

Quantifying errors in the Jarzynski estimator

Sivaraman Rajaganapathy^{✉,*,†} Cailong Hua^{✉,*‡} and Murti Salapaka^{✉§}

Department of Electrical and Computer Engineering, *University of Minnesota* Twin Cities, Minneapolis, Minnesota 55455, USA



(Received 4 May 2025; revised 25 November 2025; accepted 30 January 2026; published 26 March 2026)

Free energy is a crucial physical quantity in a wide range of domains, including systems at the nanoscale. The Jarzynski equality computes the equilibrium free energy difference between two states using the measurements of nonequilibrium work needed to move the system between the states. The Jarzynski equality uses the mean of the exponentiated work and thus holds true only in the limit of infinite data size. There are only a few studies estimating errors when employing the Jarzynski equality with a finite number of measurements; further, these estimates remain limited in scope. We provide nonrestrictive and rigorous confidence bounds on the free energy difference computable from samples of measured work without the need for any prior information on the work distribution. These bounds are applicable to a large variety of experimental conditions that take a system far from the equilibrium. The analytical claims are corroborated using simulations and experiments utilizing a toolbox developed in Python.

DOI: [10.1103/mkch-jb15](https://doi.org/10.1103/mkch-jb15)

I. INTRODUCTION

Nonequilibrium processes are commonplace in a wide range of systems in physics, chemistry, and biology. They frequently provide a convenient and often the only means to study the thermodynamic and equilibrium properties of systems at the small scale [1]. For example, the equilibrium free energy differences between states are used for characterizing many important biomolecular systems [2]. Molecular motors, such as kinesin and dynein, are known to operate in the nonequilibrium regime to convert chemical free energy from ATP molecules to useful work in carrying intracellular cargo [3–5]. Such cellular transport phenomena play a vital role in understanding diseases such as cancers, deafness, and several neurodegenerative disorders [6]. Similarly, the nanomechanics of shock-absorbing muscle proteins, such as dystrophin and utrophin, are studied via single-molecule force spectroscopy. Single-molecule force spectroscopy, while providing a fast and convenient experimental method for studying such molecules, excites them far from equilibrium [7]. Here, characterizing the equilibrium free energy landscape is crucial for unraveling the mechanisms of diseases such as Duchenne muscular dystrophy caused by the malfunctioning or absent muscle protein, dystrophin [8,9].

The Jarzynski equality links the equilibrium free energy differences between two states of a system to the nonequilibrium work required to move the system between the states [10]. The equality is given by

$$e^{-\beta \Delta F} = \langle e^{-\beta W} \rangle, \quad (1)$$

where ΔF is the free-energy difference between two states of interest, and $\beta := \frac{1}{k_B T}$ with Boltzmann constant k_B and the temperature T . The angular brackets $\langle \cdot \rangle$ denote an ensemble average obtained in the limit of an infinite number of samples. The Jarzynski estimator employing a finite number of work samples, W_i , $i = 1, 2, \dots, N \in \mathbb{N}$, is given by

$$\Delta \hat{F}_N := -\frac{1}{\beta} \ln \left[\frac{1}{N} \sum_{i=1}^N e^{-\beta W_i} \right], \quad (2)$$

which yields an approximate estimate of the free energy difference ΔF with an error, $\varepsilon_N := \Delta F - \Delta \hat{F}_N$, between the true free energy difference and the estimate, when the number of samples is N .

One of the earliest results in quantifying the quality of the Jarzynski estimator is presented in [11], wherein the bias and variance errors in the Jarzynski estimates are studied under the assumption that the work samples are taken from a Gaussian distribution; such an assumption holds primarily under near-equilibrium conditions. In [11], a numerical simulation study was used to develop empirical formulas for the bias and variance errors that depended on the knowledge of the dissipative work. In [12], the analysis of bias in the Jarzynski estimator was extended to broader classes of work distributions with decaying tails. However, their method provided estimates of the mean bias rather than statistical confidence bounds, and it relied on the accurate fitting of the tail parameters of the work distribution. The authors of [13] and [14] establish that the number of samples required for a desired accuracy increases exponentially with an increase in system size and in proportion to the average of the dissipated work. The investigation in [15] refines the analysis and provides an approximate lower bound on the percentage error in the estimate. However, an upper bound is not provided.

The analysis in [16] proposes a maximum likelihood-based technique to improve the convergence rate of the estimates

*These authors contributed equally to this work.

†Contact author: rajag038@alumni.umn.edu

‡Contact author: hua00023@umn.edu

§Contact author: murtis@umn.edu

when the distribution of the work is known and Gaussian. In [14] and [17], the use of experiments that move the system from an initial state A to a final state B and from an initial state B to a final state A are shown to improve the efficiency of the free energy calculations. Further, [17] suggests strategies, such as the use of diagnostic plots, the Bennett acceptance ratio [18], and the use of intermediate states to increase the robustness of the computations. In cases where free energies are computed via simulation studies, [19] gives a resampling technique in which some trajectories are eliminated and replaced with clones, whose work values have a greater impact on the final result.

The state-of-the-art work referenced above leaves considerable gaps in quantifying the error of the Jarzynski estimator. Here, the shortcomings of the existing work include: (1) assumptions of near-equilibrium conditions [11], (2) the prerequisite of special experimental conditions that lead to Gaussian work distributions [16], or (3) the information on experimental variables, such as the need for the true free energy difference [14,15]. While the experimental estimate can be substituted in the formulations of bounds that require the true free energy difference, the resulting approximation introduces errors that are generally hard to quantify. Plenty of systems undergo nonequilibrium state transitions that do not fall under these restrictive criteria [1]. For example, [20] reports asymmetric non-Gaussian work distributions for atomically flat surfaces undergoing friction when away from equilibrium. Recently, [21] observed for the first time, gamma-distributed work when studying adsorption free energies of the A3 peptide on a gold surface using nonequilibrium single-molecule pulling experiments.

To overcome the limitations in the existing literature and to meet the challenges involved, we provide a comprehensive method to compute the confidence error bounds on the estimates from a sample-limited Jarzynski estimator. Our method: (1) is applicable on far from equilibrium experiments, (2) can operate on a wide range of unknown work distributions, (3) has tunable confidence levels, and (4) depends only on the experimentally measured quantities, i.e., the measured work done.

This article is organized as follows: A rigorous analytical framework for quantifying the errors produced in computing free energy differences as a result of having a finite amount of data is presented in Sec. II. Two methods for determining confidence bounds with provable guarantees and a technique to combine the methods to leverage their strengths are described in Sec. III. The methods are empirically verified and experimentally validated in Sec. IV. The interpretation of the results and future directions of this work are given in Sec. V.

II. THE JARZYNSKI ESTIMATOR

The Jarzynski equality developed in [10,22,23] translates an ensemble of nonequilibrium measurements of work done to move a system between two states to the equilibrium free energy differences between the states. In this section, we briefly review the Jarzynski equality and state our assumptions. We then show that the ensemble of work involved is independent and identically distributed (i.i.d). This provides us with the framework to develop the confidence bounds.

Consider two Hamiltonian systems $\mathcal{A}$ and $\mathcal{B}$. Here, $\mathcal{A}$ represents a system of interest and $\mathcal{B}$ represents the environment. The primary focus will be on the system $\mathcal{A}$. Let $X_{\mathcal{A}}(t) := (q_{\mathcal{A}}(t), p_{\mathcal{A}}(t))$ be the states of the system $\mathcal{A}$, where $q_{\mathcal{A}}(t) \in \mathbb{R}^K$, $K \in \mathbb{N}$, are the generalized coordinates and $p_{\mathcal{A}}(t) \in \mathbb{R}^K$ are the generalized momenta of $\mathcal{A}$, and t represents time. Let $\lambda(t) \in \mathbb{R}$ be an externally controlled variable that can affect the states $X_{\mathcal{A}}(t)$. We denote a trajectory using $\{\cdot\}$; for example, $\{X_{\mathcal{A}}\}_0^\tau := \{X_{\mathcal{A}}(t) | 0 \leq t \leq \tau\}$ represents the trajectory of all the states $X_{\mathcal{A}}(t)$ from $t = 0$ to $t = \tau$. We represent the Hamiltonian by $H_{\mathcal{A}}(q_{\mathcal{A}}(t), p_{\mathcal{A}}(t), \lambda(t)) : \mathbb{R}^{2K+1} \rightarrow \mathbb{R}$. We make the following assumption about the two systems $\mathcal{A}$ and $\mathcal{B}$.

Assumption 1. (a) The energy in system $\mathcal{A}$ is finite.

(b) Energy absorptions or donations from $\mathcal{A}$ to $\mathcal{B}$ do not change the temperature $T_{\mathcal{B}}$ of $\mathcal{B}$.

(c) The dynamics of $\mathcal{A}$ and $\mathcal{B}$ are Markovian, i.e., $p(X_{\mathcal{A}}(t+dt) | \{X_{\mathcal{A}}\}_0^t) = p(X_{\mathcal{A}}(t+dt) | X_{\mathcal{A}}(t))$ for all t .

Note. While Assumption 1 provides a convenient framework for our analysis, some of its conditions may be relaxed or replaced for the Jarzynski equality to hold.

Now consider an experiment wherein $\mathcal{A}$ and $\mathcal{B}$ start at thermal equilibrium at $t = 0$, and are at a temperature $T_{\mathcal{A}} = T_{\mathcal{B}} = T$. Subsequently, the system $\mathcal{A}$ is perturbed away from the equilibrium by a protocol in which $\lambda(t)$ is varied in a fixed trajectory $\{\lambda\}_0^\tau$ from $\lambda(0) = u$ to $\lambda(\tau) = v$. During this perturbation when $0 \leq t \leq \tau$, the temperature of $\mathcal{A}$ is free to be different from that of $\mathcal{B}$, i.e., $T_{\mathcal{A}} \neq T_{\mathcal{B}} = T$. After this perturbation, the system $\mathcal{A}$ is equilibrated with $\mathcal{B}$, i.e., at $t \gg \tau$, $T_{\mathcal{A}} = T_{\mathcal{B}} = T$. At $t = 0$, when $\mathcal{A}$ is in thermal equilibrium with $\mathcal{B}$, the probability of finding the system $\mathcal{A}$ in a state $X_{\mathcal{A}}(0) = x_{\mathcal{A}}^0$ is given by $p(X_{\mathcal{A}}(0) = x_{\mathcal{A}}^0) = \frac{e^{-\beta H_{\mathcal{A}}(x_{\mathcal{A}}^0, \lambda(0))}}{Z}$, where $Z := \int e^{-\beta H_{\mathcal{A}}(X_{\mathcal{A}}, \lambda(0))} dX_{\mathcal{A}}$ is the partition function and $\beta = \frac{1}{k_B T}$. The change in $H_{\mathcal{A}}(X_{\mathcal{A}}(t), \lambda(t))$ due to a change $d\lambda$ is defined as the work done on the system and is given by

$$dW := d\lambda \frac{\partial H_{\mathcal{A}}(X_{\mathcal{A}}, \lambda)}{\partial \lambda}.$$

The work done on the system over the entire trajectory $\{\lambda\}_0^\tau$, is given by

$$W = \int_0^\tau dW = \int_0^\tau d\lambda \frac{\partial H_{\mathcal{A}}(X_{\mathcal{A}}, \lambda)}{\partial \lambda} dt.$$

We now consider N independent and identical realizations of this experiment, where

(1) The systems $\mathcal{A}_i$ and $\mathcal{B}_i$ are identical to $\mathcal{A}$ and $\mathcal{B}$, respectively, for every $i = 1, 2, \dots, N$.

