Disconnected diagrams with multi-level integration

Leonardo Giusti, *Tim Harris*, Alessandro Nada, Stefan Schaefer

UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA

DESY

Lattice 2018, 23.07.18

Outline

Motivation

Quark-line disconnected diagrams appear when we consider singlet fermion bilinears
- ⇝ flavour singlet currents, e.g. HVP
- ⇝ isosinglet channels in spectroscopy, e.g. f_0
- ⇝ hadronic matrix elements, e.g. nucleon σ-term
- ⇝ quark condensates

These are usually evaluated with a noisy estimator e.g. Hutchinson trace

$$\left\langle \dots \bar{\psi}(x)\Gamma\psi(x) \right\rangle = -\left\langle \dots \mathrm{E}_\eta \left(\eta^\dagger(x)\Gamma D^{-1}(x,y)\eta(y) \right) \right\rangle \tag{1}$$

where $\eta(x)$ are independent random white noise

The variance, σ_η^2, of the estimator

$$\mathrm{tr}\,\Gamma D^{-1} \approx \frac{1}{N}\sum_{n=1}^{N} \eta_n^\dagger(x)\Gamma D^{-1}(x,y)\eta_n(y) \tag{2}$$

is determined by off-diagonal elements

$$\sigma_\eta^2 = \frac{1}{N}\sum_{\substack{x,y \\ x\neq y}} \left|\Gamma D^{-1}(x,y)\right|^2 + \dots \tag{3}$$

[0] Hutchinson, "A stochastic estimator of the trace of the influence matrix for laplacian smoothing splines"; Bernardson, McCarty, and Thron, "Monte Carlo methods for estimating linear combinations of inverse matrix entries in lattice QCD".

Variance reduction – numerical tests

For certain currents, we need to compute a difference, e.g. EM current

$$\eta^\dagger \Gamma \left(D_{\text{light}}^{-1} - D_{\text{strange}}^{-1} \right) \eta \tag{4}$$

whose variance is suppressed due to the covariance between light and strange

Investigate using CLS $N_\mathrm{f} = 2$ O(a)-improved Wilson fermions E5 ensemble with $n_0 = 30$ configurations and study the saturation of the variance with N for

- pseudoscalar
- vector current

id	a (fm)	m_{PS} (MeV)	am_q	κ	$m_{\overline{\mathrm{MS}}}(2\,\mathrm{GeV})$ (MeV)	$\overline{N}_{\mathrm{GCR}}$
light	0.0658	450	0.0056	0.13625	32	25
strange			0.0175	0.135808	100	19

Variance reduction – mass differences

As expected variance is reduced in the difference

Furthermore…

Similar to the 'one-end' trick for TM, we can use for light-strange difference

$$\begin{aligned} D_l^{-1} - D_s^{-1} &= D_l^{-1} \left(D_s - D_l\right) D_s^{-1} \\ &= (m_s - m_l) D_l^{-1} D_s^{-1}, \end{aligned}$$

to write a new estimator for the trace

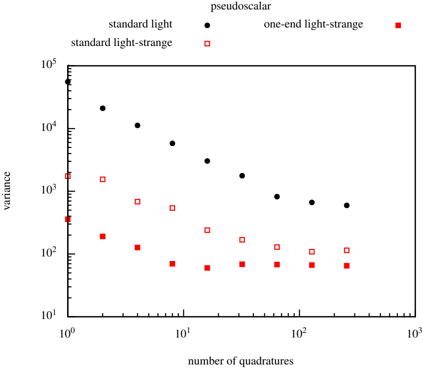
$$\mathrm{tr}\,\Gamma(D_l^{-1} - D_s^{-1}) \approx (m_s - m_l)\,\mathrm{tr}(\Gamma D_l^{-1} \eta\eta^\dagger D_s^{-1}).$$

Note that sample-wise

$$\eta^\dagger \Gamma(D_l^{-1} - D_s^{-1})\eta \neq (m_s - m_l)\,\mathrm{tr}(\Gamma D_l^{-1} \eta\eta^\dagger D_s^{-1}).$$

