Multi-level Monte Carlo integration: HVP

Leonardo Giusti

University of Milano Bicocca & INFN

Based on:

Dalla Brida, LG, Harris, Pepe, PLB 816 (2021) 136191 [arXiv:2007.02973]

LG, Harris, Nada, Schaefer, EPJC 79 (2019) 586 [arXiv:1903.10447]

Muon $g-2$ theory initiative workshop - Virtual Meeting - June 30th 2021

The bottleneck: signal/noise ratio for HVP (HLbL,...)

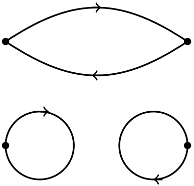
- The HVP contribution to $a_\mu = (g-2)_\mu/2$ reads

$$a_\mu^{\rm HVP} = \left(\frac{\alpha}{\pi}\right)^2 \int_0^\infty dx_0 K(x_0, m_\mu) G(x_0)$$

where

$$G(x_0) = \int d^3x \langle J_k^{\rm em}(x) J_k^{\rm em}(0)\rangle$$

with $K(x_0, m_\mu)$ being a known function



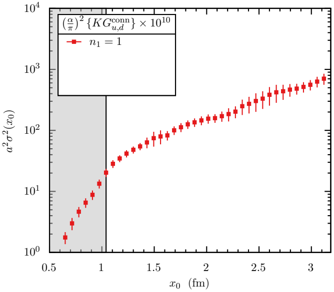
- For the light-connected contribution (by far the largest)

$$\frac{\sigma^2_{G^{\rm conn}_{u,d}}(x_0)}{[G^{\rm conn}_{u,d}(x_0)]^2} \propto \frac{1}{n_{cnfg}}\, e^{2(M_\rho - M_\pi)|x_0|}$$

where M_ρ is the lightest state in that channel.

Signal lost after 1.5-2.5 fm (depending on $m_{u,d}$) due to exp. increase of statistical error

$n_{cnfg} = n_0 = 25\,,\;\; n_{tot} = n_0 \cdot n_1$



$a = 0.065$ fm, $M_\pi = 270$ MeV

$(V/a^4) = 96 \times 48^3$

The bottleneck: signal/noise ratio for HVP (HLbL,...)

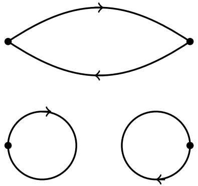
- The HVP contribution to $a_\mu = (g-2)_\mu/2$ reads

$$a_\mu^{\rm HVP} = \left(\frac{\alpha}{\pi}\right)^2 \int_0^\infty dx_0 K(x_0, m_\mu) G(x_0)$$

where

$$G(x_0) = \int d^3x \langle J_k^{\rm em}(x) J_k^{\rm em}(0)\rangle$$

with $K(x_0, m_\mu)$ being a known function



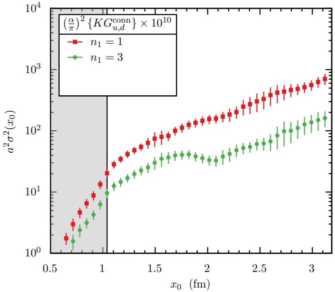
- For the light-connected contribution (by far the largest)

$$\frac{\sigma^2_{G^{\rm conn}_{u,d}}(x_0)}{[G^{\rm conn}_{u,d}(x_0)]^2} \propto \frac{1}{n_{cnfg}}\, e^{2(M_\rho - M_\pi)|x_0|}$$

where M_ρ is the lightest state in that channel.

Signal lost after 1.5-2.5 fm (depending on $m_{u,d}$) due to exp. increase of statistical error

$$n_{cnfg} = n_0 = 25\,,\;\; n_{tot} = n_0 \cdot n_1$$



$a = 0.065$ fm , $M_\pi = 270$ MeV

$(V/a^4) = 96 \times 48^3$

The bottleneck: signal/noise ratio for HVP (HLbL,...)

