Hadron Spectroscopy Ib & II

Gunnar Bali

Universität Regensburg

STRONGnet Summer School

ZiF Bielefeld, June 15, 2011

Part I: introduction and basic methods

- Motivation
- Two point functions
- Some group theory
- Smearing
- Generalized eigenvalue/variational method
- "Distillation"

Part II: all-to-all and stochastic methods

- Low mode averaging
- Stochastic all-to-all methods
- Variance reduction techniques
- The one-end-trick
- Thinning the estimates: grid noise and distillation

Many examples and more details by Christian Lang and David Richards!

Correlation matrices

We consider a correlation matrix,

$$D_{ij}(t) = \langle O_i(t) O_j^\dagger(0) \rangle , \quad i,j \in \{1,2,\ldots,M\} ,$$

between states created by different (non-orthogonal) operators $\hat{O}_j^\dagger$ (e.g. different numbers of smearing iterations).

D should be Hermitian (usually symmetric) and positive definite for any $t \geq t_{\text{min}}$ (that depends on the action and operator).

The error may be smaller for elements with more source smearing. So substituting the upper off-diagonal triangle by the lower one may be better than averaging or doing nothing.

There can be sign problems that are related to the use of $\bar{q} = q^\dagger \gamma_4$ instead of the Euclidean $q^\dagger$. For instance in Chroma the signs of mesonic correlation matrix colums j need to be flipped whenever $\hat{O}_j$ contains a Γ with the property, $\Gamma^\dagger \gamma_4 = +\gamma_4 \Gamma$. This is e.g. the case for $\Gamma = \mathbb{1}$.

Following M Lüscher, U Wolff, NPB 339 (90) 222; ALPHA: B Blossier et al, JHEP 0904 (09) 094, we define,

$$C(t) = D^{-\frac{1}{2}}(t_0)D(t+t_0)D^{-\frac{1}{2}}(t_0)\,.$$

This symmetric definition ensures orthogonality of the eigenvectors $|\psi_n(t)\rangle$:

$$C(t)|\psi_n(t)\rangle = \lambda_n(t)|\psi_n(t)\rangle\,,$$

where we order $\lambda_1(t) > \lambda_2(t) > \cdots > \lambda_M(t) > 0$ at large t. To ensure consistency over jacknifes/bootstraps, the eigenvectors should be monitored as well.

Note that, $C(0) = \mathbb{1}$: now everything in the eigenbasis of $D(t_0)$.

Also note that the original non-symmetrized definition of C Michael, NPB 259 (85) 58 yields the same eigenvalues (but different eigenvectors $|\phi_n(t)\rangle = D^{-\frac{1}{2}}(t_0)|\psi_n(t)\rangle$):

$$D^{-1}(t_0)D(t+t_0)|\phi_n(t)\rangle = \lambda_n(t)|\phi_n(t)\rangle\,.$$

Effective masses

Generalized effective masses can now be defined as,

$$E_{\mathrm{eff},n}(t) = -a^{-1} \ln \frac{\lambda_n(t)}{\lambda_n(t+a)} \stackrel{t\to\infty}{\longrightarrow} E_n \,.$$

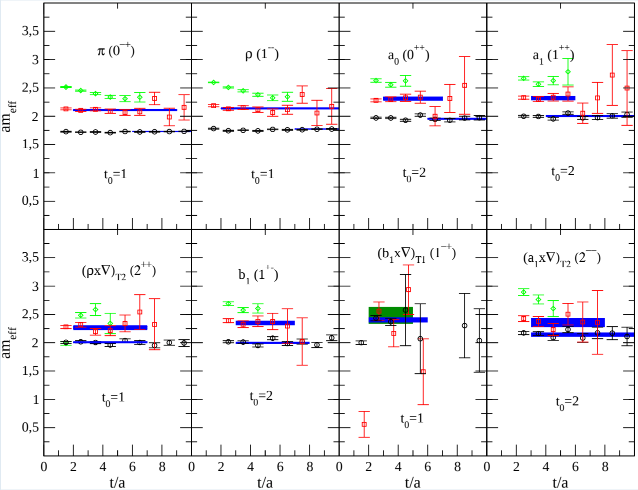
These also depend on t_0 and this should be varied.

If t_0 is too small then states with energies larger than E_M will considerably contribute and in particular excitations with small gaps relative to E_{M+1} will need a larger time distance t to plateau.

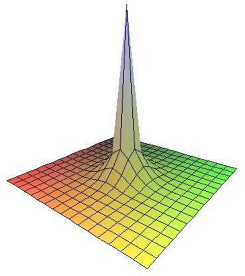
If t_0 is too big then the Mth state may have decayed within statistical errors and the rank of $D(t_0)$ may not be maximal, resulting in numerical problems. In this case the basis may need some pruning (or statistics can be increased).

