Title
The Maximum Caliber Variational Principle for Nonequilibria

Journal
Annual Review of Physical Chemistry, 71(1)

ISSN
0066-426X

Authors
Ghosh, Kingshuk
Dixit, Purushottam D
Agozzino, Luca
et al.

Publication Date
2020-04-20

DOI
10.1146/annurev-physchem-071119-040206

Peer reviewed

HHS Public Access

Author manuscript

Annu Rev Phys Chem. Author manuscript; available in PMC 2023 January 09.

Published in final edited form as:

Annu Rev Phys Chem. 2020 April 20; 71: 213–238. doi:10.1146/annurev-physchem-071119-040206.

The Maximum Caliber Variational Principle for Nonequilibria

Kingshuk Ghosh[1,*], **Purushottam D. Dixit**[2,3,*], **Luca Agozzino**[4], **Ken A. Dill**[4]

[1]Department of Physics and Astronomy, University of Denver, Denver, Colorado 80209, USA

[2]Department of Systems Biology, Columbia University, New York, NY 10032, USA

[3]Department of Physics, University of Florida, Gainesville, Florida 32611, USA

[4]Laufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook, New York 11794, USA

Abstract

Ever since Clausius in 1865 and Boltzmann in 1877, the concepts of entropy and of its maximization have been the foundations for predicting how material equilibria derive from microscopic properties. But, despite much work, there has been no equally satisfactory general variational principle for nonequilibrium situations. However, in 1980, a new avenue was opened by E.T. Jaynes and by Shore and Johnson. We review here maximum caliber, which is a maximum-entropy-like principle that can infer distributions of flows over pathways, given dynamical constraints. This approach is providing new insights, particularly into few-particle complex systems, such as gene circuits, protein conformational reaction coordinates, network traffic, bird flocking, cell motility, and neuronal firing.

Keywords

nonequilibrium; entropy; caliber; inference; network models; biological dynamics; variational principles

1. RECENT DIRECTIONS IN NONEQUILIBRIUM STATISTICAL PHYSICS

Nonequilibrium physics (NEP) is concerned with flows—often of matter, heat, or electrical current. Models often treat the macroscale, where the numbers of molecules or particles are large enough to be represented as continuous and differentiable functions of space **x** and time *t*, such as densities or concentrations $c(\mathbf{x}, t)$, and where fluctuations are small. Macroscopic flow models include Navier–Stokes hydrodynamics, Ohm's law of electrical currents, and gradient-driven flows of particles by Fick's law or of heat by Fourier's law.

kghosh@du.edu .

*These authors contributed equally to this article

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

But three factors are driving new results. First, the field is expanding beyond thermal materials: to the flows of information through the internet (1), to the flows of people moving between cities (2), to the flows of assets moving through stock markets (3), to the flocking of birds (4), to citations of scientific papers (5), to the trafficking and signaling of biochemicals and proteins or regulation inside cellular networks, to the evolutionary dynamics of changing genes and proteins in cells (6–9), to neural signals in the brain (10), to biological evolution and development, and to other applications. These situations do not involve the thermodynamic baths, physical work, or colliding particles that have historically been the core of NEP.

Second, there is growing interest in the microscale, thanks to the emergence of single-particle or few-particle experimental methods. These are cases of nonequilibrium statistical physics, where fluctuations and distributions of rates and routes are important.

Third, there is a continuing quest to find a general variational principle that underlies nonequilibrium statistical physics just as the second law of thermodynamics and the Boltzmann distribution underlie the statistical physics of matter at equilibrium.

Figure 1 gives an abbreviated history. Through the mid-1800s, there arose the phenomenological models for Newtonian viscous fluids, Fick's law of particle flow, Fourier's law of heat flow, and Ohm's law of electrical currents. The industrial revolution drove an understanding of work and heat in steam engines. A variational principle emerged, namely the second law of thermodynamics, following the discovery by Clausius around 1865 that equilibria could be predicted as a tendency toward the maximum of a quantity he called entropy, $\Delta S_{\text{Clausius}} = \delta q/T$, for heat q and temperature T. Soon thereafter followed the kinetic theory of gases, statistical thermodynamics, and the Boltzmann–Gibbs distribution law. These added great power and nuance to the second law by (*a*) harnessing models that could interpret macroscopic equilibria through the microscopic properties of molecules and materials and (*b*) rooting our understanding of the second law maximization of entropy in probability distributions over microscopic states. However, a key caveat was that these principles were restricted to equilibrium, or close to it.

There was a corresponding quest for a variational principle for nonequilibria. The question was whether dissipation rates of energy ($\mathrm{d}U/\mathrm{d}t$) or entropy ($\mathrm{d}S/\mathrm{d}t$) might tend toward minima or maxima. Examples include Onsager's minimum dissipation of energy (11) and Prigogine's minimum entropy production principle (12) and maximum entropy production principle (13). However, these putative principles have never been fully satisfactory, because they are restricted to near equilibria, requiring assumptions such as local equilibria or weak coupling to baths, and are mainly applicable at the macroscopic level, typically assuming—rather than predicting—the phenomenological relations. A major obstacle on the path to a general nonequilibrium statistical physics variational principle has been the lack of an experimental relationship serving the role that $\Delta S_{\text{Clausius}} = \delta q/T$ plays for equilibria.

Today, there are many models of flows—diffusion equations, the Langevin model, master equations, random flights, Boltzmann transport, and others. Why do we need an underlying variational principle? Current models entail simplifications—independence of particles

or time steps, near equilibria, linear force-flow relations, Gaussian noise, large particle numbers, or a focus on heat baths and particle collisions, for example. A general variational principle is needed as a road map for more challenging nonlinear few-particle complex dynamics, where typical simplifications fail, and for applications beyond thermal physics. We review here evidence that maximum caliber (Max Cal) may be such a principle.

2. MAXIMUM ENTROPY AND THE BOLTZMANN DISTRIBUTION: A MICROSCOPIC BASIS FOR THE SECOND LAW PRINCIPLE

Boltzmann's famous expression $S = k \ln W$ was an assertion that the Clausius macroscopic principle had its roots in the number W of possible microscopic arrangements of the system, establishing that Boltzmann's exponential distribution law was the microscopic manifestation of the second law variational principle. In more detail, for any probability distribution $\{p_i\} = p_1, p_2, p_3, \ldots$ over options $i = 1, 2, 3, \ldots$, we can define the mathematical entropy of that distribution to be

$$S_{\text{math}} = -\sum_i p_i \ln p_i . \quad 1.$$

This quantity can be computed for any distribution. But this mathematical entropy S_{math} is not the same thing as the physical entropy of Clausius, S_{Clausius}. S_{math} is also not what we need for making theoretical models of physical equilibria; for that, we need S_{state}, which we now describe. First, take the distribution p_i to be over microscopic states of the system. Next, we assert that only the one specific distribution $\{p_i = p_i^*\}$ that maximizes the entropy,

$$S_{\text{state}} = -k_{\text{B}} \sum_{i = \text{ states}} p_i^* \ln p_i^*, \quad 2.$$

is relevant to second law predictions of physical behaviors at equilibrium. In Equation 2, k_{B} is the Boltzmann constant. S_{state} is defined for one particular distribution, not for just any mathematical distribution. Now, for the canonical ensemble—a system is in contact with a bath that fixes the average energy of a system—the prediction procedure is to maximize S_{math} over $\{p_i\}$, subject to a constraint of average energy

$$\sum_i p_i E_i = \bar{E}, \quad 3.$$

where E_i is the energy of microscopic state i and $\bar{E}$ is the average energy of the system. The probabilities are normalized quantities, $\sum_i p_i = 1$. The result is the Gibbs–Boltzmann distribution,

$$p_i^* \propto e^{-E_i / k_{\text{B}} T}, \quad 4.$$

where p_i^* are the probabilities that satisfy these conditions. In Equation 4, T is the temperature. This distribution is at the core of equilibrium statistical physics. Equation 4 predicts the equilibrium populations of all the states $i = 1, 2, 3, \ldots$ in a model.

The lack of care in distinguishing these different meanings of entropy has been a source of confusion. S_{Clausius} is a predictor only of macroscopic equilibrium thermodynamics: that heat tends from hot bodies to cold ones, that particles tend toward places of lower concentrations, and that densities tend to equalize in materials, for example. It does not tell us about the microscale or distribution functions. S_{state} is a quantity we compute from microscopic physical models of equilibria. To relate an equilibrium model to corresponding macroscopic experiments, we equate $S_{\text{state}} = S_{\text{Clausius}}$. Below, we describe yet another entropy that is pertinent to dynamics, namely the path entropy, S_{path}.

3. MAXIMUM CALIBER IS A VARIATIONAL PRINCIPLE FOR DYNAMICAL PROCESSES

Forty years ago, here in the *Annual Review of Physical Chemistry*, E.T. Jaynes (14) introduced the principle of maximum caliber. Here, we review its status as a general principle for nonequilibria as well as some applications. It differs from alternatives in (*a*) its basis in particle trajectories, not concentrations; (*b*) its maximization of path entropies, not state entropies; (*c*) its inferences of microscopic models that have fewer confounding logical traps from experimental data constraints; and (*d*) its axiomatic foundations in probability theory (15).

Here is a brief overview. Max Cal is to dynamics what maximum entropy is to equilibrium. Whereas maximum entropy focuses on probabilities of states, Max Cal focuses on probabilities of pathways or trajectories. Let $\{\Gamma\}$ be the set of all possible trajectories of a system that is evolving in time. Γ can represent several types of dynamical processes. For example, consider a system evolving from an initial state at time T_i into a final state at time T_f (see Figure 2) or trajectories of a system that is at steady state. For the former, the individual trajectories are given by $\Gamma = \left\{x_{T_i}, x_{T_i+1}, \ldots, x_{T_f}\right\}$, which the system can take between time points T_i and T_f. Other types of trajectories are discussed below. Let P_Γ be the probability distribution defined over the ensemble $\{\Gamma\}$ of paths.

Let $J(\Gamma)$ be a functional defined on the space of paths. Examples of J include the flux of mass/heat carried by the path, the average dissipation along a path, or an average energy along a path. Suppose we want to infer the distribution P_Γ over the paths while constraining the average

$$\left\langle J \right\rangle = \sum_\Gamma P_\Gamma J(\Gamma)\,. \tag{5.}$$

Here, and in the remainder of this review, we use the uppercase P specifically to indicate probabilities over paths, whereas we use the lowercase p to indicate generic probabilities, and we will specify the meaning of the latter when we use it.

Note that there are potentially infinitely many probability distributions P_Γ that are consistent with such constraints. Analogous to the equilibrium situation, we maximize the entropy, but now over all possible paths, not states:

$$S_{\text{path}} = -\sum_{\Gamma} P_{\Gamma} \log \frac{P_{\Gamma}}{q_{\Gamma}}, \quad 6.$$

Subject to the constraint in Equation 5 and normalization (see Figure 2). Note that q_Γ is a reference or prior distribution over paths.

This constrained maximization problem is solved by introducing Lagrange multipliers. We write the unconstrained optimization function, called the caliber C,

$$C = -\sum_{\Gamma} P_{\Gamma} \log \frac{P_{\Gamma}}{q_{\Gamma}} - \gamma \left(\sum_{\Gamma} P_{\Gamma} J(\Gamma) - \left\langle J \right\rangle \right) + \alpha \left(\sum_{\Gamma} P_{\Gamma} - 1 \right). \quad 7.$$

In Equation 7, γ is a Lagrange multiplier that fixes the ensemble average $\langle J \rangle$ to the given value, and α ensures normalization. After maximization, we find

$$P_{\Gamma} = \frac{q_{\Gamma} e^{-\gamma J(\Gamma)}}{Q_d}, \quad 8.$$

where

$$Q_d = \sum_{\Gamma} q_{\Gamma} e^{-\gamma J(\Gamma)},$$

a sum of weights over paths, is the dynamical equivalent of a partition function. A key result of Max Cal is the relationship between measurable average rate quantities and the model dynamical partition function,

$$\left\langle J \right\rangle = \sum_{\Gamma} J(\Gamma) P_{\Gamma} \quad 10.$$

$$= -\frac{\partial \log Q_d}{\partial \gamma}. \quad 11.$$

In practice, Max Cal works as follows. First, one asserts what trajectories are relevant to the problem at hand. Next, on the basis of relevant constraints, one constructs a distribution over the path space (Equation 9). Each trajectory is expressed in terms of the statistical weights of the steps it takes. Using Equation 9, one can make predictions analogous to equilibrium statistical mechanics. For example, all trajectories can be summed into a dynamical partition function Q_d. Then, Equation 11 can be used to compute all the statistical weights and pathway probabilities that are consistent with the given value of $\langle J \rangle$ and other constraints if relevant. To illustrate its machinery, we provide several examples of Max Cal below.

