle (for instance, equations describing the spin expectation values) rather than satisfying a fundamental physical requirement. More specifically, the manner in which ω_{ij} depends on neighboring spin values chosen by Glauber was dictated by the necessity of describing a tendency for each spin to align itself parallel to its nearest neighbors. In summary, the assumption that ω_{ij} depends symmetrically on the two neighboring spins σ_{i-1} and σ_{i+1} as well as σ_i is a clever *ad hoc* assumption adopted by Glauber.

The expression for the transition probability ω_{ij} in Eq. (14) will be useful in the remainder of the article.

III. THE MAXCAL FORMALISM

In the MaxCal formalism [5,6,40,41], the path entropy to be maximized can be defined as [40,41]

$$H(\{p(C)\}) \stackrel{\text{def}}{=} -\sum_{\{C\}} p(C)\ln[p(C)], \tag{15}$$

where $p(C)$ denotes the probability that the dynamical process follows the path C. As mentioned in the introduction, we recall that Jaynes was concerned with continuous paths C satisfying deterministic Hamiltonian equations of motion [5,6]. However, for systems having continuous dynamical states, it is computationally hard to compute a path ensemble using the microscopic dynamics. Rather than considering continuous paths, Filyukov and Karpov focused their attention on systems having discrete dynamical states where trajectories are composed of discrete time steps [40,41]. Specifically, they assumed that the evolution of the system may be specified by a Markov chain with discrete times and a finite number of states. The trajectory C_T of a Markov chain of length T is described by the sequence of states

$$C_T = i_0 i_1 \cdots i_{T-1} i_T. \tag{16}$$

Assuming a stationary first-order Markov process [53,54], the probability $p(C_T)$ of the trajectory is given by

$$p(C_T) = p_{i_0} \omega_{i_0 i_1} \cdots \omega_{i_{T-1} i_T}, \tag{17}$$

where $\omega_{ij} = \omega_{ij}(\tau)$ denotes the conditional probability of a transition on the time interval τ from the state i to the state j while p_{i_0} is the single state i_0 occupation probability. Note that the conditional transition probabilities ω_{ij} and the stationary single state probabilities p_i satisfy the normalization constraints

$$\sum_j \omega_{ij} = 1 \text{ and } \sum_i p_i = 1, \tag{18}$$

respectively. Furthermore, the stationarity of the chain is encoded in the following constraint:

$$\sum_i p_i \omega_{ij} = p_j. \tag{19}$$

It can be shown that for a sufficiently long ergodic chain, that is to say a chain for which T approaches infinity and $\omega_{ij} > 0$ (all states are connected by a nonzero transition probability), the path entropy in Eq. (15) can be approximated by (for further details, see Refs. [39–41])

$$H(\{p(C)\}) \approx H(T) \stackrel{\text{def}}{=} T H(1), \tag{20}$$

where $H(1)$ denotes the path entropy per step defined as

$$H(1) \stackrel{\text{def}}{=} -\sum_{i,j} p_i \omega_{ij} \ln[\omega_{ij}]. \tag{21}$$

Observe that the path entropy per step reduces to the ordinary entropy of equilibrium statistical mechanics when $\omega_{ij} = p_j$, that is to say, when an instant equilibration condition is achieved.

The path entropy in Eq. (15) is also known as the *caliber*, a cross sectional area of a tube that, in part, quantifies the flow in a dynamic process. By maximizing the caliber subject to all available constraints (normalization conditions, dynamical averages, etc.), MaxCal yields the least biased probability $p(C)$ for the set of microscopic trajectories $\{C\}$ consistent with the observed information constraints. Specifically, given the knowledge of all possible microscopic trajectories explored in a specific interval of time by a system, MaxCal seeks to construct a weighted ensemble of microscopic trajectories consistent with the constrained averages obtained by measuring a small (much smaller than the number of known microscopic trajectories) number of dynamical quantities (for instance, average microscopic fluxes). In turn, this weighted ensemble of microscopic trajectories determines the time evolution of all time-dependent observables of the system. In summary, in analogy to MaxEnt, macroscopic quantities are computed in terms of derivatives of a dynamical partition function defining the normalization factor of the least biased probability $p(C)$.

IV. PATH ENTROPY MAXIMIZATION AND NEIGHBORING SPIN VALUES

In what follows, we assume that the path entropy per unit time to be maximized is given by the caliber in Eq. (21),

$$\mathcal{C} \stackrel{\text{def}}{=} -\sum_{i,j} p_i \omega_{ij} \ln[\omega_{ij}]. \tag{22}$$

