Skewed Thermodynamic Geometry and Optimal Free Energy Estimation

Steven Blaber and David A. Sivak
Dept. of Physics, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada

(Dated: October 1, 2020)

Free energy differences are a central quantity of interest in physics, chemistry, and biology. We develop design principles that improve the precision and accuracy of free energy estimators, which has potential applications to screening for targeted drug discovery. Specifically, we derive near-equilibrium approximations for moments of the excess work and find the leading- and next-to-leading order contributions to dissipation are from the Stokes’ friction tensor and the supra-Stokes’ tensor, with geodesics that respectively minimize the variance and bias of free energy estimators.

Introduction.—Free energy differences determine the equilibrium phases of thermodynamic systems as well as the relative reaction rates and binding affinities of chemical species [1]. Computational and experimental techniques which accurately and precisely predict free energy differences are therefore highly desirable. One important application is in pharmaceutical drug discovery, where computation of free energy differences can aid in the identification and design of ligands for targeted protein binding [2–5]. Current methods rely primarily on costly and time-consuming experimentation, which can be reduced through screening with efficient computational techniques [1–8].

The free energy difference between two states of a system is typically estimated by measuring the work incurred from a protocol (a dynamic variation in control parameters) that drives the system between control-parameter endpoints corresponding to each state. A unidirectional estimator calculates the free energy difference using only the work done by a protocol driving from initial to final states (forward protocol), while a bidirectional estimator also uses a protocol that drives from final to initial states (reverse protocol). The mean-work estimator [9] equates the mean work with the free energy difference and yields a biased estimate for any non-quasistatic (finite-speed) protocol. The Jarzynski estimator for free energy differences (derived from the Jarzynski equality relating the exponentially averaged work to the free energy change [10]) is unbiased for a large number of samples. The mean-work and Jarzynski estimators can be used as either uni- or bidirectional estimators; however, if bidirectional data is available the maximum log-likelihood estimate is Bennett’s acceptance ratio (BAR) [11] which yields (for a large number of samples) the minimum variance of any unbiased estimator [12, 13].

Regardless of the estimator, non-quasistatic control-parameter protocols result in excess work (work in excess of the free energy difference), which increases the bias and variance (decreases the accuracy and precision, respectively) of free energy estimates. This excess work (equaling the dissipation), and hence the error, can be reduced by following a different path through control-parameter space and varying the velocity along the path while keeping the protocol duration fixed [14–16]. Near equilibrium, this can be mapped onto the thermodynamic-geometry framework: for the mean-work estimator, the bias and variance from finite-time protocols is improved by following geodesics of a thermodynamic metric, the force-variance (FV) metric [17–21]; similarly, for BAR, the variance (but not the bias) is minimized if the protocol follows geodesics of the FV metric [22]. This has been used to improve the precision of calculated binding potentials of mean force [23–26].

The FV metric only minimizes the variance of the mean-work estimator and BAR if the relaxation time is independent of the control parameters. The friction-tensor metric [27], a generalization of the FV, provides a relatively simple prescription for reducing excess work in more general near-equilibrium processes, where minimum-dissipation protocols follow geodesics of the Riemannian metric induced by the friction tensor. For common unidirectional estimators near equilibrium, the bias and variance are proportional to the first moment (mean) of the excess work, and hence minimum-dissipation protocols defined by the friction tensor minimize both the bias and variance.

We show that for bidirectional estimators, the variance is proportional to the sum of the second moments of the excess work from forward and reverse protocols (25a), while the bias is proportional to the difference of the first moments (26a) [28]. With this in mind, we extend the thermodynamic-geometry framework to higher-order moments of the excess work and beyond the leading-order friction-tensor approximation. We find that for bidirectional estimators near equilibrium, minimum-variance protocols follow geodesics of the Riemannian metric induced by the friction tensor, while minimum-bias protocols follow geodesics of a cubic Finsler metric. For the simple model system of a Brownian particle in a quadratic trap with time-dependent stiffness (a breathing harmonic trap), the minimum-variance and minimum-bias protocols can improve variance by a factor of $3 - 4$ and bias by over a factor of $10$ (Fig. 2).

Derivation.—Consider a system in the canonical ensemble, in thermal equilibrium with a heat bath at temperature $T$. The probability distribution over microstates $x$ at control parameters $\lambda$ is $\pi(x|\lambda) =$
\[
\exp \beta [F(\lambda) - E(x, \lambda)], \text{ with energy } E(x, \lambda) \text{ and free energy } F(\lambda) \equiv -k_B T \ln \sum_x \exp \left[-\beta E(x, \lambda)\right]. \quad \text{Here } \beta \equiv \left(k_B T\right)^{-1} \text{ for Boltzmann's constant } k_B. \]
We define the work done in a single realization of an external agent changing the control parameters \(\lambda\) according to a protocol \(\Lambda\) as \(W \equiv -\int_0^t dt' f_i \lambda_i(t')\), which implies the average excess work \((W_{ex} \equiv W - \Delta F)\) is
\[
\langle W_{ex}\rangle_{\Lambda} = -\int_0^t dt' \langle \delta f_i(t') \lambda_i(t') \rangle, \quad (1)
\]
where we adopt the Einstein summation convention of implied summation over repeated indices. A dot denotes the time derivative \(\dot{\lambda}_i \equiv \frac{d\lambda_i}{dt}\), \(f_i = -\partial_{\lambda_i} U\) is the generalized force conjugate to \(\lambda_i\), and \(\delta f_i \equiv f_i - \langle f_i\rangle_{\eq}\) is the difference from the equilibrium average. Angle brackets \(\langle \cdots \rangle_{\Lambda}\) denote an average over the nonequilibrium ensemble of system responses to control-parameter protocol \(\Lambda\).