⇝ relevant for HVP

⇝ $m_s - m_l$ smaller than at physical point



Variance reduction – mass differences

As expected variance is reduced in the difference

Furthermore…

Similar to the 'one-end' trick for TM, we can use for light-strange difference

$$D_l^{-1} - D_s^{-1} = D_l^{-1}\left(D_s - D_l\right) D_s^{-1} \\ = (m_s - m_l) D_l^{-1} D_s^{-1},$$

to write a new estimator for the trace

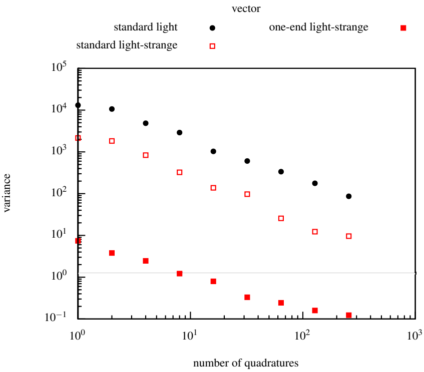
$$\mathrm{tr}\,\Gamma(D_l^{-1} - D_s^{-1}) \approx (m_s - m_l)\,\mathrm{tr}(\Gamma D_l^{-1} \eta\eta^\dagger D_s^{-1}).$$

Note that sample-wise

$$\eta^\dagger \Gamma(D_l^{-1} - D_s^{-1})\eta \neq (m_s - m_l)\,\mathrm{tr}(\Gamma D_l^{-1} \eta\eta^\dagger D_s^{-1}).$$

⇝ relevant for HVP

⇝ $m_s - m_l$ smaller than at physical point



Variance reduction – frequency splitting

Perform a frequency splitting of the estimator to separate the UV and IR

$$\mathrm{tr}\,\Gamma D_0^{-1} = \mathrm{tr}\,\Gamma(D_0^{-1} - D_1^{-1}) + \mathrm{tr}\,\Gamma(D_1^{-1} - D_2^{-1}) + \ldots + \mathrm{tr}\,\Gamma D_n^{-1} \tag{5}$$

where we compute the differences with noisy estimator for the ‘one-end’ trick

$$\mathrm{tr}\,\Gamma(D_i^{-1} - D_j^{-1}) \approx (m_j - m_i)\,\eta^\dagger D_j^{-1}\Gamma D_i^{-1}\eta \tag{6}$$

and use the hopping parameter expansion to order k for the largest mass

$$\mathrm{tr}\,\Gamma D_n^{-1} \approx \underbrace{\mathrm{tr}\,\Gamma \sum_{m=0}^{k-1} B^{-m} H^m B^{-1}}_{\text{hopping}} + \underbrace{\xi^\dagger \Gamma B^{-k} H^k D_n^{-1} \xi}_{\text{remainder}} \tag{7}$$

where $B = 4 + m + \sigma \cdot F$ contains the clover term

- use hierarchical probing or spatial dilution to compute hopping contribution exactly with finite quadratures
- remainder computed using standard noisy estimator

[0] Hasenbusch, “Exploiting the hopping parameter expansion in the hybrid Monte Carlo simulation of lattice QCD with two degenerate flavors of Wilson fermions”.

[0] Stathopoulos, Laeuchli, and Orginos, “Hierarchical probing for estimating the trace of the matrix inverse on toroidal lattices”.