A. The explicit computation

The first information constraint that we impose is the transition probability normalization constraint,

$$\sum_j \omega_{ij} = 1, \forall i, \tag{23}$$

that is,

$$\sum_j p_i \omega_{ij} = p_i, \forall i. \tag{24}$$

This constraint describes the fact that, from the state i at time t, the system has to transition to some state j at time $t + \delta t$. The second information constraint that we consider is the stationary state probability normalization constraint given by

$$\sum_{i,j} p_i \omega_{ij} = 1, \tag{25}$$

that is,

$$\sum_j p_j = 1. \tag{26}$$

The third constraint is the stationarity constraint described in terms of the following constraining relation:

$$\sum_i p_i \omega_{ij} = p_j, \forall j. \tag{27}$$

The constraint in Eq. (27) describes the fact that a system in state j at time $t + \delta t$ comes from one of the states i at time t. Before introducing the fourth constraint, we observe that within the MaxCal formalism a global constraint $\langle \sigma(t) \rangle$ is defined in terms of a path ensemble average of a dynamical quantity $\sigma(t)$ that depends on both the initial and final states i and j, respectively. Specifically [41,43],

$$\langle \sigma(t) \rangle \stackrel{\text{def}}{=} \sum_{\{C\}} p(C) \langle \sigma \rangle_C = \sum_{i,j} p_i \omega_{ij} \sigma_{ij}, \tag{28}$$

where $\langle\sigma\rangle_C$ denotes the average of σ over a steady state path $C \stackrel{\text{def}}{=} \cdots a \to b \to c \to d \cdots$ of length T,

$$\langle\sigma\rangle_C \stackrel{\text{def}}{=} \frac{1}{T}(\cdots\sigma_{ab} + \sigma_{bc} + \sigma_{cd} + \cdots). \tag{29}$$

We assume that the fourth constraint is described in terms of a path ensemble average of a dynamical quantity $\sigma = \sigma(t)$ specified by the product of two spin values $\sigma_i(t)$ and $\sigma_j(t)$. This constraint is given by

$$\langle\sigma(t)\rangle = \sum_{i,j} p_i\omega_{ij}\sigma_{ij}, \tag{30}$$

where we can generally define $\sigma_{ij}(t) \stackrel{\text{def}}{=} \sigma_i(t)\sigma_j(t)$. Since each spin σ_i is coupled to only its nearest-neighbor spins σ_{i-1} and σ_{i+1} in Glauber's stochastic Ising model, we make the following clarifying remark. For nearest-neighbor pairs of spins σ_i and σ_j with $j \in \{i \pm 1\}$ we consider $\sigma_{ij} = \sigma_i\sigma_{i+1}$, $\sigma_{ji} = \sigma_{i-1}\sigma_i$, and $\sigma_{ij} + \sigma_{ji} = \sigma_i(\sigma_{i-1} + \sigma_{i+1}) \propto \Delta E$, where ΔE is the energy difference between an old and a new microstate of the spin system as presented in Eq. (12). Finally, the fifth condition we impose is the detailed balance constraint given by

$$p_i\omega_{ij} = p_j\omega_{ji}. \tag{31}$$

Given the path entropy per unit time in Eq. (22) and the information constraints in Eqs. (24), (25), (27), (30), and (31), the caliber to be maximized becomes

$$\begin{aligned}\mathcal{C} = &-\sum_{i,j} p_i\omega_{ij}\ln[\omega_{ij}] + \sum_i \alpha_i\left(\sum_j p_i\omega_{ij} - p_i\right)\\ &+\beta\left(\sum_{i,j} p_i\omega_{ij} - 1\right) + \sum_j \gamma_j\left(\sum_i p_i\omega_{ij} - p_j\right)\\ &+\delta\left(\sum_{i,j} p_i\omega_{ij}\sigma_{ij} - \langle\sigma\rangle\right) + \sum_{i,j}\xi_{ij}\left(p_i\omega_{ij} - p_j\omega_{ji}\right),\end{aligned} \tag{32}$$

where α_i, β, γ_j, δ, and ξ_{ij} are Lagrange multipliers. Note that the variation of the caliber $\mathcal{C}$ with respect to the (unknown) stationary state probability p_i and the transition probability ω_{ij} is given by $\delta\mathcal{C}$,

$$\delta\mathcal{C} = \frac{\delta\mathcal{C}}{\delta\omega_{ij}}\delta\omega_{ij} + \frac{\delta\mathcal{C}}{\delta p_i}\delta p_i. \tag{33}$$

For the sake of completeness, we emphasize at this juncture that there are scenarios in which variations of functionals appear with respect to the Lagrange multipliers. For instance, in the presence of space-time-dependent information constraints in transport theory, the covariance functions $\mathcal{K}_{ij}$ are expressed in terms of the second functional derivative of $\ln\mathcal{Z}$ as follows [6]:

$$\mathcal{K}_{ij}(x,t;x',t') \stackrel{\text{def}}{=} \frac{\delta^2[\ln\mathcal{Z}]}{\delta\lambda_i(x,t)\delta\lambda_j(x',t')}, \tag{34}$$

where $\mathcal{Z} = \mathcal{Z}(\{\lambda_i(x,t)\})$ denotes the partition functional while λ_i are the Lagrange multipliers with $1 \leqslant i,j \leqslant m$ and m denoting the cardinality of the information constraints being chosen. Within the MaxCal formalism, Lagrange multipliers are generated by the first functional derivatives of the caliber while higher derivatives of the partition function lead to higher moments of the observables. The stationarity of $\delta\mathcal{C}$ in Eq. (33) requires that both $\frac{\delta\mathcal{C}}{\delta\omega_{ij}}$ and $\frac{\delta\mathcal{C}}{\delta p_i}$ must simultaneously vanish. Let us observe that, after some algebra, $\frac{\delta\mathcal{C}}{\delta\omega_{ij}}\delta\omega_{ij}$ is given by

$$\begin{aligned}\frac{\delta\mathcal{C}}{\delta\omega_{ij}}\delta\omega_{ij} = -\Bigg\{\sum_{i,j}[&p_i\ln\omega_{ij} + p_i - \alpha_i p_i - \beta p_i - \gamma_j p_i\\ &-\delta p_i\sigma_{ij} - p_i\xi_{ij} + p_i\xi_{ji}]\delta\omega_{ij}\Bigg\},\end{aligned} \tag{35}$$

that is,

$$\omega_{ij} = e^{\alpha_i+\beta+\gamma_j+\delta\sigma_{ij}+(\xi_{ij}-\xi_{ji})-1}. \tag{36}$$