The time derivative of the second moment of the excess-work distribution is
\[
\frac{d\langle W_{ex}^2\rangle_{\Lambda}}{dt} = 2\dot{\lambda}_i(t) \int_0^t dt' \langle \delta f_i(t') \delta f_j(t') \rangle_{\Lambda} \lambda_j(t'), \quad (2)
\]
For a sufficiently slow protocol, we replace the nonequilibrium average \(\langle \cdots \rangle_{\Lambda}\) with the equilibrium average \(\langle \cdots \rangle_{\eq}\) at fixed control parameters \(\lambda(t)\):
\[
\frac{d\langle W_{ex}^2\rangle_{\Lambda}}{dt} \approx 2\dot{\lambda}_i(t) \lambda_j(t) \int_0^t dt'' \langle \delta f_i(0) \delta f_j(t'') \rangle_{\eq}, \quad (3)
\]
where we used the stationarity of the equilibrium average, defined \(t'' \equiv t' - t\), and assumed smooth protocols to expand the control-parameter velocity to zeroth order, \(\lambda(t - t'') \approx \dot{\lambda}(t)\). Finally, we assume correlations in the conjugate forces relax quickly relative to the protocol duration and replace the integration bound \(t\) with \(\infty\) (Supplemental Material, Finite Bounds [29]), simplifying the approximation to
\[
\frac{d\langle W_{ex}^2\rangle_{\Lambda}}{dt} \approx 2\beta \zeta^{(1)}_{ij}[\lambda(t)] \dot{\lambda}_i(t) \lambda_j(t), \quad (4)
\]
for the friction tensor
\[
\zeta^{(1)}_{ij}[\lambda(t)] \equiv \beta \int_0^\infty dt'' \langle \delta f_i(0) \delta f_j(t'') \rangle_{\eq}. \quad (5)
\]
In analogy with fluid dynamics, this rank-two tensor is the \textit{Stokes’ friction}, since it produces a drag force that depends linearly on velocity.

Following parallel arguments, we approximate the third moment of the excess-work distribution as
\[
\frac{d\langle W_{ex}^3\rangle_{\Lambda}}{dt} \approx 3 \beta^2 \zeta^{(2)}_{ijk}[\lambda(t)] \dot{\lambda}_i(t) \lambda_j(t) \lambda_k(t), \quad (6)
\]
for the rank-three tensor
\[
\zeta^{(2)}_{ijk}[\lambda(t)] \equiv \beta^2 \int_0^\infty dt'' \int_0^t dt''' \langle \delta f_i(0) \delta f_j(t'') \delta f_k(t''') \rangle_{\eq}. \quad (7)
\]
(The factor of three in (6) results from grouping the index permutations \(\{ijk, jik, kij\}\) into one term, using the invariance of the sum under exchange of indices, e.g., \(\zeta^{(2)}_{ijk} \lambda_i \lambda_j \lambda_k = \zeta^{(2)}_{jik} \lambda_j \lambda_i \lambda_k\).) We call the rank-three tensor \([\text{Eq. } (7)]\) the \textit{supra-Stokes’} tensor, as it corresponds to the leading-order correction to dissipation beyond the Stokes’ friction (15a).

For fourth and higher moments, Supplemental Material, Finite Bounds [29] shows that the integration bounds must remain finite since the \(n\)-time covariance functions do not decay to zero, so there is no clear analogy to frictional drag forces. Nevertheless, we can still approximate them by
\[
\frac{d\langle W_{ex}^n\rangle_{\Lambda}}{dt} \approx n \epsilon^{(n-1)}_{\nu_1 \cdots \nu_n}[\lambda(t), t] \prod_{i=1}^n \dot{\lambda}_{\nu_i}(t), \quad (8)
\]
where we use index notation \(\nu_1, \nu_2, \nu_3 \cdots\) instead of \(i, j, k\), and define integral \(n\)-time covariance functions
\[
\epsilon^{(n-1)}_{\nu_1 \cdots \nu_n}[\lambda(t), t] \equiv \langle e^{-\beta W_{ex}} \rangle_{\lambda(t)} \prod_{i=2}^n \int_0^t dt_i \left\langle \prod_{j=2}^n \delta f_{\nu_i}(0) \delta f_{\nu_j}(t_j) \right\rangle_{\lambda(t)}, \quad (9)
\]
This implies that higher-order moments of the excess work are higher order in control-parameter velocity and are therefore smaller for slow protocols. The approximations of (4), (6), and (8) are the leading-order contributions to each moment of the excess work.

