

Fluctuation relations

Matheus Schossler^{1,*}

¹Washington University in St. Louis

Fluctuations are important for systems with a small number of degrees of freedom and have a strong effect on their measurable physical properties. This points us in the direction that fluctuations could be important for macroscopic systems as well. By properly accounting for the fluctuations (noise), one can extract information between two equilibrium states out of an ensemble of non-equilibrium states. In this paper we rederive two important fluctuation relations discovered in the past decades that relate an equilibrium property, the Helmholtz free energy, with non-equilibrium state trajectories for both classical and quantum systems: the Jarzynski equality and the Crooks fluctuation theorem. Especially, the Jarzynski equality states that the difference of the free energy between an initial and final equilibrium state is directly related to the average of the irreversible work along an ensemble of, mostly non-equilibrium, trajectories joining these two states. Applications are discussed.

I. INTRODUCTION

In the last two centuries the laws of thermodynamics has successfully explained many physical process evolving a macroscopic number of variables. In special, when macroscopic systems are in contact with a thermal reservoir at temperature T , the second law of thermodynamics can be expressed as

$$W \geq \Delta F, \quad (1)$$

where W is the work performed on the system and $F = U - TS$ is the Helmholtz free energy, U is the internal energy and S is the entropy.

The inequality (1) states that the work delivered to a system in contact to a thermal reservoir is more than its Helmholtz free energy variation, i.e., it relates a process that could be far from equilibrium to an equilibrium property, the work and the free energy respectively of the macroscopic system. The equality holds when the work is done in a reversible process. For many years this was considered whole story.

However, when the number of degrees of freedom decrease, the fluctuations become more important and affect the measurable physical properties of the system. This fact point us in the direction that fluctuations could be important for macroscopic systems as well. A proper accounting of fluctuations makes it possible to write the free energy in terms of the information extracted from non-equilibrium [1, 2]. This notes will address two important relations discovered in the past decades that relates properties of non-equilibrium and equilibrium in both classical and quantum systems in terms of equalities: the Jarzynski equality (JE) [3] and the Crooks fluctuation theorem [4].

II. CLASSICAL SYSTEMS

In this section we are going to study fluctuation relations for classical systems [1, 5]. We will assume that the dynamics

of these systems is ruled by the Hamilton equations where the initial conditions is some point of the phase space $\mathbf{z} = (\mathbf{q}, \mathbf{p})$. The Hamiltonian of the system is a combination of an unperturbed part, $H_0(\mathbf{z})$, and a perturbation $-\lambda_t Q(\mathbf{z})$,

$$H(\mathbf{z}; \lambda_t) = H_0(\mathbf{z}) - \lambda_t Q(\mathbf{z}). \quad (2)$$

Where the function $\lambda_t = \lambda(t)$ can be understood as a time dependent external field or parameter which couples the unperturbed system with an observable $Q(\mathbf{z})$. As time flows from $t = 0$ to $t = \tau$ the system is driven by λ_t . These perturbed system is an example of a nonautonomous or driven systems. The way the system is driven depends on λ , which is also called the force protocol. For each protocol and initial point of the phase space $\mathbf{z}_0 = (\mathbf{q}_0, \mathbf{p}_0)$ the equations of motion determines another point of the phase space, on time t , given by

$$\mathbf{z}_t = \varphi_{t,0}[\mathbf{z}_0; \lambda], \quad (3)$$

where φ is a function of the initial condition (or initial state) and the protocol. This is represented in the figure (1).

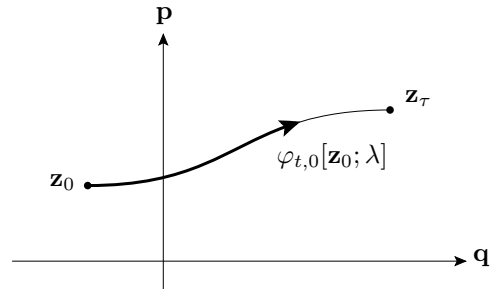


Figure 1: The initial state \mathbf{z}_0 from the phase space evolves according to equation of motions as t flows from $0 \rightarrow \tau$ under the protocol λ . Adapted from [5].