Furthermore, let us notice that $\frac{\delta\mathcal{C}}{\delta p_i}\delta p_i$ becomes, after some algebra,

$$\begin{aligned}&\frac{\delta\mathcal{C}}{\delta p_i}\delta p_i\\ &= \begin{Bmatrix} -\sum_j\omega_{ij}\ln\omega_{ij} + \alpha_i\sum_j\omega_{ij} - \alpha_i + \beta\sum_j\omega_{ij}\\ +\sum_j\gamma_j\omega_{ij} - \gamma_j + \delta\sum_j\omega_{ij}\sigma_{ij} + \sum_j\omega_{ij}(\xi_{ij} - \xi_{ji})\end{Bmatrix}\\ &\times\delta p_i,\end{aligned} \tag{37}$$

that is,

$$\begin{aligned}&\sum_j\omega_{ij}\ln\omega_{ij}\\ &= +\alpha_i\sum_j\omega_{ij} - \alpha_i + \beta\sum_j\omega_{ij} + \sum_j\gamma_j\omega_{ij} - \gamma_j\\ &+\delta\sum_j\omega_{ij}\sigma_{ij} + \sum_j\omega_{ij}(\xi_{ij} - \xi_{ji}).\end{aligned} \tag{38}$$

Substituting Eq. (36) into Eq. (38), after some manipulations, we obtain the following relation between the Lagrange multipliers α_i and γ_i:

$$\alpha_i + \gamma_i = 1, \forall i. \tag{39}$$

For the sake of notational simplicity, let us relabel the Lagrange multipliers as follows:

$$A_i \stackrel{\text{def}}{=} e^{-\alpha_i}, \quad eA_j \stackrel{\text{def}}{=} e^{\gamma_j}, \quad B \stackrel{\text{def}}{=} e^{-\beta}, \text{ and } E_{ij} \stackrel{\text{def}}{=} e^{\xi_{ij}-\xi_{ji}}, \tag{40}$$

where e denotes the Neper constant. Then, using Eqs. (36) and (39), the transition probability ω_{ij} becomes

$$\omega_{ij} = \frac{1}{B}\frac{A_j}{A_i}e^{\delta\sigma_{ij}}E_{ij}. \tag{41}$$

The quantity E_{ij} can be obtained by imposing the detailed balance constraint in Eq. (31). After some algebra, it is found that

$$\frac{\omega_{ij}}{\omega_{ji}} = \frac{p_j}{p_i} = \left(\frac{A_j}{A_i}\right)^2 e^{\delta(\sigma_{ij}-\sigma_{ji})}E_{ij}^2, \tag{42}$$

that is,

$$E_{ij} = \sqrt{\frac{p_j}{p_i}}\frac{A_i}{A_j}e^{-\frac{\delta}{2}(\sigma_{ij}-\sigma_{ji})}. \tag{43}$$

Finally, combining Eqs. (41) and (43), we obtain

$$[\omega_{ij}]_{\text{MaxCal}} = \mathcal{N}\sqrt{\frac{p_j}{p_i}}e^{-\frac{\gamma}{2}(\sigma_{ij}+\sigma_{ji})}, \tag{44}$$

where $\mathcal{N}$ and γ are defined as

$$\mathcal{N} \stackrel{\text{def}}{=} \frac{1}{B} \text{ and } \gamma \stackrel{\text{def}}{=} -\delta, \tag{45}$$

respectively. The Lagrange multiplier $B = \mathcal{N}^{-1}$ can be obtained by imposing the normalization condition

$$\sum_j \omega_{ij} = 1, \tag{46}$$

that is,

$$\sum_j W_{ij}\phi_j = B\phi_i, \tag{47}$$

with $\phi_i \stackrel{\text{def}}{=} \sqrt{p_i}$ and $W_{ij} \stackrel{\text{def}}{=} e^{-\frac{\gamma}{2}(\sigma_{ij}+\sigma_{ji})}$. Furthermore, the Lagrange multiplier δ can be obtained by imposing the path ensemble average of the dynamical variable $\sigma(t)$ in Eq. (30). After some algebraic manipulations, we finally determine

$$[\omega_{ij}]_{\text{MaxCal}} = \mathcal{N}\sqrt{\frac{p_j}{p_i}}\cosh\left[\frac{\gamma}{2}(\sigma_{ij}+\sigma_{ji})\right] \times\left\{1-\tanh\left[\frac{\gamma}{2}(\sigma_{ij}+\sigma_{ji})\right]\right\}, \tag{48}$$

where $\sigma_{ij}(t) \stackrel{\text{def}}{=} \sigma_i(t)\sigma_j(t)$. The functional form of the transition probabilities in Eq. (48) is the one inferred by the MaxCal inference algorithm given the information constraints in Eqs. (24), (25), (27), (30), and (31).

