Frequency-splitting estimators for disconnected diagrams

arXiv:1903.10447 (to appear EPJC)

Tim Harris, Leonardo Giusti, Alessandro Nada, Stefan Schaefer

Lattice 2019, Wuhan, 21.6.19

Outline

1. Analysis of single-propagator traces
   - Gauge variance
   - Stochastic variance
2. Improved estimators
   - Heavy quarks $\rightsquigarrow$ hopping expansion
   - Differences of traces $\rightsquigarrow$ split-even estimators
   - Single-propagator traces $\rightsquigarrow$ frequency-splitting estimators
3. Conclusions & Outlook
   - Application of split-even to HVP

Introduction

single-propagator traces for quark flavour r

$$t^r(x) = -\frac{1}{a^4}\,\mathrm{tr}\{\Gamma \mathbb{S}_r(x,x)\}$$

ubiquitous in lattice QCD

- ⇝ hadronic matrix elements of singlet currents
- ⇝ disconnected contribution to LO HVP
- ⇝ quark condensates
- … etc.

typically want to evaluate for many x

e.g. $\bar{t}^r(x_0) = \frac{a^3}{L^3}\sum_x t^r(x)$

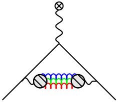


Figure: Disconnected HVP

Example: two-point function

disconnected contribution to the $\Gamma = \gamma_k$ correlator

$$C^r(x_0) = \langle \bar{t}^r(x_0)\bar{t}^r(0)\rangle$$
$$= \langle \bar{V}_k^{rr}(x_0) V_k^{r'r'}(0)\rangle$$

where $V^{rs} = \overline{\psi}^r \gamma_k \psi^s$ and $m_r = m_{r'}$

Figure: $C^r(x_0)$ correlator

Gauge variance

- fluctuations are parameterized by the variance

$$\sigma^2_{C^r}(x_0) = L^6 \langle |\bar{t}^r(x_0)|^2 |\bar{t}^r(0)|^2 \rangle - |C^r(x_0)|^2$$

$$\boxed{= L^6 \sigma^2_{\bar{t}} \cdot \sigma^2_{\bar{t}} + \ldots \quad \text{when} \quad x_0 \gg 0}$$

- the variance factorizes and is independent of x_0

$\rightsquigarrow$ determined by variance of single-propagator trace

$$\sigma^2_{\bar{t}^r} = \langle \bar{V}^{rr}(0) V^{r'r'}(0) \rangle$$

and similarly for P, S, A_k etc.

  - OPE predicts a^{-3}-divergent
  - suppressed like g_0^6 for V or g_0^4 for P

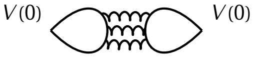


Figure: Gauge variance of t^r

Stochastic variance

Stochastic estimator

unfeasible to compute or store propagator $\mathbb{S}_r(x,x)$ for all x required for volume-averaging or momentum-projection

$\rightsquigarrow$ introduce stochastic estimator ... and stochastic variance!

introduce auxiliary fields η which satisfy

$$\langle \eta(x)\{\eta(y)\}^*\rangle = \delta_{xy}, \qquad \langle \eta(x)\rangle = 0$$

e.g. Gaussian-distributed $P[\eta] \sim e^{-\eta^\dagger \eta}$

and estimate $t^r(x)$ using the Hutchinson trace

$$\tau^r(x) = -\frac{1}{aL^3 2N_s}\sum_{i=1}^{N_s} \eta_i^\dagger(x)\Gamma\{\mathbb{S}_r\eta_i\}(x) \pm \text{c.c.}$$

introduces additional source of variance

$$\sigma^2_{\bar{\tau}^r} = \sigma^2_{\bar{t}^r} - \frac{1}{2N_s}\left\{\langle \bar{V}^{rr'}(0)V^{r'r}(0)\rangle + \sum_t \langle \bar{P}^{rr'}(t)P^{r'r}(0)\rangle\right\}$$

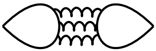
V  V

Figure: Gauge variance of t^r

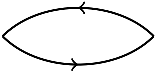
P, V  P, V

Figure: Stochastic variance of τ^r

Numerical results for single-propagator traces

- Investigate using CLS $N_\mathrm{f} = 2$ O(a)-improved Wilson fermions $m_\pi = 190, \ldots, 440$ MeV

id	L/a	N_cfg	m_π[MeV]
E5	32	100	440
F7	48	100	270
G8	64	25	190

Table: $N_\mathrm{f} = 2$ O(a)-improved Wilson fermions

$$\sigma^2_{\bar{\tau}^r} = \sigma^2_{\bar{t}^r} - \frac{1}{2N_s}\left\{\langle \bar{V}^{rs}(0)V^{sr}(0)\rangle + \sum_t \underbrace{\langle \bar{P}^{rr'}(t)P^{r'r}(0)\rangle}_{\Gamma-\text{independent}}\right\}$$