(2) The system $\mathcal{A}_i$ is in thermal equilibrium with $\mathcal{B}_i$ at temperature T at $t = 0$, for every $i = 1, 2, \dots, N$.

(3) The parameter λ_i is varied from $\lambda_i(0) = u$ to $\lambda_i(\tau) = v$ along the same trajectory $\{\lambda\}_0^\tau$. That is,

$$\{\lambda_i\}_0^\tau = \{\lambda\}_0^\tau, \text{ for all } i = 1, 2, \dots, N.$$

Then, for each realization $i = 1, 2, \dots, N$,

$$W_i = \int_0^\tau \frac{d\lambda_i}{dt} \frac{\partial H_{\mathcal{A}_i}(X_{\mathcal{A}_i}, \lambda_i)}{\partial \lambda_i} dt, \quad (3)$$

where $X_{\mathcal{A}_i}$ represents the states of the system $\mathcal{A}_i$. Further, the initial canonical distributions for each realization are the

same. That is,

$$p_i(X_{\mathcal{A}_i}(0) = x_{\mathcal{A}}^0) = p(X_{\mathcal{A}}(0) = x_{\mathcal{A}}^0) = \frac{e^{-\beta H_{\mathcal{A}}(x_{\mathcal{A}}^0, \lambda(0))}}{Z},$$

for every $i = 1, 2, \dots, N$. In addition, since the systems $\mathcal{A}_i$ and $\mathcal{B}_i$ are identical to $\mathcal{A}$ and $\mathcal{B}$ for every $i = 1, 2, \dots, N$, the transition probability functions for each realization are the same. That is,

$$p_i(X_{\mathcal{A}_i}(t + dt)|X_{\mathcal{A}_i}(t)) = p(X_{\mathcal{A}}(t + dt)|X_{\mathcal{A}}(t)),$$

for each $i = 1, 2, \dots, N$.

Thus, in each realization i , the trajectory $\{X_{\mathcal{A}_i}\}_0^\tau$ will be sampled from the same joint distribution. Since the trajectory $\{X_{\mathcal{A}_i}\}_0^\tau$ in a realization i determines the work done W_i in that realization, $p_i^{(W)}(W_i = w) = \rho(w)$, for every $i = 1, 2, \dots, N$. This implies that the random variables $W_1, W_2, \dots, W_N$ are independent and identically distributed. Under the conditions described, the Jarzynski equality states that [10,22,23], for each $i = 1, 2, \dots, N$,

$$e^{-\beta \Delta F} = \langle e^{-\beta W_i} \rangle = \int_{-\infty}^{\infty} p_i^{(W)}(w) e^{-\beta w} dw.$$

Invoking the strong law of large numbers [24],

$$P\left(\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N e^{-\beta W_i} = e^{-\beta \Delta F}\right) = 1.$$

The convergence in probability of summation $\frac{1}{N} \sum_{i=1}^N e^{-\beta W_i}$ to $e^{-\beta \Delta F}$ provides the basis for the practical Jarzynski estimator $\Delta \hat{F}_N := -\frac{1}{\beta} \ln[\frac{1}{N} \sum_{i=1}^N e^{-\beta W_i}]$ that approximates the free energy change. We define the error ε_N in the free energy estimate as $\varepsilon_N := \Delta F - \Delta \hat{F}_N$. With the framework developed here, we now describe the confidence bounds on ε_N .

III. CONFIDENCE BOUNDS

For the following derivations and discussions on confidence bounds, in addition to Assumption 1, we assume that the distribution of work has a bounded support. Therefore, there exist $W_{\min}, W_{\max} \in \mathbb{R}$ such that $W_{\min} \leq W_i \leq W_{\max}$ for all $i = 1, 2, \dots, N$, where $|W_{\min}|, |W_{\max}| < \infty$. In the previous section, we established that the measured values of work W_i , $i = 1, 2, \dots, N$ are independent and identically distributed random variables. We leverage the properties of random variables to develop confidence bounds on the error ε_N as follows. Although first-principles models may allow distributions of work with unbounded support, in practice, many experimental systems exhibit effectively bounded work due to instrumental and protocol constraints (Sec. III D). Importantly, our bounds do not depend heavily on the exact values of $W_{\min}$ and $W_{\max}$. In Sec. III D, we describe practical methods for estimating these parameters, and in Sec. IV A, we demonstrate the strong performance of our approach regardless of the specific method used to estimate the maximum and minimum work values.

A. Via mean and range

The Jarzynski estimator, as given in Eq. (2), primarily computes the approximate mean of $e^{-\beta W_i}$. To find a bound on the error ε_N , we obtain a bound for the error between an

approximate mean and the true mean of a random variable. The empirical Bernstein inequality, as given in [25], permits us to develop such a bound. We state the theorem next.

Theorem 1: Empirical Bernstein inequality. Let $X_1, X_2, \dots, X_N$ be independent and identically distributed real-valued random variables, with mean $E[X_i] = \mu$ for all $i = 1, 2, \dots, N$, each $X_i \in [0, R]$, $R > 0$. Then, the following inequality holds with probability at least $1 - \delta$:

$$|\bar{X}_N - \mu| \leq \bar{s}_N \sqrt{\frac{2 \ln(3/\delta)}{N}} + R \frac{3 \ln(3/\delta)}{N}.$$

Here,

$$\bar{X}_N := \frac{1}{N} \sum_{i=1}^N X_i,$$

denotes the empirical mean and

$$\bar{s}_N := \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - \bar{X}_N)^2},$$

denotes the empirical standard deviation of $X_1, X_2, \dots, X_N$.

As can be seen, the empirical Bernstein inequality depends on the empirical mean, the empirical variance, and the range R of the random variable. Estimating R for a system under study will be discussed in Sec. III D. We now show how to find confidence bounds on ΔF using the Bernstein inequality.

Lemma 1. Let W be the nonequilibrium work done on a system and let it follow some unknown distribution $\rho(W)$. Let $W_i \in [W_{\min}, W_{\max}]$, $i = 1, 2, \dots, N$ be the i.i.d and real-valued samples drawn from $\rho(W)$. Define $Y_i := e^{-\beta W_i}$ for every $i = 1, 2, \dots, N$, $Y := e^{-\beta W}$, and $\Delta \hat{F}_N := -\frac{1}{\beta} \ln[\frac{1}{N} \sum_{i=1}^N e^{-\beta W_i}]$. Then, for any $\delta \in (0, 1)$, the following inequality holds with probability at least $(1 - \delta)$:

$$-\frac{1}{\beta} \ln(e^{-\beta \Delta \hat{F}_N} - \Psi_N) \geq \Delta F \geq -\frac{1}{\beta} \ln(e^{-\beta \Delta \hat{F}_N} + \Psi_N),$$

where

$$\Psi_N := \bar{\sigma}_N \sqrt{\frac{2 \ln(3/\delta)}{N}} + R \frac{3 \ln(3/\delta)}{N},$$

$$R = e^{-\beta W_{\min}} - e^{-\beta W_{\max}},$$

$$\bar{\sigma}_N = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y}_N)^2}, \quad \bar{Y}_N = \frac{1}{N} \sum_{i=1}^N e^{-\beta W_i}.$$

Here, the upper bound on ΔF exists only when $\Psi_N < e^{-\beta \Delta \hat{F}_N}$.

Proof. We note that Y_i are i.i.d and real valued. Let $Z_i := Y_i - e^{-\beta W_{\max}}$, for all $i = 1, 2, \dots, N$, and $Z := Y - e^{-\beta W_{\max}}$. We observe that $Z \in (0, R)$, where $R = e^{-\beta W_{\min}} - e^{-\beta W_{\max}}$. We use the Bernstein inequality from Theorem 1 to find the bounds on $|\bar{Z}_N - E[Z]|$, where $\bar{Z}_N = \frac{1}{N} \sum_{i=1}^N Z_i$. From the Jarzynski equality, it follows that

$$\bar{Z}_N - E[Z] = \bar{Y}_N - E[Y] = \frac{1}{N} \sum_{i=1}^N e^{-\beta W_i} - e^{-\beta \Delta F}.$$

We observe that with probability $(1 - \delta)$, the following inequality holds:

$$\left| \frac{1}{N} \sum_{i=1}^N e^{-\beta W_i} - e^{-\beta \Delta F} \right| \leq \bar{\sigma}_N \sqrt{\frac{2 \ln(3/\delta)}{N}} + R \frac{3 \ln(3/\delta)}{N}.$$

Rearranging and applying the logarithm gives

$$-\frac{1}{\beta} \ln(e^{-\beta \Delta \hat{F}_N} - \Psi_N) \geq \Delta F \geq -\frac{1}{\beta} \ln(e^{-\beta \Delta \hat{F}_N} + \Psi_N),$$

where

$$\Psi_N := \bar{\sigma}_N \sqrt{\frac{2 \ln(3/\delta)}{N}} + R \frac{3 \ln(3/\delta)}{N}.$$

Since the term $e^{-\beta \Delta \hat{F}_N} - \Psi_N$ must be positive for the logarithm to be valid, we see that the upper bound is valid only when $\Psi_N < e^{-\beta \hat{F}_N}$. This completes the proof. ■

The bounds derived in Lemma 1 are valid for any distribution $\rho(W)$ and are practically computable entirely from the measured samples W_i . We observe that $\lim_N \Psi_N \rightarrow 0$ at a rate of at least $\frac{1}{\sqrt{N}}$. This implies that the upper and lower bounds given by Lemma 1 converge to the true free energy difference ΔF in the limit. In addition, when $R \gg \bar{\sigma}_N$, the convergence rate of Ψ_N is as fast as $\frac{1}{N}$. For small N , when $\Psi_N \geq e^{-\beta \Delta \hat{F}_N}$, the upper bound derived in Lemma 1 doesn't exist. Thus, for situations when $\Delta \hat{F}_N$ is large, the number N needs to be high for obtaining an upper bound. This might not be feasible in certain situations. We derive an alternate bound which overcomes this issue next.