4. MAXIMUM CALIBER PROVIDES NEW INSIGHTS

4.1. Two-State Dynamics: Predicting Route Distributions from Average Rates

Consider a single colloidal particle that can hop between two energy wells, A and B (see Figure 3). This system was studied experimentally by Phillips and colleagues (16) using two-laser traps to establish the two states. Figure 3 shows examples of the trajectories of the particle that the authors observed under different trap conditions. The dynamics is modeled using Max Cal and assuming a two-state Markov model. The figure shows the enumeration over all the paths of the statistical weights of each path. The statistical weights for what happens at each time step are u for an up-jump, a or b for staying in state A or B, respectively, and d for a down-jump. Only two of these are independent, because the weights of all the jump-to states must sum to one, and the weights of all the jump-from states must sum to one as well.

The Max Cal approach is to extract from the trajectory two of the independent quantities, say u and d, and substitute them into the Max Cal expression shown in Figure 3. With these two quantities now determined, the full path probability distribution is obtained. For example, Max Cal can predict the variance in the rate distribution from the averages. The variances agree with the values determined in the experiment.

The same approach—combining Markovian dynamics with Max Cal—has been applied to a single-molecule three-state cycle, $A \leftrightarrow B \leftrightarrow C \leftrightarrow A$ (17). It shows that the noise diminishes faster with the increase in the average motor spin rate. It also illustrates how trajectory-based modeling can capture all trajectories, not just those near equilibrium, and thus applies far from equilibrium.

4.2. Deriving a Phenomenological Law from a Variational Principle Beyond the Fick's Law Average

Here we show how a phenomenological law—Fick's law of diffusion, $\langle \mathbf{J} \rangle = -D\partial c/\partial \mathbf{x}$—can be derived from a simple microscopic model, to illustrate Max Cal as a variational principle. Diffusion-equation models in NEP usually assume that particles are sufficiently numerous that their densities or concentrations can be expressed as continuous and differentiable functions (see Figure 4). But recent colloidal experiments have explored small-numbers diffusion (see Figure 5). There are many small-numbers properties that are of interest, and useful for model building, such as the width $\langle J^2 \rangle$, of the distribution. They are readily computed using the trajectory approach of Max Cal with a model, in this case called the dog-flea model (different spatial points are represented as dogs, and the particles at those points are called fleas because they can hop between dogs).

Figure 6a shows a concentration gradient $c(x)$ along a spatial coordinate x. There are N_i particles at position x_i. In the language of the dog-flea model (18), N_i is the number of fleas on a dog located at position x_i. In a time interval Δt, r fleas jump off the dog to the right and ℓ fleas jump to the left. We are given two constraints: on average, $\langle r \rangle$ jump right and $\langle \ell \rangle$ jump left. What is the probability P_i of having specifically r and ℓ jumps of the respective types? We compute P_i by maximizing the path entropy subject to the known average rates, which are imposed using Lagrange multipliers λ_r and λ_ℓ

$$P_i = Q_i^{-1}\exp(\lambda_r r + \lambda_\ell \ell) = \sigma_r^r \sigma_\ell^\ell / Q_i, \tag{12.}$$

where we simplify by $\sigma = \exp \lambda$ and where

$$Q_i = \sum_{r,\ \ell} \Omega_i(r,\ \ell)\exp(\lambda_r r + \lambda_\ell \ell) = (1 + \sigma_\ell + \sigma_r)^{N_i} \tag{13.}$$

is the dynamical partition function. The latter comes from taking the degeneracy as $\Omega_i(r, \ell) = N_i!/[N_i - r - \ell!\, r!\, \ell!]$, assuming independence of particles.

The average flux to the right, $\langle r \rangle$, is given by

$$\langle r \rangle = \frac{\partial \ln Q_i}{\partial \lambda_r} = \frac{N_i \sigma_r}{1 + \sigma_r + \sigma_\ell}. \tag{14.}$$

Similarly, the flux to the left from the next column, having N_{i+1} particles, is given by

$$\langle \ell \rangle = \frac{\partial \ln Q_{i+1}}{\partial \lambda_\ell} = \frac{N_{i+1} \sigma_\ell}{1 + \sigma_r + \sigma_\ell}. \tag{15.}$$

Now we assume symmetry between left and right (i.e., assuming no drift), yielding $\lambda_r = \lambda_\ell$. We define $p = \sigma_r/(1 + \sigma_r + \sigma_\ell)$ and get the net flux to the right as $\langle J \rangle = \langle r \rangle - \langle \ell \rangle = (N_i - N_{i+1})p = \Delta N p$. We convert this number to a flux (i.e., the number of particles that jump per unit time per unit area) and replace numbers with concentrations using $\Delta N = -A\Delta c \Delta x$, where $A\Delta x$ is the volume containing the particles. The minus sign indicates that we are considering the forward direction to be from i to $i + 1$,while the gradient $\Delta c = c_{i+1} - c_i$ is defined in the opposite direction. This gives

$$\langle J \rangle = \frac{\langle j_i \rangle - \langle j_{i+1} \rangle}{A\Delta t} = \frac{p\Delta N}{A\Delta t} = -\frac{p\Delta x^2}{\Delta t}\left(\frac{\Delta c}{\Delta x}\right). \tag{16.}$$

Equation 16 is Fick's law as it pertains to just two adjacent planes of particles. This model gives the diffusion coefficient as $D = (p\Delta x^2)/(\Delta t)$. It explains that the reason the average flow of particles is proportional to the difference in concentrations is because there are more routes of flow of particles from more concentrated regions.

Although there are other ways to derive Fick's law, the main points here are that (*a*) this Max Cal approach derives it from an underlying general variational principle; (*b*) it gives not only the average flux, $\langle J \rangle$ (Equation 16), but also the full distribution P_J of rates (see Figure 7); and (*c*) the Max Cal approach is didactically simple.

4.3. Deriving Kirchoff's Current Law: How Does Current Divide at a Junction with Two Resistors?

Kirchoff's current relationship states that flows of particles or fluids divide at junctions in proportion to the inverse resistance beyond the junction. We summarize Jaynes's (14)

argument that Max Cal derives this principle correctly, while a putative principle of minimum entropy production rate does not. While Kirchoff's principle applies to any type of flow, for simplicity, we illustrate it in Figure 8 with two resistors in parallel, R_ℓ and R_r. The current through a resistor is given in Ohm's law by $i = V_0/R$, where V_0 is the applied voltage and R is the resistance. We assume that the resistors are connected to thermal baths of temperature T_ℓ and T_r, respectively. We ask the question, how does the total current $I = I_\ell + I_r$ split in I_ℓ through the left resistor R_ℓ and I_r through the right resistor R_r when the two resistors are in parallel?

In the minimum entropy production principle, the current split would be determined by extremizing the thermal entropy production, $\mathrm{d}S/\mathrm{d}t = \dot{S}$, with respect to how the current splits through the two paths. We write the entropy production as

$$\dot{S} = \frac{J_\ell}{T_\ell} + \frac{J_r}{T_r} = \frac{V_0 I_\ell}{T_\ell} + \frac{V_0 I_r}{T_r} = \frac{I_\ell R_\ell^2}{T_\ell} + \frac{I_r R_r^2}{T_r}, \tag{17.}$$

where J_ℓ and J_r are the rate of heat flow across the left and right resistors, respectively.

In Equation 17, the dissipation in a resistor has the usual form: force × flux ($V_0 \times I$). We have also assumed Ohm's law: $V = IR$. Minimizing $\dot{S}$ with respect to I_ℓ and I_r subject to the constraint $I = I_\ell + I_r$, we get

$$\frac{I_\ell R_\ell}{T_\ell} = \frac{I_r R_r}{T_r}, \tag{18.}$$

which is clearly wrong because the only thing that determines the current in a resistor at voltage V_0 is R; the temperature T should not appear. Of course, the resistance could depend on temperature, $R = R(T)$, but all that matters for Kirchoff's law is the resistance.

As above, we combine the dog-flea model with Max Cal. Assume there are N electrons (or fleas), of which ℓ can jump through the left branch and r can jump through the right branch in a given time interval Δt. We impose $\langle \ell \rangle$ and $\langle r \rangle$ as two constraints. These constraints basically relate to properties of the two branches. Imposing these constraints with two Lagrange multipliers λ_ℓ and λ_r, the caliber is

$$C = \sum \left[P_j \ln P_j - \lambda_\ell \, \ell \, P_j - \lambda_r r P_j \right], \tag{19.}$$

where P_j is a microtrajectory defined by a particular partitioning of ℓ and r (the normalization condition $\sum P_j = 1$ is assumed implicitly). Maximizing the caliber, we get $P_j \propto \exp(-\lambda_\ell \ell - \lambda_r r)$. Next, converting the microtrajectories in terms of macrotrajectories $[P_M(\ell, r)]$ defined by only and r, we get a combinatorial factor $N!/[\ell! r! (N - \ell - r)!]$. Using this, we get a multinomial distribution for $P_M(\ell, r)$, which yields $\langle \ell \rangle = NP_\ell$ and $\langle r \rangle = NP_r$, where $P_\ell = \exp(-\lambda_\ell)/(1 + \exp(-\lambda_\ell) + \exp(-\lambda_r))$ and $P_r = \exp(-\lambda_r)/(1 + \exp(-\lambda_\ell) + \exp(-\lambda_r))$ From this, we notice that

$$\frac{\langle \ell \rangle}{\langle r \rangle} = \frac{P_\ell}{P_r} \Rightarrow \langle \ell \rangle = \text{ constant } \times P_\ell . \quad 20.$$

Here is the connection with Ohm's law. First, Ohm's law currents are macroscopic averaged values, so $i_\ell / i_r = \langle \ell \rangle / \langle r \rangle$ in the model. Second, our intrinsic quantities correspond to resistances in Ohm's law, $P_\ell / P_r = R / R_\ell$. Hence, we recover Kirchoff's current law—namely, that the average flows partition depending on the ratio of resistances. This shows first how Max Cal uses only constraints and does not erroneously introduce extraneous factors, like temperature, as the minimum entropy production principle does. Second, unlike Kirchoff's law, which applies only to average fluxes, Max Cal plus the dog-flea model also gives the full rate distribution of the flows at junctions.

4.4. Maximum Caliber Models Few-Particle Complex Systems and Gene Circuits

Chemical reactions, biochemical networks, and gene circuits often are not simple linear systems; they can entail nonlinear elements, negative or positive feedback, oscillators, switches, gates, and logic-like elements. These underlying details are often unknown. Furthermore, the modelling challenge is compounded for these systems because data are noisy due to the small number of particles involved in gene expression. Moreover, experiments typically measure the dynamics of few proteins tagged by fluorescent markers, whereas the underlying gene expression may be governed by multiple actors. As a result, it is usually not possible to infer all underlying interactions and their rates using data. Can we instead build an effective dynamical model that is consistent with data and allows us to make predictions?

Max Cal is ideal for building such effective few-particle models with limited information, such as data on a few species. Max Cal builds models with minimal numbers of parameters while allowing maximal determination directly from the data. Below are three examples in which Max Cal captures the complex dynamics of few-particle gene circuits. In a gene circuit, proteins are synthesized from DNA and can also bind to the DNA and control the rate at which it or other proteins are made.

