Dynamical Fermion Algorithms using Multi-Step Stochastic Correction

Enno E. Scholz

BROOKHAVEN NATIONAL LABORATORY

Fourth International Workshop on Numerical Analysis and Lattice QCD

Yale University, New Haven, Ct
May 1, 2007

"realistic" QCD-simulations: 2 light quarks plus single strange quark

- efficient in light regime — tunable to simulated quark mass
- capable of odd flavor number $\longrightarrow$ approximation needed

Outline

- stochastic correction/reweighting in multi-boson algorithms
- polynomial approximation
- HMC, PHMC, RHMC
- applications of PHMC:
  - $*$ twisted mass with $N_f = 2 + 1 + 1$
  - $*$ "$N_f = 1$ QCD"

Multi-Boson Algorithms

- algorithm based on LÜSCHER's Multi-Boson representation [LÜSCHER, 1994]
  - $\det(Q^\dagger Q)^\alpha = \int \mathcal{D}[\phi^\dagger, \phi] \exp\left(- \sum_{xy} \phi^\dagger_y \left[Q^\dagger Q\right]^{-\alpha}_{yx} \phi_x \right)$
  - use polynomial approximation of
    $$P_1(x) \simeq x^{-\alpha}$$
    of order n_1 in intervall $[\epsilon, \lambda]$, such that EVs of $Q^\dagger Q$ are covered
    $$\det(Q^\dagger Q)^\alpha = \int \mathcal{D}[\phi^\dagger, \phi] \exp\left(- \sum_{xy} \sum_{i=1}^{n_1} \phi^{(i)\dagger}_y \left[(\gamma_5 Q - \rho_i^*)(\gamma_5 Q - \rho_i) \right] \phi^{(i)}_x \right)$$
  - ρ_i: "roots" of the polynom: $P_1(x^2) = c_0 \prod_{i=1}^{n_1} (x^2 - z_i) = r_0 \prod_{i=1}^{n_1} (x - \rho_i^*)(x - \rho_i)$
  - apply standard local/global heatbath, over-relaxation algorithms
- able to deal with odd number of flavors ($N_f = 1 \;\rightarrow\; \alpha = {}^1/_2, \ldots$, possible caveat: sign-problem)
- smaller quark masses $\Rightarrow$ lower EVs $\Rightarrow$ larger polynomial orders n_1 required
- large order n_1 forbids direct application (memory, long autocorrelation,. . .)
- use (multiple) stochastic correction steps

Stochastic correction step (Noisy correction)

lower n_1 (less accurate) and perform stochastic correction using P_2 of sufficient higher order n_2

$$P_1(x)P_2(x) \simeq x^{-\alpha}$$

Two-Step Multi-Boson (TSMB) algorithm [Montvay, 1995]

- first step (local updates) $[U] \to [U']$ produce configuration according to $\det(P_1(Q^2))$
- stochastic correction step
  - generate Gaussian random vector η with distribution

$$\frac{e^{-\eta^\dagger P_2(Q[U]^2)\eta}}{\int d[\eta] e^{-\eta^\dagger P_2(Q[U]^2)\eta}}$$

    use $\eta = P_2(Q[U]^2)^{-\frac{1}{2}}\eta' = \bar{P}_2(Q[U]^2)\eta'$ from simple Gaussian distributed η'

  - accept change $[U] \to [U']$ with probability

$$\min\left\{ 1,\ \exp\left[\eta^\dagger\Big(-\ P_2(Q[U']^2)\ +\ P_2(Q[U]^2)\Big)\eta\right]\right\}$$

- final distribution according to $\det(P_1(Q^2)\, P_2(Q^2))$

Reweighting

- check quality of approximation
- few low eigenvalues (close or below ϵ)
- sign problem ($R = -1$)

Reweighting factor:

$$R = \prod_i P_1(\lambda_i) P_2(\lambda_i) \lambda_i^{\alpha}, \quad \lambda_i\text{: Eigenvalue of } Q^2 \quad \text{(exact)}$$

or

$$R = \frac{1}{\det\Big(P_3(Q^2)\Big)} \quad \text{(stochastic estimate)}$$

or combine both

$$\langle A \rangle \;=\; \frac{\langle R[U] A[U] \rangle}{\langle R[U] \rangle}$$

Applications of TSMB

- $N_f = 2$ Wilson-fermions, lowest $m_q \simeq \frac{1}{4} m_s$, $a \simeq 0.20\text{fm}$, $L^3 \times T = 16^3 \times 32$ [Farchioni,. . . , EES,. . .]

