Noise, Sign Problems, and Statistics

Michael G. Endres,[1,*] David B. Kaplan,[2,†] Jong-Wan Lee,[2,‡] and Amy N. Nicholson[2,§]

[1]*Theoretical Research Division, RIKEN Nishina Center, Wako, Saitama 351-0198, Japan*

[2]*Institute for Nuclear Theory, University of Washington, Seattle, Washington 98195-1550, USA*

(Received 10 June 2011; published 8 November 2011)

We show how sign problems in simulations of many-body systems can manifest themselves in the form of heavy-tailed correlator distributions, similar to what is seen in electron propagation through disordered media. We propose an alternative statistical approach for extracting ground state energies in such systems, illustrating the method with a toy model and with lattice data for unitary fermions.

DOI: 10.1103/PhysRevLett.107.201601 PACS numbers: 11.15.Ha, 05.50.+q, 12.38.Gc, 71.10.Fd

Introduction.—One of the most challenging and interesting problems in physics is to understand the properties of a system of many strongly interacting fermions. Numerical simulation is an important tool for understanding the ground state, and the common approach is to compute the N-body correlator $C_N(\tau, \phi) = \langle 0|\Psi_N(\tau)\Psi_N^\dagger(0)|0\rangle_\phi$, where $\Psi_N^\dagger(0)$, $\Psi_N(\tau)$ are interpolating fields which create an N-body state at Euclidean time zero and annihilate it at time τ, and ϕ is a stochastic field responsible for fermion interactions. The field ϕ could be the dynamical gluon field in the case of QCD, for example, or an auxiliary field to induce short-range interactions. For large τ the averaged correlator asymptotically approaches

$$\langle C_N(\tau, \phi)\rangle \sim Ze^{-\tau E_0(N)}, \tag{1}$$

where $E_0(N)$ is the ground state energy of the system and $\sqrt{Z}$ is the amplitude for Ψ to create the ground state. Therefore, if one computes $-\frac{1}{\tau}\ln\bar{C}_N(\tau)$, where $\bar{C}_N(\tau) = \frac{1}{\mathcal{N}}\sum_i C_N(\tau, \phi_i)$ is the sample mean computed on an ensemble of $\mathcal{N}$ statistically independent ϕ fields, one expects to see a "plateau" at large τ whose height yields the ground state energy $E_0(N)$. Excited state energies and response of the ground state to probes can also be computed by modifications of this technique.

The computation of $-\frac{1}{\tau}\ln\bar{C}_N(\tau)$ can be problematic, however: it might be excessively noisy, or it may drift with τ and never find a plateau. We wish to address these problems here, defining the former as a "noise" problem, and the latter as an "overlap" problem, both of which can be related to the sign problem encountered in lattice simulations at nonzero chemical potential. In particular, referring to recent lattice simulations by the present authors of large numbers of unitary fermions, we show that the problems encountered can be manifestations of heavy-tailed distributions for $C_N(\tau, \phi)$ which make computing $\ln\langle C_N\rangle$ very difficult, and that the ideal estimator for this quantity might not simply be $\ln\bar{C}_N$, as is commonly used. We find here that a cumulant expansion in the log of the correlator is a more efficient estimator, for example. More generally, we suggest that a study of the statistics of systems exhibiting noise or an overlap problem might be exploited to greatly facilitate the extraction of useful physics from numerical simulations.

