Variance reduction for the vector current correlator

Numerical studies with $N_\mathrm{f} = 2$ O(a)-improved Wilson fermions[1,2]

Tim Harris

ETH zürich

09.12.2024
NGT Algorithm Workshop (CERN)

[1] Giusti, Harris, Nada, and Schaefer 2019.
[2] Dalla Brida, Giusti, Harris, and Pepe 2021.

The current correlator in the time-momentum representation

- The $j_\mu = \sum_f Q_f \bar{\psi}_f \gamma_\mu \psi_f$ correlator is composed of two Wick contractions

$$\mathcal{G}^{\rm conn}(x_0, y) = Z_{\rm V}^2 a^3 \sum_{\vec{x}} \sum_f Q_f^2 \,\mathrm{tr}\left\{ S_f(x,y)\gamma_k S_f(y,x)\gamma_k \right\},$$

$$\mathcal{G}^{\rm disc}(x_0, y) = -Z_{\rm V}^2 a^3 \sum_{\vec{x}} \sum_f Q_f \,\mathrm{tr}\left\{ S_f(x,x)\gamma_k \right\} \sum_g Q_g \,\mathrm{tr}\left\{ S_g(y,y)\gamma_k \right\},$$

defined in terms of the quark propagators $S_f(x,y)$ and $Q_{\rm u,s} = \frac{2}{3}$, $Q_{\rm d} = -\frac{1}{3}$.

[3] Bernecker and Meyer 2011.

The current correlator in the time-momentum representation

- The $j_\mu = \sum_f Q_f \bar{\psi}_f \gamma_\mu \psi_f$ correlator is composed of two Wick contractions

$$\mathcal{G}^{\text{conn}}(x_0, y) = Z_{\text{V}}^2 a^3 \sum_{\vec{x}} \sum_f Q_f^2 \,\text{tr}\left\{S_f(x,y)\gamma_k S_f(y,x)\gamma_k\right\},$$

$$\mathcal{G}^{\text{disc}}(x_0, y) = -Z_{\text{V}}^2 a^3 \sum_{\vec{x}} \sum_f Q_f \,\text{tr}\left\{S_f(x,x)\gamma_k\right\} \sum_g Q_g \,\text{tr}\left\{S_g(y,y)\gamma_k\right\},$$

defined in terms of the quark propagators $S_f(x,y)$ and $Q_{\text{u,s}} = \frac{2}{3}$, $Q_{\text{d}} = -\frac{1}{3}$.

- For $N_{\text{f}} = 2 + 1$ flavours, use the practitioners' decomposition

$$G = G_{u,d}^{\text{conn}} + G_s^{\text{conn}} + G_{u,d,s}^{\text{disc}},$$

dominated by the light connected part, proportional to the $I = 1$ contribution.

[3] Bernecker and Meyer 2011.

The current correlator in the time-momentum representation

- The $j_\mu = \sum_f Q_f \bar{\psi}_f \gamma_\mu \psi_f$ correlator is composed of two Wick contractions

$$\mathcal{G}^{\text{conn}}(x_0, y) = Z_{\text{V}}^2 a^3 \sum_{\vec{x}} \sum_f Q_f^2 \operatorname{tr}\{S_f(x,y)\gamma_k S_f(y,x)\gamma_k\},$$

$$\mathcal{G}^{\text{disc}}(x_0, y) = -Z_{\text{V}}^2 a^3 \sum_{\vec{x}} \sum_f Q_f \operatorname{tr}\{S_f(x,x)\gamma_k\} \sum_g Q_g \operatorname{tr}\{S_g(y,y)\gamma_k\},$$

defined in terms of the quark propagators $S_f(x,y)$ and $Q_{\text{u,s}} = \frac{2}{3}$, $Q_{\text{d}} = -\frac{1}{3}$.

- For $N_{\text{f}} = 2+1$ flavours, use the practitioners' decomposition

$$G = G_{u,d}^{\text{conn}} + G_s^{\text{conn}} + G_{u,d,s}^{\text{disc}},$$

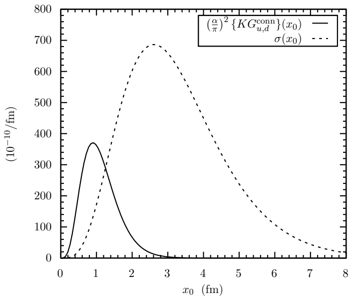
dominated by the light connected part, proportional to the $I = 1$ contribution.