- $N_f = 2$ twisted mass Wilson-fermions, exploratory studies [Farchioni,. . . , EES,. . .]

  HMC-algorithm (mass-preconditioning, multiple step-size) turned out to be more efficient

- SUSY Yang-Mills studies, e.g. Peetz, Farchioni, 2005

Multi-Step correction in MB-algorithms

MONTVAY, EES Phys.Lett. **B 623** (2005) 73

generalize TSMB to Multi-Step Multi-Boson (MSMB) algorithm:

$$P_1(x) P_2(x) \cdots P_k(x) \simeq x^{-\alpha},$$

increasing order $n_i > n_{i-1}$, $x \in [\epsilon_k, \lambda]$, $\epsilon_i \leq \epsilon_{i-1}$

$$P_i \simeq \left[x^{\alpha} P_1(x) \cdots P_{i-1}(x) \right]^{-1}, \quad i \in 2, 3, \ldots, k$$

- P_1: local updates
- $P_2, \ldots, P_k$ *nested* noisy corrections (Again: use $\bar{P}_i(x) \simeq P_i(x)^{-\frac{1}{2}}, x \in [\bar{\epsilon}_i, \lambda], \bar{\epsilon}_i \leq \epsilon_i$.)

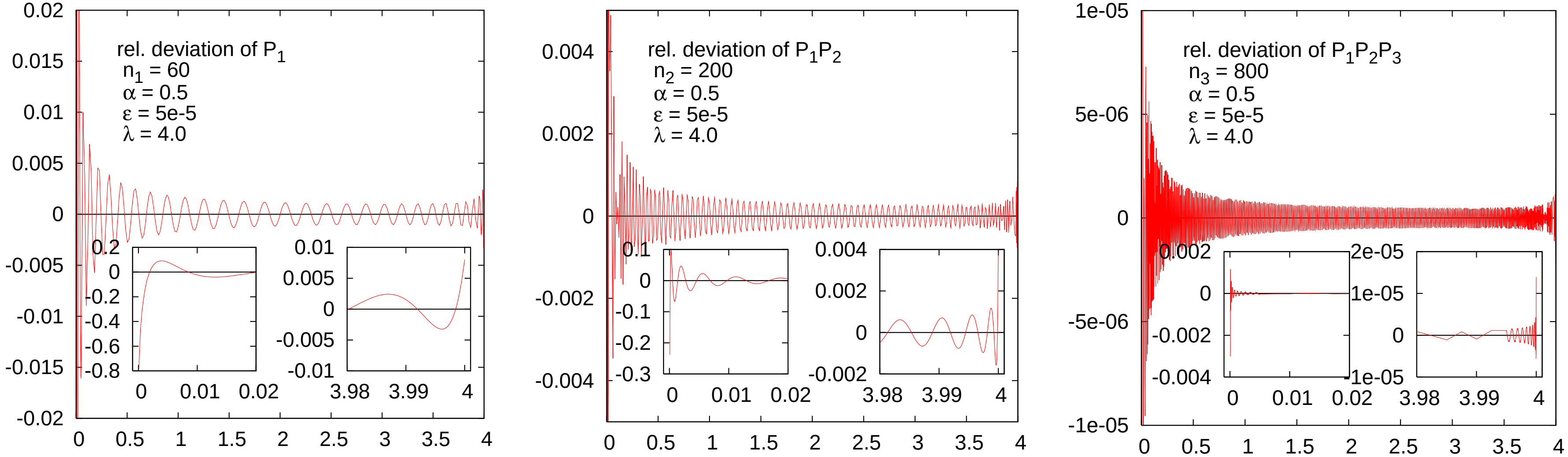
$$\text{start} \; \underbrace{\underbrace{\underbrace{\rightarrow \cdots \rightarrow}_{N_1 \text{steps} P_1} P_2 \longrightarrow \cdots \longrightarrow \underbrace{\rightarrow \cdots \rightarrow}_{N_1 \text{steps} P_1} P_2}_{N_2 \text{steps}} \longrightarrow P_3 \longrightarrow \underbrace{\cdots\cdots\cdots}_{(N_1\, N_2) \times P_1, N_2 \times P_2} \longrightarrow P_3}_{N_3 \text{steps}}$$