- stochastic variance dominated by $\langle PP\rangle$ in all channels
- stochastic variance $\ggg$ gauge variance
  $\rightsquigarrow$ perturbative suppression, large-N_c,...
- gauge variance depends strongly on channel

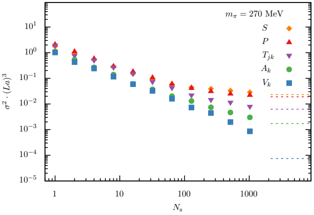


Figure: Variance vs number of sources N_s

Numerical results for single-propagator traces

- Investigate using CLS $N_f = 2$ O(a)-improved Wilson fermions $m_\pi = 190,\ldots,440$ MeV

id	L/a	N_{cfg}	m_π[MeV]
E5	32	100	440
F7	48	100	270
G8	64	25	190

Table: $N_f = 2$ O(a)-improved Wilson fermions

$$\sigma^2_{\bar{\tau}^r} = \sigma^2_{\bar{t}^r} - \frac{1}{2N_s}\left\{\langle \bar{V}^{rs}(0) V^{sr}(0)\rangle + \sum_t \underbrace{\langle \bar{P}^{rr'}(t) P^{r'r}(0)\rangle}_{\Gamma-\text{independent}}\right\}$$

- stochastic variance dominated by $\langle PP \rangle$ in all channels
- stochastic variance $\ggg$ gauge variance
  $\rightsquigarrow$ perturbative suppression, large-N_c,...
- gauge variance depends strongly on channel

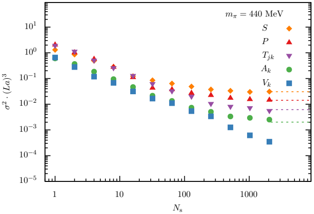


Figure: Variance vs number of sources N_s

Outline

Hopping expansion for large am_{q}

- heavy-quark regime $am_{\mathrm{q}} > 0.1$
  ⇝ free case describes stochastic variances
- use hopping representation (HPE) of propagator

$$\mathbb{S}_r = \{\mathbb{S}_r^{ee} + \mathbb{S}_r^{oo}\} \frac{1}{1-H} = \underbrace{M^n}_{\text{hopping}} + \underbrace{\mathbb{S}_r H^n}_{\text{remainder}}$$

- most of the stoch. variance comes from "hopping"

$$M^n \sim \sum_{k=0}^{n-1} H^k$$

⇝ evaluate exactly using $24(n/2)^4$ probing vectors

⇒ O(10) reduction in variance & minimal overhead

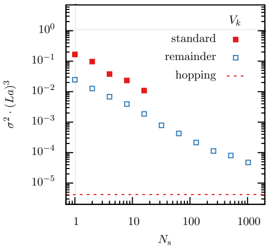


Figure: Remainder variance $am_{\mathrm{q}} = 0.3$ vs N_{s}

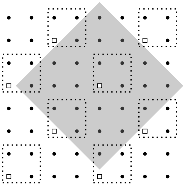


Figure: Probing vector

Stochastic estimators of differences of traces

another case of interest is the flavour differences $m_r \neq m_s$

$$t^{rs} = t^r - t^s$$

using $\mathbb{S}_r - \mathbb{S}_s = (m_s - m_r)\mathbb{S}_r\mathbb{S}_s$, two estimators

$$\theta^{rs}(x) = -\frac{(m_s - m_r)}{2a^4 N_s} \sum_{i=1}^{N_s} \eta_i^\dagger(x)\Gamma\{\mathbb{S}_r\mathbb{S}_s\eta_i\}(x) \pm \text{c.c} \qquad \text{“standard”}$$

$$\tau^{rs}(x) = -\frac{(m_s - m_r)}{2a^4 N_s} \sum_{i=1}^{N_s} \{\eta_i^\dagger\mathbb{S}_r\}(x)\,\Gamma\,\{\mathbb{S}_s\eta_i\}(x) \pm \text{c.c} \qquad \text{“split-even”}$$

