Trace estimation of matrix inverses

James C. Osborn
BU

QCDNA
Yale
May 3, 2007

Trace estimation of matrix inverses

- physics motivation
- summary of current methods
  - stochastic sources
  - pseudofermion
  - dilution
  - subtraction
- application of multigrid

Physics motivation

- chiral condensate

$$\frac{1}{V}\langle\sum_\alpha \bar{\psi}_\alpha \psi_\alpha\rangle = \frac{1}{VZ}\int dU\, d\bar{\psi}\, d\psi \left(\sum_\alpha \bar{\psi}_\alpha \psi_\alpha\right) \mathrm{e}^{-S_g(U)-\bar{\psi}D(U)\psi}$$

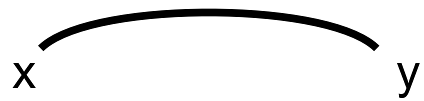
$$= \frac{-1}{VZ}\int dU \left\{\mathrm{Tr}\frac{1}{D(U)}\right\} \det[D(U)]\,\mathrm{e}^{-S_g(U)}$$

$$= -\frac{1}{V}\langle \mathrm{Tr}\frac{1}{D(U)}\rangle$$

- usually very large so easy to calculate

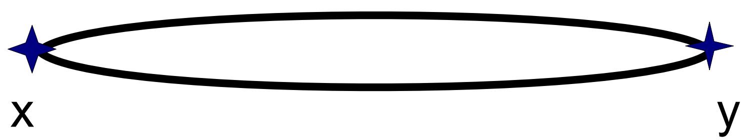
Disconnected diagrams

- quark diagrams: propagator



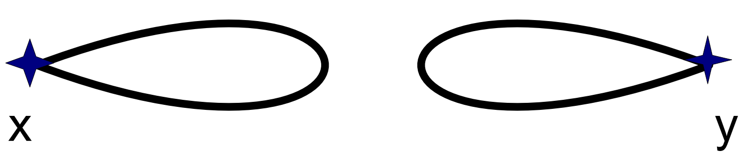
$$= \{D^{-1}\}_{y,x}$$

- connected correlator



$$= -\mathrm{Tr}_{sc}\{D^{-1}\}_{x,y} O_y \{D^{-1}\}_{y,x} O_x$$

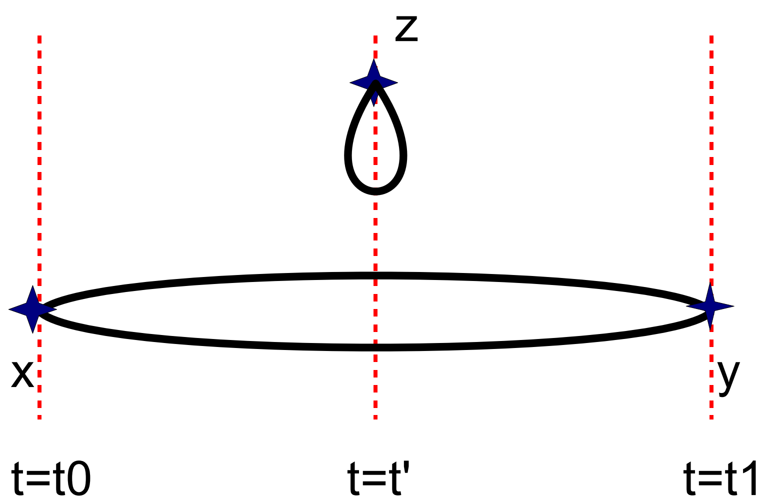
- disconnected correlator



$$= \mathrm{Tr}_{sc}[\{D^{-1}\}_{y,y} O_y]\,\mathrm{Tr}_{sc}[\{D^{-1}\}_{x,x} O_x]$$

Disconnected diagrams

- form factors



- need momentum states

$$J(\vec{p},t') = \sum_{\vec{x}} e^{i\vec{p}\cdot\vec{x}} \ \mathrm{Tr}_{sc}[O_z\{D^{-1}\}_{z,z}]_{z=\{\vec{x},t'\}}$$