4.4.1. Modeling autoactivation in a single-gene circuit.—

Figure 9 shows how autoactivation is performed in a single-gene regulatory circuit (20–22). A gene α produces a protein of type A. The protein degrades at some rate d. The gene's DNA is flanked by a promoter region. When two A protein molecules in a dimeric form (A_2) bind to the promoter, gene α then produces protein A faster than it did before.

The experimental data will appear in the form of a noisy switchlike time trace of the number $N_A(t)$ of A molecules per unit time t (see Figure 9b). It is not possible to infer all microscopic rate parameters from these data. Instead, in modeling this circuit, we aim to extract the maximum amount of information from the entire stochastic time trace (not just averages over it) to infer an effective model. We use the stochastic time trace to infer rate g of synthesis, rate d of degradation, and effective accelerated rate g^* when the promoter is activated. We are not likely to know a priori these rates, the binding affinities of A_2 to

the promoter, or other such mechanistic variables. Other modeling efforts may include such variables explicitly, but these usually require additional parameters without any experimental knowledge of their values. Max Cal can be used to extract three core quantities from the stochastic data (20, 22). Moreover, Max Cal predicts an effective feedback parameter (K) quantifying the coupling between the protein and its promoter. The model also produces bimodal distribution, a necessary condition for the model to successfully describe the switch. Moreover, the Max Cal model provides insights into how to dial the feedback parameter to alter these distributions (20, 22).

Here is how Max Cal works for this simple circuit. We write the caliber as

$$C = \sum_{\ell_\alpha, \ell_A} \left[-P_{\ell_\alpha, \ell_A} \log P_{\ell_\alpha, \ell_A} + b_\alpha \ell_\alpha P_{\ell_\alpha, \ell_A} + b_A \ell_A P_{\ell_\alpha, \ell_A} + K_A \ell_\alpha \ell_A P_{\ell_\alpha, \ell_A} \right], \quad 21.$$

where ℓ_α is the production state variable, which ranges as integer values between zero and some predefined maximum value (M),and ℓ_A is the degradation state variable, which describes the number of previously existing proteins that still exist at the end of the time interval. The corresponding Lagrange multipliers for these two constraints are b_α and b_A, and the probability of observing a particular combination of ℓ_α and ℓ_A is defined as P_{ℓ_α, ℓ_A}. The first term, therefore, is the path entropy term. The second and third terms impose the constraints on average production and degradation rates, respectively. The information that a high number of proteins (N_A) should positively correlate with the production of A is enforced by constraining the average of $\ell_\alpha \ell_A$ with a Lagrange multiplier, K_A (the last term). This is the lowest-order term in the coupling of these two variables that must be imposed to capture the essence of feedback. Based on this caliber, corresponding caliber-maximized path probabilities are

$$P_{\ell_\alpha, \ell_A} = Q_d^{-1} \binom{N_A}{\ell_A} \exp(b_\alpha \ell_\alpha + b_A \ell_A + K_A \ell_\alpha \ell_A), \quad 22.$$

$$Q_d = \sum_{\ell_\alpha = 0}^{M} \sum_{\ell_A = 0}^{N_A} \binom{N_A}{\ell_A} \exp(b_\alpha \ell_\alpha + b_A \ell_A + K_A \ell_\alpha \ell_A). \quad 23.$$

The likelihood ($\mathscr{L}$) of the experimentally observed trajectory is written in terms of P_{ℓ_α, ℓ_A} Maximization of the likelihood allows us to determine M, b_α, b_A, and K_A. Upon determining these Lagrange multipliers, we can use them to infer different rates, feedback parameters, and other features of the circuit (20). None of these parameters are directly available in experiment. The feedback parameter is defined as the Pearson correlation between the protein production variable, ℓ_α, and proteins present, ℓ_A. The value of this feedback parameter is an effective measure of how the presence of protein A influences protein production. At a molecular level this metric may be affected by the binding constant of the dimer protein molecule to the promoter site, among other variables.

4.4.2. A genetic toggle-switch circuit.—Figure 10 shows a genetic toggle switch originally designed by Gardner et al. (23). Gene α makes a protein A at rate $g*$, and gene β makes a protein B at rate $g*$. Each gene is flanked by a repressor region. When a protein molecule B binds to the repressor of α, it slows the production of protein A. Likewise, when A binds to the repressor of β, it slows the production of B. For this circuit, the experimental data are the time traces of the numbers $N_A(t)$ and $N_B(t)$ of each type of protein (see Figure 10b). The Max Cal model produces the alternating output that is the hallmark of bistability: as A molecules begin to outnumber B molecules, it becomes winner-take-all, and A molecules come to dominate; the reverse is also true. Importantly, Max Cal uses the raw input data to predict the production rate in the basal and repressed states as well as the degradation rates and the AB correlations (22, 24).

4.4.3. The repressilator, an oscillating genetic circuit.—Figure 11 shows the repressilator gene circuit of Elowitz & Leibler (25). It is a circuit of three proteins: A, B, and C. Gene α produces protein A; gene β produces protein B; and gene γ produces protein C. All are produced at rate g. Protein A can bind to the promoter region of B to slow production of protein B; similarly, protein B can slow production of protein C, and protein C can slow production of protein A. The rate of production in the repressed state is $g*$, much slower than g. Each protein has the same rate of degradation, d. The raw data traces are of each protein as a function of time (see Figure 11b). As before, we maximize the likelihood of observing the noisy oscillatory trajectory to infer the underlying effective rates g, $g*$, and d. This fully harnesses the information hidden in the dynamical data (see Figure 11). The inferred underlying parameters agree well with the underlying model that was used to generate the synthetic data (26). Max Cal also predicts the effective feedback strength (K).

4.5. Maximum Caliber Is the Method of Choice When Detailed Models Are Not Known

Traditional modelling approaches have limitations for treating few-particle dynamics and complexity (Figure 12). Mass-action (MA) models treat only mass actions (i.e., macroscale bulk average kinetics) and not fluctuations or rate distributions. MA plus noise models repair this limitation by adding time fluctuations through assumed distributions, such as in Langevin equations (27). However, the typical assumed distributions are not always valid. For example, such noise models apply only if the dynamics entails a simple unimodal distribution; they do not apply for bimodality, such as in the toggle switch. Moreover, such models treat complex dynamics by asserting nonlinear functions f—for example, of the Hill form, $f_{Hill} = \langle A\rangle^n/[\langle A\rangle^n + \text{constant}]$—which are often ad hoc, are not known independently, and require multiple parameters. Chemical master equation (CME) models (28) do explicitly account appropriately for dynamical fluctuations. However, to construct CME models, one needs to know the detailed reaction networks—that is, all the states and the arrows between them in terms of multiple species (Y in Figure 12). But often, these many species cannot be validated by experiments. Consequently, CME models involve too many parameters and a large phase space that can be computationally intractable even for systems as simple as the toggle switch or the repressilator. Although finite state projection (29) can help reduce the phase space, the combinatorics of multiple species remains a challenge. CME + MA models (30) can express the fluctuations, but they also require nonlinear functions, like f_{Hill}, and parameters that may not reflect an underlying mechanism.

Max Cal is the method of choice when information available about the underlying circuit is minimal. It is a principled way to derive the probabilities of observed trajectories, even in complex dynamics, without typical assumptions about the forms of nonlinearities. Due to its top-down nature, it keeps phase space minimal and solves the computational challenge of CME, and at the same time, it avoids the ad hoc assumptions of CME + MA. Max Cal uses all the information in the trajectory, not just the averages over it, even when experiments do not give exactly the quantities of interest. Consequently, Max Cal, combined with maximum likelihood methods for obtaining parameters, can properly handle data in the form of fluorescence signals, and not in molecule numbers, a typical situation in experiments. By incorporating a distribution for fluorescence-to-numbers conversion, Max Cal can construct the likelihood of observed trajectories in raw fluorescence. This way, Max Cal deconvolutes the uncertainty in the fluorescence (per particle) from the small-numbers noise (specific to the circuit architecture) and builds models for circuit details. The success of Max Cal plus maximum likelihood with fluorescence-to-numbers conversion has been demonstrated in different genetic circuits (20, 22, 26).

Here, we make a general point about model making in molecular kinetics. Kinetics textbooks typically describe mechanisms, which are assertions about plausible intermediate steps that occur on the dominant route from reactant to product for that process. Such mechanisms are learned about average dominant routes because they are done in beakers of large numbers of molecules. Here, we tackle a different challenge: namely, learning mechanisms in small-numbers kinetics from additional insights that come from the distributions of the path (in the form of stochastic trajectories), not just the averages. For few-particle flows, the distribution of routes, too, gives mechanistic information (i.e., the Lagrange multiplier quantities above). But Max Cal and modeling intermediate states are not mutually exclusive. The Max Cal models above can readily be augmented by whatever additional variables might be needed to address intermediate states and would bring additional insights about the route distributions around them. In other words, Max Cal can always include more information if there are data. However, in the absence of data, it avoids unnecessary assumptions while giving a minimal consistent model.

4.6. Maximum Caliber Can Infer Distributions over Network Parameters from Single-Cell Data

Inside cells are biochemical networks with rate coefficients $\mathbf{k}$ that are approximately time invariant (31). The biochemical species abundances $x(\mathbf{k}, t)$ fluctuate as a function of time following Poisson small-numbers statistics. This is called intrinsic noise. This intrinsic variability scales $\propto 1/N$, where N is the typical average abundance of a species. Another kind of variability, called extrinsic noise (32), is observed when the parameters $\mathbf{k}$ themselves differ from one cell to the next. This variability leads to different stochastic trajectories observed across cells. Can we infer a distribution $P(\Gamma)$ over trajectories of species abundance—that is, $x(\mathbf{k}, t)$—that accounts for the extrinsic noise? One challenge is that experimental techniques such as flow cytometry, immunofluorescence (33), and single-cell RNA sequencing (34) do not directly give stochastic time traces inside individual cells. They give only biochemical species abundances due to cell-to-cell variability in a given snapshot of time.

Max Cal can quantify the extrinsic variability from such time snapshot data (35). Dixit and colleagues (35, 36) recently developed an approach that infers from single-cell snapshot data the distribution over parameters $P(\mathbf{k})$.Max Cal can also infer the distribution over trajectories of abundance $P(\Gamma)$ in a cell population.

Dixit and colleagues took this approach to study the dynamical evolution of phosphorylation levels in the epidermal growth factor–activated Akt pathway, one of the most frequently mutated pathways in cancer. This approach predicts the correct distribution of phosphorylation levels in the population and also infers those biochemical rate parameters that dominate the trajectory distribution.

4.7. Maximum Caliber Estimates the Trafficking Dynamics on Networks

Consider the following problem. A biomolecule has many different metastable conformations. We want to compute the full transition matrix for all the rates k_{ab} between all pairs of states a and b. When using molecular dynamics simulations, this can be slow and challenging because transitions are rare, involving states of high free energy, which are not sampled very often. But Max Cal gives a fast way to approximately estimate this rate matrix, given information that is obtained more quickly. If a computer simulation searches the conformational space sufficiently to find and sample the metastable state populations, and if we know one or two global rate quantities, such as how fast an overall transition occurs, then Max Cal predicts the rate matrix that maximizes the path entropy (Figure 13).

Concretely, in a Markovian system with N states $\{a, b,...\}$, the path entropy can be expressed in terms of the stationary distribution $\{p_a\}$ and the transition probabilities $\{k_{ab}\}$ (37):

$$S = -\sum_{a,b} p_a k_{ab} \log k_{ab} . \quad 24.$$

Max Cal estimates the matrix of transition probabilities subject to limited rate information by maximizing the path entropy function in Equation 24.

The limited information can include the full stationary distribution $\{p_a\}$ (see, for example, 38, 39), stationary state averages $\langle E \rangle = \sum p_a E_a$, or path ensemble averages of dynamical quantities $\langle J \rangle = \sum_{a,b} p_a k_{ab} J_{ab}$. The entropy maximization can be carried out in an analytical/semianalytical manner depending on the constraints. Specifically, Dixit et al. (40) showed that the transition rates r_{ab} of Max Cal are given by

$$r_{ab} \propto \sqrt{\frac{p_b}{p_a}} e^{-\gamma J_{ab}}, \quad 25.$$

where γ is the Lagrange multiplier that enforces the value of the dynamical averag $e^{\langle J \rangle}$.