- Integrals of $G(x_0)$ determine the HVP function[3].

---

[3] Bernecker and Meyer 2011.

Variance of HVP contribution to the muon anomaly



The variance on the $K(x_0) \times G^{\text{conn}}_{u,d}(x_0)$ is dominated by large separations[4].

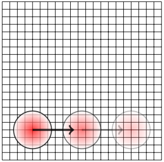
The variance on the disconnected contribution diverges.

[4] Parisi 1984; Lepage 1989.

Variance reduction by translation averaging

Averaging a localized observable over $\vec{y}$

$$\mathcal{G}_{\rm vol}(x_0, y_0) = \frac{a^3}{L^3} \sum_{\vec{y}} \mathcal{G}(x_0, y)$$



suppresses the variance

$$\sigma^2_{\rm vol}(x_0, y_0) = \frac{a^6}{L^6} \sum_{\vec{y}, y'} \Big\{ \langle \mathcal{G}(x_0, y)\mathcal{G}(x_0, y') \rangle - \langle \mathcal{G}(x_0, y) \rangle \langle \mathcal{G}(x_0, y') \rangle \Big\} \delta_{y_0, y_0'}$$

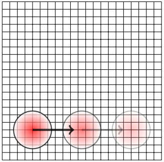
$$= \frac{a^3}{L^3} \Big[\sigma^2(x_0, y_0) + \text{const.} \Big]$$

as second term is also finite as $L \to \infty$ due to the mass gap.

[5] Meyer 2017; Lüscher 2018.

Variance reduction by translation averaging

Averaging a localized observable over $\vec{y}$



$$\mathcal{G}_{\rm vol}(x_0, y_0) = \frac{a^3}{L^3} \sum_{\vec{y}} \mathcal{G}(x_0, y)$$

suppresses the variance

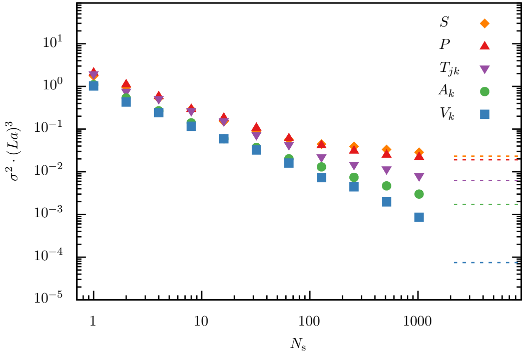
$$\sigma^2_{\rm vol}(x_0, y_0) = \frac{a^6}{L^6} \sum_{\vec{y}, y'} \Big\{ \langle \mathcal{G}(x_0, y) \mathcal{G}(x_0, y') \rangle - \langle \mathcal{G}(x_0, y) \rangle \langle \mathcal{G}(x_0, y') \rangle \Big\} \delta_{y_0, y_0'}$$
$$= \frac{a^3}{L^3} \left[\sigma^2(x_0, y_0) + \text{const.} \right]$$

as second term is also finite as $L \to \infty$ due to the mass gap.

Very costly to implement exactly for quark propagators! ⇝ see Gruber

[5] Meyer 2017; Lüscher 2018.

Approximate translation averaging for single-propagator traces I.



Approximations, e.g. using random sources[6], tend to introduce large fluctuations

$$a^3 \sum_{\vec{x}} \mathrm{tr}\left\{\gamma_k S_f(x,x)\right\} \approx \frac{1}{N_\mathrm{s}} \sum_{i=1}^{N_\mathrm{s}} a^3 \sum_{\vec{x},y} \eta^\dagger(x)\gamma_k S(x,y)\eta_i(y)\delta_{x_0,y_0}$$

---

[6] Hutchinson 1990.

Approximate translation averaging for single propagator traces II.