Example for effective masses

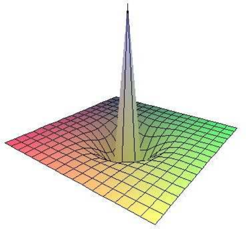
Charmonia: $a^{-1} \approx 1.73\,\mathrm{GeV}$, C Ehmann, GB



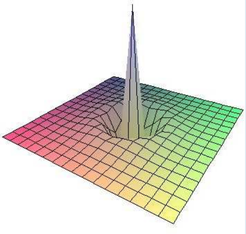
Optimized smearing functions (Coulomb gauge)



1S



2S



3S

The Gauss-Wuppertal/Jacobi smearing operator can be written as,

$$\Phi^{(n)} \approx \left(e^{\kappa \Delta t \nabla^2}\right)^n .$$

∇^2 is a scalar, Hermitian, translationally invariant, gauge covariant operator. It contains smeared transporters $\overline{U}$.
Define eigenvectors of ∇^2 at a fixed timeslice, $|v^i\rangle \in \mathbb{C}^{V_3 N_c}$:

$$\nabla^2 |v^i\rangle = \omega_i^2 |v^i\rangle , \quad \langle v^i | v^j \rangle = \delta_{ij} , \quad v^i_{\mathbf{x},a} = \langle \mathbf{x}, a | v^i \rangle .$$

From this we can define a projector onto the "LapH" subspace HSP: M Peardon et al, PRD 80 (09) 054506 of the timeslice,

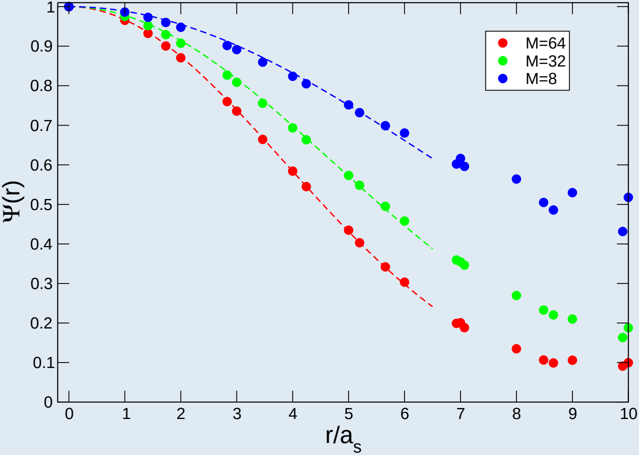
$$\triangle = \sum_i |v^i\rangle\langle v^i| \, \theta(\sigma^2 - \omega_i^2) , \quad \triangle^{\mathbf{xy}}_{ab} = \sum_i v^{i\dagger}_{\mathbf{x},a} v^i_{\mathbf{y},b} \, \theta(\sigma^2 - \omega_i^2) ,$$

where (obviously) $\triangle^2 = \triangle$. σ cuts out all eigenvectors with eigenvalues $\omega_i > \sigma$. The number of remaining eigenvalues $M(\sigma) \ll V_3 N_c$ scales at fixed $\sigma^2 \approx 1/3$ with $V_3 = L_s^3$.

The "wavefunction"

$$\Psi(\mathbf{r}) = \sqrt{\mathrm{Tr}\left(\triangle_{\mathbf{0r}}\triangle_{\mathbf{r0}}\right)}$$

(averaging the zero point over all lattice points) approaches the δ-function for $M \to L_s^3 N_c$ (Distillation becomes a basis transformation).



Mesonic two point functions

Destruction operator:

$$\hat{O}^{\mathbf{p}} = \sum_{\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{w}} e^{-i\mathbf{p}\mathbf{x}} \bar{q}^1_{\mathbf{x}} \triangle_{\mathbf{xy}} \underbrace{e^{-i\mathbf{p}\mathbf{y}} \Gamma D_{\mathbf{yz}}}_{A_{\mathbf{yz}}} \triangle_{\mathbf{zw}} q^2_{\mathbf{w}}$$

where we have suppressed colour and spin indices and A depends on $\mathbf{p}$, Γ and D. $\triangle$ depends on M. Correlation function (We allow for $\hat{\tilde{O}} \neq \hat{O}$):

$$\begin{aligned} C(t) &= \langle \tilde{O}(t) O^\dagger(0) \rangle \\ &= \pm \left\langle \bar{q}^2(t) \triangle (t) \tilde{A}(t) \triangle (t) q^1(t) \bar{q}^1(0) \triangle (0) A^\dagger(0) \triangle (0) q^2(0) \right\rangle \\ &= \pm \sum_{i,j,k,\ell} \Big\langle \langle \bar{q}^2(t) | v^i(t) \rangle \tilde{A}_{ij}(t) \langle v^j(t) | q^1(t) \rangle \\ &\qquad \times \langle \bar{q}^1(0) | v^k(0) \rangle \tilde{A}^\dagger_{k\ell}(0) \langle v^\ell(0) | q^2(t) \rangle \Big\rangle , \end{aligned}$$

where $\tilde{A}_{ij}(t) = \langle v^i(t) | \tilde{A}(t) | v^j(t) \rangle$ and $A^\dagger_{k\ell}(0) \langle v^k(0) | A^\dagger(0) | v^\ell(0) \rangle$ also depend on (not displayed) spinor indices.