B. Via the cumulative distribution function

Our approach here is to use the uncertainty quantification on the empirical cumulative distribution function (CDF) of the work measurements. The Dvoretzky-Kiefer-Wolfowitz inequality [26] quantifies the uncertainty in the empirical CDF of a random variable. This inequality was subsequently made tighter by Massart in [27]; we use this version of the inequality for our result and henceforth refer to this inequality as the Dvoretzky-Kiefer-Wolfowitz-Massart (DKWM) Inequality. The DKWM Inequality is described below.

Theorem 2: DKWM Inequality. Let $X_1, X_2, \dots, X_N$ be i.i.d random variables with a continuous cumulative distribution function $\Phi(x)$, with $x \in \mathbb{R}$. Let $\hat{\Phi}_N(x)$ be the empirical cumulative distribution function defined by

$$\hat{\Phi}_N(x) = \frac{1}{N} \sum_{i=1}^N \mathbb{I}_{(X_i \leq x)}, \quad x \in \mathbb{R},$$

with

$$\mathbb{I}_{(X_i \leq x)} := \begin{cases} 1 & X_i \leq x \\ 0 & X_i > x \end{cases},$$

representing the indicator function. Then the following inequality holds without any restrictions on α

$$P(\sqrt{N} \sup_x |\hat{\Phi}_N(x) - \Phi(x)| > \alpha) \leq 2e^{-2\alpha^2}.$$

Theorem 2 gives the interval in which the true cumulative distribution $\Phi(x)$ lies. Indeed,

$$P(\sqrt{N} \sup_x |\hat{\Phi}_N(x) - \Phi(x)| > \alpha) \leq 2e^{-2\alpha^2},$$

implies

$$P(\sqrt{N} |\hat{\Phi}_N(x) - \Phi(x)| > \alpha) \leq 2e^{-2\alpha^2},$$

holds for all x . Then, setting $\eta = 2e^{-2\alpha^2}$, we see that with probability at least $1 - \eta$,

$$\sqrt{N} |\hat{\Phi}_N(x) - \Phi(x)| \leq \alpha,$$

or

$$\hat{\Phi}_N(x) - \xi \leq \Phi(x) \leq \hat{\Phi}_N(x) + \xi,$$

with $\xi = \alpha/\sqrt{N} = \sqrt{\frac{1}{2N} \ln(2/\eta)}$.

Lemma 2. Let $W_i \in [W_{\min}, W_{\max}]$ for all $i = 1, 2, \dots, N$, be the work values sampled from an unknown distribution $\rho(W)$. Let $\Phi(W)$ be the cumulative distribution function of W , and $\hat{\Phi}_N(W)$ the empirical cumulative distribution function, constructed using W_i . Given an arbitrary $a, b \in [W_{\min}, W_{\max}]$, the following inequalities hold with probability $1 - \eta$:

$$-\frac{1}{\beta} \ln(U) \geq \Delta F \geq -\frac{1}{\beta} \ln(L),$$

where

$$U := e^{-\beta a} [\hat{\Phi}_N(a) - \hat{\Phi}_N(W_{\min}) - 2\xi] + e^{-\beta W_{\max}} [\hat{\Phi}_N(W_{\max}) - \hat{\Phi}_N(a) - 2\xi], \quad (4)$$

and

$$L := e^{-\beta W_{\min}} [\hat{\Phi}_N(b) - \hat{\Phi}_N(W_{\min}) + 2\xi] + e^{-\beta b} [\hat{\Phi}_N(W_{\max}) - \hat{\Phi}_N(b) + 2\xi], \quad (5)$$

if $L, U > 0$, and $\xi := \sqrt{\frac{\ln(2/\eta)}{2N}}$.

Proof. The Jarzynski equality gives us the following relationship:

$$e^{-\beta \Delta F} = \int_{W_{\min}}^{W_{\max}} \rho(W) e^{-\beta W} dW.$$

With a choice of an intermediate variable a such that $W_{\min} \leq a \leq W_{\max}$, we get

$$\begin{aligned} e^{-\beta \Delta F} &= \int_{W_{\min}}^a \rho(W) e^{-\beta W} dW + \int_a^{W_{\max}} \rho(W) e^{-\beta W} dW \\ &\geq e^{-\beta a} \int_{W_{\min}}^a \rho(W) dW \\ &\quad + e^{-\beta W_{\max}} \int_a^{W_{\max}} \rho(W) dW \\ &= e^{-\beta a} [\Phi(a) - \Phi(W_{\min})] \\ &\quad + e^{-\beta W_{\max}} [\Phi(W_{\max}) - \Phi(a)]. \end{aligned}$$

Invoking Theorem 2, we know that with probability $1 - \eta$,

$$\hat{\Phi}_N(W) - \xi \leq \Phi(W) \leq \hat{\Phi}_N(W) + \xi,$$

where $\xi = \sqrt{\frac{\ln(2/\eta)}{2N}}$. Thus, we have

$$\begin{aligned} e^{-\beta \Delta F} &\geq e^{-\beta a} [\hat{\Phi}_N(a) - \xi - (\hat{\Phi}_N(W_{\min}) + \xi)] \\ &\quad + e^{-\beta W_{\max}} [\hat{\Phi}_N(W_{\max}) - \xi - (\hat{\Phi}_N(a) + \xi)] \\ &= e^{-\beta a} [\hat{\Phi}_N(a) - \hat{\Phi}_N(W_{\min}) - 2\xi] \\ &\quad + e^{-\beta W_{\max}} [\hat{\Phi}_N(W_{\max}) - \hat{\Phi}_N(a) - 2\xi]. \end{aligned}$$

Applying the logarithm and rearranging the terms, we get

$$\Delta F \leq -\frac{1}{\beta} \ln(U), \quad (6)$$

with

$$\begin{aligned} U &:= e^{-\beta a} [\hat{\Phi}_N(a) - \hat{\Phi}_N(W_{\min}) - 2\xi] \\ &\quad + e^{-\beta W_{\max}} [\hat{\Phi}_N(W_{\max}) - \hat{\Phi}_N(a) - 2\xi], \end{aligned}$$

if $U > 0$. Note that the upper bound on ΔF doesn't exist for $U \leq 0$. Starting with the definition $e^{-\beta \Delta F} = \int_{W_{\min}}^{W_{\max}} \rho(W) e^{-\beta W} dW$ and following identical steps, we find the lower bound of ΔF as follows:

$$\begin{aligned} e^{-\beta \Delta F} &= \int_{W_{\min}}^b \rho(W) e^{-\beta W} dW + \int_b^{W_{\max}} \rho(W) e^{-\beta W} dW \\ &\leq e^{-\beta W_{\min}} \int_{W_{\min}}^b \rho(W) dW + e^{-\beta b} \int_b^{W_{\max}} \rho(W) dW \\ &= e^{-\beta W_{\min}} [\Phi(b) - \Phi(W_{\min})] + e^{-\beta b} [\Phi(W_{\max}) - \Phi(b)]. \end{aligned}$$

Using Theorem 2, we get

$$\begin{aligned} e^{-\beta \Delta F} &\leq e^{-\beta W_{\min}} [\hat{\Phi}_N(b) + \xi - (\hat{\Phi}_N(W_{\min}) - \xi)] \\ &\quad + e^{-\beta b} [\hat{\Phi}_N(W_{\max}) + \xi - (\hat{\Phi}_N(b) - \xi)] \\ &= e^{-\beta W_{\min}} [\hat{\Phi}_N(b) - \hat{\Phi}_N(W_{\min}) + 2\xi] \\ &\quad + e^{-\beta b} [\hat{\Phi}_N(W_{\max}) - \hat{\Phi}_N(b) + 2\xi]. \end{aligned}$$

This relationship gives us the lower bound

$$\Delta F \geq -\frac{1}{\beta} \ln(L),$$

where

$$\begin{aligned} L &:= e^{-\beta W_{\min}} [\hat{\Phi}_N(b) - \hat{\Phi}_N(W_{\min}) + 2\xi] \\ &\quad + e^{-\beta b} [\hat{\Phi}_N(W_{\max}) - \hat{\Phi}_N(b) + 2\xi]. \end{aligned}$$

Since the cumulative distribution function $\hat{\Phi}_N(W)$ is monotonic and nondecreasing, we observe that $L > 0$ for finite N . Thus, the lower bound is well defined for all $N > 0$. This completes the proof. ■

Remark. In implementing the bounds given in Lemma 2, we select $a^* = \arg \min_a [-\frac{1}{\beta} \ln(U)]$ and $b^* = \arg \max_b [-\frac{1}{\beta} \ln(L)]$, where $a, b \in [W_{\min}, W_{\max}]$, to obtain the sharpest values for the bounds.

It would appear that the upper bound given in Eq. (4) suffers from the same problem of noncomputability for small N as seen in the previous result. However, in practice, U is computable for a relatively small N . From Eq. (4), we observe that $U \geq e^{-\beta W_{\max}} [\hat{\Phi}_N(W_{\max}) - \hat{\Phi}_N(W_{\min}) - 4\xi]$. By definition of the empirical CDF, $\hat{\Phi}_N(W_{\max}) - \hat{\Phi}_N(W_{\min}) = 1$. Therefore, the upper bound is computable when $\xi < \frac{1}{4}$, or

in other words when $N > 8 \ln(2/\eta)$. For example, if η is chosen to be 0.05 (for a 95% confidence bound), $N > 29.51$, implying that the upper bound is computable after 30 samples. This is independent of the choice of a and b , as well as the nature of the CDF.

C. Combined bounds

To get the benefits of convergence to the ground truth of Lemma 1 and the improved computability offered by Lemma 2, we need to combine the two bounds. To achieve this, we make use of the Fréchet inequality [28], presented next.