The Max Cal Markov processes have found applications in a variety of areas, including understanding the dynamics of biomolecules (37–46), modeling biochemical reaction networks (47), decision theory (48), and machine learning (49). Here, we briefly describe a couple of applications in detail.

4.7.1. Estimating the dynamics of conformational change from molecular simulations.—

Molecular simulations of protein molecules are often performed to find the rates and routes of conformational change, since these actions are often determinative of biological mechanisms. As described above, the challenge is that molecular dynamics simulations are much poorer at finding and sampling rare kinetic transitions (high free energies) than they are at finding and exploring the stable and metastable states. The Max Cal equation (Equation 25) provides a useful, practical, simple, and fast way to estimate all the transition rates between such conformational states, if we know the populations of the stable and metastable states, and if we know one or two global average rate quantities (see Figure 13). A demonstration that this formula is an accurate estimator has been given on a seven-residue alanine peptide, for which the full rate matrix had been obtained by extensive computations (40).

Equation 25 has been similarly applied to a much bigger and more complex protein conformational change—namely, the allosteric transition that occurs in G protein–coupled receptor ligand-activated dynamics (50). One major issue in using Equation 25 is the determination of the dynamical average $\langle J \rangle$ from an unbiased dynamical ensemble and the consequent determination of the Lagrange multiplier γ. Most computational sampling techniques that determine equilibrium landscapes employ biased ensembles and cannot be used to estimate dynamical quantities. Recently, Meral et al. (50) provided an elegant solution. They exploited the fact that unbiased dynamical averages can be estimated from a biased ensemble in a metadynamics simulation using collective variable coordinates (51). As a result, they simultaneously estimated the equilibrium distribution and obtained unbiased dynamical averages of several quantities in Markov state models that they then used with Equation 25 to estimate the transition rates.

4.7.2. Determining reaction coordinates in molecular simulations.—

An important challenge is to find the dominant reaction coordinates in computer simulations of biomolecular changes. The conformational spaces are high dimensional; simulations sample them only sparsely; and even when a good reaction coordinate is known, the rates along that coordinate are not obtainable without sufficient sampling to get converged populations. Recently, Tiwary and colleagues (41, 42, 45) developed a clever method to determine reaction coordinates using the Max Cal equation (Equation 25). They simulate a process of interest with metadynamics, which requires first choosing some collective variables that are relevant to the problem at hand. First, they estimate the free energy profile along any particular linear combination of the collective variables. Next, for any given linear combination of reaction coordinates, they estimate the approximate rate matrix on a grid using the Max Cal equation (Equation 25). Finally, they choose as the optimal reaction coordinate the linear combination that gives the maximum spectral gap in the rate matrix (42) (see Figure 14). Optimizing over the linear combination does not require additional simulations, since enhanced simulation techniques such as metadynamics allow one to estimate the free energy profile along any linear combination given the free energy profile along any other linear combination (42).

The principle assumed here is that reaction coordinates usually describe large-scale motions, which are slower than small-scale motions (side-chain rotations, small displacements of

solvent, etc.). Thus, a pathway that has clearly separated slow motions is a good candidate reaction coordinate. This approach extends to multidimensional reaction coordinates (45) and has been applied to several examples (43, 44).

4.7.3. Using maximum caliber to correct/update Markov models, given new data.—

Consider the following common practical problem. All the microscopic rates have been estimated in a model network. Now imagine a situation in which, either because of perturbations to the experimental system or because of erroneous determination of rates, the experimental data do not agree with these predictions. For example, in simulating protein folding, due to inaccuracies in all atom force fields, the computed folding rates may disagree with experiments. Poor sampling can also lead to such inaccuracies, a well-known issue in biomolecular simulations (52). How is the rate matrix then corrected? In most cases, there is no unique answer. Max Cal gives a best way to correct the full microscopic rate matrix, to bring the model into agreement with the data (47, 53). If the original computed rates are $\{q_{ab}\}$, then we can take the new data-corrected rates to be $\{k_{ab}\}$, which are obtained by maximizing the relative entropy,

$$S = -\sum_{a,b} p_a k_{ab} \log\left[\frac{k_{ab}}{q_{ab}}\right]. \quad 26.$$

In another example (47), a growth factor membrane receptor protein undergoes a four-state biochemical cycle: unbound to ligand; bound; activated, due to phosphorylation; or degraded (Figure 15). For the normal wild-type protein, these rates are known.

Now, some mutations of the receptor—such as in some cancers (54)—cause the activated-state population p_{act} to be larger by an observable amount. We want to use this single observation to update our predicted populations and rates of the four-state model. To do this, we maximize the relative path entropy subject to the new value of p_{act}. Notably, while there are infinitely many ways to update the rate matrix to fit this information, this approach predicts that a higher phosphorylation level is most likely obtained through a decrease in the receptor internalization rate (47), a well-documented aberration in the growth factor pathway (54, 55).

4.8. Other Applications of Maximum Caliber

Max Cal has also been successfully applied to bird flocking (56), cell motility (57), and the firing of neurons (10, 58). For bird flocking, Cavagna et al. (56) constructed a Max Cal formalism that correlates observables at two consecutive time steps. This model was shown to be superior to one based on static information alone, when benchmarked using synthetic data. Tweedy et al. (57) have used Max Cal to model cell motility from cell shapes. They infer Lagrange multipliers by analyzing the time evolution of cell-shape trajectories. Their inferred Lagrange multipliers can successfully discriminate between drug-treated cells and untreated cells, and between genetically modified cells and unmodified parent cells. And Max Cal has been used to capture complex critical dynamics of neurons (10, 58). When shown natural images, retinal neurons exhibit all-or-none behavior. Using maximum entropy (59), studies have established critical behavior in statistical properties of collective neuronal

firing. These studies found that the Lagrange multipliers for the maximum entropy models were fine-tuned to criticality correspond to a spin glass model. Mora et al. (10) extended these observations to the dynamical regime; they modeled the collective behavior of neurons by constraining neuron-neuron correlations across time. They showed that incorporating dynamical information about neuronal firing using Max Cal predicted that retinal neurons are poised at dynamical criticality as well.

5. MAXIMUM CALIBER DERIVES WELL-KNOWN RESULTS OF NONEQUILIBRIUM PHYSICS, AS IT SHOULD

5.1. Why Are Markov Models So Ubiquitous in Nonequilibrium Physics?

Markov modeling is effective for a wide range of dynamical processes. In first-order Markov models, it is assumed that populations of states, and the rates between them, are adequately approximated by knowing only the properties of a given state and its adjacent kinetic state (the states at the immediately preceding and following steps). It neglects any more distant memory. Why are Markov models so ubiquitous and useful for modeling nature? The principle of Max Cal gives an explanation. Among modeling approaches, Max Cal is maximally data driven, insofar as it uses directly measured observables. And the point here is that the nature of the data that are available through any particular experiment determines the best model that captures those data. For example, in a two-state process, $A \leftrightarrow B$, if we know only the four quantities—the frequencies of transitions from A to B, A to A, B to B, and B to A—then the class of models that maximize the caliber is first-order Markovian (60–63). Models with more parameters (higher-order Markov, for example) are not warranted unless we have more information. From maximizing caliber with the given set of constraints mentioned above, it follows that the rate between two states in two consecutive time steps depends only on the previous time step. Lee & Pressé (62) provided a rigorous description when data can involve more than two steps, that is, when there is memory.

5.2. Known Near-Equilibrium Statistical Physics Results Can Be Derived from Maximum Caliber

Max Cal can derive well-known results of near-equilibrium statistical physics that relate fluxes, flows, and entropy production. These include the Green–Kubo relationship, Onsager's reciprocal relationship, and Prigogine's minimum entropy production principle (64). For example, following the experiments of Thomson on the coupling of electrical and heat flows, Onsager considered two types of flow—say a and b flow between two baths. If the fluxes J_a and J_b are linear in the forces (λ_a, λ_b), then they couple as follows:

$$J_a = L_{aa}\lambda_a + L_{ab}\lambda_b, \tag{27.}$$

$$J_b = L_{ba}\lambda_a + L_{bb}\lambda_b\,. \tag{28.}$$

Onsager gave a microscopic argument for a reciprocity relationship, $L_{ab} = L_{ba}$. It can be derived much more simply from Max Cal. First, if the particle fluxes at any given time for

any given trajectory Γ are $j_a(\Gamma, t)$ and $j_b(\Gamma, t)$, then Max Cal predicts the distribution over trajectories as

$$P(\Gamma) = \frac{p(\Gamma)}{Q_d} \exp\left(\sum_t \lambda_a(t) j_a(t) + \sum_t \lambda_b(t) j_b(t) \right), \quad 29.$$

where Q_d is the dynamical partition function, $\lambda_a(t)$ and $\lambda_b(t)$ are Lagrange multipliers, and $p(\Gamma)$ is the equilibrium distribution. Now, because Max Cal is a partition-function-based approach, we immediately have the first and second derivatives,

$$J_a(t) = \frac{\partial \log Q_d}{\partial \lambda_a(t)}, \quad 30.$$

$$\frac{\partial J_a(t)}{\partial \lambda_b(\tau)} = \frac{\partial^2 \log Q_d}{\partial \lambda_a(t) \lambda_b(\tau)}. \quad 31.$$

Just like in the Maxwell relations for equilibrium thermodynamics, the second derivative obtained is independent of the order of differentiation, from which the Onsager reciprocal relations follow immediately (see Reference 64 for details of the derivation).

These coupling relationships can be useful not only for thermoelectrics or work and heat, but also for how biochemical energy sources drive reactions, molecular motors, and ion pumps and for how energy sources can increase precision in biomolecular clocks and proofreading, as well as for how photovoltaic materials couple light to current or heat.

6. QUESTIONS, CRITICISMS, PERSPECTIVES, AND LIMITS OF THIS REVIEW

This section provides some broader context. But first we note that space limits us from reviewing other important, active, and related matters of nonequilibrium statistical physics, such as stochastic thermodynamics (65) and large-deviation theory (66).

6.1. Does Maximum Caliber Handle Dissipation Properly?

Does Max Cal properly handle dissipative processes (67, 68)? If Γ denotes a trajectory, and if the only constraint is on $\langle J \rangle$, then Max Cal predicts that the populations of the routes will be

$$P(\Gamma) \propto e^{\mu J(\Gamma)}. \quad 32.$$

But imagine a particle moving through a viscous fluid. Its average velocity could be achieved in different ways: either through a large force applied to the particle in a high-viscosity medium or through a small force applied in a low-viscosity medium. We would expect the route distribution to be different in those two cases. But Equation 32 implies that there should be no difference in the route distributions. Is this a failure of Max Cal? No. Indeed the route distributions will be different. But the single constraint on $\langle J \rangle$ is not

sufficient for such dissipative systems, where the rates $\dot{U}$ of energy input and output are also critical. In this case, we need the following (69):

$$P(\Gamma) \propto e^{\mu J(\Gamma) + \nu \dot{U}(\Gamma)}.$$

This additional constraint leads to the correct rate distribution, as shown in an exact model (69). However, note that Max Cal does not itself specify what constraints are needed for any given problem. This is a decision required of the modeler, depending on what is relevant for a given problem. The challenge resembles that of equilibrium thermodynamics: Specifying the temperature alone is not sufficient in situations where volumes and concentrations are changing at the same time; then one must also specify the pressures and chemical potentials.

6.2. Numerical Issues in Learning Lagrange Multipliers from Data

Application of Max Cal to noisy biological data often poses numerical challenges. Using data to fit several Lagrange multipliers at a time can be computationally expensive (70, 71). Fitting data can involve solving simultaneously N nonlinear equations. They are often not independent. Often, analytical expressions are not available for how $\langle J \rangle$ depends on the statistical weights. And data often have experimental errors. So obtaining Lagrange multipliers from data can involve stochastic sampling of coupled nonlinear equations. As a result, we sometimes determine constraints not as precisely fixed values but rather as distributions over the possible Lagrange multipliers (35). Another challenge in the application of Max Cal can be in determining the state space. As was shown in the gene network or Markov models for biomolecules, we impose a state space a priori. More sophisticated models do not require imposing that space (72, 73).