inserted noise in different part of trace

Application: HVP

disconnected contribution to HVP with (uds)-quarks for $\Gamma = \gamma_k$

$$C^{rs}(x_0) = L^3\langle\bar{t}^{rs}(x_0)\bar{t}^{rs}(0)\rangle$$

for $r = \text{up}, \text{down}$ and $s = \text{strange}$

Variances of differences of traces

Expressing the variances in terms of local operators

$$\sigma^2_{\bar{\theta}^{rs}} = \sigma^2_{\bar{t}^{rs}} - \frac{a^2(m_s - m_r)^2 L^6}{2N_s} \sum_{t_1,t_3} \langle \bar{S}^{rs}(t_1)\bar{V}^{ss'}(0)\bar{S}^{s'r'}(t_2)V^{r'r}(0)\rangle + \sum_{t_2}\langle \bar{S}^{rs}(t_1)\bar{P}^{ss'}(t_2)\bar{S}^{s'r'}(t_3)P^{r'r}(0)\rangle$$

$$\sigma^2_{\bar{\tau}^{rs}} = \sigma^2_{\bar{t}^{rs}} - \frac{a^2(m_s - m_r)^2 L^6}{2N_s} \sum_{t_1,t_3} \langle \bar{S}^{rs}(t_1)\bar{V}^{ss'}(0)\bar{S}^{s'r'}(t_3)V^{r'r}(0)\rangle + \langle \bar{P}^{rs}(t_1)\bar{V}^{ss'}(0)\bar{P}^{s'r'}(t_3)V^{r'r}(0)\rangle$$

- only linearly divergent in a
- gauge/stochastic are disc./conn. diagrams
- no large $\langle SPSP\rangle$ in split-even
- one sum fewer in split-even estimator $\bar{\tau}^{rs}$
  ⇝ variance of "time-diluted" estimator for free!

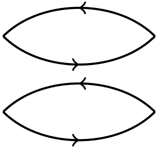


Figure: Gauge variance

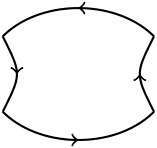
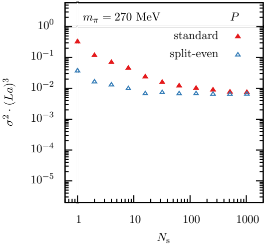


Figure: Stochastic variance

Differences of traces – standard vs split-even

- split-even stoch. variance ≪ standard stoch. variance
- for V gauge variance saturated after $N_s \gtrsim 512$

⇒ O(100) reduction in variance for V at the same cost



Why is the split-even variance reduced?

- no $\langle SPSP \rangle$ term
- similar summations to time-diluted case
- for Wilson twisted-mass fermions there is a Cauchy-Schwarz inequality $\sigma^2_{\bar{\tau}^{rs}} < \sigma^2_{\bar{\theta}^{rs}}$

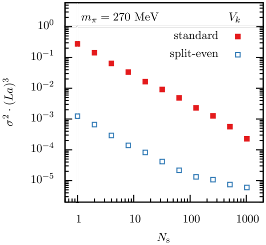


Figure: Variance of C^{rs} P and V

Differences of traces – standard vs split-even

- split-even stoch. variance $\ll$ standard stoch. variance
- for V gauge variance saturated after $N_s \gtrsim 512$

$\Rightarrow$ O(100) reduction in variance for V at the same cost

Why is the split-even variance reduced?

- no $\langle SPSP \rangle$ term
- similar summations to time-diluted case
- for Wilson twisted-mass fermions there is a Cauchy-Schwarz inequality $\sigma^2_{\bar{\tau}^{rs}} < \sigma^2_{\bar{\theta}^{rs}}$

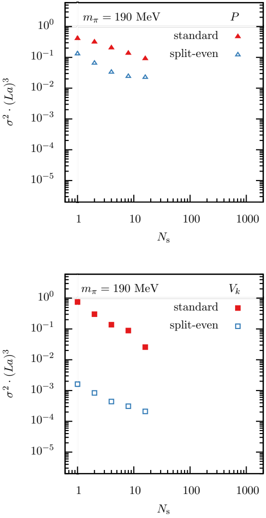


Figure: Variance of C^{rs} P and V

Frequency-splitting estimators

combining HPE and split-even estimator
$\rightsquigarrow$ frequency-splitting estimator for flavour r

$$\tau^{r}_{\mathrm{FS}}(x) = \underbrace{\tau^{rs}(x)}_{\text{light - heavy}} + \underbrace{\tau^{s}(x)}_{\text{heavy}}$$

and the generalization to N splittings

$\rightsquigarrow$ e.g. choose quark masses m_{ud}, m_{s}, m_{c}, ...