Lemma 3: Fréchet inequality. Let a random variable $X \in \mathbb{R}$ have upper bounds U_a, U_b and lower bounds L_a, L_b such that: $P(L_a \leq X \leq U_a) \geq \alpha$, and $P(L_b \leq X \leq U_b) \geq \beta$. Then, the following inequality holds:

$$P(\max[L_a, L_b] \leq X \leq \min[U_a, U_b]) \geq \max[0, \alpha + \beta - 1].$$

Proof. We know $P(\min[L_a, L_b] \leq X \leq \max[U_a, U_b])$ is bounded above by 1. By the addition law of probability, it follows that

$$\begin{aligned} P(L_a \leq X \leq U_a) + P(L_b \leq X \leq U_b) \\ - P(\max[L_a, L_b] \leq X \leq \min[U_a, U_b]) \leq 1. \end{aligned}$$

Rearranging the inequality and replacing $P(L_a \leq X \leq U_a)$ and $P(L_b \leq X \leq U_b)$ with their corresponding lower bounds, α and β respectively, yields

$$P(\max[L_a, L_b] \leq X \leq \min[U_a, U_b]) \geq \alpha + \beta - 1.$$

This expression, taken together with the non-negativity of probabilities, proves the assertion. ■

Having proven Lemma 3, we combine the bounds in Lemma 1 and Lemma 2 as follows.

Theorem 3. Let $W_i \in [W_{\min}, W_{\max}]$, $i = 1, 2, \dots, N$ be the i.i.d and real-valued nonequilibrium work measurements sampled from some unknown distribution $\rho(W)$. Then, for any $\delta, \eta \in (0, 1)$, with probability at least $(1 - \delta - \eta)$ the following holds:

$$\mathcal{U} \geq \Delta F \geq \mathcal{L}, \quad (7)$$

with the upper bound $\mathcal{U}$ and lower bound $\mathcal{L}$ defined as follows:

(i) The upper bound

$$\mathcal{U} := \min[\mathcal{U}_a, \mathcal{U}_b],$$

$$\mathcal{U}_a := \begin{cases} -\frac{1}{\beta} \ln(e^{-\beta \Delta \hat{F}_N} - \Psi_N) & e^{-\beta \Delta \hat{F}_N} > \Psi_N \\ \infty & e^{-\beta \Delta \hat{F}_N} \leq \Psi_N \end{cases},$$

and

$$\mathcal{U}_b := \begin{cases} -\frac{1}{\beta} \ln(U) & U > 0 \\ \infty & U \leq 0 \end{cases}.$$

(ii) The lower bound

$$\mathcal{L} := \max \left[-\frac{1}{\beta} \ln(e^{-\beta \Delta \hat{F}_N} + \Psi_N), -\frac{1}{\beta} \ln(L) \right].$$

Here, $\bar{Y}_N$ and Ψ_N are as defined in Lemma 1. Further U, L are as defined in Lemma 2.

Proof. The bound in Lemma 1 holds with probability at least $(1 - \delta)$, and the bound in Lemma 2 holds with probability at least $(1 - \eta)$. We observe that the bounds in Eq. (7) are the sharper of the two bounds given in Lemmas 1 and 2. Thus, for Eq. (7) to hold, we assume that both of the original bounds must be satisfied. Invoking Lemma 3 proves the theorem. ■

Note. For ease of use, given a user-specified confidence percentage $z \in (0, 100)$ in the bounds on ΔF , we set $\delta = \eta = \zeta$, where $\zeta := \frac{1-z/100}{2}$.

We have shown via Lemmas 1 and 2, and Theorem 3, confidence bounds $\mathcal{U} \geq \Delta F \geq \mathcal{L}$ on ΔF with provable guarantees. These bounds work with nonrestrictive assumptions and require only the work measurements W_i , $i = 1, 2, \dots, N$ to be known. Given the bounds on ΔF , we can obtain the bounds on the error $\varepsilon_N = \Delta F - \Delta \hat{F}_N$ as $\mathcal{U} - \Delta \hat{F}_N \geq \varepsilon_N \geq \mathcal{L} - \Delta \hat{F}_N$ with the same confidence. In the next section, we demonstrate an empirical verification and experimental validation of these bounds.

D. Estimating the range of work

The bounds we have proposed require the limits $[W_{\min}, W_{\max}]$ of the work measurements W_i . We provide the following methods to estimate the upper and lower limits of work W_i .

1. Range of measurements

Since the Bernstein inequality's formulation requires the range, a practical method for estimating this range from data is introduced in [29]. Adapting the approach to our case, we estimate the lower and upper limits from the work measurements W_i , $i = 1, 2, \dots, N$ as

$$\hat{W}_{\min} := \min_i W_i, \quad \hat{W}_{\max} := \max_i W_i.$$

We find these estimates are inaccurate for small sample sizes. The following two methods work well in a wider range of settings.

2. Instrument limits

In many experiments at the single-molecule scale, the work is done using instruments whose physical limits are well known. The physical limits can be used to estimate the maximum and minimum work that can be performed by the instrument. The knowledge of the instruments and experiments gives us (1) the maximum duration of the experiment τ_m , (2) the maximum absolute force applied, $\mathcal{F}_m := \max |\frac{\partial H_{\mathcal{A}_i}(X_{\mathcal{A}_i}, \lambda_i)}{\partial \lambda_i}|$, and (3) the maximum absolute rate with which the external variable can be changed, $v_m := \max |\frac{d\lambda_i}{dt}|$. Then from (3), for all $i = 1, 2, \dots, N$, we have

$$\begin{aligned} |W_i| &= \left| \int_0^{\tau_m} \frac{d\lambda_i}{dt} \frac{\partial H_{\mathcal{A}_i}(X_{\mathcal{A}_i}, \lambda_i)}{\partial \lambda_i} dt \right| \\ &\leq \int_0^{\tau_m} v_m \mathcal{F}_m dt \\ &= v_m \mathcal{F}_m \tau_m. \end{aligned}$$

Therefore, we can estimate the limits of the work done as

$$\hat{W}_{\min} := -v_m \mathcal{F}_m \tau_m, \quad \hat{W}_{\max} := v_m \mathcal{F}_m \tau_m.$$

3. Empirical Chebyshev inequality

A probabilistic estimate for the limits and range of work can be obtained using the empirical Chebyshev inequality. The following theorem is originally due to Saw *et al.* [30], and later modified by Kaban [31], who gives a version of the empirical Chebyshev inequality that uses only the empirical estimates. The theorem is valid for *weakly exchangeable* random variables. The definition as used by Saw *et al.* [30] is given in Definition 1.

Definition 1: Weak exchangeability. A set of random variables $X_1, X_2, \dots, X_N$ is said to be weakly exchangeable if the joint probability $P\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_N \leq x_N\}$ is invariant over all permutations of $x_1, x_2, \dots, x_N$ for all values of $x_1, x_2, \dots, x_N$.

Theorem 4: Empirical Chebyshev inequality. Let $X_1, X_2, \dots, X_N$, and X with $N \geq 2$ be weakly exchangeable samples (i.i.d implies weakly exchangeable, but the reverse is not true) from an unknown distribution, such that $P\{X_1 = X_2 = \dots = X_N = X\} = 0$. Then, for any $\epsilon > 0$,

$$P\{|X - \bar{X}_N| \geq \epsilon |\bar{X}_N|\} \leq \frac{N+1}{N} \frac{1}{\epsilon^2} \left(\frac{\bar{s}_N}{\bar{X}_N} \right)^2 + \frac{1}{N}.$$

Here, $\bar{X}_N = \frac{1}{N} \sum_{i=1}^N X_i$, is the empirical mean, and $\bar{s}_N = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - \bar{X}_N)^2}$ is the empirical standard deviation.

We note that the measurements of work W_i , $i = 1, 2, \dots, N$ are i.i.d. Setting the right-hand side of the inequality in Theorem 4 to $\gamma \in [0, 1]$, and solving for ϵ_γ yields $\epsilon_\gamma = \sqrt{\frac{N+1}{N\gamma-1} \left(\frac{\bar{s}_N}{\bar{W}_N} \right)^2}$. Here, $\bar{W}_N = \frac{1}{N} \sum_{i=1}^N W_i$ and $\bar{s}_N = \sqrt{\frac{1}{N} \sum_{i=1}^N (W_i - \bar{W}_N)^2}$. This implies that the inequality $\hat{W}_{\min} < W < \hat{W}_{\max}$ holds with probability at least $1 - \gamma$, where

$$\hat{W}_{\min} := \bar{W}_N - \epsilon_\gamma |\bar{W}_N|, \quad \hat{W}_{\max} := \bar{W}_N + \epsilon_\gamma |\bar{W}_N|.$$

In practice, we recommend choosing the probability threshold γ to be an order smaller than the probability threshold ζ chosen for the Jarzynski bound, i.e., $\gamma < \zeta/10$.

IV. METHODS AND RESULTS

We verified and validated the proposed confidence bounds empirically using synthesized data, simulations, and experiments. The first stage of our empirical results involves well-controlled simulation experiments wherein work samples are drawn from representative distributions with known variations on the distance from equilibrium (see Sec. IV A). Subsequently, we have included examples of simulated physical systems, including the spring-mass system (see Sec. IV B 1) and the particle in a time-multiplexed potential (see Sec. IV B 2). Note that, while these physical simulation experiments provide greater relevance to practical experiments, the distance from the equilibrium is harder to control in the simulated experiments. Following this, we validated our approach on a real physical system by performing non-equilibrium experiments using optical tweezers and subsequently quantifying the errors using our bounds on the experimental data obtained (see Sec. IV C).

In each of these cases, nonequilibrium work samples $W_i, i = 1, 2, \dots, N$ are collected and the free energy difference $\Delta\hat{F}_N$ computed using the Jarzynski estimator [see Eq. (2)]. Then the bounds, which are computed with only the work samples, are compared with the ground truth of the free energy difference ΔF . These steps are repeated multiple times under identical conditions to obtain the statistical behavior of the confidence bounds while keeping the sample number N fixed. To obtain the characterization for a range of sample numbers, the process is repeated for different values of N . We now describe the methods by which the three techniques generate the work samples.