6.3. Is Maximum Caliber a Principle of Inference, or of Physics?

In summary, Max Cal is a general method for drawing inferences about distributions of rates and routes in models of dynamical processes. Given a model, and given limited data—such as a few average rates or any other moments—Max Cal predicts the distribution that is consistent with the model, the data, and the rules of probability. We view this as identical in spirit to how equilibrium statistical mechanics infers distributions. We regard all of statistical mechanics as drawing inferences about models of physics.

A question is whether statistical physics can be derived from mechanics, and thus is deeper than just inference. Our view is that, in practice, it cannot. While the first law is about energies and mechanics and is grounded in physical quantities, the second law is about populations—and thus is about inferences or probability theory. One cannot derive the second law from the first. Not everything is derivable from pure mechanics alone. To paraphrase E.T. Jaynes, Boltzmann's brilliance was in recognizing that while gas behaviors could be determined, in principle, by following all billiard-ball-like collisions over time, the only practical way to compute properties was to replace the detailed mechanics with a statistical description, hence the term statistical mechanics (74). Once we accept Boltzmann's intellectual leap, and we use the expression $S = k \ln W$ or $S = -\sum p_i \ln p_i$, we are

necessarily regarding entropy variation as a task of drawing inferences about models from data.

Still, there are differences between physical and nonphysical inferences. If all we know is an average flux, Max Cal can infer the distribution, but nothing more. But, if the model we use in Max Cal is more physical, we can learn more, such as how the rates depend on properties of the particle or the flow networks, for example. Where Hamiltonian dynamics does apply, it can bring additional mechanistic insights relating forces and flows to properties of the underlying molecules. Moreover, among different types of constraints, temperature occupies a special place in statistical thermodynamics. Thermal equilibrium has an experimental ground truth that dynamics does not, namely that of the Clausius relationship, $S_{\text{Clausius}} = q/T = \langle U \rangle / T$, which determines the entropy from the mean energy. But the Clausius relationship holds only at equilibrium, and under various additional limiting restrictions. There is no corresponding ground-truth experimental foundation yet known for dynamics far from equilibrium.

6.4. How Does Maximum Caliber Remedy Key Difficulties of Nonequilibrium Physics?

Here we briefly summarize how Max Cal addresses typical challenges and issues with nonequilibrium physics.

6.4.1. Maximum caliber operates on trajectories (paths), not states.—A Max Cal model can include all relevant paths, including those very distant from equilibrium. An important but underappreciated point from Shore & Johnson (15) is that the function $-\sum p_i \ln p_i$ (*a*) is the only function that enforces the consistency of the laws of probability, but (*b*) is a valid predictor only for the single distribution function $\{p_i^*\}$ that maximizes it. The only path entropy that is invoked in Max Cal is $S_{\text{path}} = -k \sum p_\Gamma^* \ln p_\Gamma^*$, having route probabilities p_Γ^* that maximize it. There is no need to consider any deviations from that distribution; hence, it satisfies the Shore–Johnson criterion for consistent probabilistic inferences (15, 63, 75).

6.4.2. It is not limited to near equilibrium.—Max Cal does not require continuum-function starting points, such as a state entropy $S_{\text{state}} = S(U, V, N)$, which is a continuous and differentiable function of extensive variables of space and time, and which is strictly defined only for equilibria (by the Clausius relation).In order for this smoothness to hold, an assumption of local equilibrium necessarily follows. This is restrictive, and it limits consideration to processes that take only small steps, equilibrating along the way (76).

6.4.3. It applies to nonthermal systems.—Max Cal is not limited to thermal processes, or to temperature baths, so it is readily applicable to a broad range of flow problems. It is agnostic about where the model or data come from. It is not limited to materials, molecules, their collisions, Hamiltonian systems, the Liouville theorem, heat bath, or temperature. Max Cal is more general, including for stochastic-dynamic systems, gene circuits, networks, and other systems.

6.5. Maximum Caliber Gives a Way to Rationalize Different Classes of Nonequilibrium Physics

Nonequilibrium processes are often classified into broad categories, such as equilibrium, near-equilibrium relaxations (for example, resulting from changing conditions imposed on reaction networks or diffusion from high to low concentration), near-equilibrium steady state (constant ohmic current through a resistor or dragging a sphere slowly through a viscous liquid), or far from equilibrium (driving a sphere to cause turbulence in a fluid). These distinctions can be made naturally according to what constraints are relevant for formulation in Max Cal. Of course, equilibrium requires the satisfaction of detailed balance, with no additional rates imposed otherwise. Near equilibria are processes that dissipate, but where the dissipation is linearly proportional to some flow rate $\langle J \rangle$ for a fixed applied force. For example, consider ohmic heat dissipation that is proportional to voltage × current. Another example is Stokes's law, in which a sphere in a viscous liquid dissipates $\propto$ force × velocity. In these cases, near equilibrium can be defined by specifying the relevant rate $\langle J \rangle$ alone, because the linear proportionality with dissipation requires no additional dissipation constraint. Near equilibria have one or more of the corresponding features: (*a*) linear force-flow relations, (*b*) the correspondence of entropy production with dissipation, (*c*) the applicability of local equilibria assumptions, and (*d*) Green–Kubo and Onsager reciprocal relations and Prigogine's minimum entropy production (64). In contrast, far-from-equilibrium processes may require additional information for specifying both the macroscopics and microscopics—for example, particular models for the excess dissipation beyond that which is specified by $\langle J \rangle$ alone, or force constraints that cannot be interpreted as simple gradients in thermodynamic intensive variables.

7. SUMMARY: MAXIMUM CALIBER IS A GENERAL VARIATIONAL PRINCIPLE FOR DYNAMICS AND PATHWAYS

Max Cal is a principle for inferring distributions of rates and routes in dynamical models, given limited data. It can derive well-known near-equilibrium results, and it can delineate near equilibrium from far from equilibrium, but is also valid far from equilibrium. It overcomes traditional problems of NEP. With models, it can derive phenomenological laws, such as Fick's law. It is applicable to few-particle complex systems, such as gene circuits. Its logic is advantageous for building the simplest models allowed by the data. And, while not fully proven in all contexts, it gives every indication of being a general principle for nonequilibria.

ACKNOWLEDGMENTS

We are grateful to Steve Pressé, John Van Drie, Hong Qian, Jason Wagoner, Corey Weistuch, and Gili Bisker for their insightful comments. The early years of this work benefited enormously from collaborations with Rob Phillips and his students Mandar Inamdar, Effrosyni Seitaridou, and Dave Wu. We appreciate the support of the Laufer Center and of National Science Foundation grant 1205881. K.G. acknowledges support from the National Institutes of Health (award 1R15GM128162-01A1) and a PROF award from the University of Denver.

LITERATURE CITED

1. Broder A. 2000. Graph structure in the web. Comput. Netw. 33:309–20

2. Zipf GK. 1949. Human Behavior and the Principle of Least Effort. Cambridge, MA: Addison-Wesley
3. Marsili M, Maslov S, Zhang YC. 1998. Dynamical optimization theory of a diversified portfolio. Phys. A Stat. Mech. Appl. 253:403–18
4. Cavagna A, Queirós SD, Giardina I, Stefanini F, Viale M. 2013. Diffusion of individual birds in starling flocks. Proc. R. Soc. B Biol. Sci. 280:20122484
5. Peterson J, Dixit P, Dill KA. 2013. A maximum entropy framework for nonexponential distributions. PNAS 110:20380–85 [PubMed: 24297895]
6. Zeldovich K, Chen P, Shakhnovich E. 2007. Protein stability imposes limits on organism complexity and speed of molecular evolution. PNAS 104:16152–57 [PubMed: 17913881]
7. Zou T, Williams N, Ozkan S, Ghosh K. 2014. Proteome folding kinetics is limited by protein halflife. PLOS ONE 9:e112701 [PubMed: 25393560]
8. Dixit PD, Maslov S. 2013. Evolutionary capacitance and control of protein stability in protein-protein interaction networks. PLOS Comput. Biol. 9:e1003023 [PubMed: 23592969]
9. Dixit PD, Pang TY, Maslov S. 2017. Recombination-driven genome evolution and stability of bacterial species. Genetics 207:281–95 [PubMed: 28751420]
10. Mora T, Deny S, Marre O. 2015. Dynamical criticality in the collective activity of a population of retinal neurons. Phys. Rev. Lett. 114:078105 [PubMed: 25763977]
11. Onsager L. 1931. Reciprocal relations in irreversible processes. Phys. Rev. 37:405
12. Prigogine I, Kondepudi D. 1998. Modern Thermodynamics: From Heat Engines to Dissipative Structures. New York: John Wiley
13. Martyushev LM, Seleznev VD. 2006. Maximum entropy production principle in physics, chemistry and biology. Phys. Rep. 426:1–45
14. Jaynes ET. 1980. The minimum entropy production principle. Annu. Rev. Phys. Chem. 31:579–601
15. Shore J, Johnson R. 1980. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. IEEE Trans. Inf. Theory 26:26–37
16. Wu D, Ghosh K, Inamdar M, Lee H, Fraser S, et al. 2009. Trajectory approach to two-state kinetics of single particles on sculpted energy landscapes. Phys. Rev. Lett. 103:050603 [PubMed: 19792475]
17. Pressé S, Ghosh K, Phillips R, Dill KA.2010. Dynamical fluctuations in biochemical reactions and cycles. Phys. Rev. E 82:031905
18. Ghosh K, Dill K, Inamdar M, Seitaridou E, Phillips R. 2006. Teaching the principles of statistical dynamics. Am. J. Phys. 74:123–33 [PubMed: 23585693]
19. Seitaridou E, Inamdar M, Phillips R, Ghosh K, Dill K. 2007. Measuring flux distributions for diffusion in the small-numbers limit. J. Phys. Chem. B 111:2288–92 [PubMed: 17295536]
20. Firman T, Balazsi G, Ghosh K. 2017. Building predictive models of genetic circuits using the principle of maximum caliber. Biophys. J. 113:2121–30 [PubMed: 29117534]
21. Nevozhay D, Adams R, Itallie EV, Bennett M, Balázsi G. 2012. Mapping the environmental fitness landscape of a synthetic gene circuit. PLOS Comput. Biol. 8:e1002480 [PubMed: 22511863]
22. Firman T, Wedekind S, McMorrow TJ, Ghosh K. 2018. Maximum caliber can characterize genetic switches with multiple hidden species. J. Phys. Chem. B 122:5666–77 [PubMed: 29406749]
23. Gardner T, Cantor C, Collins J. 2000. Construction of a genetic toggle switch in *Escherichia coli*. Nature 403:339–42 [PubMed: 10659857]
24. Pressé S, Ghosh K, Dill K. 2011. Modeling stochastic dynamics in biochemical systems with feedback using maximum caliber. J. Phys. Chem. B 115:6202–12 [PubMed: 21524067]
25. Elowitz M, Leibler S.2000.A synthetic oscillatory network of transcriptional regulators. Nature 403:335–38 [PubMed: 10659856]
26. Firman T, Amgalan A, Ghosh K. 2019. Maximum caliber can build and infer models of oscillation in a three-gene feedback network. J. Phys. Chem. B 123:343–55 [PubMed: 30507199]
27. Van Kampen NG. 1992. Stochastic Processes in Physics and Chemistry, Vol. 1. Amsterdam: Elsevier
28. Gillespie D 1992. A rigorous derivation of the chemical master equation. Phys. A Stat. Mech. Appl. 188:404–25

