Polynomial approximations for $1+1$ dynamical quarks

T. Chiarappa and R. Frezzotti

November 2005

Following our Lat-05 proceeding contribution, hep-lat/0509154, which will be referred to as CFU (Chiarappa-Frezzotti-Urbach), the Hamiltonian for Molecular Dynamics (MD) is:

$$H_{2+1+1} = \frac{1}{2}\Pi\cdot\Pi \;+\; S_G \;+\; \Phi_h^\dagger P(\hat{S})\Phi \;+ \\ +\; \phi_1^\dagger\left[\hat{Q}_l'\left(\hat{Q}_l'\right)^\dagger\right]^{-1}\phi_1 \;+\; \phi_2^\dagger\left\{\hat{Q}_l'\left[\hat{Q}_l''\left(\hat{Q}_l''\right)^\dagger\right]^{-1}\left(\hat{Q}_l'\right)^\dagger\right\}\phi_2\,, \qquad (1)$$

where $\hat{S}=\hat{Q}_h\hat{Q}_h^\dagger$ and $\hat{}$ means that the EO-preconditioned form of the 1-flavour operators, $Q_l', Q_l'^\dagger$ and $Q_l'', Q_l''^\dagger$, and the 2-flavour operator $Q_h, Q_h^\dagger$ is taken (see eqs. 2.4, 2.5 and 2.2 in CFU). Here

$$P = P_{n,s_L}(\hat{S}) = (\sqrt{\hat{S}})^{-1}\left[1+R_{n,s_L}(\hat{S})\right]\,,$$

see eq. 3.1 in CFU, is the polynomial that we plan to use in the MD. Its roots come in complex conjugate pairs [1], thus we can write:

$$P_{n,s_L}(\hat{S}) = B(\hat{S})B(\hat{S})^\dagger\,.$$

Note that the relative fit error $R_{n,s_L}(s)=\sqrt{s}P_{n,s_L}(s)$ is *not* a polynomial in s.

Given that, one then wishes to perform the Φ_h-heatbath and the A/R Metropolis test (at the beginning and at the end of each MD-trajectory, respectively) with a controlled approximation error, which can be removed (or made negligibly small) by reweighting. This was only sketched in CFU (Sect. 3), and is discussed in detail here.

Let us introduce, as in eq. 3.2 of CFU, a second polynomial $\tilde{P}$,

$$\tilde{P}_{\tilde{n},s_L}(\hat{S}) = \left[\sqrt{\hat{S}}\;P_{n,s_L}(\hat{S})\right]^{-1}\;\left[1+\tilde{R}_{n,s_L}(\hat{S})\right] \qquad (2)$$

such that $|\tilde{R}_{\tilde{n},s_L}(s)|\ll 1$ for $s_L\le s\le 1$. We have in mind

$$\tilde{\delta} = \max_s\left|\tilde{R}_{\tilde{n},s_L}(s)\right| \ll \delta = \max_s|R_{n,s_L}(s)| < 1\,.$$

[1] Here we assume $P_{n,s_L}(s)$ never vanishes for s real. This is a plausible assumption, which however needs be proved or checked (work in progress).

For instance one might choose $\delta \sim 3 \cdot 10^{-2}$ and $\tilde{\delta} \sim 3 \cdot 10^{-4}$, or so ...

The (2-flavoured) pseudofermion Φ_h is obtained through:

$$\Phi_h = \tilde{P}\ B^\dagger\ \hat{Q}_h\ r_h \quad , \quad r_h = \text{Gaussian random vector} \tag{3}$$

and is thus distributed according to

$$p(\Phi) = c \cdot \exp\left\{ - \left| \left(\tilde{P} B^\dagger \hat{Q}_h \right)^{-1} \Phi_h \right|^2 \right\} \quad , \quad c = \text{constant} \tag{4}$$

with, after some trivial algebra and sums over spacetime and internal indices understood,

$$\left| \left(\tilde{P} B^\dagger \hat{Q}_h \right)^{-1} \Phi_h \ \right|^2 = \Phi_h^\dagger \left[1 + \tilde{R} \right]^{-1} P \ \left[1 + \tilde{R} \right]^{-1} \Phi_h \tag{5}$$

Note that this is an exact equation, involving no approximations.

Our idea is to perform the A/R Metropolis test with respect to the Hamiltonian

$$\tilde{H}_{2+1+1} = \ \ldots\ + \Phi_h^\dagger \left[1 + \tilde{R} \right]^{-1} P \ \left[1 + \tilde{R} \right]^{-1} \Phi_h \tag{6}$$

which is the same as eq. (1), but for the Φ_h sector.