- separate contributions from IR and UV
- expect good chiral scaling
- conceptually similar to multiple time-step integrators in HMC with Hasenbusch splitting

Single-propagator traces – frequency-splitting estimators

id	am_q	N_s	cost	rel. cost
FS1	0.00207	1	34	
	0.1	4	11	
				2.5
FS2	0.00207	1	34	
	0.02	1	19	
	0.06	2	13	
	0.15	3	9	
	0.3	10	6	
				6

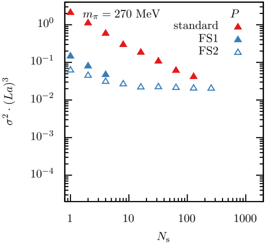
Table: FS quark mass parameters

Example: choose $m_{r_1} < m_{r_2} < \ldots$ and $N_{s,1} \lesssim N_{s,2} \lesssim \ldots$ to fulfil

$$\sigma^2_{\tau^{r_1 r_2}} \sim \sigma^2_{\tau^{r_2 r_3}} \sim \ldots \quad \text{and} \quad \mathrm{cost}(\tau^{r_1 r_2}) \sim \mathrm{cost}(\tau^{r_2 r_3}) \sim \ldots$$

- O(100) reduction in variance for V
- FS2 increased cost × 6

⇒ O(10) speed-up in vector channel



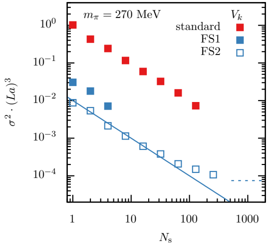


Figure: Variance of FS for P and V

Single-propagator traces – frequency-splitting estimators

id	am_q	N_s	cost	rel. cost
FS1	0.00207	1	34	
	0.1	4	11	
				2.5
FS2	0.00207	1	34	
	0.02	1	19	
	0.06	2	13	
	0.15	3	9	
	0.3	10	6	
				6

Table: FS quark mass parameters

Example: choose $m_{r_1} < m_{r_2} < \ldots$ and $N_{s,1} \lesssim N_{s,2} \lesssim \ldots$ to fulfil

$$\sigma^2_{\tau^{r_1 r_2}} \sim \sigma^2_{\tau^{r_2 r_3}} \sim \ldots \quad \text{and} \quad \text{cost}(\tau^{r_1 r_2}) \sim \text{cost}(\tau^{r_2 r_3}) \sim \ldots$$

- O(100) reduction in variance for V
- FS2 increased cost × 6

⇒ O(10) speed-up in vector channel

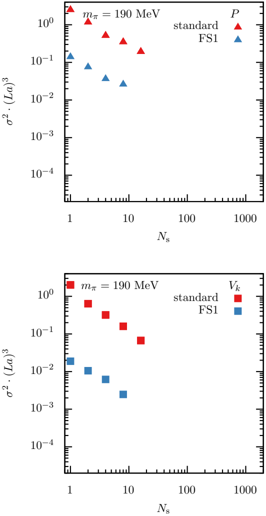


Figure: Variance of FS for P and V

Differences of traces – application to HVP

- increased $N_{cfg} = 1200$
- verified factorization for $x_0 \gg 0$
  $\rightsquigarrow$ stochastic variance factorizes immediately
- gauge variance saturated with moderate cost

$\Rightarrow$ O(100) speed-up in vector channel

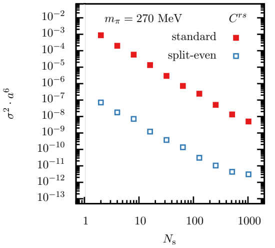


Figure: Variance of for C^{rs} vs N_s

- low scatter in data
  $\rightsquigarrow$ evidence gauge variance saturated
- good signal up to $x_0 \sim 1.4$ fm

need multi-level to tackle gauge-variance beyond 1.5 fm
$\rightsquigarrow$ work in progress (M. Dalla Brida et al.)