29. Munsky B, Khammash M. 2006. The finite state projection algorithm for the solution of the chemical master equation. J. Chem. Phys. 124:044104 [PubMed: 16460146]
30. Frigola D, Casanellas L, Sancho J, Ibanes M. 2012. Asymmetric stochastic switching driven by intrinsic molecular noise. PLOS ONE 7:e31407 [PubMed: 22363638]
31. Llamosi A, Gonzalez-Vargas AM, Versari C, Cinquemani E, Ferrari-Trecate G, et al. .2016.What population reveals about individual cell identity: single-cell parameter estimation of models of gene expression in yeast. PLOS Comput. Biol. 12:e1004706 [PubMed: 26859137]
32. Raj A, van Oudenaarden A. 2008. Nature, nurture, or chance: stochastic gene expression and its consequences. Cell 135:216–26 [PubMed: 18957198]
33. Wu M, Singh AK. 2012. Single-cell protein analysis. Curr. Opin. Biotechnol. 23:83–88 [PubMed: 22189001]
34. Saliba AE, Westermann AJ, Gorski SA, Vogel J. 2014. Single-cell RNA-seq: advances and future challenges. Nucleic Acids Res. 42:8845–60 [PubMed: 25053837]
35. Dixit PD, Lyashenko E, Niepel M, Vitkup D.2019.Maximum entropy framework for predictive inference of cell population heterogeneity and responses in signaling networks.Cell Syst. 10.1016/j.cels.2019.11.010
36. Dixit PD. 2013. Quantifying extrinsic noise in gene expression using the maximum entropy framework. Biophys. J. 104:2743–50 [PubMed: 23790383]
37. Dixit PD, Dill KA. 2014. Inferring microscopic kinetic rates from stationary state distributions. J. Chem. Theory Comput. 10:3002–5 [PubMed: 25136269]
38. Wan H, Zhou G, Voelz VA. 2016. A maximum-caliber approach to predicting perturbed folding kinetics due to mutations. J. Chem. Theory Comput. 12:5768–76 [PubMed: 27951664]
39. Zhou G, Pantelopulos GA, Mukherjee S, Voelz VA. 2017. Bridging microscopic and macroscopic mechanisms of p53-MDM2 binding with kinetic network models. Biophys. J. 113:785–93 [PubMed: 28834715]
40. Dixit PD, Jain A, Stock G, Dill KA. 2015. Inferring transition rates of networks from populations in continuous-time Markov processes. J. Chem. Theory Comput. 11:5464–72 [PubMed: 26574334]
41. Tiwary P, Berne B.2016.How wet should be the reaction coordinate for ligand unbinding? J. Chem. Phys. 145:054113 [PubMed: 27497545]
42. Tiwary P, Berne B. 2016. Spectral gap optimization of order parameters for sampling complex molecular systems. PNAS 113:2839–44 [PubMed: 26929365]
43. Tiwary P 2017. Molecular determinants and bottlenecks in the dissociation dynamics of biotin–streptavidin. J. Phys. Chem. B 121:10841–49 [PubMed: 29117680]
44. Hovan L, Comitani F, Gervasio FL. 2018. Defining an optimal metric for the path collective variables. J. Chem. Theory Comput. 15:25–32 [PubMed: 30468578]
45. Smith Z, Pramanik D, Tsai ST, Tiwary P. 2018. Multi-dimensional spectral gap optimization of order parameters (SGOOP) through conditional probability factorization. J. Chem. Phys. 149:234105 [PubMed: 30579304]
46. Dixit PD, Dill KA. 2019. Building Markov state models using optimal transport theory. J. Chem. Phys. 150:054105 [PubMed: 30736685]
47. Dixit PD. 2018. Communication: Introducing prescribed biases in out-of-equilibrium Markov models. J. Chem. Phys. 148:091101
48. Dixit PD. 2018. Entropy production rate as a criterion for inconsistency in decision theory. J. Stat. Mech. Theory Exp. 2018:053408
49. Dixit PD. 2019. Introducing user-prescribed constraints in Markov chains for nonlinear dimensionality reduction. Neural Comput. 31:980–97 [PubMed: 30883279]
50. Meral D, Provasi D, Filizola M. 2018. An efficient strategy to estimate thermodynamics and kinetics of G protein-coupled receptor activation using metadynamics and maximum caliber.J.Chem.Phys.149:224101 [PubMed: 30553249]
51. Tiwary P, Parrinello M. 2014. A time-independent free energy estimator for metadynamics. J. Phys. Chem. B 119:736–42 [PubMed: 25046020]

52. Sawle L, Ghosh K. 2016. Convergence of molecular dynamics simulation of protein native states: feasibility versus self-consistency dilemma. J. Chem. Theory Comput. 12:861–69 [PubMed: 26765584]
53. Dixit PD, Dill KA. 2018. Caliber corrected Markov modeling (C_2M_2): correcting equilibrium Markov models. J. Chem. Theory Comput. 14:1111–19 [PubMed: 29323898]
54. Herbst RS. 2004. Review of epidermal growth factor receptor biology. Int. J. Radiat. Oncol. Biol. Phys. 59:S21–26
55. Huang HJS, Nagane M, Klingbeil CK, Lin H, Nishikawa R, et al. 1997. The enhanced tumorigenic activity of a mutant epidermal growth factor receptor common in human cancers is mediated by threshold levels of constitutive tyrosine phosphorylation and unattenuated signaling. J. Biol. Chem. 272:2927–35 [PubMed: 9006938]
56. Cavagna A, Giardina I, Ginelli F, Mora T, Piovanni D, et al. .2014.Dynamical maximum entropy approach to flocking. Phys. Rev. E 89:042707
57. Tweedy L, Witzel P, Heinrich D, Insall RH, Endres RG. 2019. Screening by changes in stereotypical behavior during cell motility. Sci. Rep. 9:8784 [PubMed: 31217532]
58. Vasquez JC, Marre O, Palacios AG, Berry MJ II, Cessac B. 2012. Gibbs distribution analysis of temporal correlations structure in retina ganglion cells. J. Physiol. Paris 106:120–27 [PubMed: 22115900]
59. Tkačik G, Mora T, Marre O, Amodei D, Palmer SE, et al. 2015. Thermodynamics and signatures of criticality in a network of neurons. PNAS 112:11508–13 [PubMed: 26330611]
60. Stock G, Ghosh K, Dill K. 2008. Maximum caliber: a variational approach applied to two-state dynamics. J. Chem. Phys. 128:194102 [PubMed: 18500851]
61. Ge H, Pressé S, Ghosh K, Dill K. 2012. Markov processes follow from the principle of maximum caliber. J. Chem. Phys. 134:064108
62. Lee J, Pressé S. 2012. A derivation of the master equation from path entropy maximization. J. Chem. Phys. 137:074103 [PubMed: 22920099]
63. Pressé S, Ghosh K, Lee J, Dill K. 2013. Principle of maximum entropy and maximum caliber in statistical physics. Rev. Mod. Phys. 85:1115–41
64. Hazoglou MJ, Walther V, Dixit PD, Dill KA. 2015. Communication: Maximum caliber is a general variational principle for nonequilibrium statistical mechanics. J. Chem. Phys. 143:051104 [PubMed: 26254635]
65. Seifert U 2012. Stochastic thermodynamics, fluctuation theorems and molecular machines. Rep. Prog. Phys. 75:126001 [PubMed: 23168354]
66. Touchette H 2009. The large deviation approach to statistical mechanics. Phys. Rep. 478:1–69
67. Jack RL, Evans R. 2016. Absence of dissipation in trajectory ensembles biased by currents. J. Stat. Mech. Theory Exp. 2016:093305
68. Maes C 2018. Non-Dissipative Effects in Nonequilibrium Systems. Cham, Switz.: Springer
69. Agozzino L, Dill KA. 2019. Minimal constraints for maximum caliber analysis of dissipative steady-state systems. Phys. Rev. E 100:010105 [PubMed: 31499924]
70. Tkacik G, Schneidman E, Berry I, Michael J, Bialek W. 2006. Ising models for networks of real neurons. arXiv:q-bio/0611072 [q-bio.NC]
71. Broderick T, Dudik M, Tkacik G, Schapire RE, Bialek W. 2007. Faster solutions of the inverse pairwise Ising problem. arXiv:0712.2437 [q-bio. QM]
72. Sgouralis I, Pressé S. 2017. An introduction to infinite HMMs for single-molecule data analysis. Biophys. J. 112:2021–29 [PubMed: 28538142]
73. Sgouralis I, Pressé S. 2017. ICON: an adaptation of infinite HMMs for time traces with drift. Biophys. J. 112:2117–26 [PubMed: 28538149]
74. Jaynes ET. 1957. Information theory and statistical mechanics. Phys. Rev. 106:620
75. Dixit PD, Wagoner J, Weistuch C, Pressé S, Ghosh K, Dill KA. 2018. Perspective: Maximum caliber is a general variational principle for dynamical systems. J. Chem. Phys. 148:010901 [PubMed: 29306272]
76. Attard P. 2008. The second entropy: a variational principle for time-dependent systems. Entropy 10:380–90

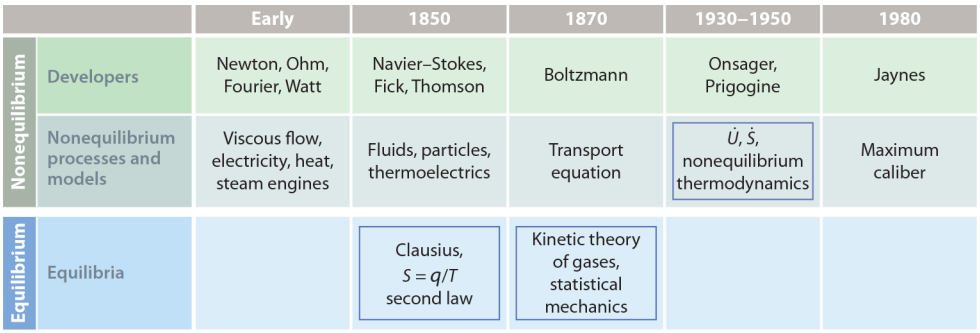

Figure 1.

Timeline of key developments in nonequilibrium physics: (*top row*) developers, (*middle row*) nonequilibrium processes and models, and (*bottom row*) equilibria. The blue boxes denote principles explored and realized.

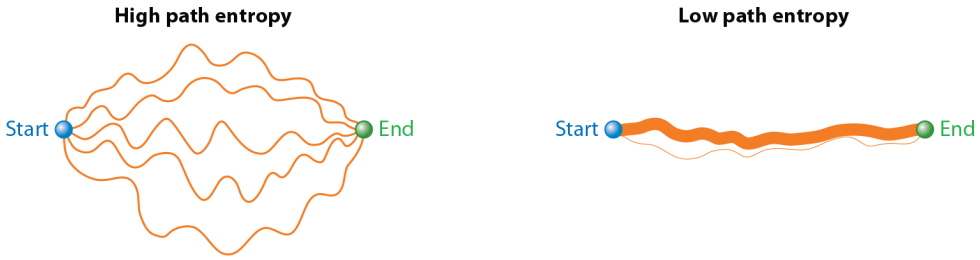

Figure 2.

Path entropy measures the uniformity of the traffic distribution through different routes. Line thickness indicates the traffic density, i.e., pathway probability.

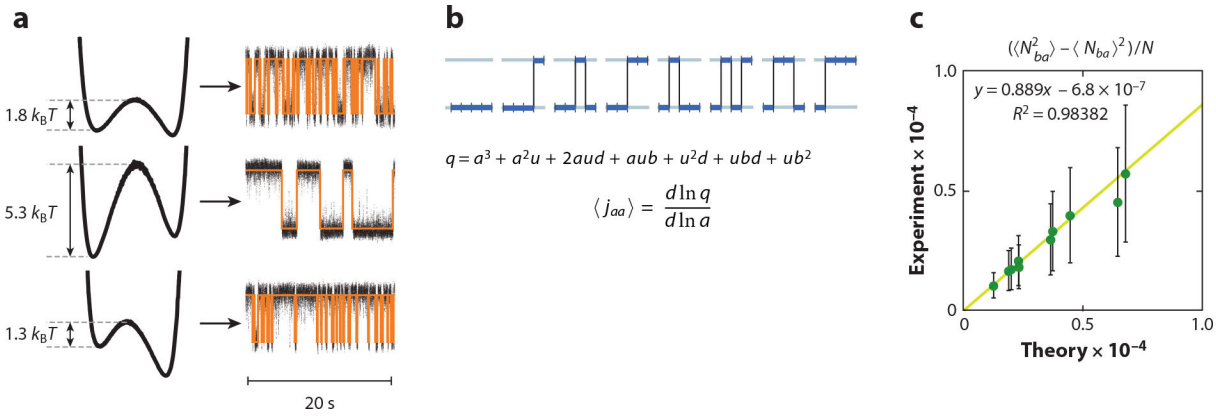

Figure 3.