To derive the next-order contribution we exploit the connection between time-reversed protocols through the Crooks relation [30, 31], which constrains the probability of forward and reverse work measurements as
\[
P_{\Lambda}(W_{ex}) e^{-\beta W_{ex}} = P_{\Lambda^t}(-W_{ex}). \quad (10)
\]
\(\Lambda^t\) is the time-reversed protocol, starting at equilibrium in the end state of the forward protocol. Integrating over \(W_{ex}\) produces Jarzynski’s equality [10]
\[
\langle e^{-\beta W_{ex}} \rangle_{\Lambda} = 1, \quad (11)
\]
while first multiplying by \(W_{ex}\) then integrating leads to
\[
\langle W_{ex} e^{-\beta W_{ex}} \rangle_{\Lambda} = -\langle W_{ex} \rangle_{\Lambda^t}. \quad (12)
\]
Taylor expanding the exponential in (11) and (12) to third order gives
\[
\langle W_{ex} \rangle_{\Lambda} \approx \frac{1}{2} \beta \langle W_{ex}^2 \rangle_{\Lambda} - \frac{1}{6} \beta^2 \langle W_{ex}^3 \rangle_{\Lambda} \quad (13a)
\]
\[
\langle W_{ex} \rangle_{\Lambda^t} \approx -\langle W_{ex} \rangle_{\Lambda} + \beta \langle W_{ex}^2 \rangle_{\Lambda} - \frac{1}{2} \beta^2 \langle W_{ex}^3 \rangle_{\Lambda}. \quad (13b)
\]
According to (8), near equilibrium the higher-order moments of the excess work are higher order in control-parameter velocity and therefore can be neglected for
slow protocols. The leading-order contributions to the sum and difference of the excess work are

\[
\langle W_{\text{ex}} \rangle_\Lambda + \langle W_{\text{ex}} \rangle_{\Lambda^\dagger} \approx \beta \langle W_{\text{ex}}^2 \rangle_\Lambda \\
\langle W_{\text{ex}} \rangle_\Lambda - \langle W_{\text{ex}} \rangle_{\Lambda^\dagger} \approx \frac{1}{2} \beta^2 \langle W_{\text{ex}}^3 \rangle_\Lambda .
\]

(14a)

(14b)

Near equilibrium, the second and third moments are approximated by (4) and (6), so the time derivatives of the first and second moments are

\[
\frac{d}{dt} \langle W_{\text{ex}} \rangle_\Lambda \approx \left( c_{ij}^{(1)} + \frac{1}{4} c_{ijkl}^{(2)} \right) \dot{\lambda}_i \dot{\lambda}_j
\]

(15a)

\[
\frac{d}{dt} \langle W_{\text{ex}}^2 \rangle_\Lambda \approx \frac{2}{\beta} \left( c_{ij}^{(1)} + \frac{3}{4} c_{ijkl}^{(2)} \right) \dot{\lambda}_i \dot{\lambda}_j
\]

(15b)

where henceforth we drop the explicit dependence of the friction and control-parameter velocity on \( \lambda \) and \( t \).

To derive an approximation for higher-order moments, we follow parallel arguments. Multiplying (10) by \( W_{\text{ex}}^n \) and integrating over \( W_{\text{ex}} \), then either adding or subtracting \( \langle (W_{\text{ex}}^n) \rangle_\Lambda \), the leading-order contributions to the sum and difference for \( n \geq 2 \) are

\[
\langle W_{\text{ex}}^n \rangle_\Lambda + \langle (W_{\text{ex}}^n) \rangle_{\Lambda^\dagger} \approx 2 \langle (W_{\text{ex}}^n) \rangle_\Lambda \\
\langle W_{\text{ex}}^n \rangle_\Lambda - \langle (W_{\text{ex}}^n) \rangle_{\Lambda^\dagger} \approx \beta \langle W_{\text{ex}}^{n+1} \rangle_\Lambda ,
\]

(16a)

(16b)

which implies

\[
\beta^{n-1} \frac{d}{dt} \langle W_{\text{ex}}^n \rangle_\Lambda \approx \left( c_{i_1 i_2 \cdots i_n}^{(n-1)} + \frac{n+1}{2} c_{i_1 i_2 \cdots i_{n+1}}^{(n)} \dot{\lambda}_{i_{n+1}} \right) \prod_{i=1}^n \dot{\lambda}_{i_n}
\]

(17)

Equations (15a), (15b) and (17) are the central results of this letter, which have applications to designing minimum-dissipation and optimal free energy estimation protocols.

Next-order contribution to excess work:—The approximation of (15a) can improve near-equilibrium estimates of the excess work. The leading-order term is the usual linear-response approximation (Eq. (13) of Ref. [27]) and is positive for all protocols. The Stokes’ friction (5) is an autocovariance function,

\[
\zeta_{ij}^{(1)} = \langle \delta f_i \delta f_j \rangle_{\lambda(t)} \hat{\tau}_{ij}^{(1)},
\]

(18)

the Hadamard (entry-by-entry) product, \( \hat{\tau} \), of the conjugate-force covariance \( \langle \delta f_i \delta f_j \rangle_{\lambda(t)} \) and integral relaxation time

\[
\tau_{ij}^{(1)} = \int_0^\infty dt' \frac{\langle \delta f_i(0) \delta f_j(t') \rangle_{\lambda(t)}}{\langle \delta f_i \delta f_j \rangle_{\lambda(t)}}.
\]

(19)

The Stokes’ friction is largest when the conjugate-force fluctuations are largest (large covariance) and most persistent (long relaxation time). In contrast to the positive contribution from the Stokes’ friction, the contribution from the supra-Stokes’ tensor (second term in (15a)) changes sign under time reversal because it is cubic in control-parameter velocity. The supra-Stokes’ tensor (7)

\[
\zeta_{ij}^{(2)} = -\langle \delta f_i \delta f_j \delta f_k \rangle_{\lambda(t)} \circ \tau_{ijk}^{(2)}
\]

(20)

is not positive semidefinite, and is the Hadamard product of the unnormalized coskewness \( \langle \delta f_i \delta f_j \delta f_k \rangle_{\lambda(t)} \) (related to skewness as covariance is related to variance) and the integral double relaxation time

\[
\tau_{ijk}^{(2)} \equiv \int_0^\infty dt'' \int_0^\infty dt''' \frac{\langle \delta f_i(0) \delta f_j(t'') \delta f_k(t''') \rangle_{\lambda(t)}}{\langle \delta f_i \delta f_j \delta f_k \rangle_{\lambda(t)}}.
\]

(21)

For a protocol with positive velocity, the contribution to dissipation from the supra-Stokes’ tensor quantifies the increase (decrease) in excess work from negatively (positively) skewed conjugate-force fluctuation.

