The RHMC Algorithm

ILFTN, Jefferson Laboratory

Michael Clark

Boston Univeristy

(Work done as part of the RBC/UKQCD collaboration)

- Motivation
- Rational Hybrid Monte Carlo
- 2+1 Domain Wall Fermions
- Improvements
- Conclusion

Introduction and Motivation

- Lattice QCD path integral

$$\langle\Omega\rangle = \frac{1}{Z}\int [dU] e^{-S(U)} [\det \mathcal{M}(U)]^{\alpha}\Omega(U)$$

where $\alpha = \frac{N_f}{4}$ $(\frac{N_f}{2})$ for staggered (Wilson) fermions, $\mathcal{M} = M^{\dagger}M$.

$\Rightarrow$ arbitrary N_f with non-integer α

- Conventional HMC fails
- Everyone wants to do 2+1 physics
- Need other algorithms

The R Algorithm (Gottlieb et al)

Rewrite fermionic determinant:

$$\det \mathcal{M}^{\alpha} = \exp\left(\alpha\, \mathrm{tr} \ln \mathcal{M}\right)$$

- Approximate trace by noisy estimator≡ pseudo-fermions
- Use integrator ($O(\delta\tau^2)$)
  - Non-reversible
  - Jacobian $\neq 1$

  ⇒ Cannot include Monte Carlo acceptance test
- R: HMD algorithm which is inexact ⇒ naïvely cheap, but extrapolation to zero stepsize (strictly) required

Polynomial Hybrid Monte Carlo

Write in pseudo-fermion notation

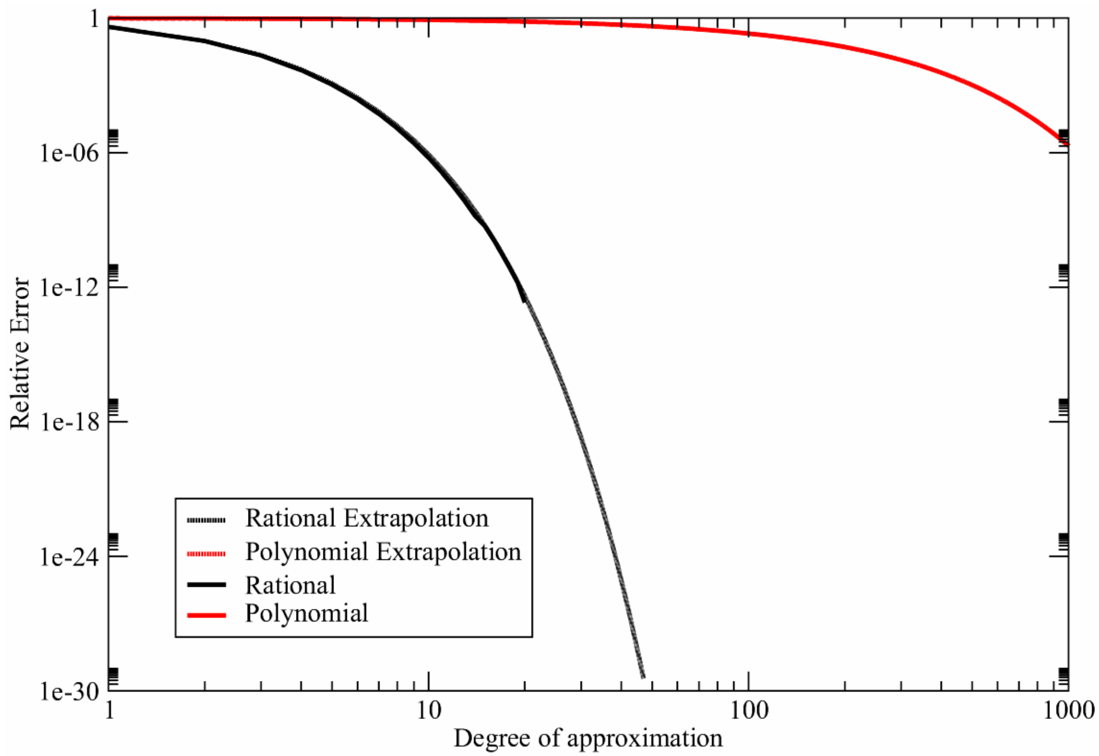
$$\begin{aligned} \det \mathcal{M}^{\alpha} &= \int \mathcal{D}\bar{\psi}\mathcal{D}\psi e^{-\bar{\psi}\mathcal{M}^{-\alpha}\psi} \\ &\approx \int \mathcal{D}\bar{\psi}\mathcal{D}\psi e^{-\bar{\psi}P(\mathcal{M})\psi}, \end{aligned}$$

where $P(\mathcal{M})$ is MiniMax polynomial approximation over spectral range.