Variance reduction – frequency splitting

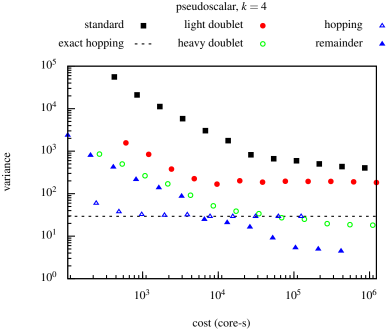
id	am_q	κ	$m_{\overline{MS}}(2\,\mathrm{GeV})$ (MeV)	$\overline{N}_{GCR}$
m_0	0.0056	0.13625	32	25
m_1	0.0876	0.133272	500	11
m_2	0.34	0.125	1914	5

Contributions depend on the intermediate masses, order of the hopping parameter expansion and channel

Can be tuned with just a few noise samples

Hierarchical probing computes the exact estimator for the hopping H^{k-1} after $k^4/8$ quadratures

Spatial dilution with even-odd blocks does the same, also for $k \neq 2^m$



Variance reduction – frequency splitting

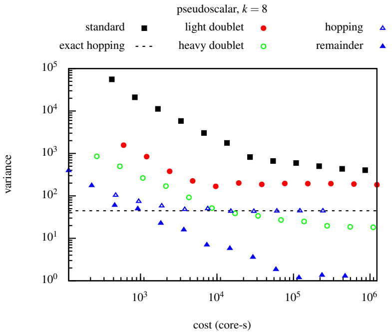
id	am_q	κ	$m_{\overline{MS}}(2\,\mathrm{GeV})$ (MeV)	$\overline{N}_{GCR}$
m_0	0.0056	0.13625	32	25
m_1	0.0876	0.133272	500	11
m_2	0.34	0.125	1914	5

Contributions depend on the intermediate masses, order of the hopping parameter expansion and channel

Can be tuned with just a few noise samples

Hierarchical probing computes the exact estimator for the hopping H^{k-1} after $k^4/8$ quadratures

Spatial dilution with even-odd blocks does the same, also for $k \neq 2^m$



Variance reduction – frequency splitting

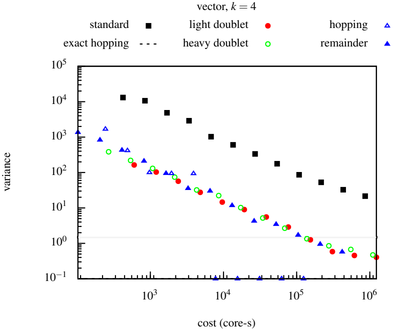
id	am_q	κ	$m_{\overline{\mathrm{MS}}}(2\,\mathrm{GeV})$ (MeV)	$\overline{N}_{\mathrm{GCR}}$
m_0	0.0056	0.13625	32	25
m_1	0.0876	0.133272	500	11
m_2	0.34	0.125	1914	5

Contributions depend on the intermediate masses, order of the hopping parameter expansion and channel

Can be tuned with just a few noise samples

Hierarchical probing computes the exact estimator for the hopping H^{k-1} after $k^4/8$ quadratures

Spatial dilution with even-odd blocks does the same, also for $k \neq 2^m$



Variance reduction – frequency splitting

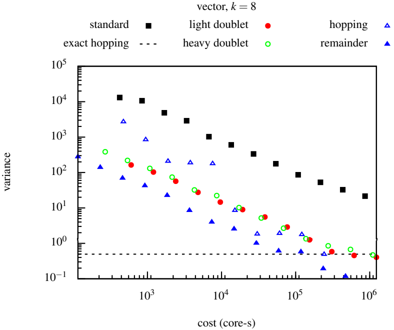
id	am_q	κ	$m_{\overline{\mathrm{MS}}}(2\,\mathrm{GeV})$ (MeV)	$\overline{N}_{\mathrm{GCR}}$
m_0	0.0056	0.13625	32	25
m_1	0.0876	0.133272	500	11
m_2	0.34	0.125	1914	5

Contributions depend on the intermediate masses, order of the hopping parameter expansion and channel

Can be tuned with just a few noise samples

Hierarchical probing computes the exact estimator for the hopping H^{k-1} after $k^4/8$ quadratures

Spatial dilution with even-odd blocks does the same, also for $k \neq 2^m$



Variance reduction – conclusions

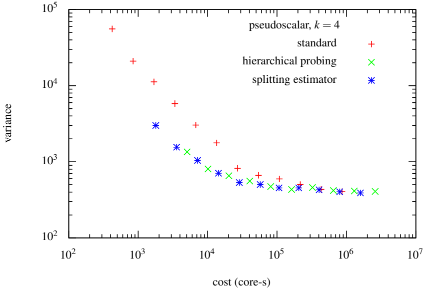
Gain of ~ 4 in the cost for the pseudoscalar, vector