An efficient estimator for $S_f(x,x)$ can be constructed via a decomposition

$$S_{f_1} = (S_{f_1} - S_{f_2}) + (S_{f_2} - S_{f_3}) + \ldots + S_{f_N}, \qquad m_1 < m_2 < \ldots < m_N.$$

using the identity

$$(S_f - S_g) = \{m_g - m_f\} S_f S_g$$

and the hopping representation for the largest mass

$$S_{f_N} = (D_{\mathrm{ee}} + D_{\mathrm{oo}})^{-1} \sum_{k=0}^{m-1} H^k + S_{f_N} H^m, \qquad H = -(D_{\mathrm{eo}} D_{\mathrm{oo}}^{-1} + D_{\mathrm{oe}} D_{\mathrm{ee}}^{-1}).$$

[7] Boucaud et al. 2008; Giusti, Harris, Nada, and Schaefer 2019.

Approximate translation averaging for single propagator traces II.

An efficient estimator for $S_f(x,x)$ can be constructed via a decomposition

$$S_{f_1} = (S_{f_1} - S_{f_2}) + (S_{f_2} - S_{f_3}) + \ldots + S_{f_N}, \qquad m_1 < m_2 < \ldots < m_N.$$

using the identity

$$(S_f - S_g) = \{m_g - m_f\} S_f S_g$$

and the hopping representation for the largest mass

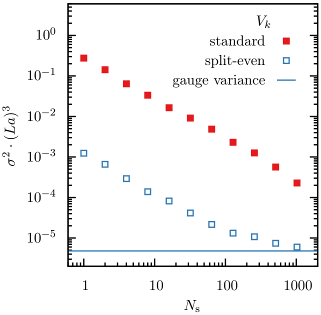
$$S_{f_N} = (D_{\rm ee} + D_{\rm oo})^{-1} \sum_{k=0}^{m-1} H^k + S_{f_N} H^m, \qquad H = -(D_{\rm eo} D_{\rm oo}^{-1} + D_{\rm oe} D_{\rm ee}^{-1}).$$

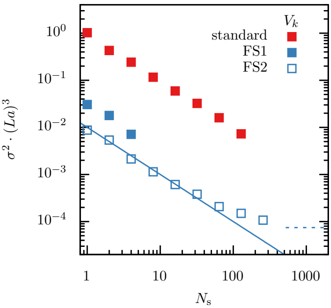
The trace of the hopping term can be computed exactly using probing vectors

$$\sum_{i=1}^{K} v^i(x) v^i(y) = \delta(x,y) \qquad \text{where} \qquad H^k(x,y) \neq 0$$

[7] Boucaud et al. 2008; Giusti, Harris, Nada, and Schaefer 2019.

Approximate translation averaging for single propagator traces III.

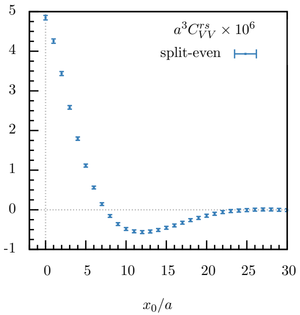




For $f = \mathrm{s} - \mathrm{ud}$ (left) and $f = \mathrm{ud}$ (right) see O(100) variance reduction

$\rightsquigarrow$ at equal cost for $f = \mathrm{s} - \mathrm{ud}$ and about $6\times$ for $f = \mathrm{ud}$

Approximate translation averaging for single propagator traces IV.



Results in a good signal for the correlator $G^{\text{disc}}_{u,d,s}$

...but with a constant error

Two-level integration (pure gauge)

If density and observable can be factorized

$$\langle O(x_1)O(x_2)\rangle \approx \frac{1}{n_0}\sum_{i=1}^{n_0}\big[O(x_1)\big]_i\big[O(x_2)\big]_i$$

in terms of the sub-averages

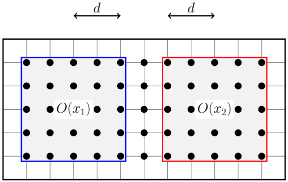
$$\big[O(x)\big]_i = \frac{1}{n_1}\sum_{k=1}^{n_1} O^{ik}(x)$$

