Multi-level integration in the presence of fermions: the case of the hadronic contribution to $(g-2)_\mu$

Leonardo Giusti

University of Milano Bicocca & INFN

Based on:

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

DWQ@25 - Virtual Workshop - December 14th 2021

Outline

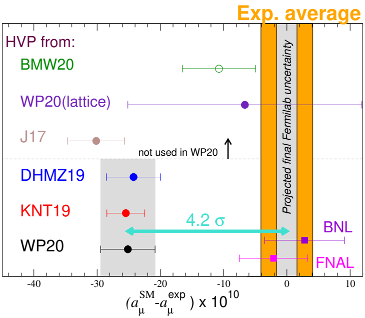
- Introduction and motivations
- Signal/noise problem in Hadronic Vacuum Polarization (HVP)
- Multi-level integration in the presence of fermions
- Split-even estimator for disconnected contributions
- First multi-level computation of HVP
- Conclusions & outlook

Introduction and motivations

[Aoyama et al. 20, Abi et al. 21]

Contribution	Section	Equation	Value $\times 10^{11}$	References
Experiment (E821)		Eq. (8.13)	116 592 089(63)	Ref. [1]
HVP LO (e^+e^-)	Sec. 2.3.7	Eq. (2.33)	6931(40)	Refs. [2–7]
HVP NLO (e^+e^-)	Sec. 2.3.8	Eq. (2.34)	−98.3(7)	Ref. [7]
HVP NNLO (e^+e^-)	Sec. 2.3.8	Eq. (2.35)	12.4(1)	Ref. [8]
HVP LO (lattice, $udsc$)	Sec. 3.5.1	Eq. (3.49)	7116(184)	Refs. [9–17]
HLbL (phenomenology)	Sec. 4.9.4	Eq. (4.92)	92(19)	Refs. [18–30]
HLbL NLO (phenomenology)	Sec. 4.8	Eq. (4.91)	2(1)	Ref. [31]
HLbL (lattice, uds)	Sec. 5.7	Eq. (5.49)	79(35)	Ref. [32]
HLbL (phenomenology + lattice)	Sec. 8	Eq. (8.10)	90(17)	Refs. [18–30, 32]
QED	Sec. 6.5	Eq. (6.30)	116 584 718.931(104)	Refs. [33, 34]
Electroweak	Sec. 7.4	Eq. (7.16)	153.6(1.0)	Refs. [35, 36]
HVP (e^+e^-, LO + NLO + NNLO)	Sec. 8	Eq. (8.5)	6845(40)	Refs. [2–8]
HLbL (phenomenology + lattice + NLO)	Sec. 8	Eq. (8.11)	92(18)	Refs. [18–32]
Total SM Value	Sec. 8	Eq. (8.12)	116 591 810(43)	Refs. [2–8, 18–24, 31–36]
Difference: $\Delta a_\mu := a_\mu^{\text{exp}} - a_\mu^{\text{SM}}$	Sec. 8	Eq. (8.14)	279(76)	

- E989 expected to reduce exp. uncertainty by 4
- To match E989 final uncertainty, 2‰ precision required for the Hadronic SM contribution
- Lattice has to improve by a factor 4–15
- *We need to develop new strategies to make lattice computations for $(g-2)_\mu$ significantly cheaper and therefore more precise and reliable*
- In particular: reduce stat. errors at large distances, access smaller lattice spacings and larger volumes, . . .



Plot courtesy of G. Colangelo

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

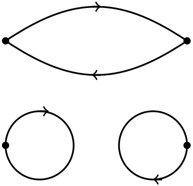
- The HVP contribution to $a_\mu = (g-2)_\mu/2$ reads

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

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

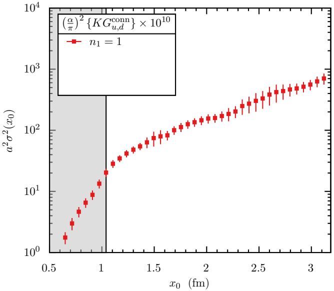


- For the light-connected contribution (by far the largest)

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

where M_ρ is the lightest state in that channel.