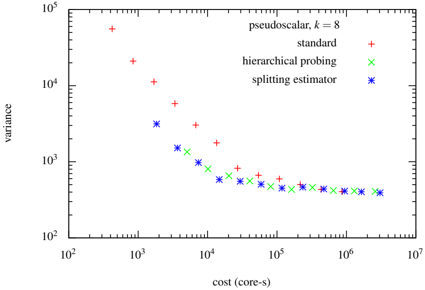
Similar to hierarchical probing for pseudoscalar, faster for vector

Parameters can be further optimized

Expect efficient estimator at smaller quark mass

Scalar and axial vector similar to pseudoscalar and vector, respectively





Variance reduction – conclusions

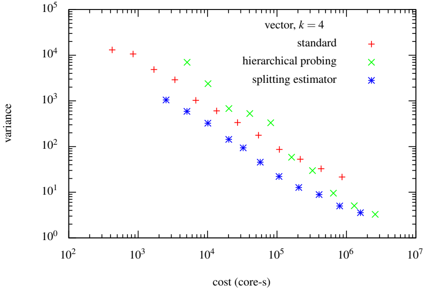
Gain of ~ 4 in the cost for the pseudoscalar, vector

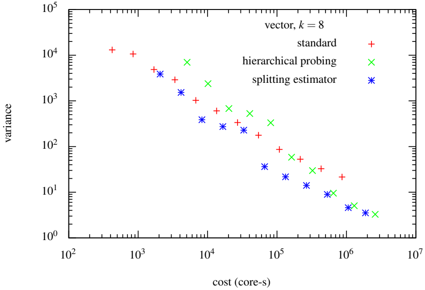
Similar to hierarchical probing for pseudoscalar, faster for vector

Parameters can be further optimized

Expect efficient estimator at smaller quark mass

Scalar and axial vector similar to pseudoscalar and vector, respectively





Recap domain decomposition

As seen in talk by A. Nada, there exists a representation

$$\langle \mathcal{O}_0(x)\mathcal{O}_1(y)\rangle_{\mathrm{QCD}} = \frac{\left\langle [\mathcal{O}_0(x)]_0 [\mathcal{O}_1(x)]_1 W_{\mathcal{N}} \right\rangle_{\mathcal{N}}}{\left\langle W_{\mathcal{N}} \right\rangle_{\mathcal{N}}}$$

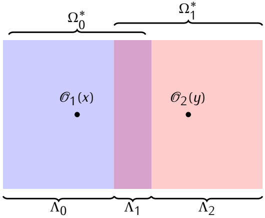
where

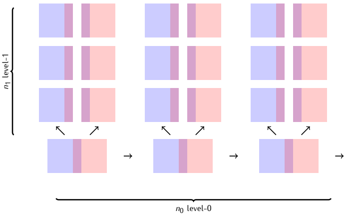
$$[\mathcal{O}_0(x)]_0 = \int_{U\in\Lambda_0} [dU_{\Lambda_0}] \mathrm{e}^{-S_{\mathrm{G}}[U_{\Lambda_0}]} \det Q_{\Omega_0^*} \mathcal{O}_0(x)$$

A factorizable multi-level estimator has a variance

$$\sigma^2 \propto \frac{1}{n_0 n_1^2}$$

To investigate multi-level scaling, our estimator's variance should be dominated by the gauge noise





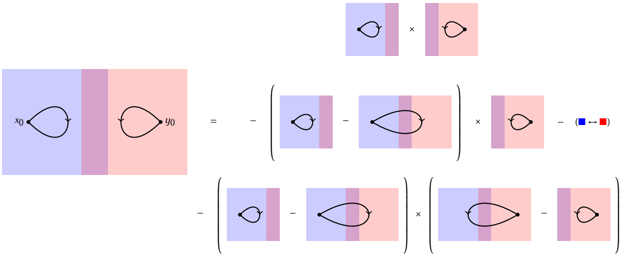
[0] Cè, Giusti, and Schaefer, "A local factorization of the fermion determinant in lattice QCD".