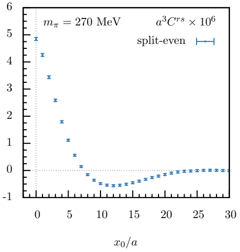


Figure: Best estimate for $C^{rs}(x_0)$

Conclusions

Further optimizations possible

- optimization of quark masses cheap procedure
  $\rightsquigarrow$ just measure the variances!
- synthesis with low-mode averaging
- employ different solvers for heavy quarks

Conclusions

- analysis of variances important to define improved estimator
- reach gauge variance for $a_\mu^{\mathrm{hvp,disc}}$ with split-even estimators
- frequency-splitting separates variance from UV and IR
- exact evaluation of hopping removes the UV variance
- method efficient toward the physical point

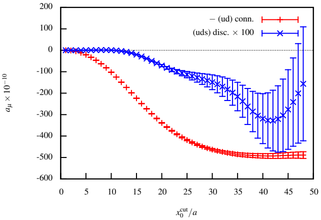


Figure: Signal-to-noise problem for a_μ^{hvp}

	P	V
$t^r(x_0)$	O(4)	O(10)
$t^{rs}(x_0)$	O(10)	O(100)

Table: Speed-up for FS/split-even

Outlook

synthesis with multi-level integration for disconnected diagrams (M. Dalla Brida et al.)

applications of split-even estimator

- use of $I = 0$ correlator for $a_\mu^{\text{hvp,disc}}$
- application to SIB, electromagnetic corrections to HVP
- hadronic matrix elements of EM current
- HLbL ...

applications of split-even estimator

- isoscalar spectroscopy
  $\rightsquigarrow$ long distance of $I = 0$ channel for HVP disc.
- disconnected single-flavour matrix elements
- ...

Back-up

Single-flavour two-point

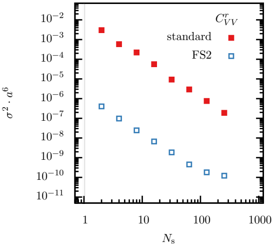


Figure: Variance of V two-point variance

Optimizing the cost

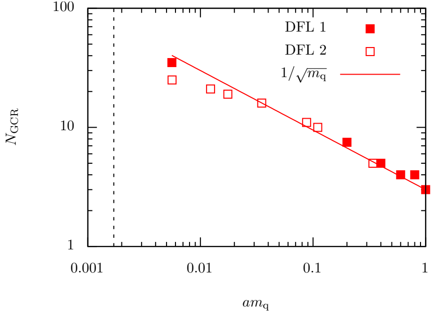


Figure: Mass-dependence of solver cost

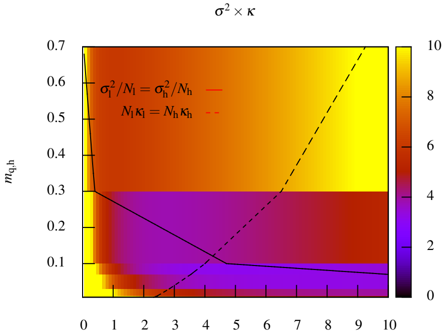


Figure: Variance for $N = 2$ in m_q-N_s plane

Hopping expansion and probing

$$D_m^{-1} = M_{2n,m} + D_m^{-1} H_m^{2n} \,,$$

$$M_{2n,m} = \frac{1}{D_{\rm ee} + D_{\rm oo}} \sum_{k=0}^{2n-1} H_m^k \,, \qquad H_m = -\left[D_{\rm eo} D_{\rm oo}^{-1} + D_{\rm oe} D_{\rm ee}^{-1} \right] ,$$

Probing vectors defined by

$$\sum_{k=0}^{K-1} v_i^k v_j^k = \delta_{ij} \qquad \text{for all} \quad i,j \quad \text{where} \quad \mathscr{M}_{ij} \neq 0,$$

then the diagonal elements of $\mathscr{M}$ are given by (no summation over i)

$$\mathscr{M}_{ii} = \sum_{k=0}^{K-1} v_i^k u_i^k, \qquad \text{where} \qquad u^k = \mathscr{M} v^k.$$

Quark-mass dependence of hopping

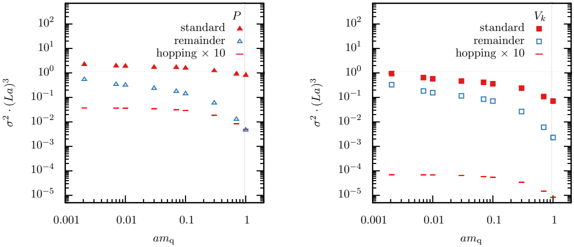


Figure: Mass-dependence of P and V variance

Variance of two-point function

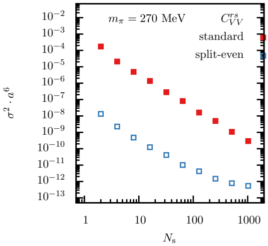


Figure: Split-even estimator for $a_\mu^{\mathrm{hvp,disc}}$