The system is assumed initially in a thermal equilibrium state with temperature $\beta = 1/k_B T$, so the probability, $\rho_{\lambda,T}(\mathbf{z})$, of the system to be at the point \mathbf{z} is given by the

* mschossler@wustl.edu

Boltzmann distribution,

$$\rho_{\lambda,T}(\mathbf{z}) = \frac{e^{-\beta H(\mathbf{z}, \lambda_t)}}{Z_{\lambda,T}}, \quad (4)$$

where $Z_{\lambda,T}$ is the canonical partition function,

$$Z_{\lambda,T} = \int d\mathbf{z} e^{-\beta H(\mathbf{z}, \lambda_t)}. \quad (5)$$

One can also calculate the free energy, $F_{\lambda_0,T}$, for the initial state from the partition function:

$$F_{\lambda_0,T} = -k_B T \ln(Z_{\lambda_0,T}). \quad (6)$$

As λ_t evolves over time, the external field inserts work in the system until $t = \tau$. At this time the system may be or maybe not in a equilibrium state depending whether or not the protocol λ varies slowly enough. In the static limit the system is always in equilibrium as t changes, but in general the intermediate states and the final state \mathbf{z}_τ are not in equilibrium. After $t = \tau$ the system relax to a equilibrium state with a defined free energy. The work delivered to the system by the protocol λ obeys the inequality (1), where

$$\Delta F = F_{\lambda_\tau,T} - F_{\lambda_0,T}. \quad (7)$$

The definition of work is used in this paper is

$$\begin{aligned} W[\mathbf{z}_0; \lambda] &= \int d\lambda_t \frac{\partial H(\mathbf{z}_t; \lambda_t)}{\partial \lambda_t} \\ &= \int_0^\tau dt \lambda_t \frac{\partial H(\mathbf{z}_t; \lambda_t)}{\partial \lambda_t} \\ &= H(\mathbf{z}_\tau; \lambda_\tau) - H(\mathbf{z}_0; \lambda_0) \end{aligned} \quad (8)$$

where $-\partial H(\mathbf{z}_t, \lambda_t) / \partial \lambda_t$ is a generalized force and $d\lambda_t$ is a differential displacement of the protocol λ . This definition of work, mentioned on literature[5] as “exclusive” work, considers the difference of the total Hamiltonian (unperturbed plus perturbed parts). In the other hand, the “inclusive” work considers the difference of the unperturbed part of the Hamiltonian only. In this paper we will use the “exclusive” work definition as work, as in the equation (8).

The work, $W[\mathbf{z}_0; \lambda]$, depends on the protocol, as mentioned, but also on the initial condition \mathbf{z}_0 . With that in mind, changes in the initial condition of the phase space implies variations of the work delivered to the system by the protocol λ , so one can calculate the average of the work or any function of the work by averaging over the all possible initial conditions using the probability distribution (4):

$$\langle f(W[\mathbf{z}_0; \lambda]) \rangle = \int d\mathbf{z}_0 \rho_{\lambda_0,T}(\mathbf{z}_0) f(W[\mathbf{z}_0; \lambda]). \quad (9)$$

In particular, one can also find the average of the exponential

of $-\beta W[\lambda]$ over the initial conditions,

$$\begin{aligned} \langle e^{-\beta W} \rangle &= \int d\mathbf{z}_0 \rho_{\lambda_0,T}(\mathbf{z}_0) e^{-\beta W} \\ &= \frac{1}{Z_{\lambda_0,T}} \int d\mathbf{z}_0 e^{-\beta H(\mathbf{z}_0, \lambda_0)} e^{-\beta W} \\ &= \frac{1}{Z_{\lambda_0,T}} \int d\mathbf{z}_\tau \left| \frac{d\mathbf{z}_0}{d\mathbf{z}_\tau} \right| e^{-\beta H(\mathbf{z}_\tau, \lambda_\tau)}, \end{aligned} \quad (10)$$