B. Statistical mechanical remarks

In what follows, several statistical mechanical remarks are presented. First, within statistical mechanics, the common wisdom is that high temperatures lead to decay of correlations [55–59]. For instance, in his 1925 Ph.D. thesis, Ising showed that the spin-spin correlation function $\langle\sigma_k\sigma_l\rangle$ in the one-dimensional Ising model with Hamiltonian $\mathcal{H}$ in the absence of an external magnetic field decays exponentially with respect to the characteristic length ξ [48],

$$\langle\sigma_k\sigma_l\rangle \stackrel{\text{def}}{=} \frac{1}{\mathcal{Z}}\sum_{\{\sigma_i\}}\sigma_k\sigma_l e^{-\beta\mathcal{H}} = e^{-\frac{|l-k|}{\xi}}, \tag{49}$$

where $|l-k|$ denotes the distance between sites k and l while ξ is defined as

$$\xi = \xi(\beta) \stackrel{\text{def}}{=} \frac{1}{|\ln[\tanh(\beta J)]|}, \tag{50}$$

with J denoting the exchange coupling constant between spins. Observe that ξ diverges as βJ approaches infinity. For $T \neq 0$, correlations decay exponentially as a function of the correlation length. A short correlation length means that distant spins are very weakly correlated. At high temperatures, $\beta J \ll 1$, ξ becomes extremely short, and $\langle\sigma_k\sigma_l\rangle$ decays exponentially. To have a grasp of what high temperatures means, let us assume that the exchange coupling constant between spins is $J \approx 1$ eV and, recalling that the Boltzmann constant equals $k_B \approx 1.38 \times 10^{-23}$ J/K, an approximate estimate of the temperature yields $T \approx 1 \times 10^4$ K $= 10$ kK. This order of temperature corresponds to the Fermi boiling point for a valence electron, a temperature that for a metal is two orders of magnitude above room temperature. Second, within statistical mechanics, the asymptotic analysis of power-series expansions of expected values of observable quantities is very important [60]. In particular, it is the case that series expansions agree well with high accuracy Monte Carlo simulations, renormalization group results, and findings for exactly solvable models [61]. In the high-temperature series expansion, the Boltzmann factor is expanded in powers of the inverse temperature. For example, it can be shown that the asymptotic high-temperature approximation of the nearest-neighbor spin-spin correlation function $\langle\sigma_k\sigma_l\rangle$ in the two-dimensional Ising model in the absence of an external magnetic field is given by [60]

$$\langle\sigma_k\sigma_l\rangle = \beta + \tfrac{5}{3}\beta^3 + O(\beta^4). \tag{51}$$

In what follows, we exploit the consequences of these two remarks in our own discussion.

C. The formal comparison

In the working hypothesis of extremely high temperatures, we observe that Glauber's transition probability ω_{ij} is proportional to the square root of the ratio between the stationary state probabilities of state i and state j. Specifically, the first-order expansion in the parameter β of Glauber's transition probability ω_{ij} in Eq. (14) leads to the following approximate expression:

$$[\omega_{ij}]_{\text{Glauber}} \stackrel{\beta\ll 1}{\approx} \frac{1}{2}\Gamma e^{-\frac{\beta\Delta E}{2}} = \frac{1}{2}\Gamma\sqrt{\frac{p_j}{p_i}}, \tag{52}$$

where, using Eq. (4), we find

$$\frac{p_j}{p_i} = e^{-\beta\Delta E}. \tag{53}$$

Furthermore, if we relax this working assumption and consider instead high temperatures, we uncover that Glauber's transition probability ω_{ij} is no longer proportional to the square root of the ratio between the stationary state probabilities of state i and state j and in this regime a new approximate expression is obtained. Specifically, the second-order expansion in the parameter β of Glauber's transition probability ω_{ij} in Eq. (10) yields

$$[\omega_{ij}]_{\text{Glauber}} = \frac{1}{2}\Gamma e^{-\frac{\beta\Delta E}{2}}\left[1-\frac{1}{2}\left(\frac{\beta\Delta E}{2}\right)^2 + O(\beta^4)\right]. \tag{54}$$

In what follows, in addition to considering very high temperature values T, we also limit our analysis to energy difference values ΔE in Eq. (54) that belong to an interval of very small Lebesgue measure proportional to δE that is centered at an energy difference value ΔE_* that sets the energy scale of the stochastic Ising model being considered. Specifically, our approximate analysis proceeds in the following manner. We restrict our attention to energy difference values ΔE

with $\Delta E_* - \delta E \leqslant \Delta E \leqslant \Delta E_* + \delta E$ and $0 \leqslant \delta E \ll 1$. In this case, the linear approximation of ΔE^2 in the neighborhood of ΔE_* is given by

$$\Delta E^2 = \Delta E_*^2 + 2\Delta E_*(\Delta E - \Delta E_*) + O\big(|\Delta E - \Delta E_*|^2\big), \tag{55}$$