- For disconnected contribution is worse since the variance of the correlator is constant in time

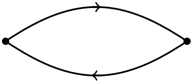


$a = 0.065$ fm , $M_\pi = 270$ MeV
$(V/a^4) = 96 \times 48^3$

Signal/noise ratio: the rôle of pions

- By defining $Q = \gamma_5 D$ and

$$W_\pi(y_0, x) = \sum_{\vec{y}} \text{Tr}\left\{Q^{-1}(y,x)[Q^{-1}(y,x)]^\dagger\right\}$$



at large time distances the pion propagator and its variance go as

$$C_\pi(y_0, x_0) = \langle W_\pi(y_0, x)\rangle \propto e^{-M_\pi |y_0 - x_0|} \qquad \sigma^2_\pi(y_0, x_0) \propto e^{-2M_\pi |y_0 - x_0|}$$

and therefore the signal/noise ratio is (almost) constant

- Indeed, when $|y - x| \to \infty$, numerical simulations confirm that

$$\text{Tr}\left\{Q^{-1}(y,x)[Q^{-1}(y,x)]^\dagger\right\} \propto e^{-M_\pi |y-x|}$$

for every background field in the representative ensemble. The size of each quark line, $\exp\{-M_\pi |y - x|/2\}$, is responsible for large fluctuations in other connected correlators

- The suppression of the propagator with the distance between source and sink, however, is also the clue for the solution ......

Signal/noise ratio: very generic problem

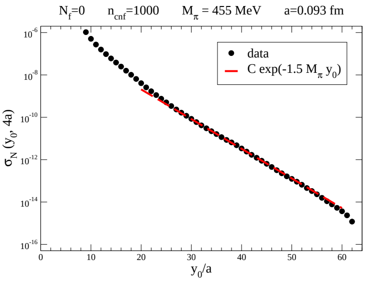
- Nucleon propagator

$$C_N(y_0, x_0) = \langle W_N(y_0, x_0)\rangle \propto e^{-M_N|y_0 - x_0|}$$

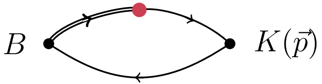
when $|y_0 - x_0| \to \infty$ goes as **[Parisi 84; Lepage 89]**

$$\sigma_N^2(y_0, x_0) \propto e^{-3M_\pi |y_0 - x_0|}$$

and analogously for other baryonic correlation functions



- Semileptonic B decays. Two (noisy) basic building blocks:
    - Mesons with (large) non-zero momentum
    - Static quark line

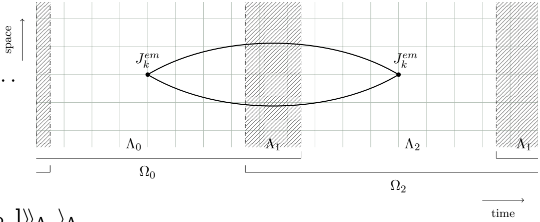


Multi-level integration

[Parisi, Petronzio, Rapuano 83; Lüscher, Weisz 01; . . . ; Meyer 02; LG, Della Morte 08 10, . . .]

- If the action and the observable can be factorized

$$S[U] = S_0[U_{\Omega_0}] + S_2[U_{\Omega_2}] + \ldots$$
$$O[U] = O_0[U_{\Omega_0}] \times O_2[U_{\Omega_2}]$$



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)

$$n_0 \to n_0 n_1^2$$

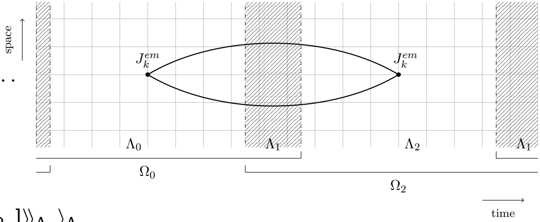
  at the cost of generating approximatively $n_0 n_1$ level-0 configurations

Multi-level integration

[Parisi, Petronzio, Rapuano 83; Lüscher, Weisz 01; ...; Meyer 02; LG, Della Morte 08 10, ...]