The use of maximum caliber plus a Markov model to represent two-state dynamics, $A \leftrightarrow B$. (*a*) Different trajectories are represented by different landscapes. Maximum caliber enumerates trajectories, weighted by path weights that are unknown at the start. (*b*) Measured averages, like the average number of times a particle remained in state A ($\langle N_{aa} \rangle$), then determine those weights and thus the relative probabilities of all the paths. (*c*) The predicted variances, given the mean, agree with experiments. Panel *c* adapted with permission from Reference 16.

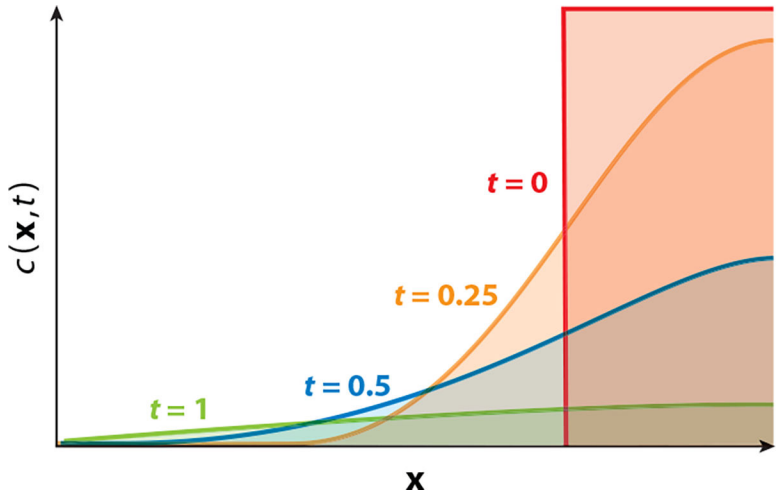

Figure 4.

Diffusion-equation modeling treats concentration $c(\mathbf{x}, t)$ as continuous and differentiable—for example, in Fick's law, $\langle \mathbf{J} \rangle = -D\partial c/\partial \mathbf{x}$.

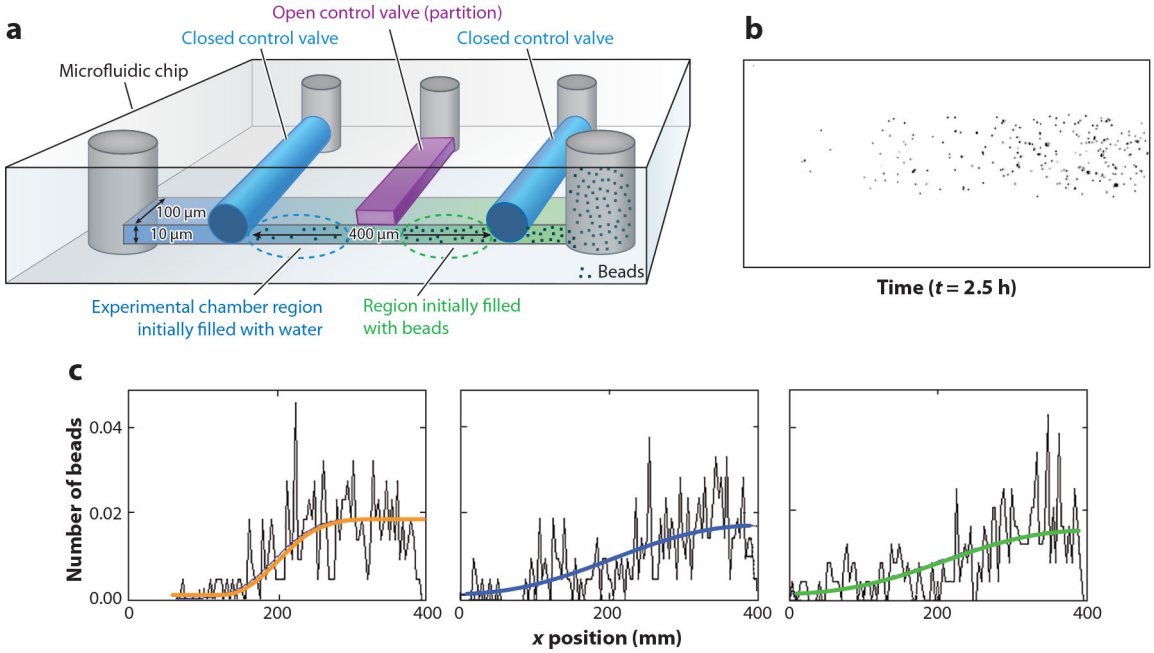

Figure 5.

Diffusion at the microscopic level of colloidal particles on a microscopic slide (18). (*a*) The microfluidic apparatus used to measure the free diffusion of a few colloidal beads to quantify fluctuations. (*b*) A snapshot from the video used to track the movements of the beads over time. (*c*) Three typical concentration profiles measured using the apparatus. These profiles indicate that few-particle flows such as the one measured in this experiment entail large fluctuations. Phenomenology like Fick's law just describes averages for large numbers of particles, and not the fluctuations indicated here. Figure adapted from Reference 18 with permission from the American Association of Physics Teachers.

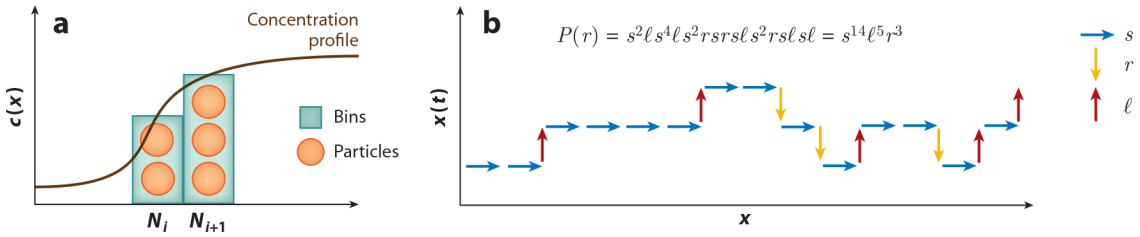

Figure 6.

(*a*) Concentration gradient $c(x)$ showing bins i and $i + 1$, illustrating the different numbers N of particles in each one. (*b*) One possible trajectory, labeled with the statistical weights of the steps.

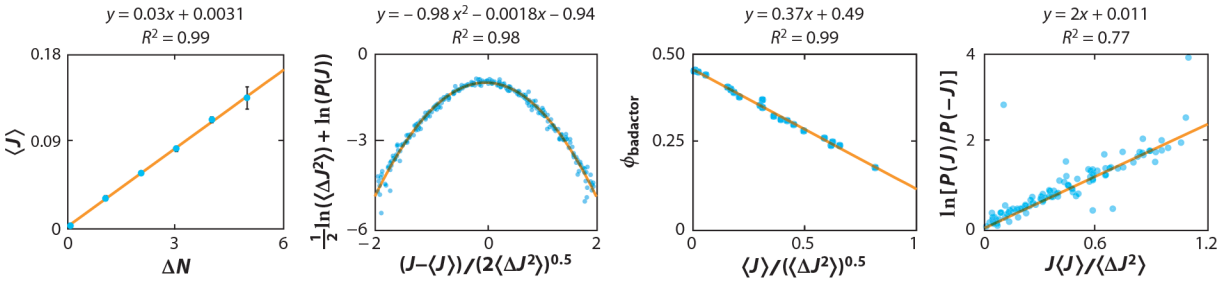

Figure 7.

Simple diffusion, from maximum caliber with the dog-flea model, successfully predicts few-particle experiments (19). First, it derives Fick's first law, $\langle J \rangle = -D\partial c/\partial \mathbf{x}$, from a variational principle. Second, it shows that Fick's law holds down to the few-particle limit. Third, it correctly predicts the full rate distribution. Fourth, it computes a Maxwell's demon–like quantity of wrong-way flows (quantified by ϕ_{badactor}), showing that they become negligible as the net flux gets larger. And fifth, it accurately gives a flux fluctuation relationship. All these predictions result when given only one quantity, equivalent to knowing the diffusion constant, D (see Reference 19 for more information). Figure adapted with permission from Reference 19.

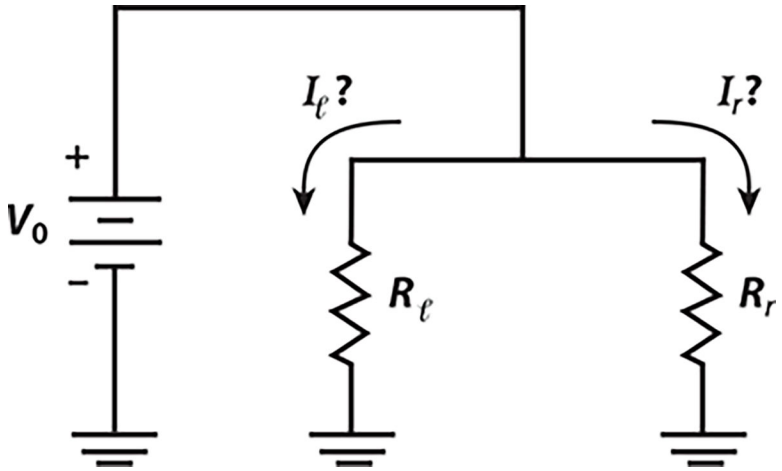

Figure 8.
Maximum caliber derives Kirchoff's current principle—namely, that current divides at junctions in proportion to the relative flow resistances of the channels.

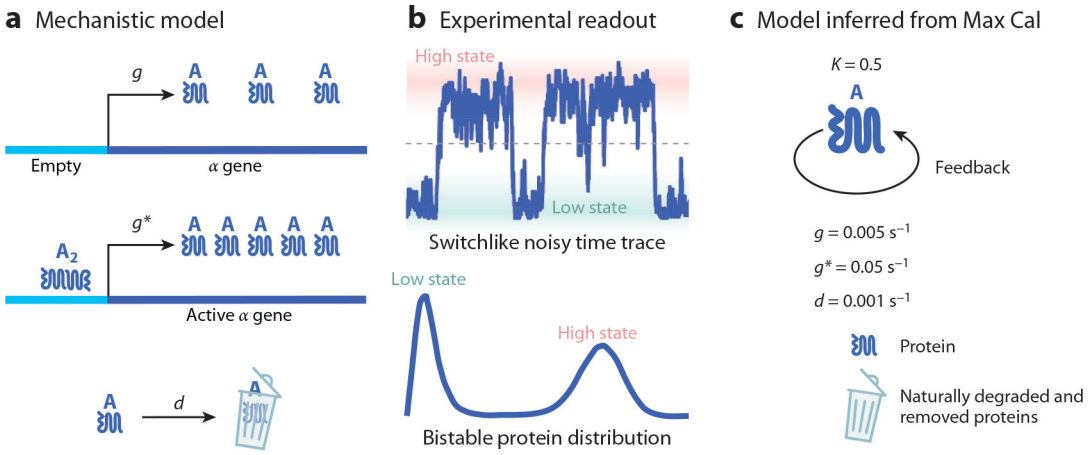

Figure 9.

Maximum caliber (Max Cal) predicts the dynamics of an autoactivation gene circuit. (*a*) Gene α produces protein A. When A_2 (dimer of A) binds to the promoter, production of protein A speeds up (20). Note that negative feedback can be readily treated similarly, except that the light blue region here called the promoter is replaced by a region called the repressor, and the effect of repression is to slow down, rather than speed up, production of A. (*b*) The experiment measures this stochastic switchlike trajectory. (*c*) Using this as input, Max Cal predicts protein production rates in the normal (g) and accelerated (g^*) states, as well as degradation rate d.