Noise and the physical spectrum.—Grand canonical simulations, such as lattice QCD at nonzero quark chemical potential, typically encounter a path integral measure neither real nor positive, leading to an exponentially hard computation [1]. In canonical simulations, where one computes correlators C_N at fixed particle number, the measure can be real and positive but the pathology reappears as a noise problem, sampled values of $C_N(\tau, \phi)$ varying wildly from the mean. This noise problem does not arise simply because of Fermi statistics; for example, constructing correlators C_N as $N \times N$ Slater determinants of one-body propagators leads to a computational cost from the determinant only scaling as N^3. Instead, the noise problem seems to arise from the existence of multiparticle states for which the energy per constituent is lower than for the states one wants to study, a problem that can occur in bosonic systems as well. This relation between noise and the physical spectrum has been quantified by Lepage [2]. For example, in QCD the expectation of a $3A$ quark correlator for a nucleus of atomic number A and mass M_A is $\langle C_A\rangle \sim e^{-M_A\tau}$, while the variance in the sample mean $\bar{C}_A$ can be estimated as

$$\sigma^2 = \frac{1}{\mathcal{N}}(\langle C_A^\dagger C_A\rangle - \langle C_A^\dagger\rangle\langle C_A\rangle) \sim \frac{1}{\mathcal{N}}e^{-3Am_\pi\tau} \tag{2}$$

for sample size $\mathcal{N}$. Since C_A corresponds to $3A$ quark propagators and $C_A^\dagger$ to $3A$ antiquark propagators, the variance is dominated by the state with $3A$ pions. Thus the signal to noise ratio scales as $\sqrt{\mathcal{N}}\exp(-\zeta\tau)$, where $\zeta = (M_A - \frac{3A}{2}m_\pi) \gg 0$, falling off exponentially with both time τ and atomic number A. The parameter ζ is the same quantity that characterizes the sign problem in the grand canonical case [3,4], and the noise and sign problems are therefore presumably closely related.

If the distribution of correlation functions C_A was Gaussian, as assumed in Ref. [2], then the small mean and large variance for C_A would imply a high degree of cancellation between the contributions from different background gauge fields. However, small mean and large variance is also consistent with a very heavy-tailed distribution, where typical contributions are very small and positive, exceptional contributions are extremely large and positive, and no cancellations are involved.

We argue here that a heavy-tail scenario could be more generic for correlation functions of large numbers of fermions. We give examples both from real simulations of unitary fermions as well as from a simple toy model. Furthermore, we show that in this case there can be statistical tools which greatly improve one's ability to discern the signal from the noise.

Unitary fermions and a mean-field description.— Nonrelativistic fermions with strong short-range interactions tuned to a conformal fixed point where the phase shift satisfies $\delta(k) = \pi/2$ for all k are called "unitary fermions." This conformal field theory is interesting to study for both its simplicity and universality, its challenges for many-body theory, and because it can be realized and studied experimentally using trapped atoms tuned to a Feshbach resonance [5]. It is also an ideal nonperturbative theory for studying fermion sign problems on the lattice, being simpler and faster to simulate than QCD. At its most basic, the lattice action is the obvious discretization of the Euclidean Lagrangian [6]

$$\psi^\dagger(\partial_\tau - \nabla^2/2M)\psi - \tfrac{1}{2}m^2\phi^2 + \phi\psi^\dagger\psi, \qquad (3)$$

where ϕ is a nonpropagating auxiliary field with m^2 tuned to a critical value, and ψ is a spin $\frac{1}{2}$ fermion with mass M; a more sophisticated action tuned to reduce discretization errors was recently presented in [7]. A simulation of this theory reveals a distribution for N-body correlators $C_N(\tau, \phi)$, which is increasingly non-Gaussian at late τ; in fact, it is $\ln C_N$ which appears to be roughly normally distributed, as shown in Fig. 1, so that $C_N(\tau, \phi)$ is approximately log-normal distributed with an increasingly large σ and long tail at late time.