This can now be factorized,

$$C(t) = \mp \left\langle \tilde{A}^{ij}_{\alpha\gamma}(t) A^{\dagger\, k\ell}_{\beta\delta}(0) \langle v^j(t)|S^1(t|0)_{\alpha\beta}|v^k(0)\rangle \langle v^\ell(0)|S^2(0|t)_{\gamma\delta}|v^i(t)\rangle \right\rangle_U$$

$$= \mp \left\langle \tilde{A}^{ij}_{\alpha\gamma}(t) A^{\dagger\, k\ell}_{\beta\delta}(0)\, \tau^{(1)}(t|0)^{jk}_{\alpha\beta}\, \tau^{(2)}(0|t)^{\ell i}_{\gamma\delta} \right\rangle_U ,$$

where the generalized propagators ("preambulators"),

$$\tau^{(n)}(t|0) = \left(\langle v^i(t)|S^n(t|0)|v^j(0)\rangle \right) ,$$

are LapH $\otimes$ spin $(4M \times 4M)$ matrices that can be obtained by inverting the Dirac operator on all $|v^j(0)\rangle$ (times the four different source spin-δs), and contracting the resulting propagators at the sink with $\langle v^i(t)|$: the colour times position indices are replaced by LapH indices i and j.
Note that the computation of the antiquark perambulator,

$$\tau(0|t)_{ij} = \langle v^i(t)|\gamma_5 S^n(t|0)\gamma_5|v^j(0)\rangle = \gamma_5 \langle v^i(t)|S^n(t|0)|v^j(0)\rangle \gamma_5$$

does not require any additional solves, due to the γ_5-Hermiticity.

Summary of Distillation

- This has been generalized to baryons etc. (straight-forward).
- This timesliceLapH-to-allLapH method is much more expensive than the standard point-to-all method. The price for the inversions scales like VV_3 (rather than V), and for mesonic contractions even like $(VV_3)^2$.
- The A_{ij} can be exchanged *a posteriori*. This will turn the method competitive when many operators are involved, in particular with derivatives at the source. Also some source "self-averaging" is built in.
- All components within the $\sum_{ijk\ell}$ have the quantum numbers of A and are gauge invariant. So different truncations can be chosen for ij and $k\ell$ (corresponding to different sink/source smearings). See also `C Lang et al, arXiv:1105.5636`.
- The smearing profiles can also be varied by introducing weight functions $f(\omega_i)$ in the contraction of a LapH index i, a possibility that could be worth exploring.

Low mode averaging

At light quark masses one may compute eigenvectors to deflate the solver.

Eigenvectors also offer the possibility of low mode averaging (LMA) T De Grand, S Schäfer CPC 159 (04) 185, L Giusti et al, JHEP 0404 (04) 013.

$$C_{\rm LMA}(t) = C_{\rm low}(t) + C^{\rm pa}(t) - C^{\rm pa}_{\rm low}(t)\,.$$

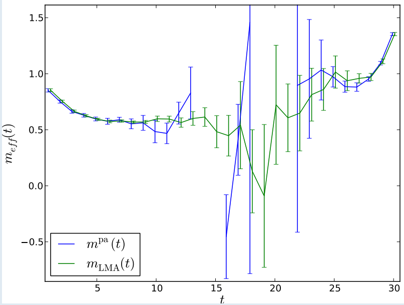
$C_{\rm low}$: contribution from low eigenmodes of $Q=\gamma_5 M$ ($Q=Q^\dagger$), all-to-all, averaged over the lattice volume.
$C^{\rm pa}$: standard point-to-all 2-point function.
$C^{\rm pa}_{\rm low}$: low mode contribution (point-to-all), needs to be subtracted since this is already included into $C^{\rm pa}$.

This does not affect the expectation value but may reduce the error, due to the self-averaging of the low-mode contribution.

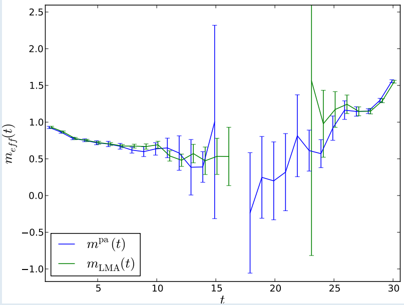
This works well for positive parity baryons and negative parity mesons GB, L Castagnini, S Collins, PoS (LATTICE2010) 096

Effective masses

N



Δ^+



Example: meson

$$Q|u^i\rangle = q_i|u^i\rangle\,,\quad \langle u^i|u^j\rangle = \delta_{ij}\,,\quad q_i \in \mathbb{R}\,,\quad Q = \gamma_5 M\,.$$

This means that,

$$Q = \sum_{i=1}^{12V} \frac{1}{q_i}|u^i\rangle\langle u^i|\,.$$

We need to truncate: $i \in \{1, 2, \ldots, m\}$ where $m \propto V$. So the number of operations increases $\propto V^2$.

The eigenvectors have position, spin and colour components: $u^i(x)_{\alpha a} = \langle x, \alpha, a|u^i\rangle$.

$$C_{\text{low}}(t) = \pm \sum_{i,j} \left\langle \frac{1}{q_i q_j}\, {}_t\langle u^j|\gamma_5\Gamma|u^i\rangle_t\, {}_0\langle u^i|\gamma_5\Gamma|u^j\rangle_0 \right\rangle_U\,,$$

where the subscripts t denote a projection of the vector onto timeslice t.