Stochastic sources

- random source vectors

$$\langle (\eta^*)_\alpha (\eta)_\beta \rangle_\eta = \delta_{\alpha,\beta}$$

- trace estimate

$$\mathrm{Tr}\, O\, D^{-1} \equiv \mathrm{Tr}\, A \approx \frac{1}{N} \sum_{i=1}^{N} \eta^{(i)\,*} \{ A\, \eta^{(i)} \}$$

- variance (Dong, Liu; Bernardson, et al.)

$$\sigma^2 = \frac{N}{N-1} \left\langle \frac{1}{N} \sum_{i=1}^{N} \left(\eta^{(i)\,*} \{ A\, \eta^{(i)} \} \right)^2 - \left(\frac{1}{N} \sum_{i=1}^{N} \eta^{(i)\,*} \{ A\, \eta^{(i)} \} \right)^2 \right\rangle_\eta$$

$$= \mathrm{Tr}\, A^2 + \left(\langle |\eta_\beta|^4 \rangle_\eta - 2 \right) \sum_\alpha \{ A_{\alpha,\alpha} \}^2$$

Random source type

- Gaussian

$$\langle |\eta_\beta|^4 \rangle_\eta = 2$$

$$\begin{aligned} \sigma^2 &= \mathrm{Tr}\, A^2 \\ &= \mathrm{Tr}\, O\, D^{-1} O\, D^{-1} \end{aligned}$$

- Z(N>2), U(1)