- if accept-reject with P_i negative: reset to last accepted with P_i

Polynomial approximation

- L_2-norm optimized turned out to be more favorable than L_∞ [Montvay, 1999]
- polynoms $P_2, \cdots, P_k$ in recursive scheme, calculating coefficients
  - ∗ arbitrary precision arithmetics [Gebert and Montvay, 2003]
  - ∗ discretization of approximation interval [Katz and Tóth, 2004]
  - ∗ allows for adaptive precision in P_k
- example from $L^3 \times T = 16^3 \times 32$, $\beta = 0.74$(DBW2), $a \simeq 0.13$fm, $N_f = 2$, $am_q \simeq 0.024$:

$$n_B = 2, \quad \epsilon = 2 \cdot 10^{-5}, \quad \lambda = 4.0, \quad n_1 = 60, \quad n_2 = 200, \quad \bar{n}_2 = 300, \quad n_3 = 800, \quad \bar{n}_3 = 900$$



Cost-formula

the cost in *Matrix-Vector Multiplications (MVM)* per cycle

Wilson-fermions: $1\text{MVM} \simeq 1344\Omega$flop

$$N_{\text{MVM}}/\text{cycle} = 6(n_B n_1 N_\phi + N_U) + F_G I_G + \sum_{i=2}^{k} 2n_B(n_k + \bar{n}_k)N_{Ck}$$

N_ϕ number of local fermion updates, N_U number of local gauge updates,
F_G, I_G frequency of global fermion updates and avg. number of MVM
N_{Ck} number of correction steps involving P_k

in $L^3 \times T = 16^3 \times 32$, $\beta = 0.74$(DBW2), $a \simeq 0.13$fm, $N_f = 2$, $am_q \simeq 0.024$

2-step: $n_B = 4$, $n_1 = 34$, $n_2 = 720$, $\bar{n}_2 = 740$, $N_{C2} = 1$,
3-step: $n_B = 2$, $n_1 = 60$, $n_2 = 200$, $\bar{n}_2 = 300$, $N_{C2} = 10$, $n_3 = 800$, $\bar{n}_3 = 900$, $N_{C3} = 1$

compare 10 cycles of 2-step with 1 of 3-step ($\approx$ same number of local updates)

$$10 \times 11{,}680 = \mathbf{116{,}800} \quad \text{vs.} \quad 10 \times 2{,}000 + 6{,}800 = \mathbf{26{,}800}$$

(plaquette-autocorr. (in cycles): $\tau_{\text{int}} \simeq 50$ vs. $\tau_{\text{int}} \simeq 10$) total gain ≈ 1.5
(but well-tuned 2-step, not too much tuning effort in 3-step)

adaptive precision

- recursive formula for "correction"-polynomials (based on expansion in orthogonal polynomials) [GEBERT, MONTVAY, 2003]

$$
\begin{aligned}
P_k^{(n_k)}(x) &= \sum_{\nu=0}^{n_k} d_\nu \Phi_\nu(x) \\
\Phi_0(x) &= 1, \quad \Phi_1(x) = x - \frac{s_1}{s_0}, \\
\Phi_{\mu+1}(x) &= (x + \beta_\mu)\Phi_\mu(x) + \gamma_{\mu-1}\Phi_{\mu-1}(x), \quad \mu = 1, 2, \ldots
\end{aligned}
$$

- coefficients independent of polynomial order n_k
- last correction-step: stop at n_i if $\Delta_i < \epsilon_{\text{precision}}$

$$
\Delta_i = \left| \sum_{\nu=0}^{i-1} d_\nu \Phi_\nu(x) - \sum_{\nu=0}^{i} d_\nu \Phi_\nu(x) \right|
$$

- allows using a very conservative correction polynomial without "wasting" too much time
- only applicable in last step (otherwise following approximations ill-defined!)