The leading-order friction-tensor approximation (first term in (15a)) endows the control-parameter space with a Riemannian metric, such that minimum-work protocols follow geodesics of the Stokes’ friction tensor. With the addition of the next-order contribution (7), the excess work can be expressed as

\[
\langle W_{\text{ex}} \rangle_\Lambda \approx \int_0^\Delta t dt' \zeta_{ij}^{\text{tot}} \dot{\lambda}_i \dot{\lambda}_j
\]

(22)

for total friction tensor

\[
\zeta_{ij}^{\text{tot}} \equiv \zeta_{ij}^{(1)} + \frac{1}{4} \zeta_{ij}^{(2)} \hat{\lambda}_k
\]

(23)

that explicitly depends on the protocol velocity. Minimum-dissipation protocols follow geodesics of the generalized cubic Finsler metric \( \zeta_{ij}^{\text{tot}} \) (Supplemental Material, Finsler Geometry [29]), an extension of the Riemannian metric \( \zeta_{ij}^{(1)} \). For small control-parameter velocities, the total friction is positive semi-definite; however, for large control-parameter velocities the approximation breaks down and no longer guarantees positive semi-definiteness, leading to the unphysical possibility of negative excess work.

To illustrate some general properties of the friction, Fig. 1 shows, for the model system of a breathing harmonic trap (Brownian particle in a harmonic trap where the control parameter is the time-dependent stiffness, see Supplemental Material, Breathing Trap [29]), the (a) force variance, (b) Stokes’ friction, (c) supra-Stokes’ contribution, and (d) total friction. The force variance and Stokes’ friction are independent of control-parameter velocity. In general, the force variance differs from the Stokes’ friction whenever integral relaxation time depends on the control parameters, as for the breathing harmonic trap. The contribution from the supra-Stokes’ tensor (c) is antisymmetric in control-parameter velocity, becoming negative for negative velocity. This antisymmetric contribution skews the total friction (d), which
depends on the control-parameter velocity and lacks any symmetry under time reversal.

Geodesics (solid curves in Fig. 1) are protocols that minimize the contribution from the corresponding metric to the excess work. For relatively fast protocols (average scaled velocities \( \dot{\lambda} \) larger than one). For this system the minimum-variance and minimum-bias protocols achieve similar amounts of bias and variance in all cases, likely due to their similar functional form (Supplemental Material, [29] Fig. 1).

Assuming small excess work, we expand the variance

\[
\langle (\delta \Delta F)^2 \rangle \approx \frac{1}{2N} \left( \langle W_{\text{ex}}^2 \rangle_{A} + \langle W_{\text{ex}}^2 \rangle_{A^t} \right) - \frac{4}{N},
\]

where for simplicity we assume an equal number of forward and reverse work measurements, and the average \( \langle \cdots \rangle \) is over a total of \( N \) work measurements. BAR saturates this bound for large \( N \). This bound demonstrates an explicit connection between the minimum variance of free energy estimators and the excess work. The variance is minimized at 0 when the excess-work distribution is a delta function at \( W_{\text{ex}} = 0 \) (achieved in the quasistatic limit). For any finite-time protocol, the average excess work is positive.

FIG. 1: Friction as a function of control-parameter velocity \( \dot{\lambda} \) and control parameter \( \lambda^* \) for the breathing harmonic trap. Scaled (a) force variance \(\langle \delta F \rangle^2 \equiv \beta^2 k^2 \langle \delta F^2 \rangle \), (b) Stokes’ friction \(\zeta(1)^{(1)} \equiv \beta^2 k^2 \zeta(1)/\gamma \), (c) supra-Stokes’ contribution \(\zeta(2)^{(1)} \equiv \beta^2 k^2 \lambda \zeta(2)/\gamma \), and (d) total friction \(\zeta^{\text{total}} \equiv \beta^2 k^3 \lambda^{\text{total}} /\gamma \), for initial spring constant \( k^* \) and damping coefficient \( \gamma \). The control parameter is the trap stiffness, given in dimensionless form as \( \lambda^* = k/k^* \), with velocity \( \dot{\lambda} \equiv \gamma k/k^2 \). Geodesics (solid curves) minimize the magnitude of the corresponding metric’s contribution to the excess work, with distinct curves representing different average protocol velocities.
Discussion.—We have developed near-equilibrium approximations for moments of the excess work, (15a), (15b), and (17), that incorporate time-reversal symmetric and antisymmetric contributions. The antisymmetric contribution to the first moment (15a) yields the next-order contribution beyond linear response, which asymmetrically skews the Riemannian metric (5) into a generalized cubic Finsler metric (23). The Stokes’ friction tensor (5) controls the leading-order contribution to the variance of both unidirectional and bidirectional free energy estimators (15b), such that minimum-variance protocols follow geodesics of the Stokes’ friction. For unidirectional estimators, these same protocols also minimize the bias; however, for bidirectional estimators such as BAR the leading-order contribution to the bias is from the supra-Stokes’ tensor (26b), and therefore minimum-bias protocols instead follow geodesics of the supra-Stokes’ tensor. From these near-equilibrium approximations, we design protocols that increase the precision and accuracy of standard nonequilibrium free energy estimators (Fig. 2).