$$\langle |\eta_\beta|^4 \rangle_\eta = 1$$

$$\sigma^2 = \sum_{\alpha \neq \beta} A_{\alpha,\beta} A_{\beta,\alpha}$$

Pseudofermion

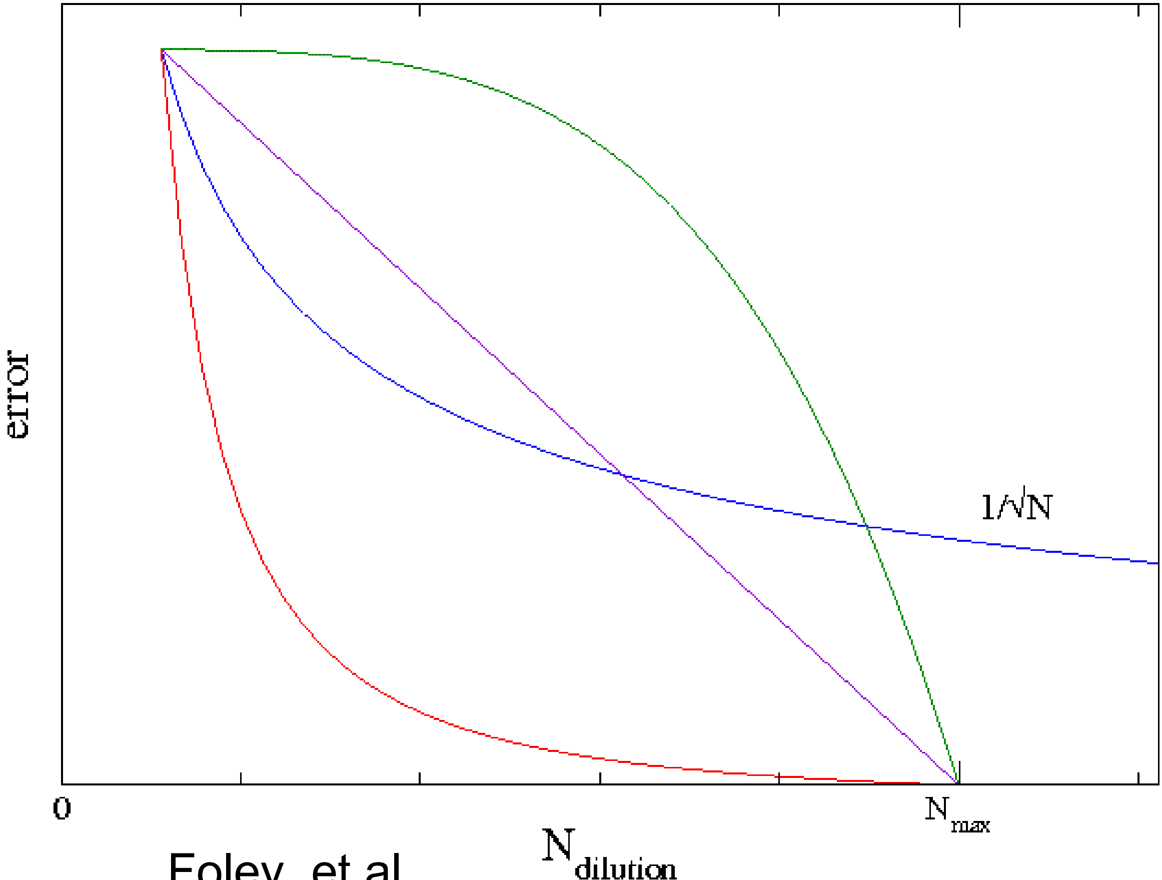
- using Gaussian source

$$\mathrm{Tr}\, O\, D^{-1} = \langle \eta^* O\, D^{-1} \eta \rangle_\eta$$
$$= (1/Z) \int d^{2\mathrm{N}} \eta \ \{ \eta^* O\, D^{-1} \eta \} \ \mathrm{e}^{-\eta^* \eta}$$
$$= (1/Z') \int d^{2\mathrm{N}} \phi \ \{ \phi^* O\, \phi \} \ \mathrm{e}^{-\phi^* D \phi}$$

- heat-bath update of Φ (Duncan, Eichten, Yoo)
  - low modes of D cause long autocorrelations
  - can be combined with low eigenvector projection

Dilution

- use separate sources for each group of matrix indices: color, spin, spatial (Wilcox; Foley, et al.)
- pure random
  error ~ 1/sqrt(N)
- "perfect" dilution
  (N_{max}=12V sources)
  error = 0
- want to approach "perfect" dilution as quickly as possible



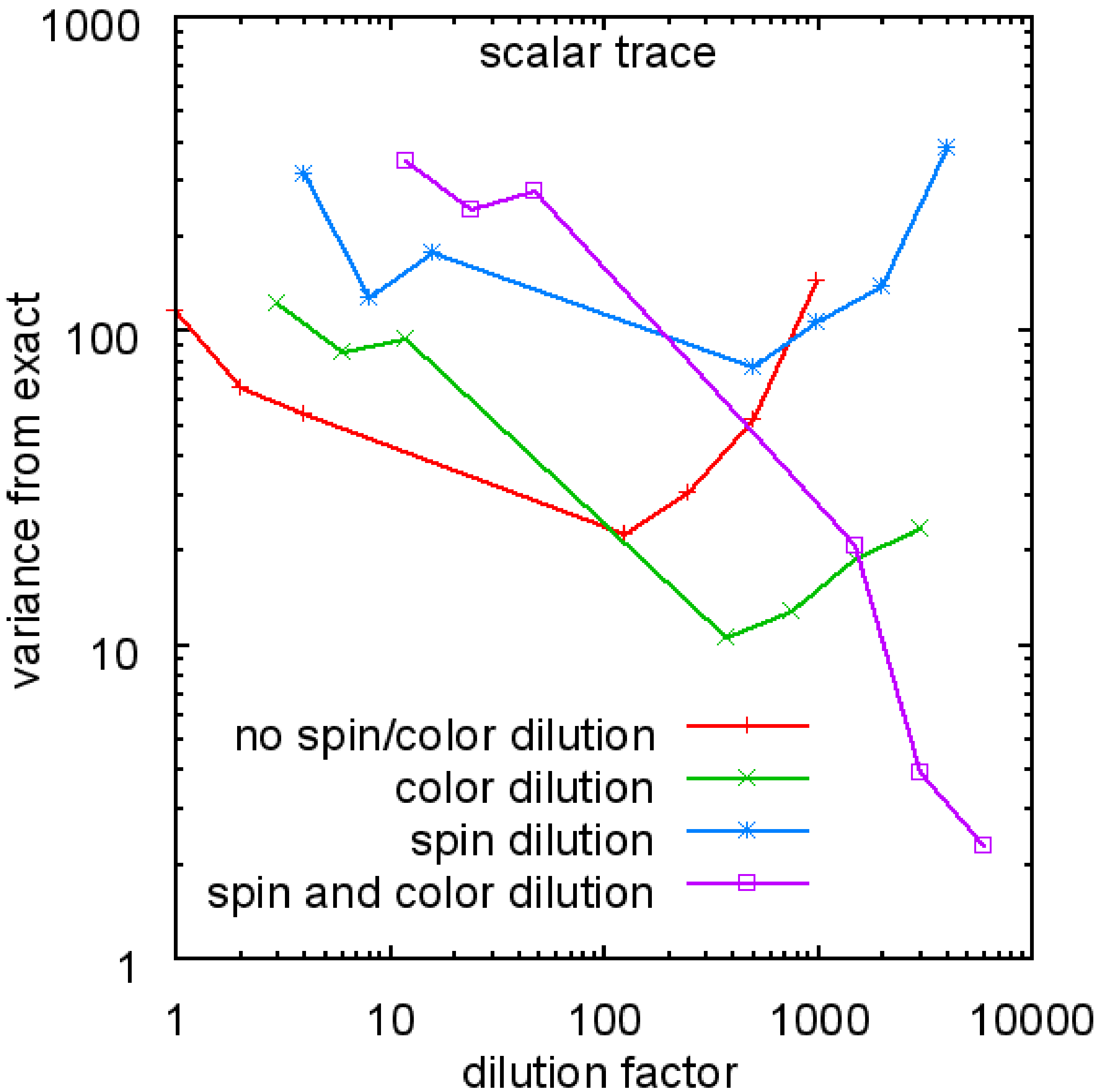
Foley, et.al.