Further improvements

- Even-Odd preconditioning of the fermion matrix
- Determinant breakup [HASENBUSCH, 1999]

  n-root-trick of CLARK, KENNEDY

  * instead of using 1 polynom with exponent α, use n_B polynoms with exponents α/n_B
  * better stochastic sampling in noisy correction step

- Mass preconditioning
  * similar to HASENBUSCH-trick in HMC-like algorithms [HASENBUSCH, 2001]
  * add a "mass-shift" $\mu_i > 0$ to each polynomial except the *final* one ($\mu_k = 0$)

$$
\begin{aligned}
P_1(x) &\simeq (x+\mu_1)^{-\alpha} \\
P_i(x) &\simeq \Big[(x+\mu_i)^{\alpha} P_1(x)\cdots P_{i-1}(x)\Big]^{-1}, \quad i \in 2,\ldots,k-1 \\
P_k(x) &\simeq \Big[x^{\alpha} P_1(x)\cdots P_{k-1}(x)\Big]^{-1}
\end{aligned}
$$

  * allows for lower n_i,
  * acceptance remains sufficiently high, if μ_i/μ_{i-1} not much smaller than 1

Hybrid Monte-Carlo algorithms and variants

- HMC and stochastic correction

$$\mathcal{H} \;=\; \frac{1}{2}P^2 \;+\; S_g[U] \;+\; (\phi^\dagger \tilde{Q}^2 \phi)$$

update-step with $\delta\tau$: using preferred integration scheme

$$\begin{aligned} P \quad &\rightarrow \quad P' \;=\; P \;-\; \delta\tau\, DS[U] \\ U \quad &\rightarrow \quad U' \;=\; \exp(i\delta\tau \sum_j \lambda_j P_j)\, U \end{aligned}$$

∗ replace $\tilde{Q}^2$ in HMC-update with mass-preconditioned HMC-update $\tilde{Q}^2 + \mu_1$ $(\mu_1 > 0)$
∗ use (multiple) stochastic correction to obtain correct $(\mu_k = 0)$ determinant

- HMC only suitable for N_f even $(\tilde{Q}^2, \tilde{Q}^4, \ldots)$

- Polynomial- or Rational-HMC for odd N_f
  - again use polynomial representation in first step [Frezzotti, Jansen, 1997-1999]

$$S_f \;=\; \phi^\dagger \, c_0 \prod_{i=1}^{n_1} (\tilde{Q} - \rho_i) \prod_{i=n_1}^{1} (\tilde{Q} - \rho_i^*) \, \phi$$

  - direct application requires high n_1: costly and rounding errors
  - again: use stochastic correction to control cost (next slides)

OR

  - use rational approximation [Clark, Kennedy, 2004, 2006]

$$S_f \;=\; \phi^\dagger \left(\alpha_0 + \sum_{k=1}^{p} \frac{\alpha_k}{\tilde{Q}^2 + \beta_k} \right) \phi$$

  - rational approx. requires lower orders

BUT

  - total numerical cost not directly given by order, matrix-inversion!!

  - cost can be controlled by some tuning effort ("light poles", "heavy poles" if L_∞ is applied)
  - sum is more convenient than product, when calculating derivative

PHMC with stochastic correction — some technical details

- use (low order) polynomial P_1 in PHMC-step (root-representation)
- calculate derivative recursively

$$\begin{aligned}
\phi_1^{(k)} &= \sqrt{c_0}\phi(\tilde{Q}-\rho_1)\cdots(\tilde{Q}-\rho_k),\ k=1,\ldots,n-1 \\
\phi_2^{(k)} &= \sqrt{c_0}\phi(\tilde{Q}-\rho_1)\cdots(\tilde{Q}-\rho_n)(\tilde{Q}-\rho_n^*)\cdots(\tilde{Q}-\rho_{k+2}^*) \\
\phi\, D\, P_1(\tilde{Q}^2)\, \phi^\dagger &= 2\mathrm{Re}\left(\sum_{k=0}^{n-1}\phi_1^{(k)}(D\tilde{Q})\phi_2^{(k)\dagger}\right)
\end{aligned}$$