- If the action and the observable can be factorized

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



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

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

- With more active blocks, at the cost of approximatively $n_0 n_1$ level-0 configurations,

$$n_0 \to n_0 n_1^{n_{\rm block}}$$

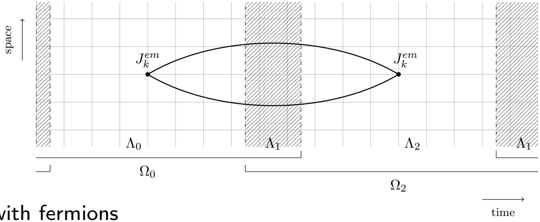
and the gain increases exponentially with the distance since $n_{\rm block} \propto |y_0 - x_0|$. For the same relative accuracy of the correlator, the computational effort would then increase approximatively linearly with the distance

Multi-level integration with fermions

[Cè, LG, Schaefer 16; Dalla Brida, LG, Harris, Pepe 20]

- Thanks to

  * Overlapping Domain Decomp.
  * Multi-Boson representation



multi-level integration also possible with fermions

- The effective action (determinant of the Dirac operator) can be decomposed as

$$\det\, D = \frac{\det(1-\omega)}{\det\, D_{\Lambda_1}\, \det D_{\Omega_0}^{-1}\, \det D_{\Omega_2}^{-1}}$$

and for 2 flavours, for instance, can be represented as

$$\{\det D^\dagger D\}^2 = \int \mathcal{D}\phi \ldots \exp\{-S_0[U_{\Omega_0},\ldots] - S_1[U_{\Lambda_1},\ldots] - S_2[U_{\Omega_2},\ldots]\}$$

- Factorization thanks to different representations of various quark-path contributions:

  * Pseudo-fermions for paths with no loops around Λ_1
  * Multi-Bosons for paths with 1–N loops (N is the number of Multi-Bosons)
  * Reweighting factor for paths with more than N loops

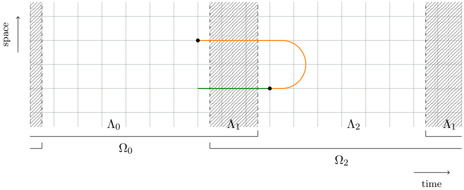
Multi-boson block factorization

- The matrix ω is

$$\omega = P_{\partial\Lambda_0} D_{\Omega_0}^{-1} D_{\Lambda_{1,2}} D_{\Omega_2}^{-1} D_{\Lambda_{1,0}}$$

  which is also:

  - similar to $\omega^\dagger$



- We can expand again $(1-\omega)^{-1}$ in series
  [Lüscher 93; Borici, de Forcrand 95; Jegerlehner 95]

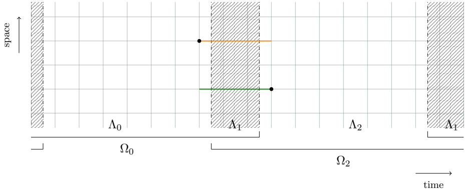
$$\frac{1}{\det[(1-\omega)^{-1}]} = \frac{1}{\det\left[\sum_{n=0}^{\infty}\omega^n\right]} \propto \prod_{k=1}^{N/2} \frac{1}{\det\{(u_k-\omega)^\dagger(u_k-\omega)\}} + \dots$$

where $u_k = e^{i\frac{2\pi k}{N+1}}$ are the roots of $P_N(\omega) = \sum_{n=0}^{N}\omega^n$

Multi-boson block factorization

- By defining the matrix

$$W_z = \begin{pmatrix} z\, P_{\partial\Lambda_0} & P_{\partial\Lambda_0} D_{\Omega_0}^{-1} D_{\Lambda_{1,2}} \\ P_{\partial\Lambda_2} D_{\Omega_2}^{-1} D_{\Lambda_{1,0}} & z\, P_{\partial\Lambda_2} \end{pmatrix}$$



we can re-write

the auxiliary multi-boson fields can be introduced on both boundaries so that for $N_f = 2$
[Lüscher 93; Borici, de Forcrand 95; Jegerlehner 95]

$$\prod_{k=1}^{N/2} \frac{1}{\det\{(u_k-\omega)^\dagger (u_k-\omega)\}} = \prod_{k=1}^{N} \left\{ \int [d\chi_k d\chi_k^\dagger] e^{-|W_{\sqrt{u_k}}\chi_k|^2} \right\}$$

where, by defining $\eta_k = P_{\partial\Lambda_0}\chi_k$ and $\xi_k = P_{\partial\Lambda_2}\chi_k$,

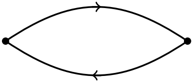
$$|W_z\chi_k|^2 = |P_{\partial\Lambda_0} D_{\Omega_0}^{-1} D_{\Lambda_{1,2}}\xi_k|^2 + |P_{\partial\Lambda_2} D_{\Omega_2}^{-1} D_{\Lambda_{1,0}}\eta_k|^2 + z(\eta_k, D_{\Omega_0}^{-1} D_{\Lambda_{1,2}}\xi_k) + \ldots$$

