Noise, sign problems, and statistics
arXiv:1106.0073 [hep-lat]
Michael Endres, D.K., Jong-Wan Lee, Amy Nicholson
... & work in progress
Physics motivation: can’t we get beyond this cartoon??

Sign problem!

This talk:

• From sign problem to noise
• Surprisingly universal features of noise
• Can these features be used to tame the noise?
The “sign” problem in the grand canonical approach: 
\[ \text{Det}(\mathcal{D}+\mu \gamma^0) \text{ complex} \]

- physics happens for \( \mu \geq m_N/3 \)...
- ...but sign problem starts at \( \mu = m_\pi/2 \)!

\[ \Delta \mu = \left( \frac{m_N}{3} - \frac{m_\pi}{2} \right) \]

Explanation (2-flavor QCD):
\[ |\text{Det}(\mathcal{D}+\mu \gamma^0)| \approx \text{isospin} \text{ chemical potential} \]

Role of phase: eliminate pion condensate for \( \mu \geq m_\pi/2 \)!

---

P.E. Gibbs, 1986

K. Splittorff
J. Verbaarschot, 2006
Canonical approach?
Compute correlator of N quarks with $\mu=0$
No sign problem...but now a noise problem

\begin{align*}
\text{nucleon correlator} \\
\text{signal: } & \sim e^{-m_N T} \\
\text{noise: } & \sim \frac{1}{\sqrt{N_{\text{conf.}}}} e^{-\frac{3}{2}m_\pi T}
\end{align*}

$$\frac{\text{signal}}{\text{noise}} \sim \sqrt{N_{\text{conf.}}} e^{-3T \left(\frac{m_N}{3} - \frac{m_\pi}{2}\right)}$$

Same factor as grand canonical

Parisi, Lepage 1980's
EXAMPLE:

Triton B.E.  
S. R. Beane et al. (NPLQCD), 

$m_\pi = 390$ MeV
Actual QCD data

Plotted: $-\frac{1}{t} \ln \frac{\sigma(t)}{\bar{x}(t)} \sim A \left( m_N - \frac{3}{2} m_\pi \right)$

Conclusion: Parisi/Lepage = qualitative estimate of noise problem
Think like a quark in a single gauge configuration

grand canonical

canonical

Am I going to be in a light pion? Or a heavy nucleon?  
Don’t know!  
Play safe: assume pion.  

Propagator: $\sim e^{-(m_\pi/2)T}$  

If nucleon, cancellations between configurations to: $\sim e^{-(m_N/3)T}$  

David B. Kaplan ~ Lattice 2011 ~ July 15, 2011
Think like a quark in a single gauge configuration

grand canonical

canonical

Am I going to be in a light pion? Or a heavy nucleon?

Don't know!

Play safe: assume pion. Propagator: \( \sim e^{-(m_\pi/2)T} \)

If nucleon, cancellations between configurations to: \( \sim e^{-(m_N/3)T} \)
Distribution of A-baryon correlators *might* look like:

- Some non-universal distribution
- Exponentially small mean \( \mu \sim e^{-A m_{NT} T} \)
- Large variance \( \sigma \sim e^{-\frac{3}{2} A m_{\pi} T} \)
Distribution of A-baryon correlators might look like:

Some non-universal distribution

Exponentially small mean $\mu \sim e^{-Am_{NT}}$

Large variance $\sigma \sim e^{-\frac{3}{2}Am_{\pi}T}$

This picture is actually not right.

Look at a simpler system where we can get large statistics: unitary fermions.
Digression: what are unitary fermions?
(Lattice 2011 talks by J. Drut, J.-W. Lee)

Nonrelativistic 2-particle scattering:

\[ A = \frac{4\pi}{M} \frac{1}{p \cot \delta - ip} \]

"unitary" fermions: \[ p \cot \delta = 0 \]

\[ \delta(p) = \frac{\pi}{2} \]
Digression: what are unitary fermions?
(Lattice 2011 talks by J. Drut, J.-W. Lee)

Nonrelativistic 2-particle scattering:

\[
A = \frac{4\pi}{M} \frac{1}{p \cot \delta - ip}
\]

“unitary” fermions:

\[
p \cot \delta = 0
\]

\[
\delta(p) = \frac{\pi}{2}
\]

- A strongly-coupled conformal system
- Studied experimentally with dilute trapped atoms @ Feshbach resonance
  (JILA, MIT, Innsbruck)
- Exhibits superfluidity

- Nonrelativistic fermions, $\mu=0$
- Short-range momentum-dependent 4-fermion interaction induced by auxiliary scalar field
- Interaction tuned to conformal fixed pt.