Using the reversibility property of the system, which is assured by the Hamilton equations, the function (3) can be inverted to express \mathbf{z}_0 as a function of \mathbf{z}_τ and the protocol in the backward direction. Because the Hamilton equations are canonical transformations, the Jacobian, $\left| \frac{d\mathbf{z}_0}{d\mathbf{z}_\tau} \right|$, is one, as a consequence

$$\begin{aligned} \langle e^{-\beta W} \rangle &= \frac{1}{Z_{\lambda_0,T}} \int d\mathbf{z}_\tau e^{-\beta H(\mathbf{z}_\tau, \lambda_\tau)} \\ &= \frac{Z_{\lambda_\tau,T}}{Z_{\lambda_0,T}}, \end{aligned} \quad (11)$$

so we get

$$\langle e^{-\beta W} \rangle = e^{-\beta \Delta F}, \quad (12)$$

this result is due Jarzynski, 1997 [3], and is called Jarzynski equality. This equality states that an equilibrium property, which is the difference of the free energies from the initial and final equilibrium states, can be found by repeating the forced experiment (through the protocol λ) many times.

III. QUANTUM SYSTEMS

Since we have a fluctuation relation for classical systems, as derived before in section II, it is natural to also think about analogies for quantum systems. In this section we will find the Crooks fluctuation theorem and the Jarzynski equality for quantum systems. Let us consider a quantum system with a time-dependent Hamiltonian $\mathcal{H}(\lambda_t)$ similar to that of equation (2) with a instantaneous basis $|\psi_{n,\gamma}^{\lambda_t}\rangle$

$$\mathcal{H}(\lambda_t) |\psi_{n,\gamma}^{\lambda_t}\rangle = E_n^{\lambda_t} |\psi_{n,\gamma}^{\lambda_t}\rangle, \quad (13)$$

where $E_n^{\lambda_t}$ are the eigenvalues of the Hamiltonian at time t , the index γ labels the g_n -fold degeneracy subspace of states with energy $E_n^{\lambda_t}$, and λ_t is the value of the protocol λ at time t . Before $t = 0$ the system is assumed to be in a thermal equilibrium state

$$\varrho(\lambda_0) = \frac{e^{-\beta \mathcal{H}(\lambda_0)}}{\mathcal{Z}(\lambda_0)}, \quad (14)$$

also known as the density operator for the quantum canonical ensemble. Because any average of an observable can be calculated using the density operator, one can call it the state

of the system. Then, at $t = 0$, the Hamiltonian is measured $E_n^{\lambda_0}$ with probability

$$p_n^0 = g_n \frac{e^{-\beta E_n^{\lambda_0}}}{\mathcal{Z}(\lambda_0)}. \quad (15)$$

After this measurement the system collapses in the state

$$\begin{aligned} \varrho_n &= \sum_{\gamma} |\psi_{n,\gamma}^{\lambda_0}\rangle \langle \psi_{n,\gamma}^{\lambda_0}| \\ &= \frac{\Pi_n^{\lambda_0} \varrho(\lambda_0) \Pi_n^{\lambda_0}}{p_n^0}, \end{aligned} \quad (16)$$

where $\Pi_n^{\lambda_0} = \sum_{\gamma} |\psi_{n,\gamma}^{\lambda_0}\rangle \langle \psi_{n,\gamma}^{\lambda_0}|$ is the projector onto the subspace spanned by the eigenstates of $\mathcal{H}(\lambda_0)$ with eigenvalue $E_n^{\lambda_0}$.