Dilution

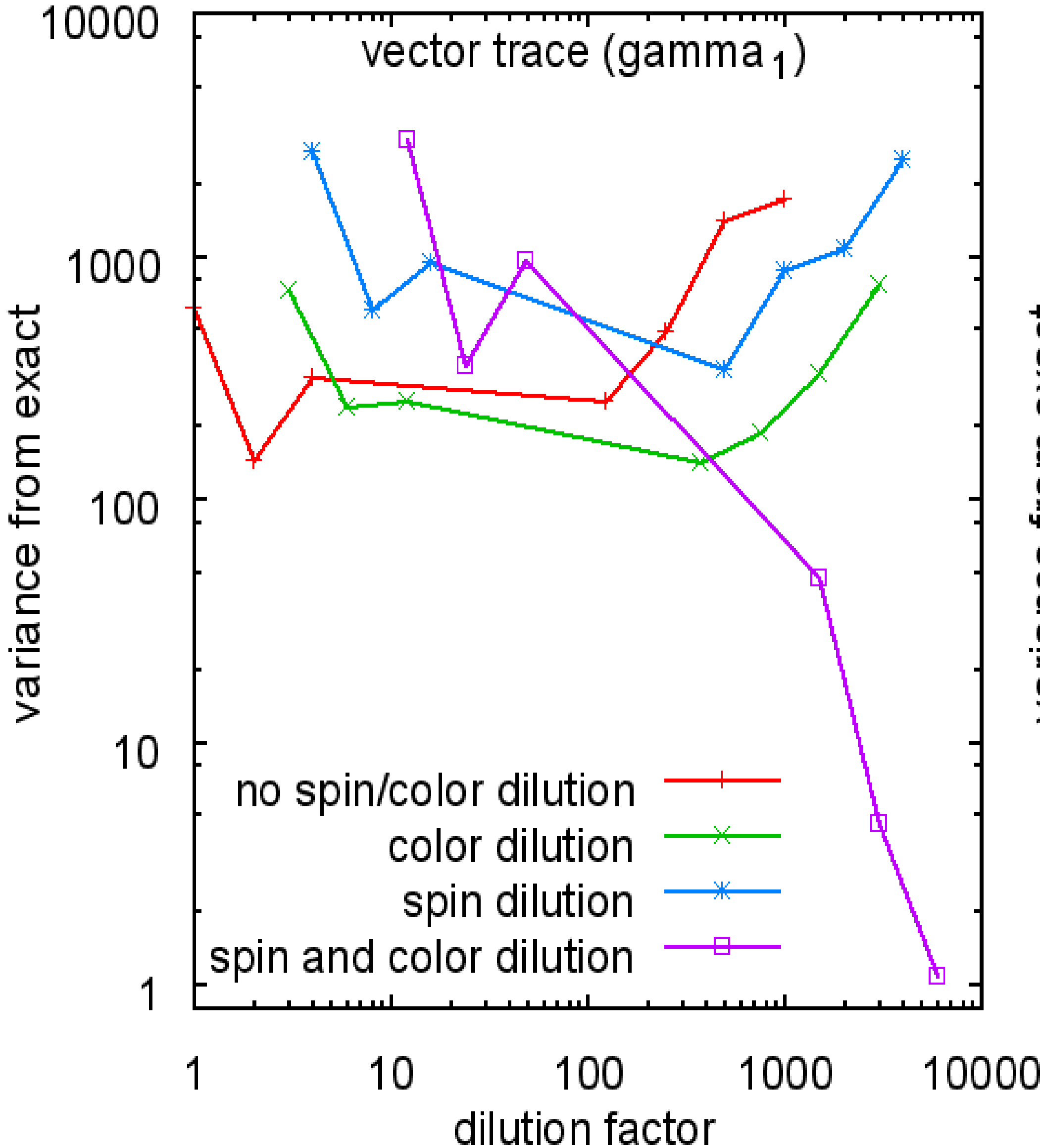
- tests on 10^3x32 Wilson Dirac Matrix
- trace on single time slice (12,000 components)
- with/without color/spin dilution
- spatial dilution tests:
  - none (dilution factor 1)
  - even/odd (dilution factor 2)
  - cubic diagonal [(0,0,0)(1,1,1)],[(0,0,1)(1,1,0)],... (dilution factor 4)
  - inner 5^3 dilution with outer dilution among inner blocks (dilution factors 125, 250, 500)