The appearance of a heavy-tailed distribution should not be surprising, since the system is similar to electrons propagating in disordered media, where heavy-tailed distributions are ubiquitous in the vicinity of the Anderson localization transition. For example, it is found that for physical quantities such as the current relaxation time or normalized local density of states, the distribution function $P(z)$ scales as $\exp(-C_d \ln^d z)$. A particularly simple way to derive these results is to use the optimal fluctuation method of Ref. [8], which is a mean-field approach. We can adapt these methods to the current problem, defining the variable $Y = \ln C_N(\tau, \phi)$ and computing its probability distribution $P(y)$ as

$$P(y) \propto \int D\phi e^{-S_\phi}\delta(Y(\tau,\phi) - y) = \int D\phi \frac{dt}{2\pi} e^{-S}, \qquad (4)$$

where $S_\phi = \int d^4x \frac{m^2}{2}\phi^2$ and $S = S_\phi - it[\ln C_N(\tau,\phi) - y]$. Using the power divergent subtraction scheme [9], we have $m^2 = M\lambda/4\pi$, where the renormalization scale λ is taken to be the physical momentum scale in the problem—in this case, $\lambda = k_F \equiv (3\pi^2 N/V)^{1/3}$, $N/2$ being the number of fermions with a single spin orientation. We proceed now to evaluate this integral using a mean-field expansion; it is not evident that there is a small parameter to justify this expansion, but the leading order result is illuminating and fits the numerical data well. We expand about $\phi(x) = \phi_0$, $t = t_0$, and use the fact that for large τ the nth functional derivative of $\ln C_N(\tau, \phi)$ with respect to $\phi(x)$ equals the 1-loop Feynman diagram with n insertions of $\psi^\dagger\psi$ in the presence of a chemical potential $\mu = k_F^2/(2M)$. The equations for ϕ_0 and t_0 are given by

$$t_0 = -i\frac{m^2\phi_0}{\langle n(x)\rangle_c} = -i\frac{Vm^2\phi_0}{N}, \qquad \phi_0 = -\frac{y - \ln Z + \tau E_0(N)}{N\tau}, \qquad (5)$$

where $E_0(N) = 3NE_F/5$ is the total energy of N free degenerate fermions ($N/2$ of each spin), and Z is the overlap of the source and sink with the free fermion state. The leading term in the mean-field expansion for $P(y)$ can therefore be expressed as $P(y) \propto \exp[-\frac{(y-\bar{y})^2}{2\sigma^2}]$ with

$$\bar{y} = \ln Z - \tau E_0(N), \qquad \sigma^2 = \frac{40}{9\pi}E_0(N)\tau. \qquad (6)$$

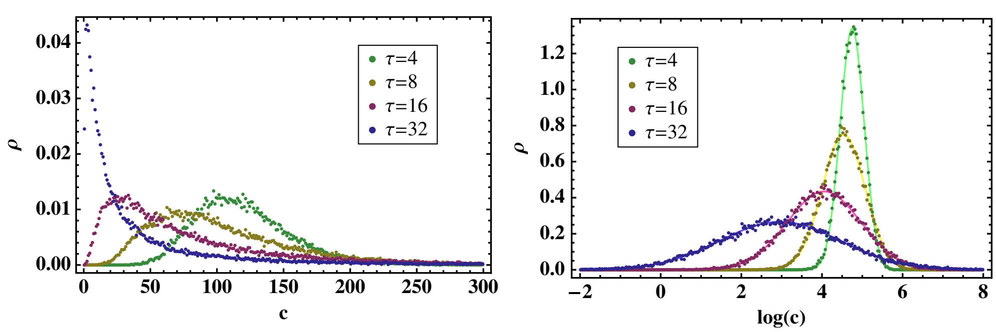

FIG. 1 (color online). Distribution histograms for $c = C_N(\tau, \phi)$ and $\ln(c)$ for $N = 4$ unitary fermions at several times τ, taken from Ref. [11]. Curves fitting $\ln(c)$ are Gaussian, implying that c is approximately log-normal distributed, with σ^2 increasing with time.