A. Synthesized data

The free energy difference ΔF between two states and the work W required to move the system between these states are related by

$$W = \Delta F + w_d, \quad (8)$$

where w_d denotes the dissipated work [32]. If the states are fixed, then ΔF is a constant. However, the dissipated work is stochastic and follows some unknown distribution $w_d \sim \rho_{w_d}(w_d)$. For the verification, we used the truncated normal, a bimodal distribution (the mixture of two truncated normals), and the beta distribution as possible candidates for $\rho_{w_d}(w_d)$. We generate work samples W_i by fixing ΔF and adding to it randomly generated samples of dissipated work w_d from a fixed candidate distribution for $\rho_{w_d}(w_d)$. To ensure the Jarzynski equality in Eq. (1) applies to these samples, the distribution ρ_{w_d} is selected such that $\langle e^{-\beta w_d} \rangle = 1$. There are multiple advantages to synthesizing data in this manner. The ground truth free energy ΔF is known. The distribution of work $\rho(W)$ is very flexible. It is possible to quantify and control the distance from equilibrium by adjusting the mean of the dissipated work $\bar{w}_d$.

Figure 1 shows the performance of the three confidence bounds described in Sec. III to verify the bounds function correctly under varied conditions, three candidate distributions $\rho(W)$ for the nonequilibrium work were chosen: the truncated normal distribution, the bimodal distribution, and the beta distribution. While keeping the free energy difference ΔF fixed to $10 k_B T$, the mean of the dissipated work $\langle w_d \rangle$ was varied from $2 k_B T$ to $20 k_B T$. Thus, the ratio $\frac{\langle w_d \rangle}{\Delta F}$, which represents the distance from equilibrium, was varied from 0.2 to 2.0. This was done for each of the candidate distributions. For the truncated normal distribution, the standard deviation of the dissipated work was set to twice the mean, and the lower and upper truncation ranges were optimized to ensure that $\langle e^{-\beta w_d} \rangle = 1$, with the upper range constrained to remain symmetric about $\langle w_d \rangle$. For the bimodal distribution, consisting of two truncated normal distributions, the mean dissipated work of the left component was fixed at zero, while the mean of the right component was shifted to the right by $2\langle w_d \rangle$. The two components shared the same standard deviation, which served as the optimized parameter. For the beta distribution, we used $\rho_{\text{Beta}}(x) = x^{\alpha-1} (1-x)^{\beta-1} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$ with $\alpha = 0.4, \beta = 0.5$, where Γ denotes the Gamma function. Since the beta distribution is defined on $[0,1]$, it was scaled to a broader range of dissipated work, with the ranges determined by

optimization and the upper bound constrained to remain symmetric about $\langle w_d \rangle$. Further, the number of samples of work N , was changed from 20 to 10^4 . For each case, where the distribution of work, the ratio $\frac{\langle w_d \rangle}{\Delta F}$, and N are fixed, $M = 500$ batches of work samples were generated. Upper and lower confidence bounds are computed for each batch of work samples. Thus, the empirical observations for each case are derived from $M = 500$ independent calculations of the bounds. We chose a confidence level of 95%.

In Fig. 1(a), we show using box plots, batches of combined bounds computed for the cases where N is varied, for the truncated normal distribution of work and $\frac{\langle w_d \rangle}{\Delta F} = 0.5$. Boxes are drawn such that the center line represents the median value of a batch. The edges of the boxes represent the interquartile range, which is 50% of the population from both sides of the median. The whiskers show the minimums and maximums after excluding outliers. The outliers are data points which lie outside of three standard deviations from the median, assuming a Gaussian distribution for the data. The graph on the right in Fig. 1(a) shows the histograms of the $M = 500$ upper bounds, free energy estimates $\Delta\hat{F}_N$, and the lower bounds computed for the case where $N = 864$. In this example, all of the upper and lower bounds contain the true ΔF , which is shown by the dotted line passing through $10 k_B T$. Since we are using confidence bounds, in some instances, they may fail to contain the true ΔF . We track such 'fail' cases and have included them in our data. This empirical study shows that the probability of seeing such bound failures is less than 5%, as is guaranteed and expected from confidence bounds determined at 95% (see SM Fig. 4 in [33]). The box plots in this figure start only from $N = 20$, because the bounds are not computable for a smaller number of samples as alluded to earlier.

In Fig. 1(b), the behavior of the three bounds is summarized for all the cases tested. The graphs in the first column show the distributions of nonequilibrium work used. The second column shows the performance of the bounds using mean, the third column the bounds using CDF, and the fourth column the combined bounds. The graphs in the second, third, and fourth columns are determined as follows. The upper and lower boundaries of the shaded region of a particular color represent the modes of the $M = 500$ values of the upper and lower bounds computed for a particular case. We observe that the bounds using mean and the combined bounds converge to the true ΔF as the number of samples of work increases. Further, all the bounds are sharper and converge faster when the experiment is closer to the equilibrium, that is, when $\langle w_d \rangle$ is smaller. Out of the three bounds, we note that the combined bounds have the benefits of the other two—the convergence property of the bounds using mean, and the better computability ($N \sim 50$) of the bounds using CDF. We remark that all of the bounds, for each of the cases, were computed strictly from the work values W_i , without any knowledge of true ΔF or the nature of the work distributions. The range of the work values was assumed to be known, obtained using instrument limits. The performance of the bounds on the synthetic data using range estimation using measurements and using the empirical Chebyshev inequality is shown in SM Figs. 9–12 in [33]. Notably, our confidence bounds achieve at least a 95% success rate with either estimation method.

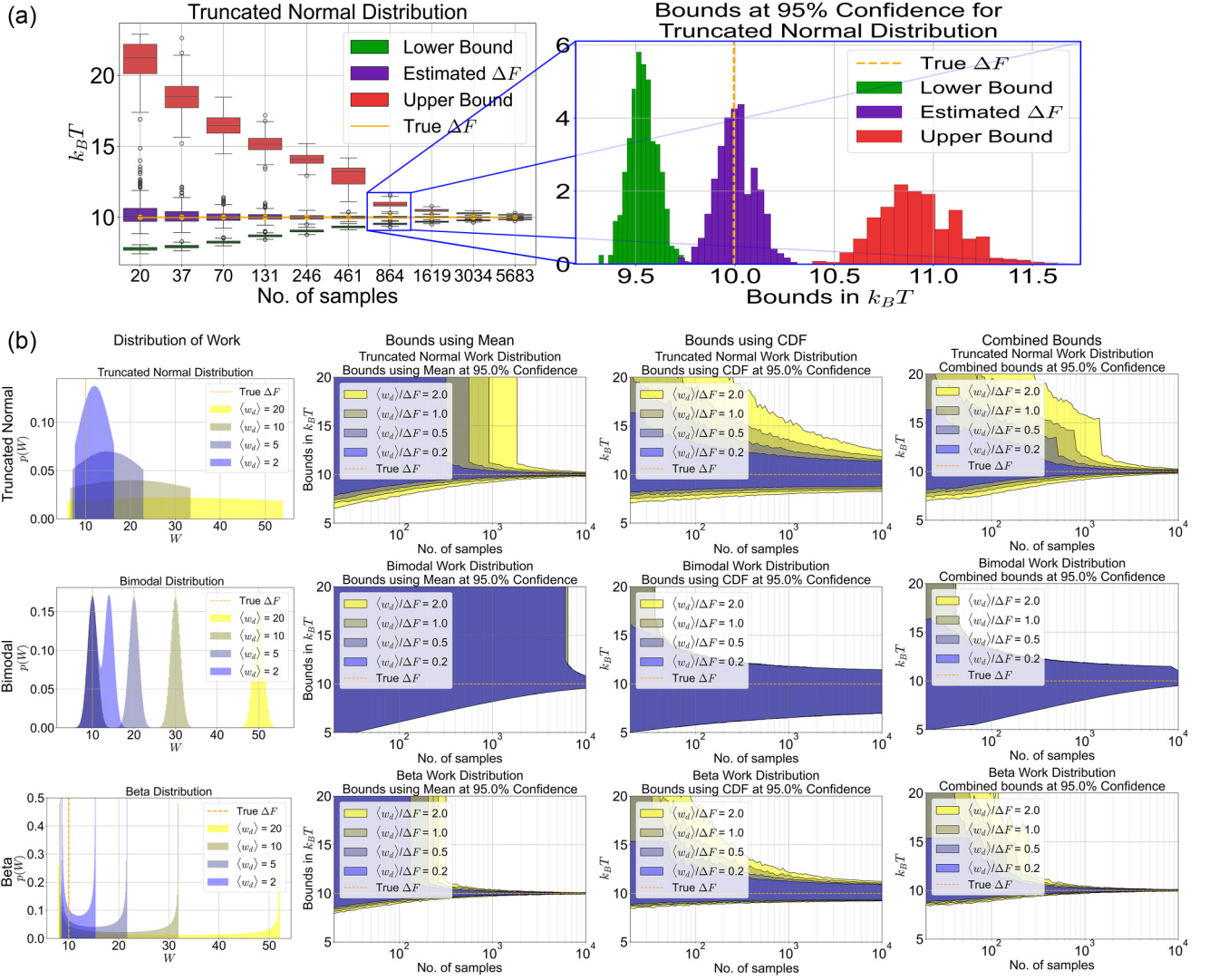


FIG. 1. A summary of the performance and characteristics of the bounds on the synthetic data. (a) A detailed view of the performance results. The box plot (left) showcases the performance of the combined bounds at 95% confidence level on work distributions that are truncated normals, with $\langle w_d \rangle / \Delta F = 0.5$ as the number of samples is increased. Each box plot represents one case of the synthetic data validation, wherein each case comprises of a collection of 500 computations of the bounds. An example case ($N = 864$) is shown in more detail on the right. The summarized view of the bounds from the synthetic data study, with varying distributions, varying $\langle w_d \rangle$, and increasing N , is shown in (b). The performance curves have been obtained using 500 computations for each case and interpolating the modes of the data obtained in each case.