The point-to-all low mode contribution can be obtained using (Note that $u^i(x) = \langle x|u^i\rangle$ is a spin-colour vector),

$$C^{\rm pa}_{\rm low}(t) = \pm \sum_{i,j} \left\langle \frac{1}{q_i q_j} \,{}_t\langle u^j|\gamma_5\Gamma|u^i\rangle_t \, u^i(0)^\dagger \gamma_5 \Gamma u^j(0) \right\rangle_U .$$

It is straight-forward to add momenta and smearing functions. The latter however cannot be factorized: unlike the LapH vectors, the eigenvectors have a colour component.

What about eigenmodes of M?

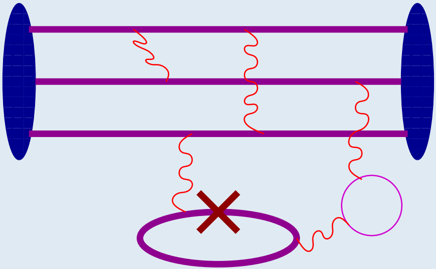
Left $\langle \ell^i|$ and right $|r^i\rangle$ eigenvectors of an eigenvalue $\lambda_i \in \mathbb{C}$ need to be distinguished. These fulfill the biorthonormality relations $\langle \ell^i|r^i\rangle = \delta_{ij}$ and $M^{-1} = \sum_i \frac{1}{\lambda_i}|r^i\rangle\langle \ell^i|$. Moreover, $\langle r_i|\gamma_5$ and $\gamma_5|\ell_i\rangle$ are left and right eigenvectors, respectively, with eigenvalue λ_i^*. It turns out that this converges badly L Castagnini et al, PoS (LATTICE2010) 096: the dynamics appears to be driven by eigenmodes of the Hermitian Dirac operator Q.

Often "all-to-all" is necessary:

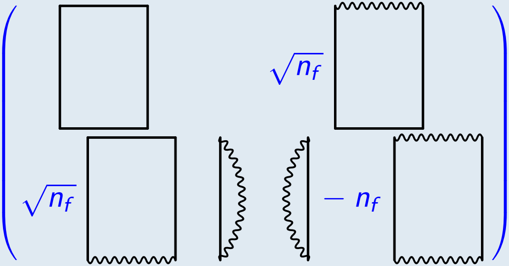
nucleon structure:

$\langle N^\dagger(t)|J_\mu(t/2)|N(0)\rangle$:

(Example: $J_\mu = \psi^\dagger\gamma_\mu\psi$)



decays/scattering:



Stochastic methods K Bitar et al, NPB 313 (89) 348:

Generate a set of random noise vectors $|\eta^\ell\rangle$, $\ell = 1, \ldots, n$ where

$$\frac{1}{n}\sum_\ell |\eta^\ell\rangle\langle\eta^\ell| = \overline{|\eta\rangle\langle\eta|}_n = \overline{|\eta\rangle\langle\eta|} = \mathbb{1} + \mathcal{O}(1/\sqrt{n}),$$

$$\overline{\langle\eta|} = \mathcal{O}(1/\sqrt{n}).$$

Often: $\eta^\ell(x)_{\alpha a} \in Z = \mathbb{Z}_2 \otimes i\,\mathbb{Z}_2/\sqrt{2}$ S Dong, K-F Liu, PLB 328 (94) 130. Other choices: $Z = \mathbb{Z}_2, \mathbb{Z}_3, \mathrm{U}(1), \mathrm{SU}(3)$.

By solving

$$M|s^\ell\rangle = |\eta^\ell\rangle$$

for the $|s^\ell\rangle$ one can construct an unbiased estimate:

$$M_E^{-1} = \overline{|s\rangle\langle\eta|} = M^{-1} + M^{-1}\underbrace{\left(\overline{|\eta\rangle\langle\eta|} - \mathbb{1}\right)}_{\mathcal{O}(1/\sqrt{n})}$$

$\Rightarrow$ $n \ll 12V$ solver applications only !

On each configuration an estimate A_E of A has a stochastic error $\Delta_{\rm stoch}A = \mathcal{O}(1/\sqrt{n})$. We define:

$$\sigma^2_{A,\rm stoch} := \frac{\langle(\Delta_{\rm stoch}A)^2\rangle_U}{N} \propto \frac{1}{Nn} \quad \text{for} \quad n, N \text{ large}\,,$$

where N is the number of gauge configurations. The configuration average $\langle A_E\rangle_U$ carries the statistical error $\sigma_{A,\rm gauge}$:

$$\sigma^2_{A,\rm gauge} \geq \sigma^2_{A,\rm stoch}\,.$$

Both sides scale $\propto 1/N$.

$\sigma_{A,\rm gauge} \simeq \sigma_{A,\rm stoch} \Rightarrow$ increase n.

$\sigma_{A,\rm gauge} \gg \sigma_{A,\rm stoch} \Rightarrow$ reduce n and increase N (or the source positions).

The optimal choice depends on the observable A.

Increasing n is usually not the smartest thing to do.

It is better to reduce the coefficient of the $1/\sqrt{n}$ term.