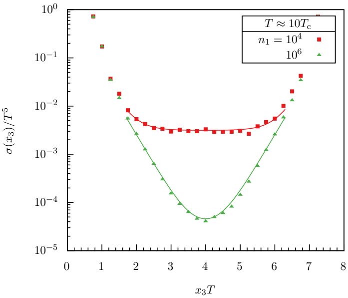
the variance should behave as

$$\sigma^2_{2\text{lvl}} = \frac{1}{n_0}\left[c_2\frac{1}{n_1^2} + c_1\frac{1}{n_1}\mathrm{e}^{-2m_O d} + c_0\mathrm{e}^{-4m_O d}\right]$$

in contrast to

$$\sigma^2_{1\text{lvl}} = \frac{1}{n_0}c_0'.$$





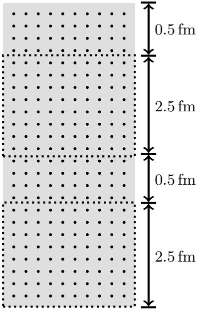
[8] Lüscher and Weisz 2001; Meyer 2003; García Vera and Schaefer 2016.

Two-level integration (QCD)

Use the decomposition of Cè, Giusti & Schaefer

Perform $n_0 = 25$ cycles with $n_1 = 10$ sub-updates

$\rightsquigarrow n_0 n_1 = 250$ gauge field updates



[9] Cè, Giusti, and Schaefer 2016; Cè, Giusti, and Schaefer 2017.
[10] Dalla Brida, Giusti, Harris, and Pepe 2021.

Two-level integration (QCD)

Use the decomposition of Cè, Giusti & Schaefer

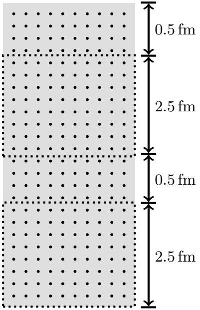
Perform $n_0 = 25$ cycles with $n_1 = 10$ sub-updates

$\rightsquigarrow n_0 n_1 = 250$ gauge field updates

Study the variance on $G(x_0)$ without factorizing it

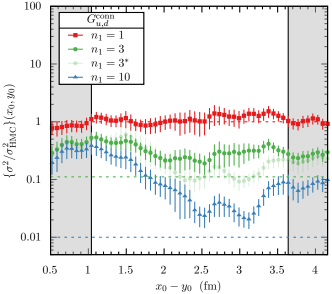
$$G(x_0) \approx \frac{1}{n_0 n_1^2} \sum_{i=1}^{n_0} \sum_{k,l=1}^{n_1} G^{ikl}(x_0)$$

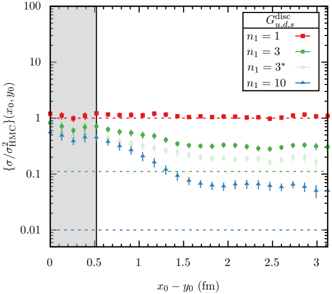
$\rightsquigarrow$ perform $n_0 n_1^2 = 2500$ measurements



[9] Cè, Giusti, and Schaefer 2016; Cè, Giusti, and Schaefer 2017.
[10] Dalla Brida, Giusti, Harris, and Pepe 2021.

Variance reduction on the current correlator

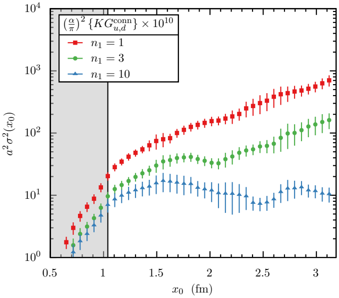


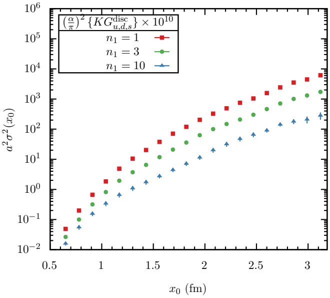


For the connected correlator, fix y_0 such that $t_{\mathrm{bnd}} - y_0 = 0.5\,\mathrm{fm}$