The addition of the supra-Stokes’ tensor has several physical implications. Notably, it accounts for time-reversal asymmetric dissipation, which arises from skewed conjugate-force fluctuations. Since the supra-Stokes’ contribution is antisymmetric, it cancels out for equal numbers of forward and reverse protocols, which has implications for the design principles of molecular machines: many molecular machines (e.g., kinesin walking toward a microtubule’s plus end [32] or ATP synthase synthesizing ATP [33]) achieve directed behavior; the supra-Stokes’ tensor quantifies the leading-order energetic cost of directed operation compared to a coequal forward and reverse process.

In some cases the minimum-work protocols can be calculated exactly [34, 35], or minimum-variance protocols solved numerically [36, 37], which would yield more accurate and precise estimators; however, to date these optimizations are limited to simple systems and do not permit straightforward generalization. The present formalism can be applied to more general settings than exact calculations and makes fewer approximations than Shenfeld et al. [22], who only examine the force-variance metric and do not consider minimizing bias.

Here we considered continuous protocols, but free energies are often estimated from sampling discrete states [38–41]. Additionally, for discrete state sampling in replica-exchange [42] or parallel-tempering [43] simulations it is important to consider the acceptance probability between states [44–52]. The force-variance metric has been applied in this context [22], and the linear-response framework of Ref. 27 has been generalized to discrete control [53]. As Ref. 22 has shown, spacing the states along geodesics of the force variance not only reduces the variance of the free energy estimator but also increases the acceptance probability. We expect that analogous principles hold for designing minimum-variance and minimum-bias protocols for discrete state sampling using the milder approximations of our approach.

Acknowledgments.—The authors thank Steve Large, Emma Lathouwers, and Joseph Lucero (SFU Physics), Miranda Louwerse (SFU Chemistry), John Chodera and Josh Fass (Memorial Sloan Kettering Computational Biology), and Gavin Crooks (Berkeley Institute for Theoretical Science) for feedback on the manuscript. This work is supported by an SFU Graduate Deans Entrance Scholarship (SB), an NSERC Discovery Grant and Discovery Accelerator Supplement (DAS), and a Tier-II Canada Research Chair (DAS).

FIG. 2: (a) Scaled variance $\langle (\beta \delta A/F)^2 \rangle / N$ (25a) and (b) scaled bias $\langle \hat{\beta} \delta A/F \rangle / N$ (26a) as functions of scaled protocol duration $\Delta t / \tau^{(1)}$ (for slowest relaxation time $\tau^{(1)} = \gamma/(2k_B))$, for the breathing harmonic trap. Ratios of designed and naive (c) variances $\langle (\delta A/F)^2 \rangle_{\text{des}}/(\delta A/F)^2_{\text{naive}}$ and (d) biases $\langle \delta A/F \rangle_{\text{des}}/(\delta A/F)^2_{\text{naive}}$. Subscript ‘naive’ denotes the constant-velocity protocol, and ‘des’ the designed protocols, which are the force-variance-optimized ‘force variance’, the minimum-variance ‘accurate’, and the minimum-bias ‘precise’ protocols. Dash-dotted lines denote the limits for short (black) and long duration (colors). Final control-parameter value (spring constant) is $k_t/k_1 = 1/16$. 