Disconnected two-point function with quenched multilevel

Use splitting estimator with hopping order $k = 4$ and hopping parameters

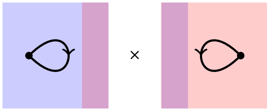
β	n_0	n_1	L/a	T/a	c_{SW}	κ	m_{PS} (MeV)	a (fm)
6.2	50	16	32	96	1.61375	0.1352 0.128 0.115	580	0.068

An unbiased estimator for the disconnected contraction is

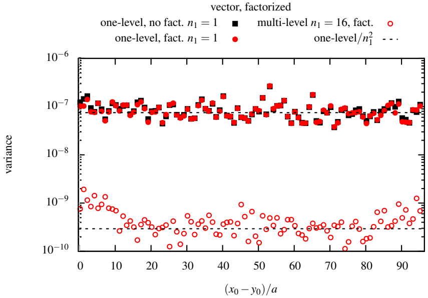


- first line is fully factorized $\rightsquigarrow$ multi-level
- remainder terms are computed on $n_0 \times n_1$ configurations

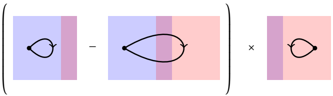
Multilevel error scaling: factorized contribution



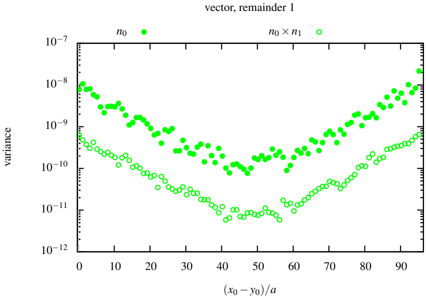
- variance scales as $1/n_0 n_1^2$
- frozen region $x_0 - y_0/a < 8$



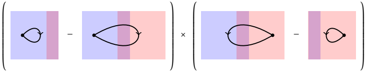
Multilevel error scaling: remainder 1

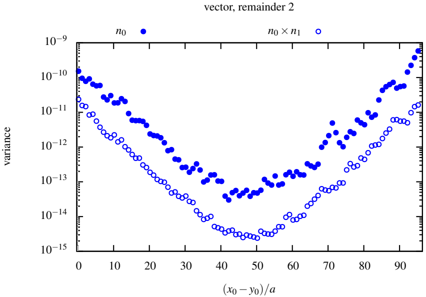


- variance scales as $1/n_0 n_1$



Error scaling: remainder 2





- variance is highly suppressed,
- scales as $1/n_0 n_1$ and

Conclusions

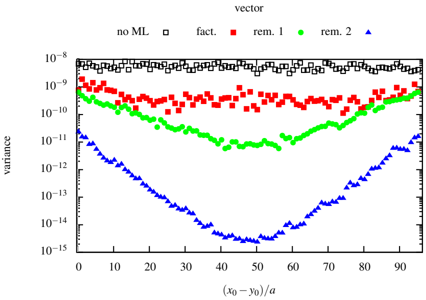
Variance reduction

- 'one-end' type trick for Wilson for light-strange
- frequency-splitting of loop to split UV and IR

Multi-level disconnected diagrams

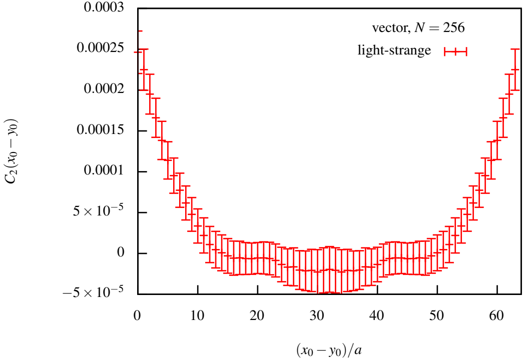
- observed expected error scaling
- need to investigate dependence of gain on distance