The stochastic error

$$\left[\Delta M^{-1}_{XZ}\right]^2 := \left[\Delta_{\mathrm{stoch}} M^{-1}_{XZ}\right]^2 = \sum_Y \left[M^{-1} - M_E^{-1}\right]_{XY} \left[M^{-1} - M_E^{-1}\right]^\dagger_{YZ} \,,$$

$$\left[\Delta M^{-1}\right]^2 = M^{-1}\mathbb{O}\left[M^{-1}\mathbb{O}\right]^\dagger \,,$$

where

$$\mathbb{O} = \mathbb{1} - \overline{|\eta\rangle\langle\eta|} = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$$

is an off-diagonal $12V \times 12V$ matrix. $[X = (x, \alpha a)]$. This means that,

$$\left[\Delta M^{-1}_{XZ}\right]^2 \propto \frac{1}{n} \sum_{Y \neq X,Z} M^{-1}_{XY} M^{-1\dagger}_{YZ}$$

$$\left[\Delta\left(\mathrm{Tr}\,\Gamma M^{-1}\right)\right]^2 \propto \frac{1}{n} \sum_{x,y} \bar{q}_y \Gamma \gamma_5 q_y \bar{q}_x \Gamma \gamma_5 q_x \quad \text{minus diagonal terms}$$

This is a sum over a mesonic two point function $c_M(y-x)$!

The stochastic error $\Delta\,\mathrm{Tr}\,\Gamma M^{-1} \propto [(V/n)\sum_{y\neq 0} c(y)]^{1/2}$ (plus non spin-colour-diagonal terms at $y = 0$.)

$c(y)$ is the point-point correlation function of $\hat{O}_M = \bar{q}\Gamma\gamma_5 q$.

Biggest contributions are from the "neighbourhood", where $c(y)$ is large. Intuitively this is clear from $M_E^{-1} - M^{-1} = M^{-1}(\overline{|\eta\rangle\langle\eta|} - \mathbb{1})$ but above is gauge invariant. Exercise: repeat this for a mesonic two-point-function with and without one-end-trick.

Hopping parameter expansion (HPE)

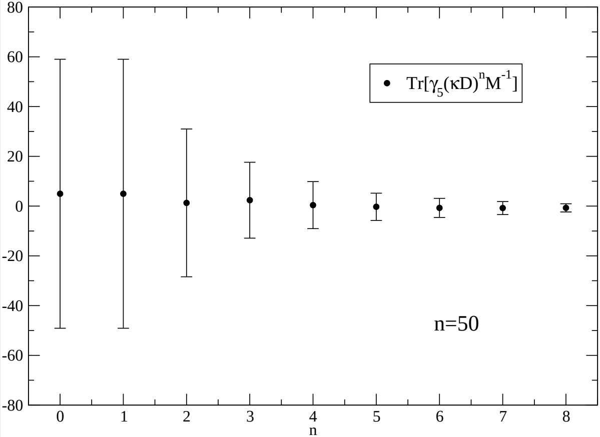
C Thron et al, PRD 57 (98) 1642; C Michael et al, NPPS 83 (00) 185. For static-light mesons: SESAM: GB et al, PRD 71 (05) 114513.

The first few terms of the hopping parameter expansion of $\mathrm{Tr}(\Gamma M^{-1}) = \mathrm{Tr}[\Gamma(\mathbb{1} - \kappa \not{D})^{-1}]$ vanish identically but still contribute to the noise. For the Wilson action, $\mathrm{Tr}(\Gamma M^{-1}) = \mathrm{Tr}(\Gamma\kappa^n \not{D}^n M^{-1})$, $n = 4, 8$, depending on Γ, where estimating the latter yields smaller errors. The $n = 0$ term for $\Gamma = \mathbb{1}$ can easily be calculated and corrected for.

This only works for ultra-local actions. No Neuberger Fermions!

Coined "unbiased subtraction method" the first few non-vanishing $\kappa \not{D}$ orders have been calculated analytically for the clover action M Deka et al, PRD 79 (09) 094502.

Heavy quarks (charm) C Ehmann, GB, PoS (LATTICE2008) 114



Partitioning S Bernardson et al, CPC 78 (93) 256; J Viehoff et al, NPPS 63 (98) 269; W Wilcox, arXiv:hep-lat/9911013

(also known as the spin-explicit-method (SEM) or dilution)
Decompose $\mathcal{R} =$ volume $\otimes$ colour $\otimes$ spin into n_p subspaces:

$$\mathcal{R} = \oplus_{j=1}^{n_\mathrm{p}} \mathcal{R}_j \,.$$

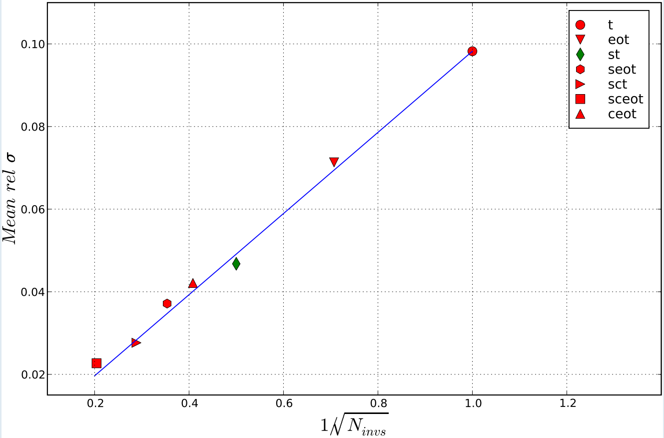
Set $|\eta^\ell_{|j}\rangle$ to zero outside of the domain $\mathcal{R}_j$.
Calculate restricted solutions,

$$M|s^\ell_{|j}\rangle = |\eta^\ell_{|j}\rangle \,.$$

Now: $M_E^{-1} = \sum_j \overline{|s_{|j}\rangle\langle\eta_{|j}|}$
This can be used to black out large off-diagonal error terms.