- The dependence of the full bosonic action from the links in Λ_0 and Λ_2 is thus factorized. The (small) direct coupling, *due to quarks looping up to N times around the boundaries*, is replaced by a block-local interaction of links with $N/2$ multi-boson fields per flavour

Signal/noise ratio for HVP: multi-level solution

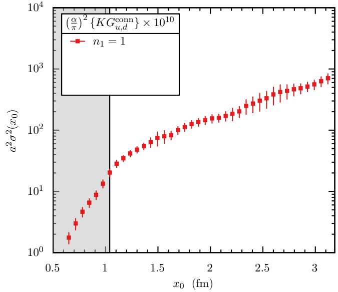
- Wilson glue with $O(a)$-improved Wilson quarks

$$\beta = 5.3\,, \quad (T/a) \times (L/a)^3 = 96 \times 48^3$$



$$a = 0.065 \text{ fm}\,, \quad M_\pi = 270 \text{ MeV}$$

$$n_0 = 25\,, \qquad n_1 = 10\,, \qquad n_{tot} = n_0 \cdot n_1$$



- Domain Decomposition adopted:

$$\Lambda_0 : x_0/a \in [0, 39]\,, \quad \Lambda_1 : x_0/a \in [40, 47] \cup [88, 95]$$

$$\Lambda_2 : x_0/a \in [48, 87]$$

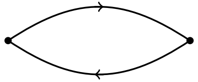
Signal/noise ratio for HVP: multi-level solution

- Wilson glue with $O(a)$-improved Wilson quarks

$$\beta = 5.3\,, \quad (T/a)\times(L/a)^3 = 96\times 48^3$$

$$a = 0.065\ \mathrm{fm}\,, \quad M_\pi = 270\ \mathrm{MeV}$$

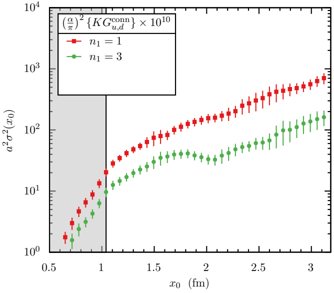
$$n_0 = 25\,, \qquad n_1 = 10\,, \qquad n_{tot} = n_0 \cdot n_1$$



- Domain Decomposition adopted:

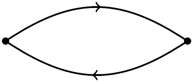
$$\Lambda_0 : x_0/a \in [0,39]\,, \quad \Lambda_1 : x_0/a \in [40,47] \cup [88,95]$$

$$\Lambda_2 : x_0/a \in [48,87]$$



Signal/noise ratio for HVP: multi-level solution

- Wilson glue with $O(a)$-improved Wilson quarks

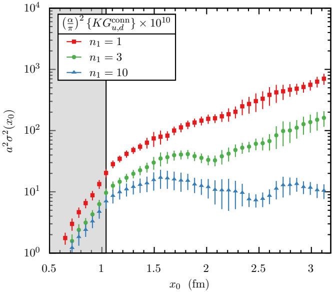


$$\beta = 5.3\,, \quad (T/a) \times (L/a)^3 = 96 \times 48^3$$

$$a = 0.065 \text{ fm}\,, \quad M_\pi = 270 \text{ MeV}$$

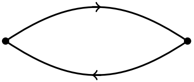
$$n_0 = 25\,, \qquad n_1 = 10\,, \qquad n_{tot} = n_0 \cdot n_1$$

- 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 $G^{\mathrm{conn}}_{u,d}$



Signal/noise ratio for HVP: multi-level solution

- Wilson glue with $O(a)$-improved Wilson quarks



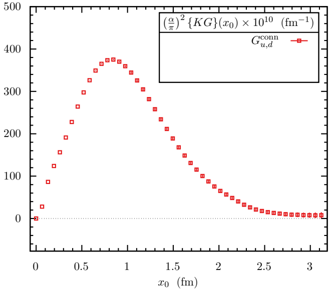
$$\beta = 5.3\,, \quad (T/a) \times (L/a)^3 = 96 \times 48^3$$

$$a = 0.065 \text{ fm}\,, \quad M_\pi = 270 \text{ MeV}$$

$$n_0 = 25\,, \qquad n_1 = 10\,, \qquad n_{tot} = n_0 \cdot n_1$$

- 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 $G_{u,d}^{\rm conn}$



Split-even estimator of disconnected contribution

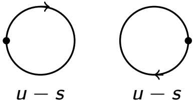
[LG, Harris, Nada Schaefer 19]