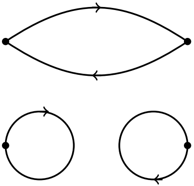
- The HVP contribution to $a_\mu = (g-2)_\mu/2$ reads

$$a_\mu^{\rm HVP} = \left(\frac{\alpha}{\pi}\right)^2 \int_0^\infty dx_0 K(x_0, m_\mu) G(x_0)$$

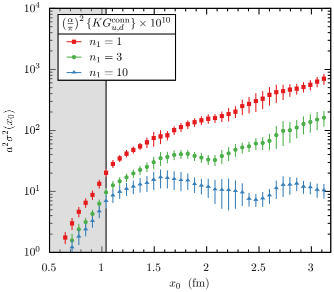
where

$$G(x_0) = \int d^3x \langle J_k^{\rm em}(x) J_k^{\rm em}(0)\rangle$$

with $K(x_0, m_\mu)$ being a known function



$n_{cnfg} = n_0 = 25\,,\;\; n_{tot} = n_0 \cdot n_1$



$a = 0.065$ fm , $M_\pi = 270$ MeV

$(V/a^4) = 96 \times 48^3$

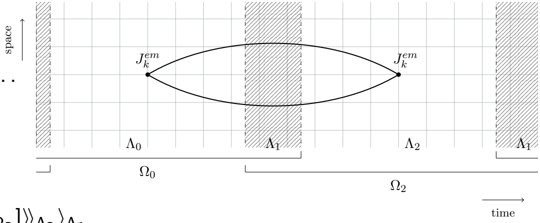
- Sharp rise of σ^2 with x_0 when computed by a standard 1-level integration (red points) is automatically flattened out by the 2-level integration (blue-points)

- Accurate computations can be obtained at large distances: no need for any modeling of the long-distance behaviour of $G_{u,d}^{\rm conn}$

Why/How does it work ?

- If the action and the observable can be factorized

$$
\begin{aligned}
S[U] &= S_0[U_{\Omega_0}] + S_2[U_{\Omega_2}] + \dots \\
O[U] &= O_0[U_{\Omega_0}] \times O_2[U_{\Omega_2}]
\end{aligned}
$$



then

$$\langle O[U] \rangle = \langle \langle\langle O_0[U_{\Omega_0}]\rangle\rangle_{\Lambda_0} \times \langle\langle O_2[U_{\Omega_2}]\rangle\rangle_{\Lambda_2} \rangle_{\Lambda_1}$$

where

$$\langle\langle O_0[U_{\Omega_0}]\rangle\rangle_{\Lambda_0} = \frac{1}{Z_{\Lambda_0}} \int \mathcal{D}U_{\Lambda_0}\, e^{-S_0[U_{\Omega_0}]}\, O_0[U_{\Omega_0}]$$

- Two-level integration:

  - n_0 configurations U_{Λ_1}
  - n_1 configurations U_{Λ_0} and U_{Λ_2} for each U_{Λ_1}

- If $\langle\langle\cdot\rangle\rangle_{\Lambda_i}$ can be computed efficiently with a statistical error comparable to its central value, then the prefactor in the signal/noise ratio changes as (until S/N problem solved)

$$n_0 \to n_0 n_1^2$$

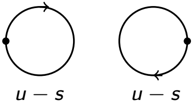
at the cost of generating approximatively $n_0 n_1$ level-0 configurations

Split-even estimator of disconnected contribution

- Advantage of multi-level sets in when variances are due to fluctuations of gauge field

- The disconnected Wick contraction reads

$$t(x) = \mathrm{Tr}\left[\gamma_k\{D^{-1}_{m_u}(x,x) - D^{-1}_{m_s}(x,x)\}\right]$$
$$= (m_s - m_u)\,\mathrm{Tr}\left[\gamma_k D^{-1}_{m_u} D^{-1}_{m_s}(x,x)\right]$$

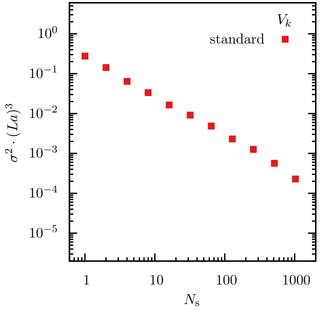


- Standard stochastic estimator $[\langle\eta(x)\eta^\dagger(y)\rangle = \delta_{xy}]$

$$\theta(x) = \frac{(m_s - m_u)}{N_s}\sum_{i=1}^{N_s} \mathrm{Im}\left[\eta_i^\dagger(x)\gamma_k\{D^{-1}_{m_u} D^{-1}_{m_s}\eta_i\}(x)\right]$$

is expensive. It requires $O(10^4)$ random fields η for its σ^2 to be dominated by gauge fluctuations