Dilution

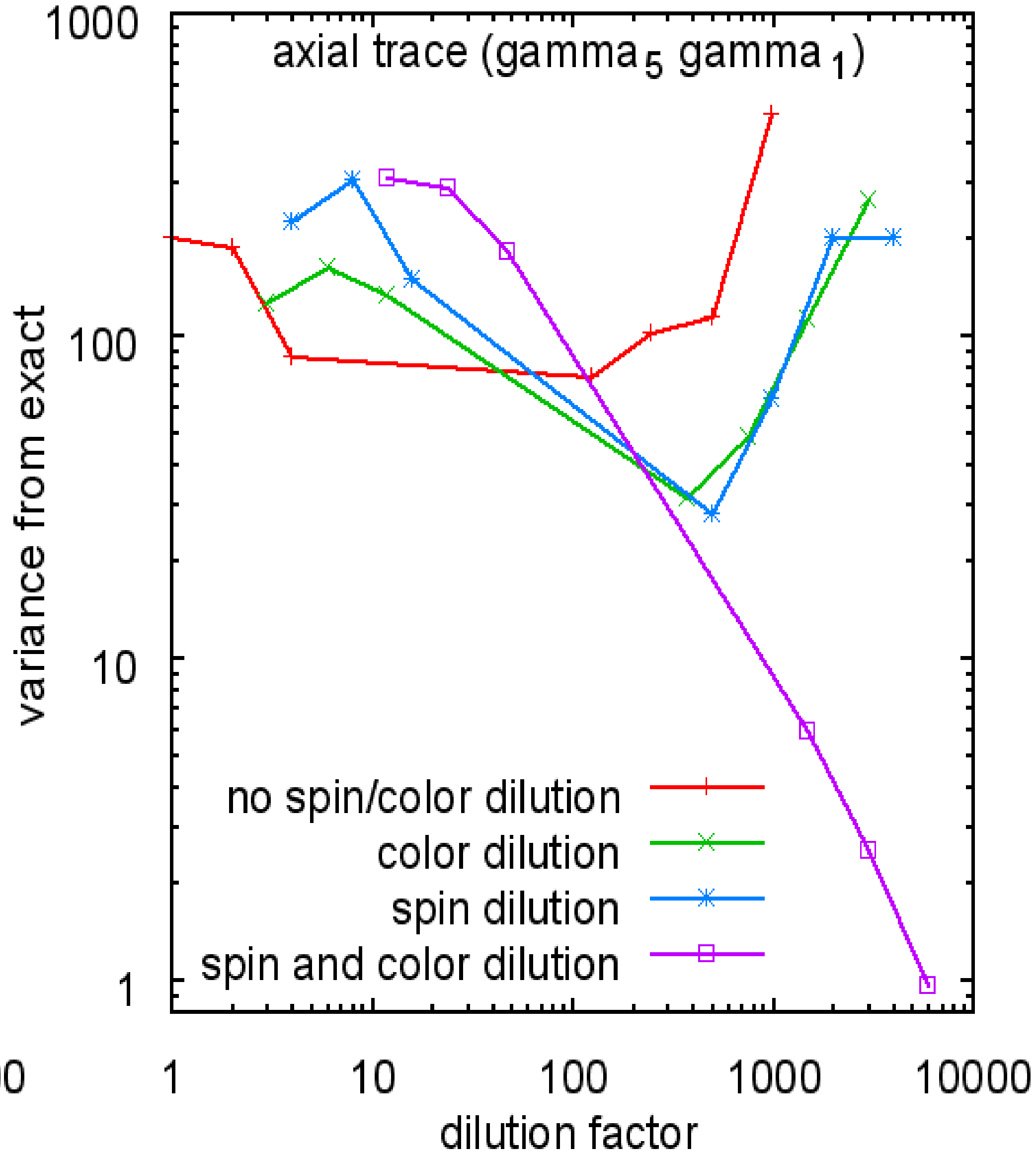
- compare dilution (spin,color,spatial) to exact trace
- all points at fixed amount of work
- exact trace: dilution factor = 12,000
- spatial dilution generally helps
- need spin/color dilution for >1000 sources