- *Advantage of multi-level sets in when variances are due to fluctuations of gauge field. If not, estimator needs to be first improved. This has been the case for the disconnected contribution*

- The disconnected Wick contraction reads

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

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

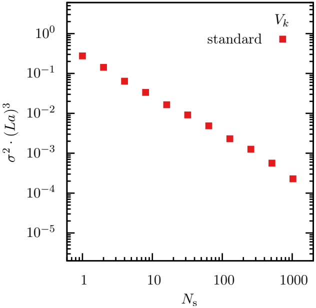


- Standard stochastic estimator [$\langle\eta(x)\eta^\dagger(y)\rangle = \delta_{xy}$]

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

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



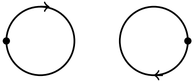
Split-even estimator of disconnected contribution

[LG, Harris, Nada Schaefer 19]

- *Advantage of multi-level sets in when variances are due to fluctuations of gauge field. If not, estimator needs to be first improved. This has been the case for the disconnected contribution*
- The disconnected Wick contraction reads

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

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

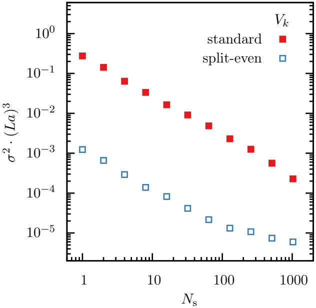


- Split-even stochastic estimator [$\langle \eta(x)\eta^\dagger(y)\rangle = \delta_{xy}$]

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

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) **[ETM Coll. 08, 12]**

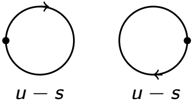


Split-even estimator of disconnected contribution

[LG, Harris, Nada Schaefer 19]

- *Advantage of multi-level sets in when variances are due to fluctuations of gauge field. If not, estimator needs to be first improved. This has been the case for the disconnected contribution*
- The disconnected Wick contraction reads

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

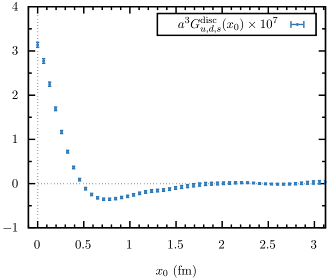


- Split-even stochastic estimator [$\langle\eta(x)\eta^\dagger(y)\rangle = \delta_{xy}$]

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

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 and other quantities by CLS



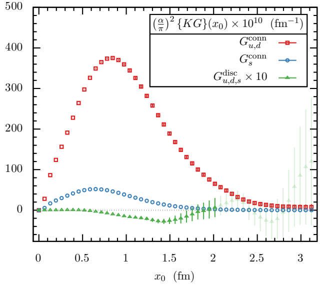
First multi-level computation of HVP

- Wilson glue with $O(a)$-improved Wilson quarks

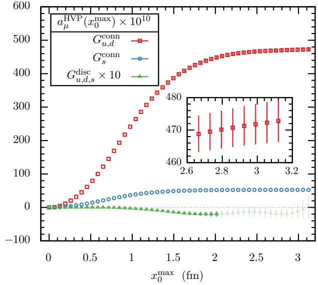
$$\beta = 5.3\,, \quad (T/a) \times (L/a)^3 = 96 \times 48^3$$

$$a = 0.065\ \mathrm{fm}\,, \quad M_\pi = 270\ \mathrm{MeV}$$

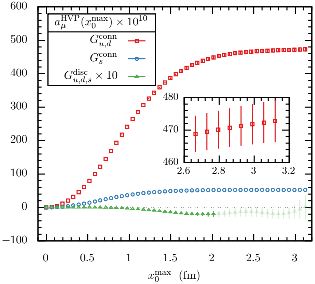
$$n_0 = 25\,, \qquad n_1 = 10\,, \qquad n_{tot} = n_0 \cdot n_1$$



- With 2-level integration achieved 1% precision with just $n_0 \cdot n_1 = 250$ configurations
- The contribution to the variance from the long distance part becomes negligible
- With lighter quarks, the gain due to the 2-level integration is even more dramatic since $(M_\rho - M_\pi)$ increases significantly



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, baryons, . . .