It is sensible to choose the same random vector components within each subspace (if they have the same dimension). This allows for hand-coding of, e.g., the spin structure (SEM).

Often not all columns of M^{-1} are required (e.g. time partitioning for 3-point functions). Spin partitioning sometimes is justified by the error reduction. Mostly it does not do any harm either:



Comparison of partitioning patterns in mesonic three point functions R Evans, S Collins, GB, PRD 82 (10) 094501.

The partitioning pattern can be adapted to the problem:
staggered spin dilution (SSD).

There is also the possibility of "recursive noise subtraction" (RNS).

These methods are introduced in `C Ehmann, GB, PoS (LATTICE2008) 114`.

Truncated eigenmode approach (TEA) H Neff et al, PRD 64 (01) 114509; GB et al, NPPS 140 (05) 609; PRD 71 (05) 114513; A O'Cais et al, NPPS 140 (05) 844; CPC 172 (05) 145.

Calculate the m lowest eigenvalues and eigenvectors of $\mathrm{Q} = \gamma_5 M$, q_i and $|v^i\rangle$. Projection operator:

$$\mathbb{P} = \sum_{i=1}^{m} |v^i\rangle\langle v^i| \,.$$

With

$$M|s_\perp^\ell\rangle = |\eta_\perp^\ell\rangle = \gamma_5 \left(\mathbb{1} - \mathbb{P}\right) \gamma_5 |\eta^\ell\rangle$$

one obtains,

$$M_E^{-1} = \overline{|s_\perp\rangle\langle\eta_\perp|} + \sum_{i=1}^{m} |v^i\rangle q_i^{-1} \langle v^i|\gamma_5 \,.$$

Deflation is included for free and with the CG algorithm, the solution does not need to be projected back.

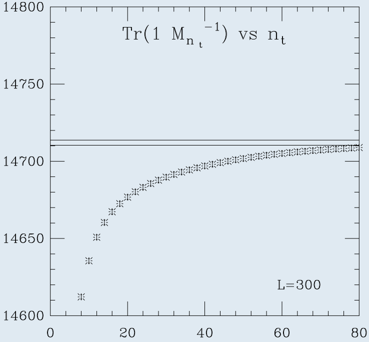
Truncated solver method (TSM)

S Collins et al, PoS (LAT2007) 141

Obtain approximate solutions $|s^{\ell}_{n_t}\rangle$ after n_t solver iterations (before convergence), and estimate the difference stochastically to obtain an unbiased estimate of M^{-1}:

$$M_E^{-1} = \overline{|s_{n_t}\rangle\langle\eta|}_{n_1} + \overline{(|s\rangle - |s_{n_t}\rangle)\langle\eta|}_{n_2} \quad \text{with} \quad n_2 \ll n_1 \, .$$

n_2/n_1 can be optimized to minimize the cost for a given error.

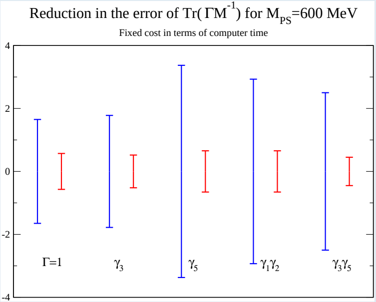


Do ∃ other factorizations of M^{-1} into an expensive contribution with a small error and a cheap contribution with a larger error?

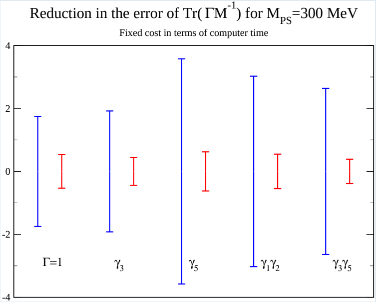
Iterative schemes to fight $\sqrt{V/n}$ problem?

Reduction of the stochastic error at fixed cost

Results for $\mathrm{Tr}(\Gamma M^{-1})$ on 1 configuration Sara Collins et al, PoS (LATTICE2008) 161; CPC 181 (10) 1570:



(a) Partitioning, HPE, TSM



(b) Partitioning, HPE, eigenmodes, TSM

- Significant gain for all Γs.
- Using different combinations of methods allows one to obtain similar gains at different quark masses.