Assuming the system is isolated after the first measurement, the system evolves by a unitary transformation

$$\varrho_n(t) = U_{t,0}[\lambda] \varrho_n U_{t,0}^\dagger[\lambda], \quad (17)$$

where $U_{t,0}[\lambda]$ is the evolution operator under the action of a protocol λ . At time τ the Hamiltonian is measured again yielding the eigenvalue $E_m^{\lambda_\tau}$ with conditional probability

$$\begin{aligned} p_{m|n}[\lambda] &= \sum_{\gamma} \langle \psi_{m,\gamma}^{\lambda_\tau} | \varrho_n(\tau) | \psi_{m,\gamma}^{\lambda_\tau} \rangle \\ &= \text{Tr} \Pi_m^{\lambda_\tau} \varrho_n(\tau). \end{aligned} \quad (18)$$

We will make a pause in the derivation of fluctuation relations to briefly introduce the concept of probability density function (PDF). Let X be a stochastic variable with probability p_i for each realization x_i . The PDF is a function whose value at any given sample in the sample space can be interpreted as providing the probability p of x to occur. The PDF of X is defined as

$$P_X(x) = \sum_i p_i \delta(x - x_i). \quad (19)$$

This PDF, of a discrete stochastic variables X , is interesting because allows one to represent a discrete probability distribution in terms of a continuous distribution, and it carries all statistical information about X .

The quantum version of work is similar to (8), but since work is not an observable one cannot be represented directly in terms of a single observable $\mathcal{H}(\lambda)$. Instead it must be calculated from the eigenvalues of two copies of $\mathcal{H}(\lambda)$. The realizations of the stochastic variable work, W , are

$$w[n \rightarrow m; \lambda] = E_m^{\lambda_\tau} - E_n^{\lambda_0}. \quad (20)$$

Although we can express the realizations of W in terms of the eigenvalues of $\mathcal{H}(\lambda)$, it is not right to say that $W = \mathcal{H}(\lambda_\tau) - \mathcal{H}(\lambda_0)$ because states evolves on time in the Schrodinger picture, in addition to the time dependence evolution of the Hamiltonian due the time dependent protocol λ_t .

One must also includes the Schrodinger evolution of states in the definition of W . However, formulating the Hamiltonian in the Heisenberg dynamical picture the states no longer change in time. Therefore the natural definition of W is

$$W = \mathcal{H}^H(\lambda_\tau) - \mathcal{H}^H(\lambda_0), \quad (21)$$

where the superscript H denotes Heisenberg picture,

$$\mathcal{H}_\tau^H(\lambda_\tau) = U_{\tau,0}^\dagger[\lambda] \mathcal{H}(\lambda_\tau) U_{\tau,0}[\lambda], \quad (22)$$

which coincides with the Schrodinger picture only for $\tau = 0$. The two energy measurements of the Hamiltonian are random variables, as it is for any observable. As a consequence the PDF for the work W can be found through the generalization of equation (19) for two variables,

$$P_W[w; \lambda] = \sum_{m,n} \delta(w - [E_m^{\lambda_\tau} - E_n^{\lambda_0}]) p_{m|n}[\lambda] p_n^0, \quad (23)$$

where $p_{m|n}[\lambda] p_n^0$ is the probability to obtain both $E_n^{\lambda_0}$ and $E_m^{\lambda_\tau}$. Other useful definition is the characteristic function (CF)

$$G[u; \lambda] = \int dw e^{iuw} P_W[w; \lambda], \quad (24)$$

which is the Fourier transform of $P_W[w; \lambda]$. Now we are going to show a important relation for this CF:

$$G[u; \lambda] = \langle e^{iu\mathcal{H}_\tau^H(\lambda_\tau)} e^{-iu\mathcal{H}(\lambda_0)} \rangle \quad (25)$$