that is,

$$\Delta E^2 \approx 2\Delta E_* \Delta E - \Delta E_*^2. \tag{56}$$

Within this set of working hypotheses, combining Eqs. (54) and (56) and recalling from Eqs. (10) and (12) that $\Delta E \stackrel{\text{def}}{=} 2\sigma_i J(\sigma_{i-1} + \sigma_{i+1})$, we obtain the following *approximate* expression for Glauber's transition probability:

$$[\omega_{ij}]_{\text{Glauber}} = \frac{1}{2}\Gamma e^{-\frac{\beta\Delta E}{2}}\left[1 - \frac{1}{2}J\Delta E_*\beta^2\sigma_i(\sigma_{i-1} + \sigma_{i+1}) + \frac{\Delta E_*^2}{8}\beta^2 + O(\beta^4)\right]. \tag{57}$$

Using Eq. (53) and introducing the notations

$$\tilde{\Omega} \stackrel{\text{def}}{=} \frac{\Delta E_*^2}{8}\beta^2 \text{ and } \tilde{\gamma} \stackrel{\text{def}}{=} J\Delta E_*\beta^2, \tag{58}$$

we finally determine that the new approximation for Glauber's ω_{ij} becomes

$$[\omega_{ij}]_{\text{Glauber}} = \frac{1}{2}\Gamma\sqrt{\frac{p_j}{p_i}}\left[1 - \frac{1}{2}\tilde{\gamma}\sigma_i(\sigma_{i-1} + \sigma_{i+1}) + \tilde{\Omega} + O(\beta^4)\right], \tag{59}$$

that is,

$$[\omega_{ij}]_{\text{Glauber}} \stackrel{\beta^2 \ll 1}{\approx} \frac{1}{2}\Gamma\sqrt{\frac{p_j}{p_i}}\left[1 - \frac{1}{2}\tilde{\gamma}\sigma_i(\sigma_{i-1} + \sigma_{i+1}) + \tilde{\Omega}\right]. \tag{60}$$

We emphasize that the functional form obtained in Eq. (60) is not exact and it is only *approximately* valid in the limit of very high temperatures T and narrowly distributed energy changes ΔE in the spin system being considered. Finally, we remark that the parameter $\tilde{\gamma}$ in Eq. (60) is a quadratic function of the parameter β as evident from Eq. (58). The approximate expressions in Eqs. (52) and (60) will be compared with their analogs obtained within the MaxCal platform. First, observe that in the absence of a path-dependent dynamical information constraint, one sets the Lagrange multiplier γ in Eq. (48) equal to zero. In this case, the expression of the transition probability inferred by MaxCal reduces to

$$[\omega_{ij}]_{\text{MaxCal}} = \mathcal{N}\sqrt{\frac{p_j}{p_i}}. \tag{61}$$

Furthermore, if we assume as a working hypothesis that the Lagrange multiplier γ in Eq. (48) is nonvanishing but very small, that is, $0 \neq \gamma \ll 1$, considering the first order series expansion in γ of the transition probability ω_{ij} in Eq. (48), we obtain

$$[\omega_{ij}]_{\text{MaxCal}} = \mathcal{N}\sqrt{\frac{p_j}{p_i}}\left[1 - \frac{1}{2}\gamma(\sigma_{ij} + \sigma_{ji}) + O(\gamma^2)\right], \tag{62}$$

that is,

$$[\omega_{ij}]_{\text{MaxCal}} \stackrel{\gamma \ll 1}{\approx} \mathcal{N}\sqrt{\frac{p_j}{p_i}}\left[1 - \frac{1}{2}\gamma(\sigma_{ij} + \sigma_{ji})\right]. \tag{63}$$

On the one hand, upon comparison of Eqs. (52) and (61), we can essentially identify the normalization factors $\mathcal{N}$ and Γ and conclude that the inferred transition probability inferred by the MaxCal formalism in the absence of a path-dependent dynamical information constraint exhibits the same square-root dependence that appears in Glauber's approximate expression (first-order expansion in β, extremely high temperatures) of the transition probability in the limiting case of extremely high temperature values. On the other hand, comparison of Eqs. (60) and (63), in addition to identification of the normalization factors $\mathcal{N}$ and Γ (more specifically, $\mathcal{N} \leftrightarrow \frac{1+\tilde{\Omega}}{2}\Gamma$), allows one to link the quantity γ in Eq. (48) to the quantity $\tilde{\gamma}$ in Eq. (58) (more specifically, $\gamma \leftrightarrow \frac{1}{1+\tilde{\Omega}}\tilde{\gamma}$) and exploit the relation $\sigma_{ij} + \sigma_{ji} = \sigma_i(\sigma_{i-1} + \sigma_{i+1})$ in the working hypothesis of nearest-neighbor interactions of each spin σ_i with pairs of spins σ_j with $j \in \{i \pm 1\}$. We then uncover that the MaxCal formalism infers an approximate expression (first-order expansion in γ) of the transition probability that is functionally identical to the approximate expression (second-order expansion in β, high temperatures) obtained from Glauber's analysis. The fact that we have equated a relation obtained from a first-order expansion in γ [MaxCal, Eq. (63)] with a relation derived from a second-order expansion in β [Glauber, Eq. (60)] is remarkably consistent since we have identified the MaxCal Lagrange multiplier γ with $\tilde{\gamma} = \tilde{\gamma}(\beta) \propto \beta^2$ in Eq. (58).