- $\bar{P}_1 \simeq P_1^{-\frac{1}{2}}$ to generate boson-fields from Gaussian distribution
- $P_2, \bar{P}_2, \cdots, P_k, \bar{P}_k$ in correction-steps
- Sexton-Weingarten integration scheme with multiple time-steps
- determinant break-up, preconditioning

Application I: twisted mass with split-doublet

CHIARAPPA, ..., EES, ..., Eur. Phys. J. **C 50** (2007) 373; hep-lat/0606011

- exploratory studies of $N_f = 2$ TM-fermions revealed interesting phase-structure of Wilson-fermions
- automatic $\mathcal{O}(a)$-improvement
- first results large study of $N_f = 2$ using HMC presented by ETMC [BOUCAUD et al., hep-lat/0701012]
- state-of-the-art of "realistic" QCD: include strange quark

  * use untwisted strange quark $\longrightarrow$ no $\mathcal{O}(a)$-improvement
  * use 2nd doublet unsplit $\longrightarrow$ only quenched QCD possible

- use mass-splitting in 2nd doublet [FREZZOTTI, ROSSI, 2004]

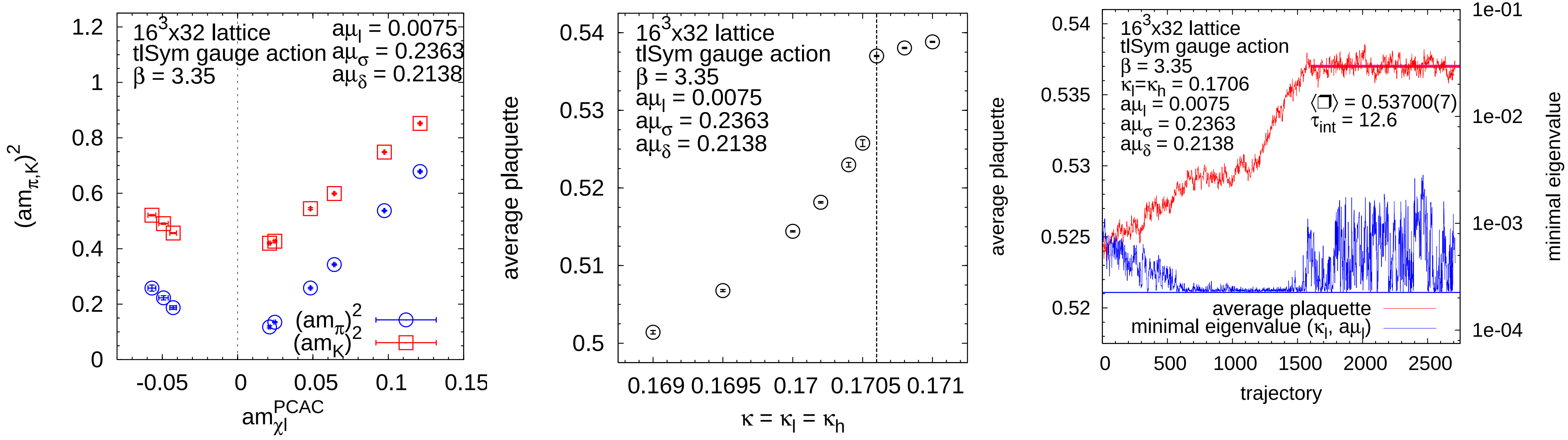
$$Q^{\mathrm{TM}}_{x,y} = \delta_{x,y}\Big[\mu_\kappa + i\gamma_5\tau_1 a\mu_\sigma + \tau_3 a\mu_\delta\Big] - \frac{1}{2}\sum_{\mu=\pm 1}^{4}\delta_{x,y+\hat{\mu}}(\gamma_\mu + r)U_{y\mu}$$