Let us start from the definition (24) and equation (23)

$$\begin{aligned} G[u; \lambda] &= \sum_{m,n} \int dw e^{iuw} \delta(w - [E_m^{\lambda_\tau} - E_n^{\lambda_0}]) p_{m|n}[\lambda] p_n^0 \\ &= \sum_{m,n} e^{iu[E_m^{\lambda_\tau} - E_n^{\lambda_0}]} p_{m|n}[\lambda] p_n^0 \end{aligned}$$

using the definition of $p_{m|n}[\lambda]$ we can rewrite the equation above as

$$\begin{aligned} G[u; \lambda] &= \sum_{m,n} e^{iu[E_m^{\lambda_\tau} - E_n^{\lambda_0}]} \sum_{\gamma} \langle \psi_{m,\gamma}^{\lambda_\tau} | U_{t,0}[\lambda] \\ &\quad \times \Pi_n^{\lambda_0} \varrho(\lambda_0) \Pi_n^{\lambda_0} U_{t,0}^\dagger[\lambda] | \psi_{m,\gamma}^{\lambda_\tau} \rangle \\ &= \sum_{m,n,\gamma} \langle \psi_{m,\gamma}^{\lambda_\tau} | U_{t,0}[\lambda] e^{-iu\mathcal{H}(\lambda_0)} \\ &\quad \times \Pi_n^{\lambda_0} \varrho(\lambda_0) \Pi_n^{\lambda_0} U_{t,0}^\dagger[\lambda] e^{iu\mathcal{H}(\lambda_\tau)} | \psi_{m,\gamma}^{\lambda_\tau} \rangle, \end{aligned} \quad (26)$$

replacing the sum over $\{m, \gamma\}$ by the trace, and using the

relation

$$\begin{aligned} \sum_n \Pi_n^{\lambda_0} \varrho(\lambda_0) \Pi_n^{\lambda_0} &= \sum_n \Pi_n^{\lambda_0} \frac{e^{-\beta \mathcal{H}(\lambda_0)}}{\mathcal{Z}(\lambda_0)} \Pi_n^{\lambda_0} \\ &= \sum_n \frac{e^{-\beta E_n^{\lambda_0}}}{\mathcal{Z}(\lambda_0)} \Pi_n^{\lambda_0} \\ &= \frac{e^{-\beta \mathcal{H}(\lambda_0)}}{\mathcal{Z}(\lambda_0)} = \varrho(\lambda_0), \end{aligned} \quad (27)$$

one can write

$$G[u; \lambda] = \text{Tr } U_{t,0}[\lambda] e^{-iu\mathcal{H}(\lambda_0)} \varrho(\lambda_0) U_{t,0}^\dagger[\lambda] e^{iu\mathcal{H}(\lambda_t)}. \quad (28)$$

The cyclic property of the trace, and equation (22) allow us to rewrite this equation as

$$\begin{aligned} G[u; \lambda] &= \text{Tr } \varrho(\lambda_0) U_{t,0}^\dagger[\lambda] e^{iu\mathcal{H}(\lambda_t)} U_{t,0}[\lambda] e^{-iu\mathcal{H}(\lambda_0)} \\ &= \text{Tr } \varrho(\lambda_0) e^{iu\mathcal{H}_t^\dagger(\lambda_t)} e^{-iu\mathcal{H}(\lambda_0)} \end{aligned} \quad (29)$$

proving that a CF can be written as a quantum correlation function (25). This result was first shown by Talkner [6], 2007.

The last supposition is that $\mathcal{H}(\lambda)$ is invariant under time reversal, as a consequence the quantum micro-reversibility for nonautonomous (driven) systems holds:

$$U_{t,\tau}[\lambda] = \Theta^\dagger U_{\tau-t,0}[\tilde{\lambda}] \Theta, \quad (30)$$

where Θ is the time reversal operator, and $\tilde{\lambda}$ is the protocol λ in the time backward direction. Using the previous equation and properties of the trace used before to find (29), one can show

$$\mathcal{Z}(\lambda_0) G[u; \lambda] = \mathcal{Z}(\lambda_\tau) G[-u + i\beta; \tilde{\lambda}]. \quad (31)$$