For the disconnected, fix x_0 and y_0 symmetrically

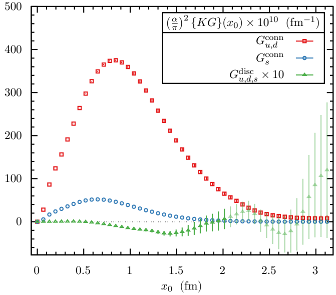
Variance reduction on the integrand

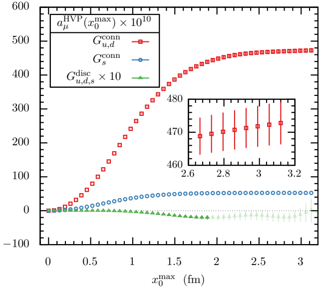




The variance on $G^{\text{conn}}_{u,d}(x_0)$ stops increasing after $x_0 \approx 1.5\,\text{fm}$

Final estimate for integrand





The error grows slowly on the partial sum of $K(x_0) \times G^{\rm conn}_{u,d}(x_0)$

$x_0^{\rm max}$/fm	$n_1 = 1$	10
2.5	446(26)	467.0(8.4)
3.0	424(38)	473.4(8.6)

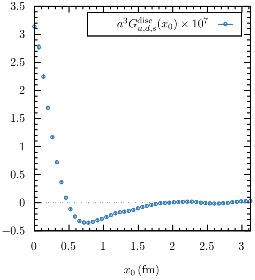
Summary

- Translation averaging

⇝ typically expensive for propagators

⇝ good estimators exist for $S(x, x)$

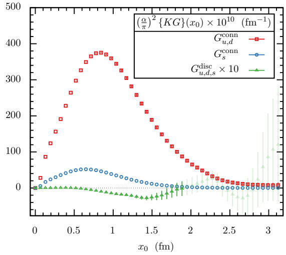
⇝ for large $|x - y|$ need low modes (see Gruber)



- Updates using overlapping domain decomposition

⇝ variance reduction works out in QCD

⇝ did not account for factorization of observable...



References I

L. Giusti, T. Harris, A. Nada, and S. Schaefer. “Frequency-splitting estimators of single-propagator traces”. In: Eur. Phys. J. C 79.7 (2019), p. 586.

M. Dalla Brida, L. Giusti, T. Harris, and M. Pepe. “Multi-level Monte Carlo computation of the hadronic vacuum polarization contribution to $(g_\mu - 2)$”. In: Phys. Lett. B 816 (2021), p. 136191.

D. Bernecker and H. B. Meyer. “Vector Correlators in Lattice QCD: Methods and applications”. In: Eur. Phys. J. A 47 (2011), p. 148.

G. Parisi. “The Strategy for Computing the Hadronic Mass Spectrum”. In: Phys. Rept. 103 (1984), pp. 203–211.

G. P. Lepage. “The Analysis of Algorithms for Lattice Field Theory”. In: Boulder ASI 1989:97-120. 1989, pp. 97–120.

H. B. Meyer. “Lorentz-covariant coordinate-space representation of the leading hadronic contribution to the anomalous magnetic moment of the muon”. In: Eur. Phys. J. C 77.9 (2017), p. 616.

References II

- M. Lüscher. “Stochastic locality and master-field simulations of very large lattices”. In: EPJ Web Conf. 175 (2018). Ed. by M. Della Morte, P. Fritzsch, E. Gámiz Sánchez, and C. Pena Ruano, p. 01002.
- 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.
- P. Boucaud et al. “Dynamical Twisted Mass Fermions with Light Quarks: Simulation and Analysis Details”. In: Comput. Phys. Commun. 179 (2008), pp. 695–715.
- M. Lüscher and P. Weisz. “Locality and exponential error reduction in numerical lattice gauge theory”. In: JHEP 09 (2001), p. 010.
- H. B. Meyer. “Locality and statistical error reduction on correlation functions”. In: JHEP 01 (2003), p. 048.
- M. García Vera and S. Schaefer. “Multilevel algorithm for flow observables in gauge theories”. In: Phys. Rev. D 93 (2016), p. 074502.

References III

M. Cè, L. Giusti, and S. Schaefer. "Domain decomposition, multi-level integration and exponential noise reduction in lattice QCD". In: Phys. Rev. D93.9 (2016), p. 094507.

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.

L. Giusti, T. Harris, A. Nada, and S. Schaefer. "Multi-level integration for meson propagators". In: PoS LATTICE2018 (2018), p. 028.