We remark that the MaxCal algorithm allows us to make plausible inferences but not logical deductions. Such inferences rely on the nature of the chosen information constraints used in the algorithm. The validation of this type of modeling scheme can be checked only *a posteriori*. If discrepancies between the inferred predictions and experimental observations are recorded, a different set of information constraints has to be chosen. In our analysis, we do not recover the exact functional form of Glauber's transition probability. Our correspondence between the MaxCal and Glauber's solutions is only *approximately* valid in the limit of very high temperatures and narrowly distributed energy changes in the spin system being investigated. A more refined attempt to recover the exact expression would also require the clever introduction of some sort of information constraint that captures the *ad hoc* assumption employed by Glauber, that is, the symmetric dependence of the transition probability on the two neighboring spins σ_{i-1} and σ_{i+1} as well as σ_i.

V. CONCLUSIVE REMARKS

In this article, we employed the MaxCal variational principle as an inference algorithm used to predict dynamical properties of complex nonequilibrium stationary statistical systems in the presence of incomplete information. Specifically, we maximized the path entropy over discrete time step trajectories subject to normalization, stationarity, and detailed balance constraints together with a path-dependent dynamical information constraint reflecting a suitably chosen average global behavior of the complex system. Furthermore, we considered a path-dependent information constraint defined

in terms of the average of the product of two neighboring spin values $\sigma_i(t)$ and $\sigma_j(t)$ as specified in Eq. (30). A general expression for the transition probability associated to the stationary random Markov processes describing the nonequilibrium stationary system was computed and reported in Eq. (48),

$$[\omega_{ij}]_{\mathrm{MaxCal}} = \mathcal{N}\sqrt{\frac{p_j}{p_i}}\cosh\left[\frac{\gamma}{2}(\sigma_{ij}+\sigma_{ji})\right] \times\left\{1-\tanh\left[\frac{\gamma}{2}(\sigma_{ij}+\sigma_{ji})\right]\right\}. \quad (64)$$

The expression in Eq. (64) was then compared to the well-known Glauber hyperbolic tangent rule for the transition probability that characterizes the stochastic Ising model as recast by us in Eq. (14),

$$[\omega_{ij}]_{\mathrm{Glauber}} = \Gamma\frac{e^{-\frac{\beta\Delta E}{2}}}{e^{\frac{\beta\Delta E}{2}}+e^{-\frac{\beta\Delta E}{2}}}. \quad (65)$$

The comparison was presented in two limiting cases. In the first case, we uncovered that, in the absence of a path-dependent dynamical information constraint, ω_{ij} in Eq. (64) exhibits the same square-root dependence that appears in Glauber's approximate expression (first-order expansion in β, extremely high temperatures) of the transition probability ω_{ij} in Eq. (65) in the limiting case of extremely high temperature values. This first comparison was performed by considering Eqs. (52) and (61). In the second case, we uncovered that in the presence of a path-dependent dynamical information constraint the MaxCal formalism infers an approximate expression (first-order expansion in γ) of the transition probability ω_{ij} in Eq. (64) whose functional structure is identical to that of the approximate expression (second-order expansion in β, high temperatures) obtained from Glauber's ω_{ij} in Eq. (65). This second comparison was performed by considering Eqs. (60) and (63).

In summary, the main findings of our scientific activity reported in this manuscript can be outlined as follows:

(i) We established a quantitative link between the MaxCal formalism [5,6,40,41] and the stochastic Ising model as originally presented by Glauber [49]. Specifically, a connection between the transition probability inferred by MaxCal and Glauber's hyperbolic tangent rule is proposed in two limiting scenarios.

(ii) We advanced the line of research based on the MaxCal inference algorithm, especially the one advocated by Dixit [44], by extending the applicability of his use of the MaxCal formalism to a dynamical constraint in the form of a path ensemble average of the product of two neighboring spin values.

(iii) We significantly elaborated on the very intriguing preliminary remark concerning the relation between MaxCal and Glauber's dynamics as recently reported in Ref. [45]. Our elaboration led to an important advancement in the conceptual understanding of the above mentioned preliminary consideration.

We think our work presented herein is a valid addition to our continuing effort of providing a unifying theoretical framework of a statistical mechanical nature for describing and understanding complex systems of arbitrary nature in the presence of incomplete information [62]. As pointed out by Feynman in his Nobel lecture [63], using different mathematical approaches for describing the same physical result can provide a better starting point for subsequent reasoning. This statement seems to be especially well suited for the connection between the MaxCal formalism and the conventional nonequilibrium statistical mechanical thinking. MaxCal could potentially be used to understand what is not yet understood. In light of these considerations, it is our sincere hope that other scientists will find our work presented here relevant and worthy of further refinement.

ACKNOWLEDGMENTS

C.C. is grateful to Adom Giffin and Domenico Felice for discussions on the MaxCal formalism and the Glauber dynamics, respectively. Constructive criticism from two anonymous referees leading to an improved version of this manuscript are sincerely acknowledged by the authors. Finally, C.C. acknowledges the hospitality of the Air Force Research Laboratory (AFRL) where part of his contribution to this work was completed.