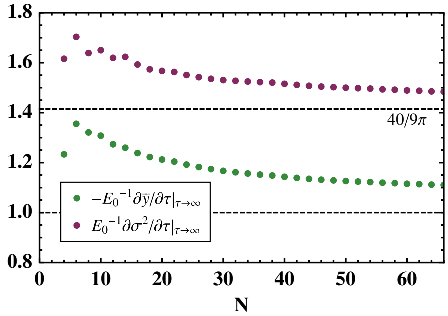

FIG. 2 (color online). The quantities $-(1/E_0)(\partial\bar{y}/\partial\tau)$ and $(1/E_0)(\partial\sigma^2/\partial\tau)$ as a function of N for unitary fermions at late times on a lattice of size $L = 10$, compared to the mean-field prediction of Eq. (6) (dashed lines).

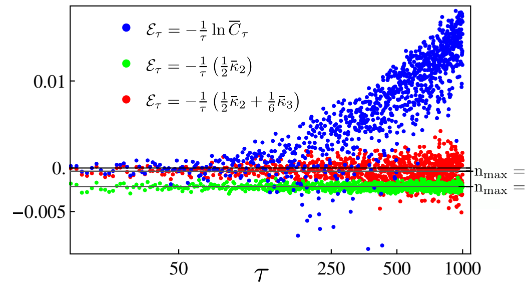

FIG. 3 (color online). Simulation of the energy $\mathcal{E}_\tau$ for the model Eq. (8) with $g = 1/2$. The exact answer is $\mathcal{E}_\tau = 0$ (black line); exact values of Eq. (7) truncated at order $n = 2, 3$ are indicated.

This describes a log-normal distribution for the N-fermion propagator $C_N(\tau, \phi)$, with both mean and variance growing in magnitude with time in units of the energy of N free degenerate fermions. In Fig. 2 we plot the quantities $-\frac{1}{E_0}\frac{\partial\bar{y}}{\partial\tau}$ and $\frac{1}{E_0}\frac{\partial\sigma^2}{\partial\tau}$ as a function of N obtained from correlator distribution data for unitary fermions at late τ, and find that the gross features of the results are compatible with the mean-field estimates of unity and $40/9\pi$ obtained from Eq. (6).

Toy model.—It would be useful to devise an algorithm to reliably estimate energies without having to exhaustively sample the long tail of the correlator distribution, yet without making incorrect assumptions about the exact functional form of that tail. An approach we suggest here is to exploit the general relationship between stochastic variables X and $Y = \ln X$:

$$\ln\langle X\rangle = \sum_{n=1}^{\infty}\frac{\kappa_n}{n!}, \tag{7}$$

where κ_n is the nth cumulant of Y. This relation can be proved by noting that the generating function for the κ_n is $\ln\phi_Y(t)$, where $\phi_Y(t) = \langle e^{Yt}\rangle = \langle X^t\rangle$ is the moment generating function for Y, and evaluating at $t = 1$, assumed to be within the radius of convergence. The motivation for investigating Eq. (7) is that if the distribution $P(X)$ were exactly log-normal, the above sum would end after the second term, as $\kappa_{n>2}$ would all vanish; therefore, by replacing the κ_n by sample cumulants $\bar{\kappa}_n$ and truncating the sum at $n = n_{max}$, one might hope to have a reliable estimator for $\ln\langle X\rangle$, provided that the systematic error from truncating Eq. (7) and the statistical error from sampling $\bar{\kappa}_n$ can be simultaneously minimized.

Distributions with log-normal-like tails arise naturally in products of stochastic variables. The propagator $C_N(\tau, \phi)$ for unitary fermions can be expressed in a transfer matrix formalism as the product of τ matrices—one per time hop—each of which is the direct product of N $V \times V$ matrices of the form $e^{-K/2}(1 + g\varphi)e^{-K/2}$, where K is a constant matrix (the spatial kinetic operator), φ is a random diagonal matrix with $O(1)$ entries corresponding to stochastic ϕ fields living on the time links, and g is a coupling constant [identified with $1/m^2$ in Eq. (3)] that has been tuned to a particular critical value that is $O(1)$. Unfortunately, little seems to be known about products of random matrices beyond dimension two [10]. Therefore, we analyze instead a toy model where we define a "correlator" C_τ as a product of random numbers, and an "energy" $\mathcal{E} = \lim_{\tau\to\infty}\mathcal{E}_\tau$, where