Remembering the definition (24) we get

$$\begin{aligned} \mathcal{Z}(\lambda_0) \int dw e^{i u w} P_W[w; \lambda] &= \\ &= \mathcal{Z}(\lambda_\tau) \int dw e^{-i u w - \beta w} P_W[w; \tilde{\lambda}], \end{aligned} \quad (32)$$

replacing the quantum canonical partition function, $\mathcal{Z}(\lambda_t) = \text{Tr } e^{-\beta \mathcal{H}(\lambda_t)} = e^{-\beta F(\lambda_t)}$, one finally proves the quantum version of the Crooks fluctuation theorem:

$$\frac{P_W[w; \lambda]}{P_W[-w; \tilde{\lambda}]} = e^{\beta(w - \Delta F)}. \quad (33)$$

This result relates that the PDF for work to an equilibrium property, the free energy. It was first discovered for classical systems by Crooks, in 1998[4], then found for quantum systems by Tasaki[7] and Kurchan[8], in 2000. From (33) we have

$$\int dw P_W[w; \lambda] e^{-\beta w} = e^{-\beta \Delta F} \int dw P_W[-w; \tilde{\lambda}]. \quad (34)$$

The integrals on the left hand and right hand side are the average of the exponent and one, respectively. This implies the Jarzynski equality for quantum systems,

$$\langle e^{-\beta w} \rangle_\lambda = e^{-\beta \Delta F}. \quad (35)$$

This result is very similar as the one for classical systems from section II: the free energy of equilibrium quantum states can be found by averaging the work delivered to the system through the protocol λ on all possible paths which is largely covered by non-equilibrium states.

IV. APPLICATIONS

The first application of Jarzynski equality is naturally a verification of second law of thermodynamics. This also also give us confidence about the validity of this fluctuation relation. Using the Jensen's inequality,

$$\langle e^X \rangle \geq e^{\langle X \rangle}, \quad (36)$$

in the equation (12) we find

$$\langle W \rangle \geq \Delta F, \quad (37)$$

which is the average of the equation (1), proving the second law.

Now we are going to show an application from Jarzynski equality to rederive the drag coefficient for the Brownian motion, one of Einstein's first major scientific contributions. We will follow the paper by Gittes[9], 2017. Let us suppose, in addition to the drag force and the random force, ξ , due collisions, a constant force, f , applied to a particle in a Brownian movement. In each process the protocol is one for t from 0 to τ , and zero otherwise. So the Langevin equation becomes

$$m \frac{dv(t)}{dt} = -\gamma v(t) - \xi + \lambda f. \quad (38)$$

Therefore, the work of the external force is

$$W = f \Delta x, \quad (39)$$

where Δx is the displacement of the particle. Then we repeat the process in order to calculate the average in the Jarzynski equality.

We will assume that the extra force is totally dissipated by the drag friction, so there is no change of free energy of the system between $t = 0$ and $t = \tau$,

$$\Delta F = 0, \quad (40)$$

and

$$f = \gamma \bar{v}. \quad (41)$$

Under these assumptions the Jarzynski equality tell us

$$\langle e^{-\beta W} \rangle = 1. \quad (42)$$

In the regime of a small force f we should expect have a good approximation to the above average in terms of the first terms of the cumulant expansion

$$\langle e^{-\beta W} \rangle = e^{-\beta \bar{W} + (\beta^2/2) \overline{\Delta W^2} + \dots} \quad (43)$$

The first term is just

$$\bar{W} = f \bar{\Delta x}. \quad (44)$$

The second is the variance $\overline{\Delta W^2} = f^2 \overline{\Delta x^2}$. For f small we can assume a free Brownian motion because the displacement due the extra force, $\bar{\Delta x}$, is much less than $\sqrt{\overline{\Delta x^2}}$ (see figure 2)

$$\overline{\Delta x^2} = 2D\tau. \quad (45)$$

With this approximation

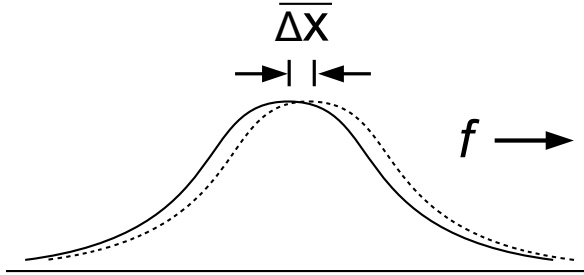


Figure 2: The displacement $\bar{\Delta x}$ is much less than the root-means-square of the position of a the particle. Figure taken from Gittes[9].