In Fig. 2, we have provided a performance comparison of our bounds with the method given by Gore-Ritort-Bustamante in [11]. We refer to this approach as the ‘GRB’ method in

this document. The GRB method is applicable only when the distribution of work is normal. It can be seen that the bounds produced by the GRB method, while sharper for distributions

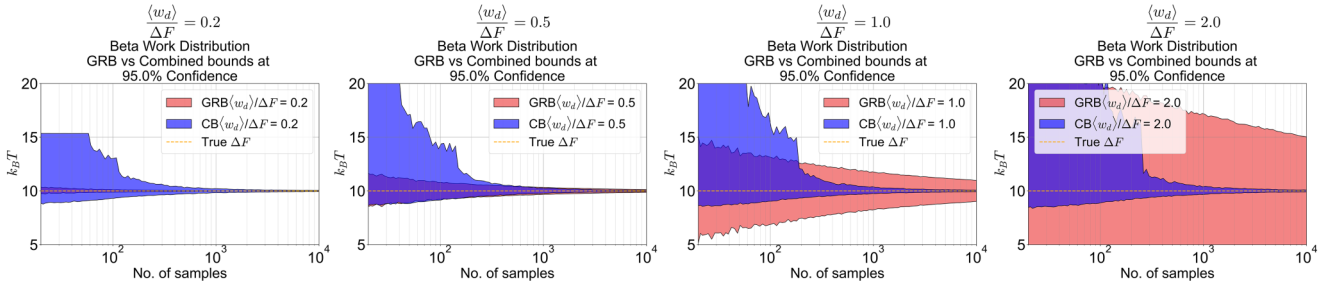


FIG. 2. Comparison of our combined bounds with the GRB method on synthetic data with truncated normal distributions of work. The performance curves have been obtained using 500 computations for each case and interpolating the modes of the data obtained in each case.

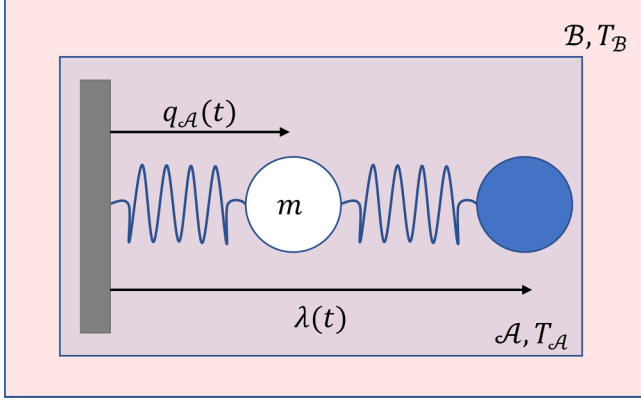


FIG. 3. Schematic of the spring-mass system.

of work close to equilibrium ($\langle w_d \rangle / \Delta F = 0.2, 0.5$), are inaccurate close to 50% of the time (see SM Fig. 5 in [33]). Our bounds give a predictable and tunable confidence level. For distributions of work far from equilibrium ($\langle w_d \rangle / \Delta F = 1.0, 2.0$), the GRB bounds are more conservative.

B. Simulations

Here, we generate nonequilibrium work samples W_i using Monte Carlo simulations. To this end, we consider two model systems. The first is the spring-mass system, an idealized model with deterministic dynamics, for which a closed-form expression for the free energy difference can be derived, as described in [34]. The second is the particle in a time-multiplexed potential, whose dynamics are governed by Langevin dynamics as discussed in [35]. The detailed descriptions of both systems are provided next.

1. The spring-mass system

Figure 3 shows a schematic of the spring-mass system. It consists of a mass m connected to wall and a controllable handle (represented by a blue circle) via two identical springs, with identical spring constants k and rest lengths l_o . The position of the mass with respect to the wall is given by $q_A(t)$ and the position of the controllable handle is $\lambda(t)$. The momentum of the mass is given by $p_A(t) = m\dot{q}_A(t)$. The wall, the two springs, and the mass comprise the system of interest $\mathcal{A}$ and are placed in an environment $\mathcal{B}$. The temperature of $\mathcal{A}$ and $\mathcal{B}$ are T_A and T_B respectively. In a nonequilibrium experiment, $\mathcal{A}$ and $\mathcal{B}$ are equilibrated to temperature T , and the handle is moved according to a schedule $\{\lambda\}_0^\tau$. The equations that model the states (q_A, p_A) are given by

$$\begin{aligned} \dot{q}_A(t) &= p_A(t)/m, \\ \dot{p}_A(t) &= -2k(q_A(t) - l_o) + k\lambda(t). \end{aligned} \quad (9)$$

When the handle position is changed in the form of a ramp, i.e., $\lambda(t) = (L_1 - L_0)\frac{t}{\tau} + L_0$; $0 \leq t \leq \tau$, where L_0, L_1 are the initial and final positions of $\lambda(t)$, the free energy change between the start and end points is given by [34] as

$$\Delta F = \frac{k}{4}(L_1^2 - L_0^2).$$

In simulations, the work done on the system is computed using the relationship [32]

$$W_i = \int_0^\tau k[\lambda(t) - 2(q_A(t) - l_o)]\dot{q}_A(t)dt.$$

Here, W_i is the work done on the i -th instance with the initialization $(q_A^i(0), p_A^i(0))$ sampled from the canonical distribution of $\mathcal{A}$.

The confidence bounds on the work values obtained from the simulated spring-mass system are shown in Fig. 4. The work values are normally distributed. The mean of the dissipated work $\langle w_d \rangle$ is varied by changing the speed with which $\lambda(t)$ in Fig. 3 is changed. The faster speeds lead to higher dissipated work and thus are farther from the equilibrium. Similar to the analysis on the synthesized data, the bounds corresponding to each case were computed $M = 500$ times on freshly generated simulated data sets with new initializations. The modes of these bounds are plotted as the upper and lower boundaries for the regions displayed in Fig. 4. The graphs show that the bounds contain and converge to the true ΔF , with the combined bounds becoming computable around 30 samples of work and converging to within 10% of ΔF around 300 to 3000 samples, depending on the dissipated work. The confidence bounds were found to contain the true value 100% of the time for all the cases in this example (see SM Fig. 6 in [33]).

We have also included a comparison with the GRB method for the spring-mass system. Across all three cases ($\langle w_d \rangle / \Delta F = 0.01, 0.17$, or 0.44), we observe that while the GRB method produces sharper bounds, these are inaccurate, with success rates falling below 20%, as shown in Fig. 5. Moreover, the GRB method, as originally formulated, does not consider cases where $\langle w_d \rangle \ll k_B T$ [11]. In particular, for $\langle w_d \rangle / \Delta F = 0.01$, the GRB method fails to provide a meaningful bound for small N .

2. Particle in a time-multiplexed potential

Radiation pressure from lasers can be used to trap transparent microspheres in optical tweezers. The potential created by such a system applies a force that holds the microsphere at the focus of the laser. The force can be approximated as $f(x) = k(x - x_o)$, where k is the spring constant of the optical trap, x the position of the micro-sphere, and x_o the focus of the laser [36]. Figure 6(a) shows the schematic of such a particle in an optical trap, and the resulting parabolic potential $\mathcal{U}(x)$ is shown in Fig. 6(b).

By multiplexing the focus of the laser rapidly between two locations, say L_r and L_l , the time-averaged potential can be shown to form a double well as shown in Figs. 6(c) and 6(d). The depth of the two potential wells can be controlled by the amount of time the laser spends in the two locations. Let t_r and t_l be the amount of time the laser spends in the right and left wells, respectively, before switching to the other well. We define the duty ratio as $\mathcal{D} := \frac{t_r}{t_r + t_l}$ and a Boolean variable $r(t)$ to indicate if the laser focus is at location L_r . The value of $r(t)$ is 1 if the laser is at location L_r at time t and 0 if the laser is at

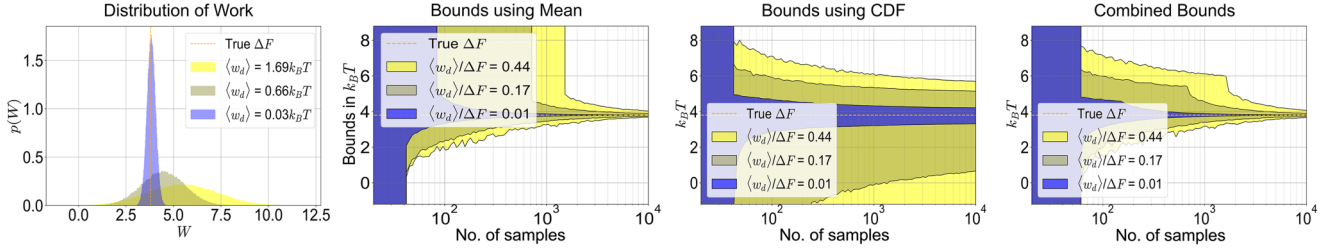


FIG. 4. Distributions of work and the corresponding bounds obtained from the simulated spring-mass system.

L_l at t . The potential $\tilde{\mathcal{U}}(x, r)$ is described by [35]

$$\tilde{\mathcal{U}}(x, r) = \begin{cases} \frac{1}{2}k(x - L_r)^2 + \mathcal{U}_b, & \text{if } |x - L_r| \leq \omega, r = 1, \\ \frac{1}{2}k(x - L_l)^2 + \mathcal{U}_b, & \text{if } |x - L_l| \leq \omega, r = 0, \\ \frac{1}{2}k\omega^2 + \mathcal{U}_b, & \text{otherwise,} \end{cases} \quad (10)$$

where ω is the width of the potential well, $\mathcal{U}_b$ is a constant. A particle in such a laser trap can be modeled as a Brownian particle in a time-multiplexed double-well potential. The dynamics of the particle are then described by

$$c \frac{dx}{dt} = v(t) - \frac{\partial \tilde{\mathcal{U}}(x(t), r(t))}{\partial x}. \quad (11)$$

Here, c is the coefficient of viscosity (determined experimentally by the step response method [35,37]), and $v(t)$ is a noise term that implicitly models the particle's collisions with the particles in its environment. Specifically, $v(t)$ is a Gaussian random variable such that $\langle v(t) \rangle = 0$ and $\langle v(t), v(t') \rangle = \frac{2c}{\beta} \delta(t - t')$ for all $t, t' \in \mathbb{R}$, with $\delta(\cdot)$ being the Dirac-delta function.