- Less severe sign problem than QCD; no gauge symmetry; nonrelativistic (quenched)
- Have simulated up to $N=70$ fermions on $14^3 \times 64$ lattice
- ~1% accuracy in energies
- up to 2 billion configurations for auxiliary field
Simulation:

\[ \phi \text{ lives on time links} \]
\[ \text{(couples to } \psi^* \psi) \]

Generate an ensemble of random \( \phi \) fields, compute average of an N-particle correlator \( C_N(T; \phi) \) from \( t=0 \) to \( t=T \)

Extract ground state energy:

\[ E_N = \lim_{T \to \infty} \left[ -\frac{1}{T} \ln \langle C_N(T, \phi) \rangle_\phi \right] \]
Correlators are products of many transfer matrices in background random $\Phi$.
Example of conventional effective mass plot

\[ \frac{m_{\text{eff}}(\tau)}{E_{\text{Free}}} \]

\( \tau \)

\( N = 46 \) fermions
\( L = 12 \)
40 M configs
Example of conventional effective mass plot

- noise
- drift
- worse for larger N

$N = 46$ fermions
$L = 12$
$40 \text{ M configs}$
Effective mass plot with standard technique

Example of conventional effective mass plot

\[ m_{\text{eff}}(\tau)/E_{\text{Free}} \]

- noise
- drift
- worse for larger N

**N = 46 fermions**
**L = 12**
**40 M configs**

Is there any information here?
Look at raw correlator probability distributions:

N=4 correlator distribution at different times ($\tau$):

- $\tau=4$ (green)
- $\tau=8$ (brown)
- $\tau=16$ (purple)
- $\tau=32$ (blue)

Long tails at late times
...but look at distribution for LOG of correlator:

Distribution of \( \log[\text{correlator}] \) with Gaussian fit for different values of \( \tau \):
- \( \tau = 4 \)
- \( \tau = 8 \)
- \( \tau = 16 \)
- \( \tau = 32 \)
Correlators seem to flow toward a log-normal distribution (which is described by only two parameters).

Noise and drift in measurement due to problems sampling long tail for computing $<C>$
**Digression: statistics & the RG**
Digression: statistics & the RG

Probability distribution:

\[ P(x) \]
Digression: statistics & the RG

Probability distribution:

\[ P(x) \]

Moment generating function:

\[ \phi(t) = \langle e^{-tx} \rangle = 1 - t\langle x \rangle + \frac{t^2}{2}\langle x^2 \rangle + \ldots \]
Digression: statistics & the RG

**Probability distribution:**

\[ P(x) \]

**Moment generating function:**

\[ \phi(t) = \langle e^{-tx} \rangle = 1 - t\langle x \rangle + \frac{t^2}{2}\langle x^2 \rangle + \ldots \]

**Cumulant generating function:**

\[ -\ln \phi(t) = t\langle x \rangle + \frac{t^2}{2} (\langle x^2 \rangle - \langle x \rangle^2) + \ldots \]

\[ = \sum_{n=1}^{\infty} \frac{t^n}{n!} \kappa_n \]
Digression: statistics & the RG

Probability distribution:

\[ P(x) \]

Moment generating function:

\[ \phi(t) = \langle e^{-tx} \rangle = 1 - t\langle x \rangle + \frac{t^2}{2} \langle x^2 \rangle + \ldots \]

Cumulant generating function:

\[ -\ln \phi(t) = t\langle x \rangle + \frac{t^2}{2} (\langle x^2 \rangle - \langle x \rangle^2) + \ldots \]

\[ = \sum_{n=1}^{\infty} \frac{t^n}{n!} \kappa_n \]
**Digression: statistics & the RG**