The reason is that for $\tilde{H}_{2+1+1}$ one can compute $\Delta \tilde{H}_{2+1+1} = \tilde{H}^{end}_{2+1+1} - \tilde{H}^{start}_{2+1+1}$, at the end and start points of each MD-trajectory, with very good precision, as explained below. Moreover for $\tilde{\delta}$ sufficiently small the acceptance should not be significantly affected by $\tilde{H}_{2+1+1} \neq H_{2+1+1}$, where H_{2+1+1} drives the MD. In view of eq. (5), we expect that having $\tilde{\delta} = \max_s \left| \tilde{R}_{\tilde{n}, s_L}(s) \right| \simeq 10^{-3} \leftrightarrow 10^{-4}$ can be sufficient.

The values of $\tilde{H}_{2+1+1}$ at the start and end of each MD-trajectory can be computed through

$$\tilde{H}^{start}_{2+1+1} = \left| \left(\tilde{P} B^\dagger \hat{Q}_h \right)^{-1} \Big|_{U_{start}} \Phi_h \ \right|^2 = r_h^\dagger r_h \tag{7}$$

$$\tilde{H}^{end}_{2+1+1} = \left| \left(\tilde{P} B^\dagger \hat{Q}_h \right)^{-1} \Big|_{U_{end}} \Phi_h \ \right|^2 = \Phi_h^\dagger B^\dagger B\ \chi_h \tag{8}$$

where

$$\chi_h = \left(1 + \tilde{R} \right)^{-2} \Phi_h = \left(\tilde{P} P \hat{S} P \tilde{P} \right)^{-1} \Phi_h$$

is given by

$$\chi_h = \left(1 - A + A^2 - \ldots \right) \Phi_h \quad , \quad A = \tilde{P} P \hat{S} P \tilde{P} - 1 \tag{9}$$

and the truncation of the first expression in eq. (9) is justified by $|A| \simeq 2\tilde{\delta} \ll 1$.

For instance, using χ_h truncated to the second order, $\chi_h^{(2)} = (1 - A + A^2)\Phi_h$ should lead to an error $\epsilon\left(\tilde{H}^{end}_{2+1+1}\right) \simeq c\sqrt{N^h_{dof}}\tilde{\delta}^3$, with c to be found by numerical experiment (expect small as the series given in eq. (9) has alternating signs), $N^h_{dof} = \left(\frac{L}{a}\right)^3 \frac{T}{a} \cdot 2 \cdot 3 \cdot 4 = 24 N_{sites}$ can be $10^5 \leftrightarrow 10^8$, so one should have $\tilde{\delta} \sim 10^{-4}$ to keep the error $\epsilon\left(\tilde{H}^{end}_{2+1+1}\right) \leq 10^{-8}$.

Note: if one insists in performing the A/R Metropolis test with respect to H_{2+1+1} in eq. (1), the Φ-heatbath requires $\Phi_h = B^{-1}(\hat{S})r_h$ which is computationally difficult, since one has to solve $\sqrt{\hat{S}}B^\dagger B\Phi_h = \sqrt{\hat{S}}B^\dagger r_h$.

When the (P)HMC algorithm with the A/R Metropolis test based on $\tilde{H}_{2+1+1}$ (and MD based onH_{2+1+1}) is employed, it delivers -after thermalisation- a sample of gauge configurations (U) that will be distributed according to

$$\mu^{2+1+1}_{PHMC}(U) \propto e^{-S_G}\Big|_U \ \det\left[\left(\hat{Q}''_l(\hat{Q}''_l)^\dagger\right)\Big|_U\right] \ \det\left[\left((1+\tilde{R})\ P^{-1}\ (1+\tilde{R})\right)\Big|_U\right] \tag{10}$$

while we would like to have a distribution probability

$$\mu^{2+1+1}_{PHMC}(U) \propto e^{-S_G}\Big|_U \ \det\left[\left(\hat{Q}''_l(\hat{Q}''_l)^\dagger\right)\Big|_U\right] \ \det\left[\left(\sqrt{\hat{S}}\right)\Big|_U\right] \tag{11}$$

A reweighting procedure cab be hence employed to get a stochastic estimate of the correction weight

$$W_h[U] = \det\left[\left(\sqrt{\hat{S}}\ P(\hat{S})\ (1+\tilde{R})^{-2}\right)\Big|_U\right] \tag{12}$$

which can also be written in the computationally suggestive form

$$W_h[U] = \left\{\det\left[\tilde{P}(\hat{S})\Big|_U\right]\right\}^{-1}\left\{\det\left[1+\tilde{R}(\hat{S})\Big|_U\right]\right\}^{-1} \tag{13}$$

The two factors in $W_h[U]$ in eq. (13) can be estimated either separately or together, see below. The separate estimate is of particular interest in those cases where one expects $W_h[U]$ to be almost constant as a functional of U, e.g: for sufficiently small δ (thus $\sqrt{\hat{S}}P$ close to 1) and $\tilde{\delta}$.