By trapping such a particle with some initial duty ratio $\mathcal{D}_i$ and switching the duty ratio to a final value $\mathcal{D}_f$, we can perform work on it. If the time to switch from $\mathcal{D}_i$ to $\mathcal{D}_f$ is much shorter than the time required for the particle to equilibrate with its environment, we can perform nonequilibrium work. Sekimoto [38] provides a way to compute the work done on the particle described by

$$W = \mathcal{U}(x(t_j), \mathcal{D}(t_j^+)) - \mathcal{U}(x(t_j), \mathcal{D}(t_j^-)). \quad (12)$$

Here, t_j is the time at which the duty ratio is changed, with $t_j^- = t_j - dt$, and $t_j^+ = t_j + dt$ the times just before and after the change. In (12), the time-averaged potential $\mathcal{U}(x, \mathcal{D})$

derived from (10) is used and is given by

$$\mathcal{U}(x, \mathcal{D}) = \begin{cases} \frac{\mathcal{D}}{2}k(x - L_r)^2 + \frac{1-\mathcal{D}}{2}k\omega^2 + \mathcal{U}_b, & \text{if } |x - L_r| \leq \omega, \\ \frac{1-\mathcal{D}}{2}k(x - L_l)^2 + \frac{\mathcal{D}}{2}k\omega^2 + \mathcal{U}_b, & \text{if } |x - L_l| \leq \omega, \\ \frac{1}{2}k\omega^2 + \mathcal{U}_b, & \text{otherwise.} \end{cases} \quad (13)$$

Monte Carlo simulations of the system can be done using Eq. (11) and the samples of work generated using Eq. (12). The free energy difference between the potentials created by the duty ratios $\mathcal{D}_i$ and $\mathcal{D}_f$ can be computed analytically (see SM Sec. II [33]) and can be used to verify the confidence bounds on the estimate.

We perform simulations in which the duty ratio is varied from $\mathcal{D}_i = 0.5$ to $\mathcal{D}_f = 0.3$. The analytically computed free energy difference (see SM Sec. II [33]) and the corresponding work distribution are shown in yellow in Fig. 7(a). Moreover, the combined bounds consistently contained the true free energy value in 100% of cases, as shown in Fig. 7(b) and SM Fig. 7 in [33].

C. Experiments

We use optical tweezers to perform nonequilibrium experiments. The optical tweezers consist of a 1064 nm trapping laser and a 850 nm detection laser. An acousto-optic deflector (AOD) is used to control the position of the focus of the trapping laser in the x-y plane. The detection laser, in conjunction with a quadrature photodiode (QPD), is used to measure the location of a microsphere caught in the optical trap. For the experiment protocol, we used $\mathcal{D}_i = 0.5$ and $\mathcal{D}_f = 0.3$, with the time to switch from one duty ratio to the other kept below 1 μ s. The multiplexing period $t_r + t_l$ of the laser was set to

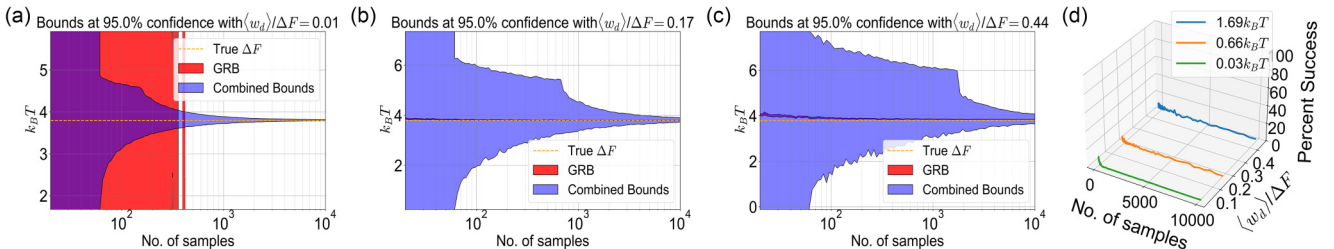


FIG. 5. Comparison of our combined bounds with the GRB method on data from the simulated spring-mass system, together with the corresponding success rate of the GRB method.

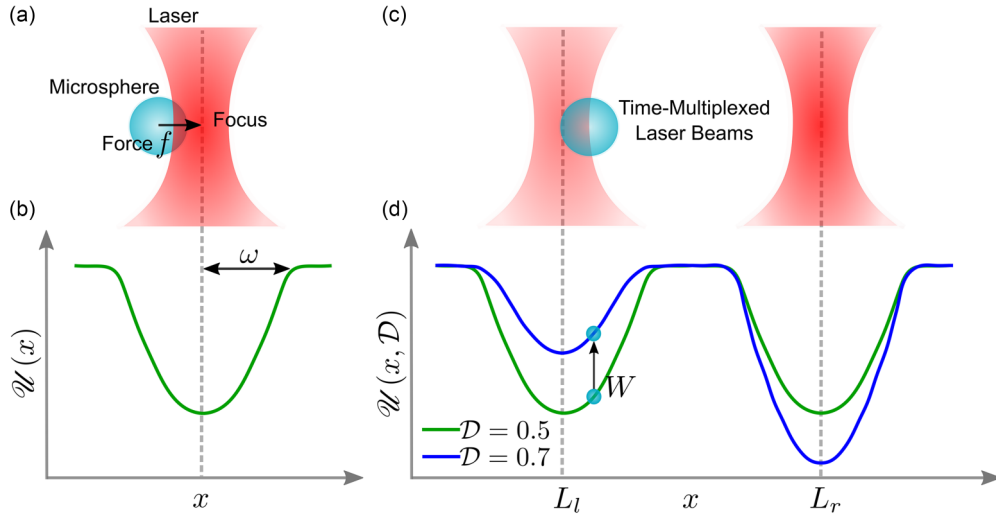


FIG. 6. Time-multiplexed potentials created by optical tweezers. (a) A microsphere trapped by a laser beam experiencing a restoring force f toward the focus. (b) The parabolic potential created by the laser trap of width ω . (c) The resultant of a laser beam rapidly multiplexed between two different foci, L_l and L_r , and (d) the resultant time-averaged potential. Changing the duty ratio $\mathcal{D}$ from 0.5 to 0.7 results in a reshaping of the potential and W work is done on the microsphere.

20 μs . The two foci of the laser were at $L_l = -500\text{ nm}$ and $L_r = 500\text{ nm}$. The width ω of the potential wells was determined to be 80 nm, and the spring constant at full power of the laser was estimated to be 13.3 pN/nm. A dilute solution of polystyrene beads with an average diameter of 1 μm in water was used at room temperature ($T = 297\text{ K}$). The average time constant of the beads was experimentally found to be 1 ms using a step test.

In experiments using optical tweezers, the work values obtained in moving a 1 μm particle in an optical potential, from a duty ratio of $\mathcal{D}_i = 0.5$ to one at $\mathcal{D}_f = 0.3$, is shown in Fig. 7(a). The histogram in Fig. 7(a), which represents the distribution of experimental work values, appears nonsmooth and deviates from the corresponding distribution obtained from simulations. This discrepancy arises from nonidealities in the shape of the energy landscape generated by the optical tweezers. As indicated by Eq. (12), the work done depends directly on differences in the spatial energy distribution of the tweezers at two duty ratios. In SM Fig. 3 in [33], we observe that for a duty ratio of 0.3, the potential energy landscape of the optical wells is distorted, with the right well showing significant deviation from an ideal parabolic shape due to physical limitations of the setup. Figure 7(c) shows the modes of the 95% confidence bounds computed on $M = 500$ batches of work values (with different sample numbers N). The combined bounds are the sharpest and contain the true ΔF in all of the cases tested here (see SM Fig. 8 in [33]). The true ΔF is computed by using a large batch of 10^7 work values and is ensured to match the expected analytical results [dashed lines in Fig. 7(a)]. The ratio $\frac{\langle w_d \rangle}{\Delta F}$ was computed to be 1.04, indicating the experiment was done in the nonequilibrium setting.

V. DISCUSSION

Free energy differences are an important quantity in a wide range of applications, including the study of systems

at the nanoscale. The experiments conducted to measure this quantity are often easier to do under the nonequilibrium setting. While the Jarzynski equality provides a way to link the equilibrium free energy differences to the nonequilibrium work measurements, it is asymptotic. Is it possible to quantify the errors in the estimates when only a finite number of work samples are available? In this article, we provided two methods to obtain confidence intervals on the finite sample estimates and, consequently, the errors. We fused the two methods to combine the advantages of each technique—of computability at small sample sizes and convergence to the ground truth. The confidence bounds developed in this article are the first of their kind and have several advantages over previously proposed solutions. They require only the measured work samples. We demonstrated that they can successfully be used on a wide range of work distributions without having any prior knowledge the nature of distributions or the true free energy difference. This can be seen in Fig. 1, wherein the combination of our techniques produces 95% confidence bounds that are within 20% of the true ΔF using 500 to 1500 samples of work. For a fixed number of samples of work, our bounds become less sharp with increasing distance from the equilibrium. However, the width of the bounds grows slowly with the distance from the equilibrium. For example, at $N = 1000$ samples and $\langle w_d \rangle / \Delta F = 0.2$ when the true $\Delta F = 10 k_B T$, we observe the width of the combined bounds are $0.73 k_B T$ for the truncated normal distribution, $3.78 k_B T$ for the bimodal distribution, and $0.38 k_B T$ for the beta distribution of work. The corresponding widths for $N = 1000$ samples, $\langle w_d \rangle / \Delta F = 2.0$, and true $\Delta F = 10 k_B T$ are: $6.64 k_B T$ for the truncated normal distribution, $3.84 k_B T$ for the bimodal distribution, and $0.74 k_B T$ for the beta distribution of work. Therefore, for a 10-fold increase in the distance from equilibrium, the width of bounds increases by a factor of 9.10 for the truncated normal, 1.02 for the bimodal distribution, and 1.95 for the beta distribution. The comparison with the GRB method showcases this robustness to the distance from

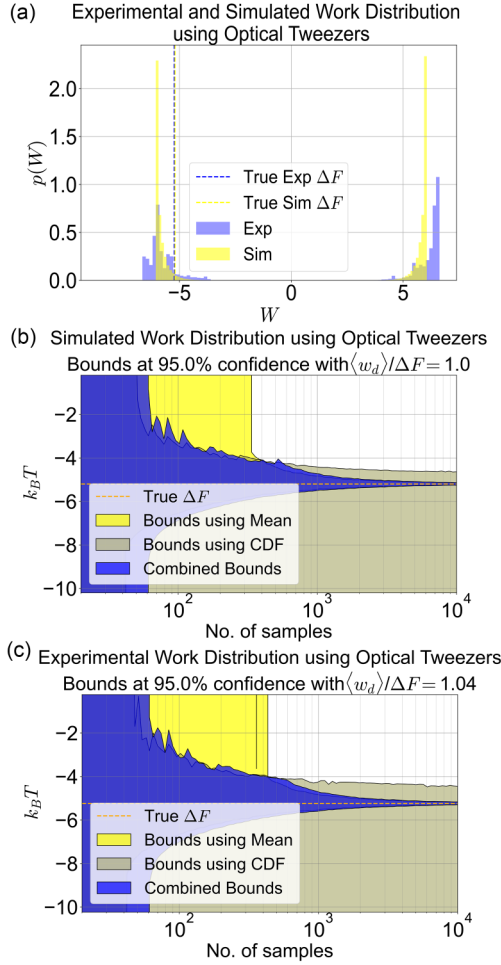


FIG. 7. (a) Distribution of the work done on a microsphere by changing the shape of the double well optical potential, from an initial duty ratio of 50% to a final duty ratio of 30%. (b) Confidence bounds and the true ΔF applied to the simulated work measurements. (c) Confidence bounds and the true ΔF applied to the experimental work measurements.

equilibrium even better. In Fig. 2, we observe that for $N = 1000$, $\langle w_d \rangle / \Delta F = 0.2$, and true $\Delta F = 10 k_B T$, the width of the GRB bound is $0.054 k_B T$. At $N = 1000$ and $\langle w_d \rangle / \Delta F = 2.0$ the width of the GRB bound is $14.12 k_B T$, which is an increase by a factor of 261, where our bounds increased by a factor of 9.10. Note also that the GRB bounds do not have a tunable confidence level, the confidence is set to a default level of 50% and the bounds are guaranteed only when the distribution of work is normal.