$$C_\tau = \prod_{i=1}^{\tau}(1 + g\varphi_i), \qquad \mathcal{E}_\tau = -\frac{1}{\tau}\ln\langle C_\tau\rangle, \tag{8}$$

where $0 \le g \le 1$ and the φ_i are independent and identically distributed random numbers with a uniform distribution on the interval $[-1, 1]$. The exact value for the energy is obviously $\mathcal{E}_\tau = 0$ since the statistical average of the correlator is $\langle C_\tau\rangle = 1$. The cumulants of the variable $Y = \ln(C_\tau)$ are given by

$$\kappa_1 = \tau\left[\frac{1}{2}\log(1 - g^2) + \frac{\tanh^{-1}(g)}{g} - 1\right],$$
$$\frac{\kappa_n}{n!} = \tau\left[\frac{(-1)^n}{n} - \mathrm{Li}_{1-n}\left(\frac{1+g}{1-g}\right)\frac{[2\tanh^{-1}(g)]^n}{n!}\right]$$

for $n \ge 2$; for $g < 1$ one finds that the $\kappa_n/n!$ rapidly decrease as n increases.

In Fig. 3 we show the results of a simulation where we compute $\mathcal{E}_\tau$ for $g = \frac{1}{2}$ and $\tau = 1, \ldots, 1000$. At each value of τ we independently generated an ensemble of values for C_τ of size $\mathcal{N} = 50\,000$. From that ensemble we computed $\mathcal{E}_\tau$ by (i) using the conventional estimator $\mathcal{E}_\tau = -\frac{1}{\tau}\ln\bar{C}_\tau$ (blue), which shows a striking systematic error for $\tau \gtrsim 50$, and statistical noise increasing up to $\tau \simeq 500$ but decreasing beyond that; (ii) using Eq. (7) truncated at $n_{max} = 2$ using sample cumulants $\bar{\kappa}_n$ (green), showing a τ-independent systematic error with smaller but slowly growing statistical error; (iii) Eq. (7) truncated at $n_{max} = 3$ (red) with a negligible constant systematic error but a larger statistical error. Evidently, one trades systematic

TABLE I. $\mathcal{E}$ determined from 250 blocks of 50 000 configurations each for the model Eq. (8) with $\tau = 1000$, $g = 1/2$.

Method	$\mathcal{E}$	Statistical error	Systematic error
Conventional	0.014 932	0.002 485	$\cdots$
$\kappa_{n\le2}$	−0.002 159	0.000 304	−0.002 165
$\kappa_{n\le3}$	−0.000 412	0.001 618	−0.000 324
$\kappa_{n\le4}$	−0.000 647	0.008 379	0.000 050
$\kappa_{n\le5}$	−0.001 794	0.037 561	3.34×10^{-6}
$\kappa_{n\le6}$	0.010 943	0.147 739	-1.22×10^{-6}

error for statistical error by truncating Eq. (7) at increasingly large $n_{\max}$.

Table I displays results of a simulation of 1.25×10^7 ϕ configurations blocked into 250 blocks of 50 000 each, for the model Eq. (8) at $\tau = 1000$ and $g = 1/2$. We give the conventional estimate $\mathcal{E}_\tau = -1/\tau \ln \bar{C}_\tau$ and estimates based on the cumulant expansion Eq. (7) truncated at various $n_{\max}$, where the exact value is $\mathcal{E} = 0$. For each method we give the computed value for $\mathcal{E}$ and the statistical error; for the cumulant expansion we also give the exact systematic error from truncating Eq. (7) at $n = n_{\max}$ using our analytic expressions for κ_n. These numbers show how the conventional method gives a wrong answer with deceptively small statistical error. For the cumulant expansion one sees again the trade of systematic error for statistical error with increasing $n_{\max}$. Table I shows that the combined error is minimized for $n_{\max} = 3$, justified by noting that the $n_{\max} = 4$ result with statistical errors encompasses the $n_{\max} = 3$ result; we suggest this as a practical algorithm for determining where to truncate the cumulant expansion in general. Figure 4 shows how this works in a real simulation for 50 trapped unitary fermions [11], where truncating the expansion at $n_{\max} = 4$ is supported by the data.