One-end literature

- One-end-trick
  M Foster, C Michael, PRD 59 (99) 074503
- Spin-explicit OET
  C McNeile, C Michael, PRD 73 (06) 074506
- Sequential use in 3-point functions
  ETMC: S Simula et al, PoS (LAT2007) 371;
  UKQCD: P Boyle et al, JHEP 0807 (08) 112;
  R Evans et al, PRD 82 (10) 094501
- Sequential use in 4-point functions
  CP-PACS: S Aoki et al, PRD 76 (07) 094506
- OET in baryons
  χQCD: A Li et al, PRD 82 (10) 114501;
  L Castagnini et al, in preparation

Define noise $\eta^\ell(x)_{\alpha a} \in Z$ that is zero for any $t \neq t_0$.

$$\frac{1}{n}\sum_{\ell=1}^{n} |\eta^\ell\rangle\langle\eta^\ell| = \mathbb{1}_{t_0} + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \approx \sum_{\mathbf{x},\alpha,a} |x,\alpha,a\rangle\langle x,\alpha,a|\,,$$

where $x_4 = t_0$. Consider the (not gauge averaged) pion two-point function ($t_0 = 0,\ y = (\mathbf{y}, t)$),

$$\begin{aligned} c(t) &= \sum_{\mathbf{xy}} \mathrm{Tr}\, M^{-1}(y|x)[M^{-1}(x|y)]^\dagger \approx c_E(t) \\ &= \sum_{\mathbf{y}} \frac{1}{n}\sum_{\ell=1}^{n} \mathrm{Tr}\,\langle y|M^{-1}|\eta^\ell\rangle\langle\eta^\ell|M^{-1\dagger}|y\rangle \\ &= \sum_{\mathbf{y}} \frac{1}{n}\sum_{\ell=1}^{n} \mathrm{Tr}\,\langle y|s^\ell\rangle\langle s^\ell|y\rangle = \sum_{\mathbf{y},\beta,b} \frac{1}{n}\sum_{\ell=1}^{n} |s^\ell(y)_{\beta b}|^2, \end{aligned}$$

where $M|s^\ell\rangle = |\eta^\ell\rangle$. $c_E(t)$ differs from $c(t)$ by terms of $\mathcal{O}(1/\sqrt{n})$. Since the noise is unbiased, $C(t) = \langle c(t)\rangle_U = \langle c_E(t)\rangle_U$.

Without the OET we would have needed two sets of sources $|\eta_1^\ell\rangle$ and $|\eta_2^\ell\rangle$:

$$c_E^{\mathrm{trad}}(t) = \sum_{\mathbf{y}} \frac{1}{n^2} \sum_{\ell,k=1}^{n} \mathrm{Tr}\,\langle y|s_1^\ell\rangle\langle\eta_1^\ell|\eta_2^k\rangle\langle s_2^k|y\rangle$$

$$= \sum_{\mathbf{y}} \frac{1}{n^2} \sum_{\ell,k=1}^{n} \mathrm{Tr}\,\langle y|M^{-1}\overline{|\eta_1\rangle\langle\eta_1|}\,\overline{|\eta_2\rangle\langle\eta_2|}M^{-1\dagger}|y\rangle\,.$$

Each product with $\overline{|\eta\rangle\langle\eta|}$ involves a sum over $12V_3$ randomly oscillating components of moduli $\mathcal{O}(1/\sqrt{n})$.

This means that the OET error scales $\propto \sqrt{V_3/n}$ while the traditional error is $\propto \sqrt{V_3^2/n}$. Source self-averaging yields a factor $\propto 1/\sqrt{V_3}$.

For baryons the OET error is $\propto \sqrt{V_3^2/n}$ while without the OET (LHPC: R Edwards et al, PoS (LAT2007) 108) it will scale $\propto \sqrt{V_3^3/n}$.

NB: the error can be reduced by a constant factor by recycling random sources: $\frac{1}{n^2}\sum_{\ell,k}^{n}\langle\eta_1^\ell|\eta_2^k\rangle \mapsto \frac{1}{n(n-1)}\sum_{\ell\neq k}^{2n}\langle\eta^\ell|\eta^k\rangle$, $\{|\eta\rangle\} = \{|\eta_1\rangle\}\cup\{|\eta_2\rangle\}$.
J Foley et al, CPC 172 (05)145

The OET can be made spin-explicit, defining,

$$\eta_\alpha^\ell(x)_{\beta a} = \delta_{\alpha\beta}\tilde{\eta}^\ell(x)_a \,,$$

where $|\tilde{\eta}\rangle$ is a (spin-independent) noise colour vector in the timeslice $t_0 = x_4$. With solutions,

$$M|s^\ell_{\Phi,\alpha}\rangle = \Phi|\eta^\ell_\alpha\rangle \quad \text{and} \quad M|s^\ell_{\Phi,\mathbf{p},\alpha}\rangle = e^{i\mathbf{px}}\Phi|\eta^\ell_\alpha\rangle \,,$$

we can contract,

$$\begin{aligned} c^{\mathbf{p}}_{\Gamma,\Phi}(y) &= \sum_{\mathbf{x}}[M^{-1}\Phi](y|x)e^{i\mathbf{px}}\Gamma\left[[\Phi M^{-1}](x|y)\right]^\dagger \\ &\approx \frac{1}{n}\sum_{\ell,\alpha,\beta}\left\langle y|s^\ell_{\Phi,\mathbf{p},\alpha}\right\rangle \Gamma_{\alpha\beta}\left\langle s^\ell_{\Phi,\beta}|y\right\rangle \,. \end{aligned}$$

This can now be contracted with $e^{-i\mathbf{py}}$, smearing and a Γ at the sink and averaged over gauge configurations.