**Probability distribution:**

\[ P(x) \]

**Moment generating function:**

\[
\phi(t) = \langle e^{-tx} \rangle = 1 - t\langle x \rangle + \frac{t^2}{2}\langle x^2 \rangle + \ldots
\]

**Cumulant generating function:**

\[
-\ln \phi(t) = t\langle x \rangle + \frac{t^2}{2} (\langle x^2 \rangle - \langle x \rangle^2) + \ldots
\]

\[
= \sum_{n=1}^{\infty} \frac{t^n}{n!} \kappa_n
\]

**Analogue with path integral**

Like \( e^{-S} \)
Digression: statistics & the RG

Probability distribution:

\[ P(x) \]

Moment generating function:

\[ \phi(t) = \langle e^{-tx} \rangle = 1 - t\langle x \rangle + \frac{t^2}{2} \langle x^2 \rangle + \ldots \]

Cumulant generating function:

\[ -\ln \phi(t) = t\langle x \rangle + \frac{t^2}{2} (\langle x^2 \rangle - \langle x \rangle^2) + \ldots \]

\[ = \sum_{n=1}^{\infty} \frac{t^n}{n!} \kappa_n \]
Digression: statistics & the RG

Probability distribution:

\[ P(x) \]

Moment generating function:

\[ \phi(t) = \langle e^{-tx} \rangle = 1 - t\langle x \rangle + \frac{t^2}{2}\langle x^2 \rangle + \ldots \]

Cumulant generating function:

\[ -\ln \phi(t) = t\langle x \rangle + \frac{t^2}{2} \left( \langle x^2 \rangle - \langle x \rangle^2 \right) + \ldots \]

\[ = \sum_{n=1}^{\infty} \frac{t^n}{n!} \kappa_n \]
Digression: statistics & the RG

Probability distribution:

\[ P(x) \]

Moment generating function:

\[ \phi(t) = \langle e^{-tx} \rangle = 1 - t\langle x \rangle + \frac{t^2}{2}\langle x^2 \rangle + \ldots \]

Cumulant generating function:

\[ -\ln \phi(t) = t\langle x \rangle + \frac{t^2}{2}(\langle x^2 \rangle - \langle x \rangle^2) + \ldots \]

\[ \sum_{n=1}^{\infty} \frac{t^n}{n!} \kappa_n \]

Like effective action \( W[J] \)

Like partition function \( Z[J] \)

Analogue with path integral

Like \( e^{-S} \)

\( n^{th} \) cumulant - like \( n \)-pt. operators in effective action, increasing dimension
The Central Limit Theorem as RG flow

$P(x) = \text{some probability distribution with zero mean, unit variance.}$

Characterize by cumulants:

$P(0, 1, \kappa_3, \kappa_4, \ldots ; x)$
The Central Limit Theorem as RG flow

$P(x) = \text{some probability distribution with zero mean, unit variance.}$

Characterize by cumulants:

$P(0, 1, \kappa_3, \kappa_4, \ldots ; x)$

Average pairwise: 
$x_1 \rightarrow \frac{x_1 + x_2}{\sqrt{2}}, \quad x_2 \rightarrow \frac{x_3 + x_4}{\sqrt{2}}, \ldots$

(and rescale)
The Central Limit Theorem as RG flow

\( P(x) = \) some probability distribution with zero mean, unit variance.

Characterize by cumulants:
\[
P(0, 1, \kappa_3, \kappa_4, \ldots; x)
\]

Average pairwise:
(and rescale)
\[
x_1 \rightarrow \frac{x_1 + x_2}{\sqrt{2}}, \quad x_2 \rightarrow \frac{x_3 + x_4}{\sqrt{2}}, \ldots
\]

Cumulants get rescaled:
\[
\kappa_n \rightarrow 2^{(1-n/2)} \kappa_n
\]
The Central Limit Theorem as RG flow

$P(x) = \text{some probability distribution with zero mean, unit variance.}$

Characterize by cumulants:

$P(0, 1, \kappa_3, \kappa_4, \ldots; x)$

Average pairwise: $x_1 \rightarrow \frac{x_1 + x_2}{\sqrt{2}}, x_2 \rightarrow \frac{x_3 + x_4}{\sqrt{2}}, \ldots$

(and rescale)

Cumulants get rescaled:

$\kappa_n \rightarrow 2^{(1-n/2)} \kappa_n$

Repeat: $P \Rightarrow P(0, 1, 0, 0, \ldots; x)$

... $P$ flows to normal distribution
The Central Limit Theorem as RG flow

$P(x) = \text{some probability distribution with zero mean, unit variance.}$

Characterize by cumulants:

$P(0, 1, \kappa_3, \kappa_4, \ldots; x)$

Average pairwise:

(\text{and rescale})

Cumulants get rescaled:

$\kappa_n \rightarrow 2^{(1-n/2)} \kappa_n$

Repeat:

$P \Rightarrow P(0, 1, 0, 0, \ldots; x)$

$\ldots P \text{ flows to normal distribution}$
In our case: correlator $C(\Phi)$ driven toward log-normal distribution

$\log[C(\Phi)]$ driven toward normal distribution

If cumulants $\kappa_n$ of $\log[C(\Phi)]$ behave as irrelevant operators, is there equivalent of effective field theory approach?
In our case: correlator $C(\Phi)$ driven toward log-normal distribution

$\log[C(\Phi)]$ driven toward normal distribution

If cumulants $\kappa_n$ of $\log[C(\Phi)]$ behave as irrelevant operators, is there equivalent of effective field theory approach?

**YES**, truncate exact relation:

$$\ln\langle C \rangle = \sum_{n=1}^{\infty} \frac{\kappa_n}{n!}$$

$\kappa_n$ are computed from finite sample.
Back to real data for N=46 unitary fermions (40 M configs.)

Conventional effective mass plot already shown:

\[ m_{\text{eff}}(\tau)/E_{\text{Free}} \]

**N= 46 fermions**

**L=12**

**40 M configs**
Applying the cumulant expansion for the same data
  • conventional effective mass plot (gray)
  • cumulant expansion (blue) to order $n$

$$\ln \langle C \rangle = \sum \frac{\kappa_n}{n!}$$
Applying the cumulant expansion for the same data

- conventional effective mass plot (gray)
- cumulant expansion (blue) to order $n$

\[
\ln\langle C \rangle = \sum \frac{\kappa_n}{n!}
\]
Applying the cumulant expansion for the same data

- conventional effective mass plot (gray)
- cumulant expansion (blue) to order $n$

\[
\ln \langle C \rangle = \sum \frac{\kappa_n}{n!}
\]
Applying the cumulant expansion for the same data

- conventional effective mass plot (gray)
- cumulant expansion (blue) to order n

\[ \ln \langle C \rangle = \sum \frac{\kappa_n}{n!} \]
Applying the cumulant expansion for the same data

- conventional effective mass plot (gray)
- cumulant expansion (blue) to order \( n \)

\[ \ln\langle C \rangle = \sum \frac{\kappa_n}{n!} \]
Determination of ground state energy for $N=46$ from cumulant expansion of $\log[C]$. 

![Graph showing $E_{N_k}/E_{\text{Free}}$ for different orders of the cumulant expansion.](image)

$N=46, L=12$ 

Order of cumulant expansion 

$E_{N_k}/E_{\text{Free}}$ values for different orders of the cumulant expansion.
Determination of ground state energy for N=46 from cumulant expansion of log[C]

Order of cumulant expansion

$N=46, L=12$

$E_{N_k}/E_{\text{Free}}$
Are distributions approaching log-normal appearing in QCD?
Are distributions approaching log-normal appearing in QCD?

Apparently yes, although time dependence seems to be different

NPLQCD data

distribution of log $C_{\Lambda}$

Each curve: 100,000 samples
Why do almost log-normal distributions arise? Typically, multiplicative stochastic processes.

- Fracturing of materials
- Flow of oil through porous rock
- ...