Expected combination of gains from both variance reduction and multi-level



Backup

Variance reduction – light-strange two-point function



Variance reduction – hopping parameter expansion

The hopping parameter expansion is based upon a polynomial approximation to D^{-1},

$$
\begin{aligned}
D_{\mathrm{sw}}^{-1} &= (\underbrace{(D_{\mathrm{oo}} + D_{\mathrm{ee}})}_{(2\kappa)^{-1}B} + \underbrace{(D_{\mathrm{eo}} + D_{\mathrm{oe}})}_{(2\kappa)^{-1}H})^{-1} \\
&= 2\kappa(1 - \kappa B^{-1}H)^{-1}B^{-1} \\
&= \sum_{n=0}^{k-1} \underbrace{\kappa^n B^{-n} H^n B^{-1}}_{\text{cheap}} + \underbrace{\kappa^k B^{-k} H^k D_{\mathrm{sw}}^{-1}}_{\text{reduced variance}}
\end{aligned}
\tag{8}
$$

[0] Dong and K. F. Liu, "Stochastic Estimation with Z_2 Noise"; Q. Liu, Wilcox, and Morgan, "Polynomial Subtraction Method for Disconnected Quark Loops".

Variance reduction – hierarchical probing

Probing for a sparse matrix can compute a trace with fewer quadratures

$$\begin{pmatrix} 1 & 0 & 7 & 9 \\ 0 & 4 & 4 & 9 \\ 7 & 9 & 1 & 0 \\ 4 & 9 & 0 & 4 \end{pmatrix} \quad \text{probing with} \quad \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \tag{9}$$

Hierachical probing for lattices chooses Hadamard vectors $\{h_i \odot \eta\}$ allows nesting

- contribution from d leading diagonals eliminated with $2d^4$ quadratures
- $\operatorname{tr} H^n$ exactly estimated with $n^4/8$ quadratures

Both hierarchical probing and hopping parameter expansion work well for large masses

[0] Stathopoulos, Laeuchli, and Orginos, "Hierarchical probing for estimating the trace of the matrix inverse on toroidal lattices".

References

M. Hutchinson. “A stochastic estimator of the trace of the influence matrix for laplacian smoothing splines”. In: *Communications in Statistics - Simulation and Computation* 19.2 (1990), pp. 433–450. eprint: https://doi.org/10.1080/03610919008812866. URL: https://doi.org/10.1080/03610919008812866.

S. Bernardson, P. McCarty, and C. Thron. “Monte Carlo methods for estimating linear combinations of inverse matrix entries in lattice QCD”. In: *Computer Physics Communications* 78.3 (1994), pp. 256–264. ISSN: 0010-4655. URL: http://www.sciencedirect.com/science/article/pii/0010465594900043.

M. Hasenbusch. “Exploiting the hopping parameter expansion in the hybrid Monte Carlo simulation of lattice QCD with two degenerate flavors of Wilson fermions”. In: *Phys. Rev.* D97.11 (2018), p. 114512. arXiv: 1805.03560 [hep-lat].

A. Stathopoulos, J. Laeuchli, and K. Orginos. “Hierarchical probing for estimating the trace of the matrix inverse on toroidal lattices”. In: (2013). arXiv: 1302.4018 [hep-lat].

M. Cè, L. Giusti, and S. Schaefer. “A local factorization of the fermion determinant in lattice QCD”. In: *Phys. Rev.* D95.3 (2017), p. 034503. arXiv: 1609.02419 [hep-lat].

S. J. Dong and K. F. Liu. “Stochastic Estimation with Z_2 Noise”. In: *Phys.Lett. B* 328 (1994), pp. 130–136. eprint: hep-lat/9308015.

Q. Liu, W. Wilcox, and R. Morgan. “Polynomial Subtraction Method for Disconnected Quark Loops”. In: (2014). arXiv: 1405.1763 [hep-lat].