* Electronic address: dsivak@sfu.ca
† Electronic address: sblaber@sfu.ca

[1] V. Gapsys, S. Michielssens, J. H. Peters, B. L. de Groot, and H. Leonov, Calculation of Binding Free Energies (Springer New York, New York, NY, 2015), pp. 173–209.
[2] C. E. Schindler, H. Baumann, A. Blum, D. Böse, H.-P. Buchstaller, L. Burgdorf, D. Cappel, E. Chekler, P. Czodrowski, D. Dorsch, et al., J. Chem. Inf. Model (2020).
[3] B. Kuhn, M. Tichý, L. Wang, S. Robinson, R. E. Martin, A. Kuglstatter, J. Benz, M. Giroud, T. Schirmeister,
[4] M. Ciordia, L. Pérez-Benito, F. Delgado, A. A. Trabanco, and G. Tresadern, J. Chem. Inf. Model 56, 1856 (2016).
[5] L. Wang, Y. Wu, Y. Deng, B. Kim, L. Pierce, G. Krilov, D. Lupyan, S. Robinson, M. K. Dahlgren, J. Greenwood, et al., J. Am. Chem. Soc. 137, 2695 (2015).
[6] M. Aldeghi, V. Gapsys, and B. L. de Groot, ACS Cent. Sci. 4, 1708 (2018).
[7] J. D. Chodera, D. L. Mobley, M. R. Shirts, R. W. Dixon, K. Branson, and V. S. Pande, Curr. Opin. Struc. Biol. 21, 150 (2011).
[8] F. Weinhold, J. Chem. Phys. 63, 2479 (1975).
[9] J. Schön, J. Chem. Phys. 105, 10072 (1996).
[10] J. D. Chodera and M. R. Shirts, J. Chem. Phys. 113, 7035 (2000).
[11] D. A. Sivak and G. E. Crooks, Phys. Rev. Lett. 99, 100602 (2007).
[12] D. K. Shenfeld, H. Xu, M. P. Eastwood, R. O. Dror, and D. E. Shaw, Phys. Rev. E 80, 046705 (2009).
[13] T. T. Pham and M. R. Shirts, J. Chem. Phys. 135, 034114 (2011).
[14] T. T. Pham and M. R. Shirts, J. Chem. Phys. 136, 124120 (2012).
[15] R. Abel, et al., J. Med. Chem 60, 2485 (2017).
[16] S. Park and W. Im, J. Chem. Theory Comput. 10, 2719 (2014).
[17] S. Kim, Y. W. Kim, P. Talkner, and J. Yi, Phys. Rev. E 86, 041130 (2012).
[18] See Supplemental Material for Model Details, Discussion of Finite Integration Bounds and Finsler Geometry.
[19] G. E. Crooks, Phys. Rev. E 60, 2721 (1999).
[20] G. E. Crooks, Phys. Rev. E 61, 2361 (2000).
[21] G. E. Crooks, Phys. Rev. E 60, 2721 (1999).
[22] T. Schmiedl and U. Seifert, Phys. Rev. Lett. 98, 108301 (2007).
[23] A. Gomez-Marin, T. Schmiedl, and U. Seifert, J. Chem. Phys. 129, 024114 (2008).
[24] P. Geiger and C. Dellago, Phys. Rev. E 81, 021127 (2010).
[25] A. P. Solon and J. M. Horowitz, Phys. Rev. Lett. 120, 180605 (2018).
[26] D. A. Pearlman and P. A. Kollman, J. Chem. Phys 90, 2460 (1989).
[27] N. Lu and D. A. Kofke, J. Chem. Phys 111, 4414 (1999).
[28] R. J. Radmer and P. A. Kollman, J. Comput. Chem 18, 902 (1997).
[29] D. Wu and D. A. Kofke, J. Chem. Phys 123, 084109 (2005).
[30] Y. Sugita, A. Kitao, and Y. Okamoto, J. Chem. Phys 113, 6042 (2000).
[31] D. J. Earl and M. W. Deem, Phys. Chem. Chem. Phys 7, 3910 (2005).
[32] D. A. Kofke, J. Chem. Phys 117, 6911 (2002).
[33] D. A. Kofke, J. Chem. Phys 119, 4119 (2005).
[34] A. P. Solon and J. M. Horowitz, Phys. Rev. Lett. 120, 180605 (2018).
[35] R. M. Dirks, H. Xu, and D. E. Shaw, J. Chem. Theory Comput. 8, 162 (2012).
[36] J. L. MacCallum, M. I. Muniyat, and K. Gaalswyk, J. Stat. Mech.: Theory Exp. 2019, 083212 (2019).
Supplemental Material for “Skewed Thermodynamic Geometry and Optimal Free Energy Estimation”

Steven Blaber and David A. Sivak*
Dept. of Physics, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada†
(Dated: October 1, 2020)

BREATHING HARMONIC TRAP

A simple non-trivial system for designing and testing minimum-variance and minimum-bias protocols is the breathing harmonic trap, since its correlation functions can be calculated analytically and it has a non-Gaussian work distribution. Consider a colloidal particle in a harmonic trap with variable stiffness. The particle position \(x\) obeys the overdamped Langevin equation,

\[
\dot{x} = -kx + \sqrt{2}\gamma k_B T \eta,
\]

(S1)
in a trap of time-dependent strength \(k\), damping coefficient \(\gamma\), environmental temperature \(T\), Boltzmann’s constant \(k_B\), and Gaussian white noise \(\eta\). The control parameter is the trap strength \(\lambda = k\), so the conjugate force is \(f \equiv -\partial U / \partial k = -\frac{1}{2}x^2\). The joint probability distribution of the particle position and work obeys \([1, 2]\) \(p(x, w, t) = \pi(x|k(t))\delta(w) = [2\pi/(\beta k(t))]^{-1/2} \exp\{-\beta k(t)x^2/2\} \delta(w)\), with \(\delta(w)\) the Dirac delta function.

The force variance \(\langle \delta f^2 \rangle_{k(t)}\) and the supra-Stokes’ coefficients \(\zeta^{(2)}\) and \(\zeta^{(3)}\), and the third integral covariance function \(\langle \delta w^3 \rangle\) are

\[
\langle \delta f^2 \rangle_{k(t)} = \frac{1}{2k^2} \quad \text{(S3a)}
\]

\[
\zeta[k(t)] = \frac{\beta \gamma}{4k^3} \quad \text{(S3b)}
\]

\[
\zeta^{(2)}[k(t)] = \frac{\beta^2 \gamma^2}{2k^5} \quad \text{(S3c)}
\]

\[
C^{(3)}[k(t), t] \approx \beta^2 \gamma \frac{3t}{16k^6} + \frac{51\gamma}{32k^4} \quad \text{(S3d)}
\]

respectively. In (S3d), we neglect terms of order \(t^{-n}\) for \(n \geq 1\). Since we only consider one-dimensional control, we can analytically determine the minimum-variance (precise) and minimum-bias (accurate) protocols that minimize \(\text{Eq. (15b)}\) and \(\text{Eq. (26b)}\), respectively. The force-variance-optimized protocol proceeds at velocity \(\dot{k} \propto k\), minimum-variance protocol at \(\dot{k} \propto k^{3/2}\), and minimum-bias protocol at \(\dot{k} \propto k^{5/3}\). Figure 1 plots these protocols alongside the naive (constant-velocity) protocol.