Similar physics in electron propagation in random media

Try mean field treatment for probability distribution (inspired by Smolyarenko, Altschuler, 1997)
Mean field argument: distribution for $y = \log[C_N(\Phi, T)]$

$$P(y) = \int [D\phi] e^{-\int d^4x \frac{1}{2} m^2 \phi^2} \delta(\ln C_N[\phi, T] - y)$$

N particle correlator

time separation $T$
Mean field argument: distribution for $y = \log[C_N(\Phi, T)]$

$$
P(y) = \int [D\phi] e^{-\int d^4 x \frac{1}{2} m^2 \phi^2} \delta(\ln C_N[\phi, T] - y)
$$

Perform semiclassical expansion in $\Phi$. 

N particle correlator

time separation $T$
Mean field argument: distribution for $y = \log[C_N(\Phi, T)]$

$$P(y) = \int [D\phi] e^{-\int d^4x \frac{1}{2} m^2 \phi^2} \delta(\ln C_N[\phi, T] - y)$$

Perform semiclassical expansion in $\Phi$.

Leading order result:

- Log Normal distribution (with corrections at higher order);
- $\mu$, $\sigma^2$ scale with $N$ and $T$ as seen in data
Useful to have an analytically soluble toy model: 1 particle, one spatial site
Useful to have an analytically soluble toy model: 1 particle, one spatial site

\[(1 + g\phi(x, t))\]
Useful to have an analytically soluble toy model:

1 particle, one spatial site

Toy model:

\[ C(T) = \prod_{i=1}^{T} (1 + g\phi_i) \]

\( \phi_i \in [-1, 1] \) uniform dist.
Exact answer for the “energy”:

\[ E(T) \equiv -\frac{1}{T} \ln \langle C(T) \rangle_{\phi} = 0 \]

Compare with simulation (finite sample size), g=1/2
Toy model:

\[ C(T) = \prod_{i=1}^{T} (1 + g\phi_i) \]

\( \phi_i \in [-1, 1] \) uniform dist.

Exact answer for the “energy”:

\[ E(T) \equiv -\frac{1}{T} \ln \langle C(T) \rangle_\phi = 0 \]

Compare with simulation (finite sample size), \( g=1/2 \)

Two strategies:

- Conventional:

\[ E \to -\frac{1}{T} \ln \left[ \frac{1}{N} \sum_{i=1}^{N} C(T, \phi_i) \right] \]

- New “EFT” approach: use identity

\[ \ln \langle C \rangle = \sum_n \frac{\kappa_n}{n!} \quad \kappa_n = \text{cumulants of } \ln C(T, \phi) \]

estimate \( \kappa_n \) from sample for low \( n \).
Effective mass plot for toy model

Simulation: sample size N=50,000
(each dot)

Analytic results:
exact: $E=0$

$\kappa_1, \kappa_2, \kappa_3$

$\kappa_1, \kappa_2, \kappa_3$

$\kappa_1, \kappa_2$
Intriguing observation about toy model:
improvement if reweighted by mean field solution
(Endres)

sample size = 5000

conventional weighting

reweighted: shifted by mean field solution
Intriguing observation about toy model:

Improvement if reweighted by mean field solution

(Endres)

Model too simplistic

...but is there some way to use m.f.t. to reweight a real field theory?
Conclusions:
Conclusions:

- Noise distribution in multi-fermion simulations can have nearly universal structure
Conclusions:

• Noise distribution in multi-fermion simulations can have nearly universal structure

• Amenable to analysis similar to EFT -- characterized by hierarchy of correlations...can extract information
Conclusions:

• Noise distribution in multi-fermion simulations can have nearly universal structure

• Amenable to analysis similar to EFT -- characterized by hierarchy of correlations...can extract information

• Possible clues to effective reweighting from mean field expansion of noise distribution
Conclusions:

- Noise distribution in multi-fermion simulations can have nearly universal structure
- Amenable to analysis similar to EFT -- characterized by hierarchy of correlations...can extract information
- Possible clues to effective reweighting from mean field expansion of noise distribution
- Might similar structure be hiding in other noisy systems (e.g., disconnected diagrams)?