  * lower bound on EV, automatic $\mathcal{O}(a)$ improvement
  * including strange-quark without sign-problem ($N_f = 2 + (1+1)$; (u,d)+(c,s))
  * $\mu_\delta \neq 0$: use PHMC (HMC not applicable)

Dynamical Simulations

- tree-level Symanzik gauge action, two lattice spacings (fixed physical volume: $aL \simeq 2.4$fm)
  - $a \simeq 0.20$fm ($\beta = 3.25$, $L^3 \times T = 12^3 \times 24$)
    $a\mu_l = 0.01$, $a\mu_\sigma = 0.315$, $a\mu_\delta = 0.285$, $\kappa \in [0.1740, 0.1755]$ (7 values, 10 runs)
  - $a \simeq 0.15$fm ($\beta = 3.35$, $L^3 \times T = 16^3 \times 32$)
    $a\mu_l = 0.0075$, $a\mu_\sigma = 0.2363$, $a\mu_\delta = 0.2138$, $\kappa \in [0.1690, 0.1710]$ (9 values)
- varied $\kappa (= \kappa_l = \kappa_h)$ to find κ_{crit}, explore phase-structure
- lattice-spacing, light-doublet similar to previous studies (DBW2 and pure Wilson gauge action)
- performed at ZAM Jülich (IBM-p690) at NIC Jülich and PC-Cluster at Hamburg

$a \approx 0.15\mathrm{fm},\ L^3 \times T = 16^3 \times 32$



- minimal $m_\pi \simeq 450\text{MeV}$
- minimal $m_K \simeq 850\text{MeV}$
- sharp rise in $\langle \square \rangle$
  - **?** weak *first order phase transition*

    or
  - **?** *cross-over*
  - $*$ not distinguishable in finite volume

- not an algorithmic imperfection
  - $*$ transition low $\longrightarrow$ high plaquette phase
    positive $\longrightarrow$ negative quark mass phase
  - $*$ "crossing near origin"
  - $*$ lowest EV fluctuating in high plaquette phase

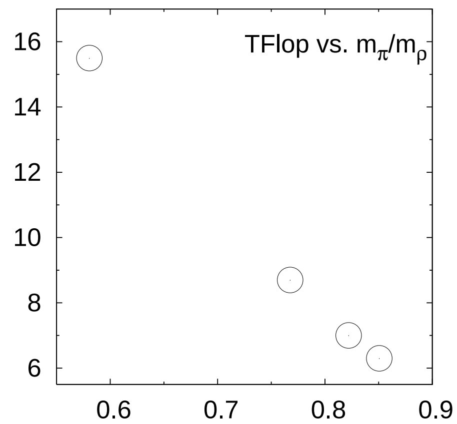
cost of the PHMC-algorithm

- 1 stochastic correction step
- Sexton-Weingarten-int. with multiple time-steps, $\delta\tau = 0.4$, $N_T = 2$, $N_G = 5$, $N_Q = 2$, $n_B = 2$
  (Omelyan-integrator with $\lambda_{\text{Omelyan}} = {}^1/_6$)

$$N_{\text{MVM}} \approx \sum_{\text{doublets}} \left\{ 2n_B(n_2 + \bar{n}_2) \;+\; N_T\Big[2n_B(n_1+\bar{n}_1) + n_B(3+2N_Q)(4n_1 - 1)\Big]\right\}$$

$L^3 \times T = 16^3 \times 32$

m_π/m_ρ	$m_{\pi,K}$/MeV	n_1	$\bar{n}_1$	n_2	$\bar{n}_2$	cost/Tflop
0.85	1080	70	110	120	160	
	1210	50	80	90	130	6.4τ
0.58	450	220	320	800	930	
	852	60	100	120	160	15.6τ



$*$ $\tau \simeq \mathcal{O}(1)$ (depends on observable)

$*$ larger step-size may improve cost/autocorrelation (alpha-Collaboration, Meyer et al., 2006)

$*$ may use mixed HMC (unsplit) and PHMC (split-doublet) ("best of both worlds")