As a simple test of the approximations of Eqs. (15a), (15b), and (17), Fig. 2 plots the sum and difference of the first three moments of the excess work from forward and reverse protocols. We choose the forward protocol to be a decrease in the control parameter (decreasing \(k\)), and therefore the reverse increases the control parameter (increasing \(k\)). In all cases the exact calculations agree with the approximation for slow protocols (large \(\Delta t / \tau_{\dot{k}}^{(1)}\) for slowest relaxation time \(\tau_{\dot{k}}^{(1)} \equiv \gamma/(2k(t))\), and overestimate for fast protocols (small \(\Delta t / \tau_{\dot{k}}^{(1)}\)). For large \(\Delta t / \tau_{\dot{k}}^{(1)}\), the first and second moments are approximated by \(\dot{k}(t)\) and \(\dot{k}(t)\) so the sums \((n = 1 \text{ and } 2 \text{ } (+))\) are both approximated by the Stokes’ friction \((5)\) and decrease as \(1/\Delta t\), while the differences \((n = 1 \text{ and } 2 \text{ } (-))\) are proportional to the supra-Stokes’ tensor \((7)\) and decrease as \(1/(\Delta t)^2\). The third moment is approximated by \(\dot{k}(t)\), so the sum \((n = 3 \text{ } (+))\) is approximated by the third integral covariance \((9)\) and decays as \(1/(\Delta t)^2\) (due to the term linear in \(t\) in (S3d)) while the difference \((n = 3 \text{ } (-))\) is proportional to the supra-Stokes tensor and decays as \(1/(\Delta t)^2\).
FIG. 1: The control parameter (spring constant $k$) as a function of time for three different protocols in the model system of a breathing harmonic trap. The control parameter drives the system between relatively strongly and weakly confined states ($k_f/k_i = 1/16$). ‘Naive’ denotes the constant-velocity protocol, ‘Force Variance’ the force-variance-optimized protocol proceeding according to $\dot{k} \propto k$, ‘Accurate’ the minimum-variance protocol with $\dot{k} \propto k^{3/2}$, and ‘Precise’ the minimum-bias protocol with $\dot{k} \propto k^{5/3}$.

FIG. 2: The sum (black, +) and the difference (red, −) of the moments of excess work for forward and reverse protocols of the breathing harmonic trap, as a function of protocol duration $\Delta t$ (scaled by the slowest relaxation time $\tau_f^{(1)} = \gamma/(2k_f)$). Solid lines are the near-equilibrium approximations, given by Eqs. (15a), (15b), and (17). Dashed ($n = 1$), dotted ($n = 2$), and dash-dotted ($n = 3$) curves show exact results. The protocol $k(t)$ is linear with $k_f/k_i = 1/2$. The coefficients $\alpha_1^+ = 1/2$, $\alpha_1^- = 2$, $\alpha_2^+ = 1/4$, $\alpha_2^- = 1/3$, $\alpha_3^+ = 1/2$ and $\alpha_3^- = 1/4$ are chosen such that any moment approximated by the same friction coefficient collapses onto a single curve. In all cases the exact calculations agree with the approximation for large $\Delta t/\tau_f^{(1)}$.

FINITE INTEGRATION BOUNDS

The infinite integration bound on the friction assumes that correlations relax quickly relative to the protocol speed. For the first two moments, despite the finite integration bound generally yielding a more accurate approximation, the infinite bound considerably simplifies the approximation and allows for straightforward protocol optimization.

The effect of finite integration bounds is significant for two calculations: approximation of fourth- and higher-order moments of the excess work, and next-order approximations (including both leading- and next-to-leading-order contributions) for the moments. In the former, one must treat the bound as finite since the $n$-time covariance functions do not decay to zero for some subspace of large time arguments.
In more detail, consider the four-time covariance (kurtosis) \( \langle \delta f_i(0) \delta f_j(t_2) \delta f_k(t_3) \delta f_\ell(t_4) \rangle_{\lambda(t)} \). When any one of the four times \( \{0, t_2, t_3, t_4\} \) significantly differs from the others, the conjugate forces decorrelate, and the kurtosis decays to zero; however, when \( t_3 \sim t_4 \gg t_2 \sim 0 \), any variables separated by significant time decorrelate, and the kurtosis approaches \( \langle \delta f_i(0) \delta f_j(t_2) \rangle_{\lambda(t)} \langle \delta f_k(t_3) \delta f_\ell(t_4) \rangle_{\lambda(t)} \). This limit represents a plane in the \( (t_1, t_2, t_3) \) parameter space where the kurtosis asymptotes to a finite value even for large time arguments. Integrating the above and multiplying by three (to account for permutations of the indices 1, 2, 3, 4) yields the integral four-time covariance in the limit \( t_3 \sim t_4 \gg t_2 \sim 0 \),

\[
C^{(3)}_{ijk}[\lambda(t), t] = 3t \, C^{(1)}_{ij}[\lambda(t), t] C^{(1)}_{k\ell}[\lambda(t), t] = 3t \, \zeta^{(1)}_{ij}[\lambda(t)] \zeta^{(1)}_{k\ell}[\lambda(t)] , \quad t \to \infty .
\]

(S4a)

Since this is the only case that remains finite as \( t \to \infty \), the second line is the approximation (to highest order in \( t \)) for the integral four-time covariance. Parallel arguments hold for the higher-order moments. When approximating the average excess work (22) with both the Stokes’ and supra-Stokes’ tensors (23), finite integration bounds on the Stokes’ friction may be necessary. For finite integration bounds on the Stokes’ friction we replace (23) with

\[
C_{ij}[\lambda(t), \dot{\lambda}(t), t] \equiv C^{(1)}_{ij}[\lambda(t), t] + \frac{1}{4} C^{(2)}_{ijk}[\lambda(t)] \dot{\lambda}_k(t) .
\]

(S5)

For systems with weakly skewed conjugate-force fluctuation, the contribution from the finite bound on the first term can be comparable in magnitude to the contribution from the supra-Stokes’ tensor.