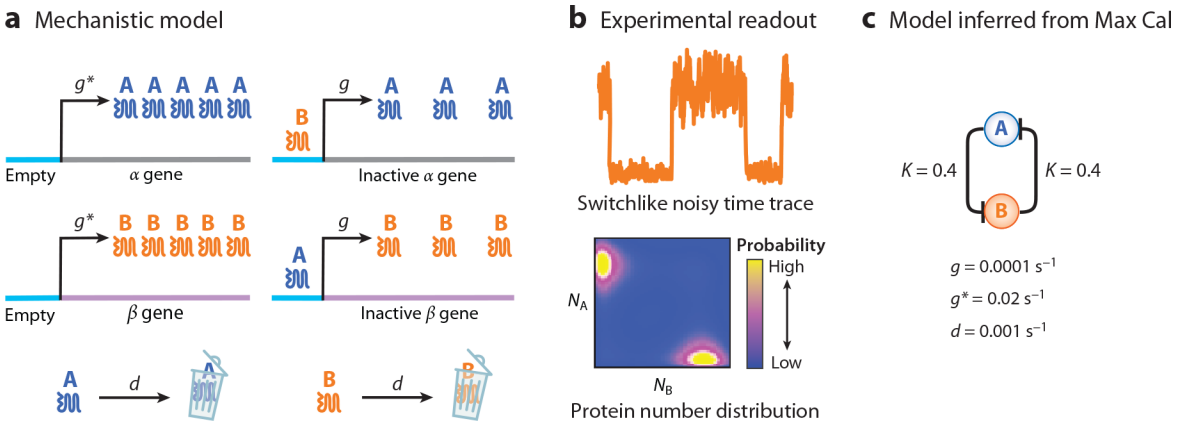

Figure 10.

Maximum caliber (Max Cal) gives rate distributions in a toggle-switch gene circuit. (*a*) Gene α produces protein A. Gene *β* produces protein B. The binding of A represses production of B, and the binding of B represses production of A. (*b*) The effect is bistability (winner-take-all): When either species, A or B, enters into small excess, it then grows further to fully dominate the population. The stochastic trajectory is experimentally measured. Bottom part of panel *b* adapted with permission from Reference 22. (*c*) In the absence of anything else, Max Cal uses the stochastic trajectory for both proteins to infer different rate parameters *g*, *g**, *d*, etc.

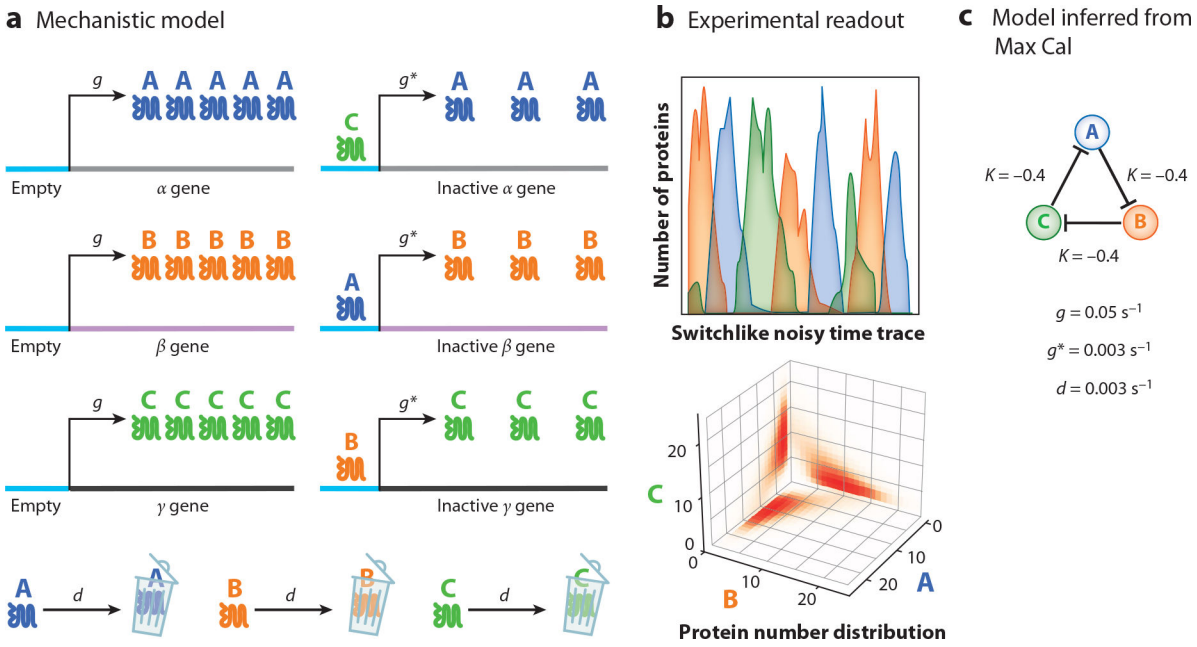

Figure 11.

Maximum caliber (Max Cal) gives the few-particle dynamics of the repressilator gene circuit. (*a*) Genes α, β, and γ produce proteins A, B, and C, respectively. The binding of A represses production of B, the binding of B represses production of C, and the binding of C represses production of A. (*b*) The effect is an oscillatory time trace. The distribution of populations of A, B, and C contains less information than the stochastic time trajectory, typically measured in experiments. (*c*) In the absence of anything else, Max Cal uses the stochastic trajectory for all three proteins to infer different rate parameters *g*, *g**, and *d* and feedback strength *K*. Panel *c* adapted with permission from Reference 26.

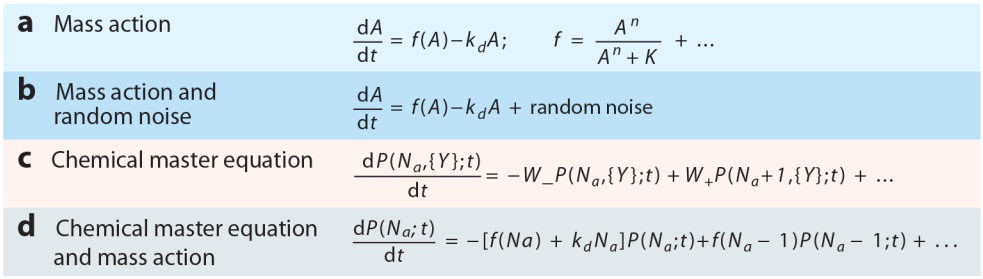

Figure 12.

Traditional dynamical models of gene networks. (*a*) Mass-action (MA) model describes averages (A) with arbitrary nonlinear function f to model feedback (k_d is the degradation rate). (*b*) MA + random noise model adds random noise around the MA equation, yielding a Langevin-type equation. (*c*) Chemical master equation (CME) model describes time evolution of probability distributions (P) in terms of transition probabilities (W) determined by invoking a set of auxiliary species ($\{Y\}$) that are often not seen in experiments. (*d*) CME + MA is a coarse-grained model for the time evolution of probability when substituting the phenomenological f functions used in MA for W.

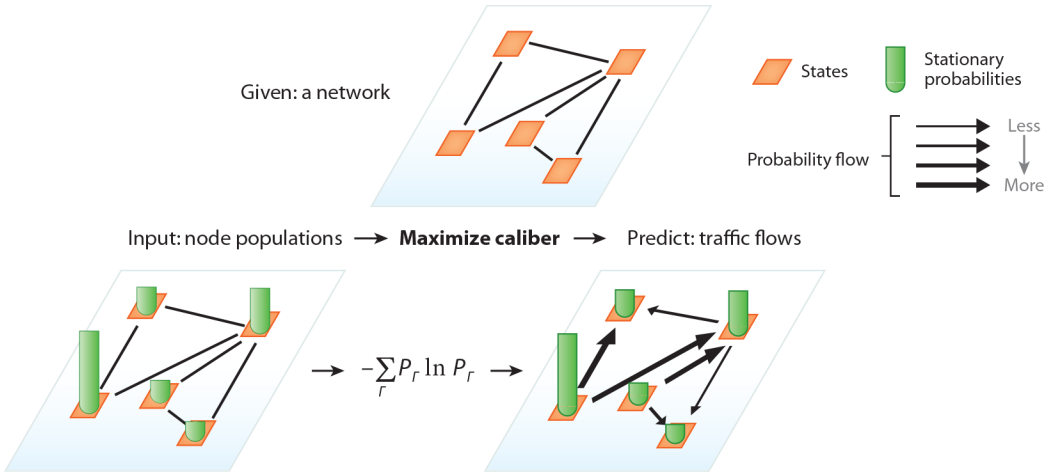

Figure 13.

Maximum caliber gives a fast estimate of traffic flows on networks. If we know the steady-state populations of the mobile elements that flow through a network, then the maximum caliber equation (Equation 25) estimates the full transition-rate matrix, as that which maximizes the path entropy.

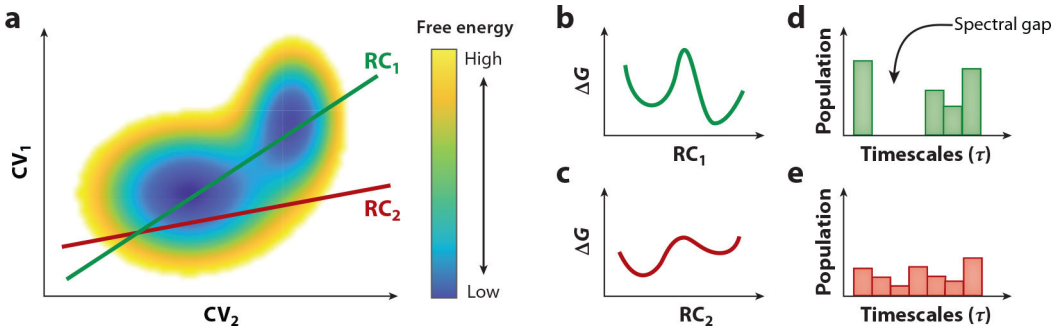

Figure 14.

Maximum caliber can identify good RCs. A complex potential energy landscape (panel *a*) can be projected onto any RC; ΔG is free energy (panels *b* and *c*). Maximum caliber allows us to quickly estimate the approximate kinetics along any RC and identify the RC that leads to the highest separation in timescales (panels *d* and *e*). In this simple example, RC_1 is a better reaction coordinate than RC_2.

Abbreviations: CV, collective variable; RC, reaction coordinate.

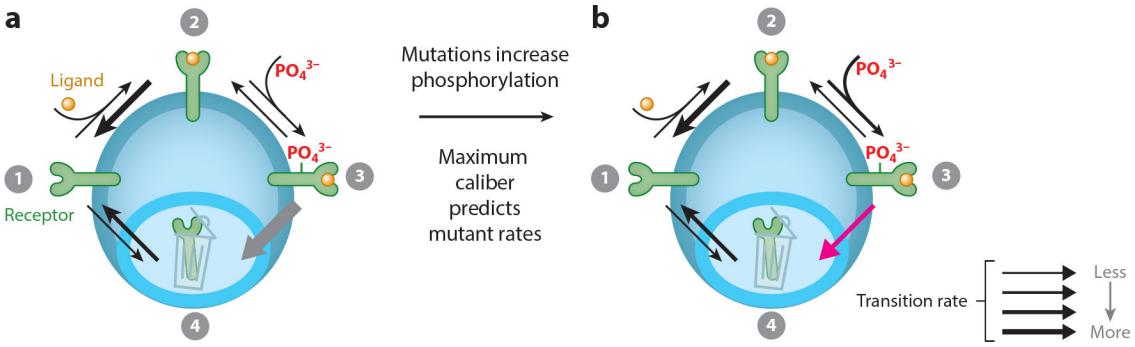

Figure 15.

Model of the growth factor–activated pathway. The four states (①–④) of the receptor are as follows. (①) Ligand-free receptors (*green*) on the cell surface can be (②) bound by a ligand (*yellow*) and then (③) phosphorylated. (④) All receptors can be internalized and degraded, albeit at different rates. The arrows indicate transition rates, and the most altered transition rate, as predicted by maximum caliber, is shown in (*a*) gray and (*b*) pink.