Discussion.—Heavy-tail distributions are likely to be ubiquitous in N-body simulations, and perhaps even in other types of noisy calculations; for example, there seems to be evidence for similar phenomena in multibaryon computations in lattice QCD [12]. With such distributions theoretical statistical means can deviate wildly from sample means for any realizable sample size and render conventional estimates of expected fluctuations irrelevant. We have shown that there are more efficient estimators for ground state energies using the cumulants of the log of the correlator instead of the conventional effective mass, at least for positive correlators. This method is presumably only effective for nonpositive data provided extended tails to the distribution are asymmetric.

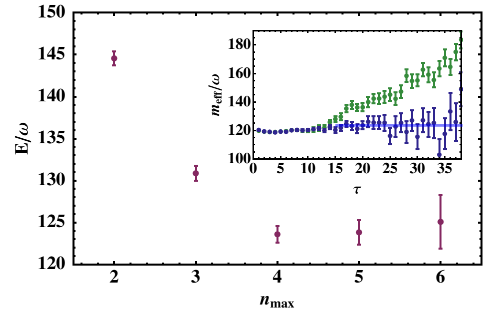

FIG. 4 (color online). Energy for 50 unitary fermions in a harmonic trap of frequency ω, 10^6 configurations; fits performed using Eq. (7) truncated at order $n_{\max}$. Inset: Conventional effective mass $m_{\mathrm{eff}}(\tau) = \log \bar{C}(\tau)/\bar{C}(\tau+1)$ (green, rising) and corresponding fitted cumulant effective mass with $n_{\max} = 4$ (blue, flat).

This work was supported in part by U.S. DOE Grant No. DE-FG02-00ER41132. M. G. E is supported by the Foreign Postdoctoral Researcher program at RIKEN.

*endres@riken.jp
†dbkaplan@uw.edu
‡jwlee823@u.washington.edu
§amynn@u.washington.edu

[1] M. Troyer and U.-J. Wiese, Phys. Rev. Lett. **94**, 170201 (2005).
[2] G. Peter Lepage, in *Proceedings of the 1989 Theoretical Advanced Study Institute in Elementary Particle Physics, Boulder, CO, 1989*, edited by T. DeGrand and D. Toussaint (World Scientific, Singapore, 1990).
[3] P. E. Gibbs, Report No. PRINT-86-0389 (GLASGOW), 1985 [inspirehep.net/record/18426].
[4] K. Splittorff and J. J. M. Verbaarschot, Phys. Rev. Lett. **98**, 031601 (2007).
[5] T.-L. Ho, Phys. Rev. Lett. **92**, 090402 (2004).
[6] J.-W. Chen and D. B. Kaplan, Phys. Rev. Lett. **92**, 257002 (2004).
[7] M. G. Endres, D. B. Kaplan, J.-W. Lee, and A. N. Nicholson, Proc. Sci. LATTICE2010 (2010) 182.
[8] I. E. Smolyarenko and B. L. Altshuler, Phys. Rev. B **55**, 10 451 (1997).
[9] D. B. Kaplan, M. J. Savage, and M. B. Wise, Phys. Lett. B **424**, 390 (1998).
[10] A. D. Jackson, B. Lautrup, P. Johansen, and M. Nielsen, Phys. Rev. E **66**, 066124 (2002).
[11] M. G. Endres, D. B. Kaplan, J.-W. Lee, and A. N. Nicholson, Phys. Rev. A **84**, 043644 (2011).
[12] W. Detmold, K. Orginos, and M. J. Savage, for the NPLQCD Collaboration (private communication).