The distributions we selected for the synthetically generated work samples reflect the nature of distributions of work observed in real systems. In Fig. 4, we provide the distributions of work done on a spring-mass system whose schematic is shown in Fig. 3. The spring-mass system is a proxy for force spectroscopy of single molecules of proteins using atomic force microscopes and produces normally distributed work. Our bounds work under this setting with a growth in the width of the bounds by a factor of ~ 20 when the distance to equilibrium increases by a factor of 44, corroborating our observations on synthetic data. Using optical tweezers to work on a particle in a time-multiplexed well (see Fig. 6), we show a real example where the nonequilibrium work done is beta distributed. Such a multimodal distribution cannot be approximated by a normal distribution, and hence, methods developed for normal distributions are not applicable. In contrast, our method can be successfully applied, as seen in Fig. 7. Our method produces 95% confidence bounds whose width converges to less than 20% of the true ΔF at around $N = 400$ samples.

First-principle models (for example, the spring-mass model discussed in [34]) admit distributions of work with unbounded support. However, in real-world experimental systems, such as instrument-imposed constraints, such as limits on applied forces, velocities, and extensions, impose an effective practical limit on the realizable values of work. Thus, our confidence bounds remain meaningful in those contexts.

The techniques provided in this article are an important step toward improving the reproducibility of research and increasing confidence in measurements made at the nanoscale. Future directions of this research are to improve the sharpness of the bounds. This might be possible by further optimizing the intermediate steps or by using the bounds to detect dominant events which contribute more to errors in the estimates. The developments in this article are available as an user-friendly toolbox in Python with tunable confidence levels.

ACKNOWLEDGMENTS

We would like to thank Mishfad Veedu, James Melbourne, Harish Doddi, and Rachit Shrivastava for discussions and assistance with experiments on the optical tweezers. This project was funded by the NSF Electrical, Communications, and Cyber Systems grant (Grant No. 1809194) awarded to M.S.

DATA AVAILABILITY

The data that support the findings of this article are openly available [39].

- [1] C. Bustamante, J. Liphardt, and F. Ritort, The nonequilibrium thermodynamics of small systems, *Phys. Today* **58**(7), 43 (2005).
- [2] N. Kolimi, A. Pabbathi, N. Saikia, F. Ding, H. Sanabria, and J. Alper, Out-of-equilibrium biophysical chemistry: The case for multidimensional, integrated single-molecule approaches, *J. Phys. Chem. B* **125**, 10404 (2021).
- [3] R. Ait-Haddou and W. Herzog, Brownian ratchet models of molecular motors, *Cell Biochem. Biophys.* **38**, 191 (2003).
- [4] A. I. Brown and D. A. Sivak, Theory of nonequilibrium free energy transduction by molecular machines, *Chem. Rev.* **120**, 434 (2020).
- [5] C.-B. Li and S. Toyabe, Efficiencies of molecular motors: A comprehensible overview, *Biophys. Rev.* **12**, 419 (2020).

- [6] H. L. Sweeney and E. L. Holzbaur, Motor proteins, *Cold Spring Harbor Perspect. Biol.* **10**, a021931 (2018).
- [7] G. Hummer and A. Szabo, Free energy surfaces from single-molecule force spectroscopy, *Acc. Chem. Res.* **38**, 504 (2005).
- [8] S. Rajaganapathy, J. L. McCourt, S. Ghosal, A. Lindsay, P. M. McCourt, D. A. Lowe, J. M. Ervasti, and M. V. Salapaka, Distinct mechanical properties in homologous spectrin-like repeats of utrophin, *Sci. Rep.* **9**, 5210 (2019).
- [9] K. Bushby, R. Finkel, D. J. Birnkrant, L. E. Case, P. R. Clemens, L. Cripe, A. Kaul, K. Kinnett, C. McDonald, S. Pandya, J. Poysky, F. Shapiro, J. Tomezsko, and C. Constantin, Diagnosis and management of Duchenne muscular dystrophy, part 1: Diagnosis, and pharmacological and psychosocial management, *Lancet Neurol.* **9**, 77 (2010).
- [10] C. Jarzynski, Nonequilibrium equality for free energy differences, *Phys. Rev. Lett.* **78**, 2690 (1997).
- [11] J. Gore, F. Ritort, and C. Bustamante, Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements, *Proc. Natl. Acad. Sci. USA* **100**, 12564 (2003).
- [12] M. Palassini and F. Ritort, Improving free-energy estimates from unidirectional work measurements: Theory and experiment, *Phys. Rev. Lett.* **107**, 060601 (2011).
- [13] R. C. Lua and A. Y. Grosberg, Practical applicability of the Jarzynski relation in statistical mechanics: A pedagogical example, *J. Phys. Chem. B* **109**, 6805 (2005).
- [14] C. Jarzynski, Rare events and the convergence of exponentially averaged work values, *Phys. Rev. E* **73**, 046105 (2006).
- [15] N. Y. Halpern and C. Jarzynski, Number of trials required to estimate a free-energy difference, using fluctuation relations, *Phys. Rev. E* **93**, 052144 (2016).
- [16] M. Arrar, F. M. Boubeta, M. E. Szretter, M. Sued, L. Boechi, and D. Rodriguez, On the accurate estimation of free energies using the Jarzynski equality, *J. Comput. Chem.* **40**, 688 (2019).
- [17] A. Pohorille, C. Jarzynski, and C. Chipot, Good practices in free-energy calculations, *J. Phys. Chem. B* **114**, 10235 (2010).
- [18] C. H. Bennett, Efficient estimation of free energy differences from Monte Carlo data, *J. Comput. Phys.* **22**, 245 (1976).
- [19] N. M. Roussey and A. Dickson, Enhanced Jarzynski free energy calculations using weighted ensemble, *J. Chem. Phys.* **153**, 134116 (2020).
- [20] J. R. Gomez-Solano, C. July, J. Mehl, and C. Bechinger, Non-equilibrium work distribution for interacting colloidal particles under friction, *New J. Phys.* **17**, 045026 (2015).
- [21] Z. Kuang, K. M. Singh, D. J. Oliver, P. B. Dennis, C. C. Perry, and R. R. Naik, Gamma estimator of Jarzynski equality for recovering binding energies from noisy dynamic data sets, *Nat. Commun.* **11**, 5517 (2020).
- [22] C. Jarzynski, Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach, *Phys. Rev. E* **56**, 5018 (1997).
- [23] C. Jarzynski, Nonequilibrium work relations: Foundations and applications, *Eur. Phys. J. B* **64**, 331 (2008).
- [24] D. Bertsekas and J. N. Tsitsiklis, *Introduction to probability* (Athena Scientific, 2008), Vol. 1.
- [25] J.-Y. Audibert, R. Munos, and C. Szepesvari, Tuning bandit algorithms in stochastic environments, in *Algorithmic Learning Theory*, edited by M. Hutter, R. A. Servedio, and E. Takimoto (Springer, Berlin, Heidelberg, 2007), Vol. 4754, pp. 150–165.
- [26] A. Dvoretzky, J. Kiefer, and J. Wolfowitz, Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator, *Ann. Math. Stat.* **27**, 642 (1956).
- [27] P. Massart, The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality, *Ann. Probab.* **18**, 1269 (1990).
- [28] L. Rüschendorf, Fréchet-bounds and their applications, in *Advances in Probability Distributions with Given Marginals: Beyond the Copulas* (Springer, 1991), pp. 151–187.
- [29] V. Mnih, C. Szepesvári, and J.-Y. Audibert, Empirical Bernstein stopping, in *Proceedings of the 25th International Conference on Machine Learning* (Association for Computing Machinery, New York, 2008), pp. 672–679.
- [30] J. G. Saw, M. C. K. Yang, and T. C. Mo, Chebyshev inequality with estimated mean and variance, *Am. Stat.* **38**, 130 (1984).
- [31] A. Kabán, Non-parametric detection of meaningless distances in high dimensional data, *Stat. Comput.* **22**, 375 (2012).
- [32] C. Jarzynski, Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale, *Time: Poincaré Seminar 2010*, Progress in Mathematical Physics, Vol. 63 (Birkhäuser, Basel, 2013), pp. 145–172.
- [33] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/mkch-jb15> for additional derivations, algorithm details, and supplemental figures.
- [34] H. Híjar and J. M. O. de Zárate, Jarzynski's equality illustrated by simple examples, *Eur. J. Phys.* **31**, 1097 (2010).
- [35] S. Talukdar, S. Bhaban, and M. V. Salapaka, Memory erasure using time-multiplexed potentials, *Phys. Rev. E* **95**, 062121 (2017).
- [36] Edited by J. Baillieul and T. Samad, *Encyclopedia of Systems and Control* (Springer International Publishing, Cham, 2021).
- [37] K. Visscher and S. M. Block, Versatile optical traps with feedback control, in *Methods in Enzymology* (Elsevier, 1998), Vol. 298, pp. 460–489.
- [38] K. Sekimoto, Kinetic characterization of heat bath and the energetics of thermal ratchet models, *J. Phys. Soc. Jpn.* **66**, 1234 (1997).
- [39] SalapakaLab-SIMBioSys, Non-equilibrium-toolbox, GitHub repository (2026), <https://github.com/SalapakaLab-SIMBioSys/Non-Equilibrium-Toolbox>.