$$e^{-\beta \bar{W} + (\beta^2/2) \overline{\Delta W^2}} \approx 1, \quad (46)$$

what implies

$$\overline{\Delta W^2} \approx 2k_B T \bar{W}. \quad (47)$$

Inserting (44), and (45), we have

$$2Df^2\tau = 2k_B T f \bar{\Delta x}, \quad (48)$$

Using (41), we obtain

$$D = \frac{k_B T}{\gamma}. \quad (49)$$

Which is the diffusion coefficient for a particle in a Brownian motion.

CONCLUSION

In this notes we have seen the importance of fluctuations and how it connects with equilibrium properties. Taking into account the trajectories of both classical and quantum systems over non-equilibrium states we were able to connect an intrinsic none-equilibrium quantity, the work W , with an equilibrium one, the variation of the Helmholtz free energy. This results overcome the second law of thermodynamics, in the sense that they are equalities relating equilibrium and none-equilibrium, instead of an inequality, respectively. We have proved the Jarzynski equality for classical and quantum systems, and the Crooks fluctuation theorem for quantum systems. Although, the demonstration of the Crooks relation for classical systems does not require other concepts than those written in this paper and can be found in the various references.

In addition we also have shown a directly application of the classical Jarzynski equation to the Brownian movement of a particle, which was the first example of fluctuation-dissipation relation derived. The JE allowed us to calculate the diffusion coefficient for this system. In the last decades many other applications of fluctuation relations, including such the Jarzynski equality, have been found: foundations of machine learning [10], geophysical fluid dynamics with turbulent fluctuation (air-sea interaction) [11], single-molecule experiments, detection of quantum entanglement, combinatorial optimization, etc [1]. In 10 the link between a measurable quantities and hyperparameters in the stochastic gradient descent algorithm is made by stationary fluctuation-dissipation relations. Because the Jarzynski equality makes use of highly none-equilibrium trajectories in its core, we hypothesize that it could be used to deeper our understanding on algorithms that takes longer training to learn which requires large data sets, e.g., deep learning algorithms.

-
- [1] C. Jarzynski, Annual Review of Condensed Matter Physics **2**, 329 (2011).
 - [2] E. Sevick, R. Prabhakar, S. R. Williams, and D. J. Searles, Annual Review of Physical Chemistry **59**, 603 (2008), arXiv:0709.3888.
 - [3] C. Jarzynski, Physical Review Letters **78**, 2690 (1997), arXiv:9610209 [cond-mat].
 - [4] G. E. Crooks, Journal of Statistical Physics **90**, 1481 (1998).
 - [5] M. Campisi, P. Hänggi, and P. Talkner, Reviews of Modern Physics **83**, 771 (2011), arXiv:1012.2268.
 - [6] P. Talkner, E. Lutz, and P. Hänggi, Physical Review E -

- Statistical, Nonlinear, and Soft Matter Physics **75**, 7 (2007), arXiv:0703189 [cond-mat].
- [7] H. Tasaki, (2000), arXiv:0009244 [cond-mat].
- [8] J. Kurchan, , 1 (2000), arXiv:0007360 [cond-mat].
- [9] F. Gittes, American Journal of Physics **86**, 31 (2018), arXiv:1704.07805.
- [10] S. Yaida, 7th International Conference on Learning Representations, ICLR 2019 , 1 (2018), arXiv:1810.00004.
- [11] A. Wirth and F. Lemarié, Earth System Dynamics **12**, 689 (2021).