1. Separate estimates of the two factors in $W_h[U]$:

$$\left(\det\left[\tilde{P}(\hat{S})\Big|_U\right]\right)^{-1} = \int D\eta_1\ e^{-\eta_1^\dagger\eta_1}\ \exp\left\{\eta_1^\dagger\left(1-\tilde{P}(\hat{S})\Big|_U\right)\eta_1\right\} \tag{14}$$

$$\left(\det\left[(1+\tilde{R}(\hat{S}))\Big|_U\right]\right)^{-1} = \int D\eta_2\ e^{-\eta_2^\dagger\eta_2}\ \exp\left\{\eta_2^\dagger\left(1-\bar{P}(\hat{S})\Big|_U\right)\eta_2\right\} \tag{15}$$

with $\bar{P}$ a high degree polynomial in $\hat{S}$ approximating $1+\tilde{R}$

$$\bar{P}(\hat{S}) = \left(1+\tilde{R}(\hat{S})\right)\ \left(1+\bar{R}(\hat{S})\right) \tag{16}$$

and $\bar{\delta} = \max_s \left|\bar{R}(s)\right|$ very small, say $\bar{\delta} \leq 10^{-6}$, or so.
The idea here is that the evaluation of eq. (14) may be sufficient in most cases, and one can always check *a posteriori* that the correction in eq. (15) does not change the results within the statistical errors (including the η_1-noise due to stochastic reweighting).

2. Another method for estimating $W_h[U]$ is the following:

$$W_h[U] = \int D\eta \; e^{-\eta^\dagger \eta} \; \exp\left\{\eta^\dagger \left(1 - (\tilde{P}\mathcal{P}\tilde{P})\Big|_U\right) \eta\right\} \tag{17}$$

with a fourth high degree polynomial $\mathcal{P}$ such that

$$\mathcal{P}(\hat{S}) = C^\dagger(\hat{S})C(\hat{S}) \equiv \left(\sqrt{\hat{S}}\; P(\hat{S})\right) \; \left(1 + \mathcal{R}(\hat{S})\right) \; , \tag{18}$$

with $\max_s \left|\mathcal{R}(\hat{S})\right|$ so small that $(1 + \mathcal{R}(\hat{S})$ equals 1 on all gauge configurations, up to deviations negligible compared to the statistical errors.

Control of roundoff

1. $\Phi_h = \tilde{P}(\hat{S}) \; B^\dagger(\hat{S}) \; \hat{Q}_h \; r_h = B^\dagger(\hat{S}) \; \tilde{P}(\hat{S}) \; \hat{Q}_h \; r_h$
   - need monomial ordering (bit-reversal ?) for applying $B^\dagger(\hat{S})$
   - can use Clenshaw recursion for application of $\tilde{P}(\hat{S})$
2. Molecular Dynamics with $P = B^\dagger B$
   - need of course monomial ordering (bit-reversal ?) for B and $B^\dagger$
3. $\tilde{H}^{end}_{2+1+1} = \Phi^\dagger_h B^\dagger B \; \chi_h \quad , \quad \chi_h = (1 - A + A^2 - ...)\Phi_h$
   - use Clenshaw recursion for both $B^\dagger B = P$ and $A = \tilde{P}P\hat{S}P\tilde{P}$ (four times: two for P and two for $\tilde{P}$)
4. Reweighting
   - use Clenshaw recursion for $\left\{ \begin{array}{cc} 1 - \tilde{P} \;\; \text{and} \;\; 1 - \bar{P} & \text{eqs. (14,15)} \\ 1 - \tilde{P}\bar{P}\tilde{P} & \text{eq. (17)} \end{array} \right\}$
     some standard tricks can be applied in the evaluation
     $\eta_1^\dagger(1 - \tilde{P})\eta_1 \quad , \quad \eta_2^\dagger(1 - \bar{P})\eta_2 \qquad$ eqs. (14,15)
     $\eta^\dagger(1 - \tilde{P}\mathcal{P}\tilde{P})\eta_1 \quad , \quad \mathcal{P} = C^\dagger C \qquad$ eq. (17)
     which can be discussed further when we arrive at that stage ...