For each momentum $\mathbf{p} \neq \mathbf{0}$ and each smearing function Φ four solves are required.

Summary of OET

- $|\eta\rangle$ and $|s\rangle$ are temporally separated (less noise).
- Only one set of random sources needed, no noise–noise correlations.
- Scaling improved by $\sqrt{V}$, relative to the naive method.
- Making OET spin-explicit costs a factor four but allows for all 16 Γs.
- No t self-averaging.
- Loss of generality: for each momentum/smearing new solves are needed.
- Note that there is no use in combining the OET with the HPE.

The “thinning” idea

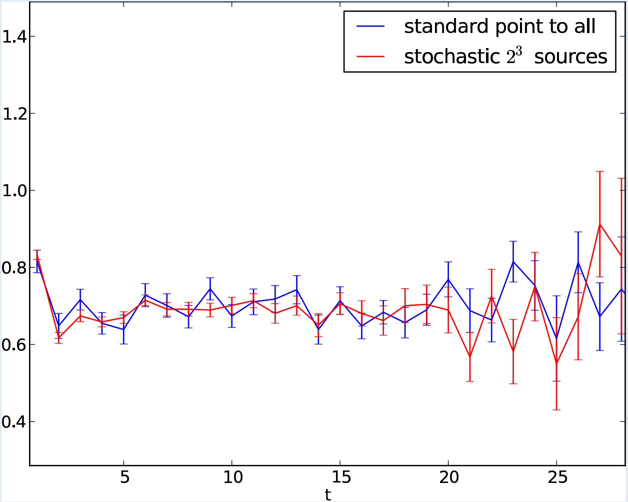
The OET error scaling (ignoring the benefit of self-averaging) is $\propto \sqrt{V_3/n}$ for mesons and $\propto \sqrt{V_3^2/n}$ for baryons. The V_3 factors are due to the number of non-zero entries of the stochastic noise vectors.

Reducing the number of non-zero entries to M points yields $\sqrt{M/n}$ and $\sqrt{M^2/n}$ behaviour, respectively, while self-averaging (for M not taken overly small) largely remains unaffected, in particular at light quark masses. L Castagnini et al, in preparation.

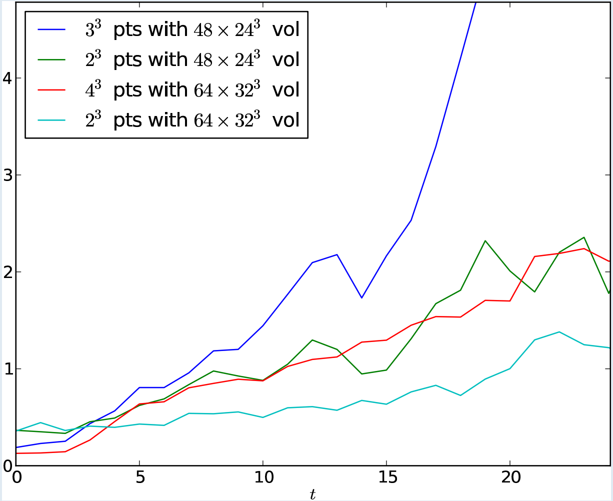
This looks like partitioning, however there is no exponential fall-off with the distance: only with respect to self-averaging it matters what points are being selected.

Grid noise was combined with low mode substitution (rather than averaging) in χQCD: A Li et al, PRD 82 (10) 114501.

Nucleon effective masses on $V = 32^3 64$ at equal cost



Error ratios for the nucleon effective mass



Noise thinning using LapH basis instead of a regular grid

It seems possible to reduce the computational overhead of the distillation method by stochastically estimating the preambulators within the LapH space HSC: C Morningstar et al, arXiv:1104.3870.

Introduce spin-explicit noise vectors in LapH space:

$$|\eta_\alpha^\ell\rangle = \sum_{i=1}^{M} \eta_i^\ell e_\alpha |v^i(0)\rangle \,,$$

where $\eta_i^\ell \in Z$, $\ell \in \{1,\ldots,n\}$, e_α is a unit spin vector in direction α and $|v^i(0)\rangle$ are LapH basis vectors on timeslice 0.

Now solve,

$$M|s_\alpha^\ell\rangle = |\eta_\alpha^\ell\rangle \,.$$

Estimates of the preambulators are now given by,

$$\tau_E(t|0)_{\alpha\beta}^{ik} = \frac{1}{n}\sum_{\ell=1}^{n}\langle v^i(t)|s_\alpha^\ell\rangle\langle\eta_\beta^\ell|v^k(0)\rangle \,.$$

Summary & Outlook

- All-to-all methods are needed in particular at small m_π where many hadrons become unstable and in general isosinglet contributions should become more important.
- Note that OET is a timeslice-to-all, distillation a timesliceLapH-to-allLapH method.
- Combinations of (new?) methods can easily save large factors of computer time.
- Efficient solvers for multiple right hand sides are needed.
- Scaling $n \propto V$ or $n \propto V_3$: can this be overcome?
- The number of low eigenmodes of Q scales like V but $4/m_\pi^{\rm phys}$ is almost 6 fm. Similarly the LapH space can become large for such volumes. Is there any "inexact" eigen/domain method?