Application II: "$N_f = 1$ QCD"

- polynomial approximation allows to study $N_f = 1$
- no spontaneous chiral symmetry breaking: $\mathrm{U}(1)$-anomaly
- applicability of partially quenched chiral perturbation theory: $\mathrm{SU}(3|2)$ (one sea, two valence quarks)?
- $N_f = 1$ results helpful to resolve fourth root (staggered) issue
- study relation to $\mathcal{N} = 1$ SYM *(orientifold planar equivalence)*

in collaboration with (publication in preparation):

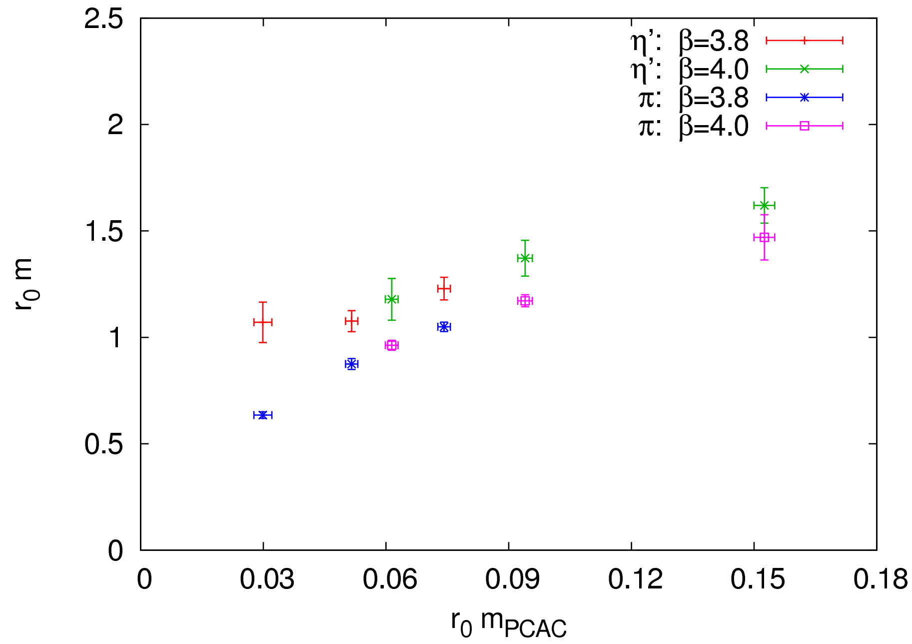
ECT*
Trento

institut für theoretische physik
Universität Münster

DESY
Hamburg

- tree-level Symanzik gauge action with Wilson fermions
- $L^3 \times T = 12^3 \times 24$ ($\beta = 3.8$, $a \simeq 0.19$fm) and $16^3 \times 32$ ($\beta = 4.0$, $a \simeq 0.13$fm)
- PHMC with one correction step + reweighting/negative sign
- $\delta\tau$: 1.5 – 1.8 (3–6 trajectories)
- polynomials at lightest masses ($m_\pi \approx 240\text{MeV}, 400\text{MeV}$):

$L^3 \times T$	$[\epsilon, \lambda]$	n_1	$\bar{n}_1$	n_2	$\bar{n}_2$
$12^3 \times 24$	$[3.25 \cdot 10^{-6}, 2.6]$	350	550	1400	1600
$16^3 \times 32$	$[1.20 \cdot 10^{-5}, 2.4]$	250	370	1000	1150



Summary

- multiple stochastic correction allows for good cost control
  can be implemented with any basic updating algorithm, we used
  - Multi-Boson ($\longrightarrow$ TSMB, MSMB)
  - Polynomial-HMC (direct cost control compared to Rational-HMC)

- polynomial formulation allows to be flexible in flavor-number

- mass-preconditioning (Hasenbusch-trick) easy to implement

- recent applications of PHMC with stochastic correction:
  - exploratory study of $N_f = 2 + 1 + 1$ twisted mass: feasible approach
  - "$N_f = 1$ QCD" vs. SUSY Yang-Mills . . . (work in progress)