Figures 3 shows the first and second moments of the excess work for forward and reverse protocols, compared to different approximations. For the mean excess work, the supra-Stokes’ tensor contributes with opposite sign from the Stokes’ friction for forward (decreasing \( k \)) protocols and with same sign for reverse (increasing \( k \)). Since the Stokes’-friction approximation overestimates the excess work in both cases \( (a,c) \), adding the supra-Stokes’ tensor reduces the accuracy of the approximation for the reverse excess work. This effect is an artifact of the infinite integral bound; indeed, if the bound is kept finite \( (C^{(1)} \) rather than \( \zeta^{(1)} \), the first integral autocovariance overestimates for forward protocols and underestimates for reverse protocols, and adding the supra-Stokes’ tensor improves the approximation in both cases. Finally, we note that the Stokes’ friction does not always overestimate for both forward and reverse protocols. For the second moment in \( (b,d) \), the Stokes’ friction overestimates for forward protocols and underestimates for reverse protocols, and the supra-Stokes’ tensor improves the approximation when either the Stokes’ friction or first integral covariance are used.

Optimization of the (finite-bound) integral \( n \)-time covariance functions is not in general analytically tractable, but still could be done numerically.

**FINSLER GEOMETRY**

The addition of the supra-Stokes’ tensor (7) comes at the cost of a more complex geometric structure. In contrast to Riemannian geometry, Finsler geometry is not restricted to a quadratic norm. In general, the inner products are not characterized by points, but rather by points and directions. Despite this, several useful concepts from Riemannian geometry (notably curvature, length, and geodesics) generalize. There are therefore standard procedures to find the geodesics despite the more complex landscapes induced by the generalized cubic Finsler metric (23).

Finsler geometry has several applications in both physics and biology [3]. In the thermodynamic context, the flexion tensor [4]

\[
F_{ijk}[\lambda(t)] \equiv \left. \frac{\partial^3 \ln \pi(x|\lambda)}{\partial \lambda_i \partial \lambda_j \partial \lambda_k} \right|_{\lambda(t)}
\]

(S6)

is a Finsler metric arising as the third-order contribution to the near-equilibrium expansion of the relative entropy (Kullback-Leibler divergence) of a probability distribution relative to the equilibrium distribution [3]:

\[
D(\pi(x|\lambda)||\pi(x'|\lambda)) \equiv \int_{-\infty}^{\infty} dx \, \pi(x|\lambda) \ln \frac{\pi(x'|\lambda)}{\pi(x|\lambda)} \approx \mathcal{I}_{ij} \Delta \lambda_i \Delta \lambda_j + F_{ijk} \Delta \lambda_i \Delta \lambda_j \Delta \lambda_k .
\]

(S7a)

(S7b)
FIG. 3: The first and second moments of excess works for forward (a,b) and reverse (c,d) protocols of the breathing harmonic trap as a function of protocol time $\Delta t$ (scaled by the slowest relaxation time $\tau_{f}^{(1)} = \gamma/(2k_{f})$). The dashed, dotted, and dash-dotted lines are different forms of the near-equilibrium approximation. The protocol $k(t)$ is linear with $k_{f}/k_{i} = 1/2$.

The coskewness tensor (20) is related to the flexion tensor by

$$F_{i j k}[\lambda(t)] = \langle \delta f_{i} \delta f_{j} \delta f_{k} \rangle_{\lambda(t)} + \sum_{\sigma_{i j k}} \left\langle \frac{\partial^{2} \ln \pi(x|\lambda)}{\partial \lambda_{i} \partial \lambda_{j}} \frac{\partial \ln \pi(x|\lambda)}{\partial \lambda_{k}} \right\rangle_{\lambda(t)},$$

(S8)

where the sum is over permutations of the indices $\sigma_{i j k} = \{ 123, 132, 321 \}$. In weak gravitational lensing [5, 6] the flexion tensor is the third-order correction to the shapes of images, which is described as flexing the shape of images from an ellipse towards a banana shape. Similarly, the supra-Stokes' tensor (product of coskewness tensor and integral double relaxation time (20)) skews the thermodynamic geometry of minimum-dissipation protocols, as demonstrated in Fig. 1 of the main text.

* Electronic address: dsivak@sfu.ca
† Electronic address: sblaber@sfu.ca

[1] A. Imparato, L. Peliti, G. Pesce, G. Rusciano, and A. Sasso, Phys. Rev. E 76, 050101 (2007).
[2] A. P. Solon and J. M. Horowitz, Phys. Rev. Lett. 120, 180605 (2018).
[3] P. L. Antonelli, R. S. Ingarden, and M. Matsumoto, The theory of sprays and Finsler spaces with applications in physics and biology, vol. 58 (Springer Science & Business Media, 2013).
[4] E. Sellentin, M. Quartin, and L. Amendola, Mon. Not. R. Astron. Soc. 441, 1831 (2014).
[5] D. M. Goldberg and D. J. Bacon, Astrophys. J 619, 741 (2005).
[6] D. Bacon, D. Goldberg, B. Rowe, and A. Taylor, Mon. Not. R. Astron. Soc. 365, 414 (2006).