[1] J. W. Gibbs, On the equilibrium of heterogeneous substances, *Transactions of the Connecticut Academy of Arts and Sciences*, Vol. 3 (1874–78), pp. 108–248 and 343–524.

[2] C. E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. **27**, 379 (1948).

[3] E. T. Jaynes, Information theory and statistical mechanics, Phys. Rev. **106**, 620 (1957).

[4] E. T. Jaynes, Information theory and statistical mechanics. II, Phys. Rev. **108**, 171 (1957).

[5] E. T. Jaynes, The minimum entropy production principle, Annu. Rev. Phys. Chem. **31**, 579 (1980).

[6] E. T. Jaynes, Macroscopic prediction, in *Complex Systems: Operational Approaches in Neurobiology, Physics, and Computers*, edited by H. Haken (Springer-Verlag, Berlin, 1985), p. 254.

[7] E. T. Jaynes, Predictive statistical mechanics, in *Frontiers of Nonequilibrium Statistical Physics*, edited by G. T. Moore and M. O. Scully (Plenum Press, New York, 1986), p. 33.

[8] S. Chibbaro, L. Rondoni, and A. Vulpiani, On the foundations of statistical mechanics: Ergodicity, many degrees of freedom and inference, Commun. Theor. Phys. **62**, 469 (2014).

[9] R. C. Dewar, Maximum entropy production as an inference algorithm that translates physical assumptions into macroscopic predictions: Don't shoot the messenger, Entropy **11**, 931 (2009).

[10] S. F. Gull and G. J. Daniell, Image reconstruction from incomplete and noisy data, Nature (London) **272**, 686 (1978).

[11] J. Skilling and R. K. Bryan, Maximum entropy image reconstruction: General algorithm, Mon. Not. R. Astron. Soc. **211**, 111 (1984).