Why random noise much larger than gauge one?
Computable and understandable in QFT



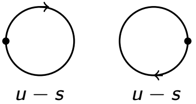
Split-even estimator of disconnected contribution

- Advantage of multi-level sets in when variances are due to fluctuations of gauge field

- The disconnected Wick contraction reads

$$t(x) = \mathrm{Tr}\left[\gamma_k\{D_{m_u}^{-1}(x,x) - D_{m_s}^{-1}(x,x)\}\right]$$

$$= (m_s - m_u)\,\mathrm{Tr}\left[\gamma_k D_{m_u}^{-1} D_{m_s}^{-1}(x,x)\right]$$

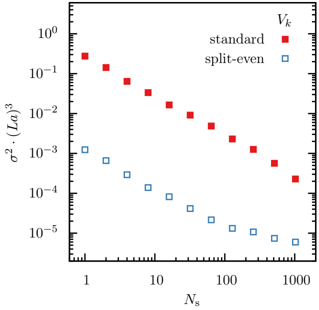


- Split-even stochastic estimator $[\langle\eta(x)\eta^\dagger(y)\rangle = \delta_{xy}]$

$$\tau(x) = \frac{(m_s - m_u)}{N_s}\sum_{i=1}^{N_s} \mathrm{Im}\left[\{\eta_i^\dagger D_{m_u}^{-1}\}(x)\,\gamma_k\,\{D_{m_s}^{-1}\eta_i\}(x)\right]$$

requires $O(10^2)$ random fields η to hit gauge noise. Gain: 2 orders of magnitude. Definition suggested by the QFT analysis of the variance.

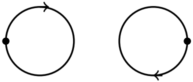
Used in the past for pseudoscalar density in TMQCD (one-end trick)



Split-even estimator of disconnected contribution

- Advantage of multi-level sets in when variances are due to fluctuations of gauge field
- The disconnected Wick contraction reads

$$t(x) = \mathrm{Tr}\left[\gamma_k\{D^{-1}_{m_u}(x,x) - D^{-1}_{m_s}(x,x)\}\right]$$
$$= (m_s - m_u)\,\mathrm{Tr}\left[\gamma_k D^{-1}_{m_u} D^{-1}_{m_s}(x,x)\right]$$

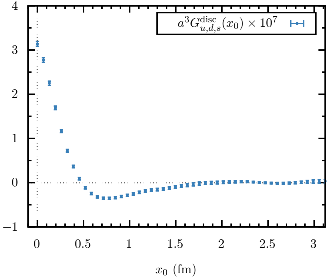


- Split-even stochastic estimator [$\langle\eta(x)\eta^\dagger(y)\rangle = \delta_{xy}$]

$$\tau(x) = \frac{(m_s - m_u)}{N_s}\sum_{i=1}^{N_s} \mathrm{Im}\left[\{\eta_i^\dagger D^{-1}_{m_u}\}(x)\,\gamma_k\,\{D^{-1}_{m_s}\eta_i\}(x)\right]$$

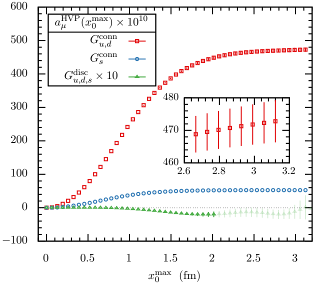
combined with multi-level integration is a solution for a precise computation of the disconnected contribution

It is already being applied in production phase for HVP by CLS (Mainz)



Conclusions & Outlook

- Permille precision and accuracy on HVP is the challenge for lattice QCD



- Our strategy: new integration and estimators (better "machine" and "experiment")

- Multi-level integration reduces the variance exponentially:
  - with the time-distance of the currents
  - when pion mass gets lighter (physical point)

- Next step: R&D $\Longrightarrow$ production. Significant human and numerical resources needed

- Analogous variance-reduction pattern expected to work out also for lattice calibration, electromagnetic corrections, HLbL, ...