Dilution



vector trace (gamma 1)
variance from exact
dilution factor
10000
1000
100
10
1
1
10
100
1000
10000
no spin/color dilution
color dilution
spin dilution
spin and color dilution



axial trace (gamma 5 gamma 1)
variance from exact
dilution factor
1000
100
10
1
1
10
100
1000
10000
no spin/color dilution
color dilution
spin dilution
spin and color dilution

Subtraction

- unbiased subtraction

$$\mathrm{Tr}\left(\Gamma D^{-1}\right) = \left\langle \eta^* \Gamma \left(D^{-1} - S_{tr}\right) \eta \right\rangle_\eta$$

$$S_{tr} = S - \Gamma^{-1} \frac{1}{N_{max}} \mathrm{Tr}\, \Gamma S$$

- want to reduce variance (off diagonal elements) without changing trace

Subtraction

- hopping parameter expansion (Wilson)

$$D = (1 - \kappa M)/2\kappa$$

$$S = 2\kappa(1 + \kappa M + \kappa^2 M^2 + \ldots)$$

  – cheap
  – works for small kappa

- eigenvalue projection

$$S = P_{ev} D^{-1} P_{ev}$$

  – setup requires finding low eigenvectors
  – works well especially for small masses

Multigrid application

James Brannick

Richard Brower

Mike Clark

James Osborn

Claudio Rebbi

...

in collaboration with D. Keyes & TOPS

(see Brannick's talk)

Subtraction

- multigrid subtraction

$$S = V - cycle = T_f + R_{fc} D_c^{-1} R_{cf}$$

$$\mathrm{Tr}\, \Gamma D^{-1} = \langle \eta^* \Gamma (D^{-1} - S) \eta \rangle_\eta + \mathrm{Tr}\, \Gamma S$$

$$\begin{aligned} \mathrm{Tr}\, \Gamma S &= \mathrm{Tr}\, \Gamma T_f + \mathrm{Tr}\, \Gamma R_{fc} D_c^{-1} R_{cf} \\ &= \mathrm{Tr}\, \Gamma T_f + \mathrm{Tr}\, R_{cf} \Gamma R_{fc} D_c^{-1} \end{aligned}$$

- can afford setup costs

Multilevel subtraction

- (geometric) adaptive smooth aggregation
- 4d Wilson fermions, 8^4 lattice, 4^4 blocks

Γ	Tr ΓD^{-1}	Tr $\Gamma(D^{-1}-S)$	Tr ΓS	var ΓD^{-1}	var $\Gamma(D^{-1}-S)$	var ΓS
1	1503	18	1485	1213	121	1187
γ_5	0.00684	0.00033	0.00651	1213	121	1187

subtracted variance reduced ~10x

variance shifted to coarse operator

- 2d Schwinger, 32^2 lattice, 4x4 blocks
  variance reduced ~100x

Summary

- trace estimation appears in many areas of lattice QCD
- usually either very easy (condensate) or very hard (most other cases) and usually ignored
- need better methods