[12] G. Minerbo, A maximum entropy algorithm for reconstructing a source from projected data, Comput. Graphics Image Process. **10**, 48 (1979).
[13] D. M. Collins, Electron density images from imperfect data by iterative entropy maximization, Nature (London) **298**, 49 (1982).
[14] R. K. Bryan, M. Bansal, W. Folkhard, C. Nave, and D. A. Marvin, Maximum-entropy calculation of the electron density at 4 A resolution of Pfl filamentous bacteriophage, Proc. Natl. Acad. Sci. USA **80**, 4728 (1983).
[15] S. Sibisi, J. Skilling, R. G. Brereton, E. D. Lane, and J. Staunton, Maximum entropy signal processing in practical NMR spectroscopy, Nature (London) **311**, 446 (1984).
[16] T. Mora, S. Deny, and O. Marre, Dynamical Criticality in the Collective Activity of a Population of Retinal Neurons, Phys. Rev. Lett. **114**, 078105 (2015).
[17] C. Cafaro and S. A. Ali, Jacobi fields on statistical manifolds of negative curvature, Physica D **234**, 70 (2007).
[18] C. Cafaro, Works on an information geometrodynamical approach to chaos, Chaos, Solitons Fractals **41**, 886 (2009).
[19] C. Cafaro and S. Mancini, Quantifying the complexity of geodesic paths on curved statistical manifolds through information geometric entropies and Jacobi fields, Physica D **240**, 607 (2011).
[20] G. Henry and D. Rodriguez, On the instability of two entropic dynamical models, Chaos, Solitons Fractals **91**, 604 (2016).
[21] C. Cafaro, Information geometry, inference methods and chaotic energy levels statistics, Mod. Phys. Lett. B **22**, 1879 (2008).
[22] C. Cafaro and S. A. Ali, Can chaotic quantum energy levels statistics be characterized using information geometry and inference methods? Physica A **387**, 6876 (2008).
[23] D.-H. Kim, S. A. Ali, C. Cafaro, and S. Mancini, Information geometric modeling of scattering induced quantum entanglement, Phys. Lett. A **375**, 2868 (2011).
[24] D.-H. Kim, S. A. Ali, C. Cafaro, and S. Mancini, Information geometry of quantum entangled wave-packets, Physica A **391**, 4517 (2012).
[25] A. Giffin, C. Cafaro, and S. A. Ali, Application of the maximum relative entropy method to the physics of ferromagnetic materials, Physica A **455**, 11 (2016).
[26] K. Ghosh, K. A. Dill, M. M. Inamdar, E. Seitaridou, and R. Phillips, Teaching the principles of statistical dynamics, Am. J. Phys. **74**, 123 (2006).
[27] E. Seitaridou, M. M. Inamdar, R. Phillips, K. Ghosh, and K. Dill, Measuring flux distributions for diffusion in the small-numbers limit, J. Phys. Chem. B **111**, 2288 (2007).
[28] D. Wu, K. Ghosh, M. Inamdar, H. L. Lee, S. Fraser, K. Dill, and R. Phillips, Trajectory Approach to Two-State Kinetics of Single Particles on Sculpted Energy Landscapes, Phys. Rev. Lett. **103**, 050603 (2009).
[29] A. Cavagna, I. Giardina, F. Ginelli, T. Mora, S. Piovani, R. Tavarone, and A. M. Walczak, Dynamical maximum entropy approach to flocking, Phys. Rev. E **89**, 042707 (2014).
[30] J. T. Dougherty, Foundations of non-equilibrium statistical mechanics, Philos. Trans. R. Soc. London A **346**, 259 (1994).
[31] R. Dewar, Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states, J. Phys. A: Math. Gen. **36**, 631 (2003).
[32] D. Gonzalez, S. Davis, and G. Gutierrez, Newtonian dynamics from the principle of maximum caliber, Found. Phys. **44**, 923 (2014).
[33] S. Davis and D. Gonzalez, Hamiltonian formalism and path entropy maximization, J. Phys. A: Math. Theor. **48**, 425003 (2015).
[34] G. Stock, K. Ghosh, and K. A. Dill, Maximum caliber: A variational approach applied to two-state dynamics, J. Chem. Phys. **128**, 194102 (2008).
[35] J. Lee and S. Presse, A derivation of the master equation from path entropy maximization, J. Chem. Phys. **137**, 074103 (2012).
[36] H. Ge, S. Presse, K. Ghosh, and K. A. Dill, Markov processes follow the principle of maximum caliber, J. Chem. Phys. **136**, 064108 (2012).
[37] M. Otten and G. Stock, Maximum caliber inference of nonequilibrium processes, J. Chem. Phys. **133**, 034119 (2010).
[38] M. J. Hazoglou, V. Walther, P. D. Dixit, and K. A. Dill, Maximum caliber is a general variational principle for nonequilibrium statistical mechanics, J. Chem. Phys. **143**, 051104 (2015).
[39] S. Presse, K. Ghosh, J. Lee, and K. A. Dill, Principles of maximum entropy and maximum caliber in statistical physics, Rev. Mod. Phys. **85**, 1115 (2013).
[40] A. A. Filyukov and V. Y. Karpov, Method of the most probable path of evolution in the theory of stationary irreversible processes, J. Eng. Phys. **13**, 416 (1967).
[41] A. A. Filyukov and V. Y. Karpov, Description of steady transport processes by the method of the most probable path of evolution, J. Eng. Phys. **13**, 326 (1967).
[42] A. A. Filyukov, Compatibility property of steady systems, J. Eng. Phys. **14**, 429 (1968).
[43] P. D. Dixit and K. A. Dill, Inferring microscopic kinetic rates from stationary state distributions, J. Chem. Theory Comput. **10**, 3002 (2014).
[44] P. D. Dixit, Stationary properties of maximum entropy random walks, Phys. Rev. E **92**, 042149 (2015).
[45] P. D. Dixit, A. Jain, G. Stock, and K. A. Dill, Inferring transition rates on networks with incomplete knowledge, arXiv:1504.01277.
[46] K. Huang, *Statistical Mechanics* (John Wiley and Sons, New York, 1987).
[47] H. W. Huang, Time-dependent statistics of the Ising model in a magnetic field, Phys. Rev. A **8**, 2553 (1973).
[48] E. Ising, Beitrag zur theorie des ferromagnetismus, Z. Phys. D **31**, 253 (1925).
[49] R. J. Glauber, Time-dependent statistics of the Ising model, J. Math. Phys. **4**, 294 (1963).
[50] O. Penrose, A mean-field equation of motion for the dynamic Ising model, J. Stat. Phys. **63**, 975 (1991).
[51] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, U.K., 2005).
[52] W. Janke, Monte Carlo simulations in statistical physics: From basic principles to advanced applications, in *Order, Disorder and Criticality: Advanced Problems of Phase Transition Theory*, edited by Yu. Holovatch (World Scientific, Singapore, 2012), Vol. 3, pp. 93–166.
[53] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North Holland, Amsterdam, 1981).

[54] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (John Wiley and Sons, New York, 2006).

[55] L. Gross, Decay of correlations in classical lattice models at high temperature, Commun. Math. Phys. **68**, 9 (1979).

[56] C. Cammarota, Decay of correlations for infinite range interactions in unbounded spin systems, Commun. Math. Phys. **85**, 517 (1982).

[57] M. Lehto, H. B. Nielsen, and M. Ninomiya, A correlation decay theorem at high temperature, Commun. Math. Phys. **93**, 483 (1984).

[58] T. Kennedy and E. H. Lieb, An itinerant electron model with crystalline or magnetic long range order, Physica A **138**, 320 (1986).

[59] J. Bricmont and A. Kupiainen, High temperature expansions and dynamical systems, Commun. Math. Phys. **178**, 703 (1996).

[60] A. J. Guttmann, Asymptotic analysis of power-series expansions, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and J. L. Lebowitz (Academic Press, London, 1989), Vol. 3, pp. 1–234.

[61] A. Wipf, High-temperature and low-temperature expansions, Lect. Notes Phys. **100**, 173 (2013).

[62] C. Cafaro, S. A. Ali, and A. Giffin, Thermodynamic aspects of information transfer in complex dynamical systems, Phys. Rev. E **93**, 022114 (2016).

[63] R. P. Feynman, The development of the space-time view of quantum electrodynamics, in *Nobel Lectures, Physics 1963-1970* (Elsevier, Amsterdam, 1972).