- Pseudo-fermion heatbath easily realised since $P(\mathcal{M}) = p^{\dagger}(\mathcal{M})p(\mathcal{M})$
- Use standard MD leapfrog $\Rightarrow$ exact
- Significant Error $|P(\mathcal{M}) - \mathcal{M}^{-\alpha}| >$ CG residual
  - Reweight acceptance test or observable
- High degree polynomial
  - costly in memory (for non-analytic force)
  - potential for rounding errors

Optimal rational approximations

- Very accurate
- Generated using Remez algorithm
- Real non-degenerate roots (poles are always +ve)



- Partial fractions - $r(x) = \sum_{k=1}^{n} \frac{\alpha_k}{x+\beta_k}$
- Evaluate using multi-shift solver

Rational Hybrid Monte Carlo

- Rewrite determinant

$$\det \mathcal{M}^{\alpha} = \int \mathcal{D}\bar{\phi}\mathcal{D}\phi e^{-\bar{\phi}\mathcal{M}^{-\alpha}\phi}$$
$$\approx \int \mathcal{D}\bar{\phi}\mathcal{D}\phi e^{-\bar{\phi}r^2(\mathcal{M})\phi},$$

  with $r(x) = x^{-\alpha/2}$
- $\Delta = |1 - x^{\alpha/2}r(x)| < \epsilon \Rightarrow$ conventional Metropolis
- RHMC:
  - Hybrid Molecular Dynamics Trajectory
    * Momentum refreshment heatbath ($P(\pi) \propto e^{-\pi^*\pi/2}$).
    * Pseudo-fermion heatbath ($\phi \propto r(\mathcal{M})^{-1}\xi$, where $P(\xi) \propto e^{-\xi^*\xi/2}$).
    * MD trajectory with $\tau/\delta\tau$ steps.
  - Metropolis Acceptance Test $P_{\text{acc}} = \min(1, e^{-\delta H})$

Rational Hybrid Monte Carlo

- MD trajectory
  - Double inversion from $r^2(\mathcal{M})$
  - Use low degree approx $\bar{r} \approx \mathcal{M}^{-\alpha} \approx r^2$
  - Pseudo-fermion force

$$S'_{\mathsf{pf}} = -\sum_{i=1}^{\bar{n}} \bar{\alpha}_i \phi^\dagger (\mathcal{M} + \bar{\beta}_i)^{-1} \mathcal{M}' (\mathcal{M} + \bar{\beta}_i)^{-1} \phi.$$

- CG cost per trajectory $\approx$ R alg, $(2 + \tau/\delta\tau) \approx (\tau/\delta\tau)$
- Alternative multiple pseudo-fermions - Nroots (hep-lat/0409134)

$$\begin{aligned} \det \mathcal{M}^\alpha &= (\det \mathcal{M}^{\alpha/n})^n \\ &= \prod_{k=1}^{n} \int \mathcal{D}\bar{\phi}_k \mathcal{D}\phi_k \exp(-\bar{\phi}_k \mathcal{M}^{-\alpha/n} \phi_k) \end{aligned}$$

- Use multiple timescales

2+1 Domain Wall Fermions

- One flavour determinant $\frac{\det\sqrt{M_{\mathsf{f}}^\dagger M_{\mathsf{f}}}}{\det\sqrt{M_{\mathsf{pv}}^\dagger M_{\mathsf{pv}}}} = \det\left(\sqrt{\mathcal{M}^{DW}}\right)$ with

$$\mathcal{M}^{DW} = (M_{\mathsf{pv}}^\dagger)^{-1} M_{\mathsf{f}}^\dagger M_{\mathsf{f}} (M_{\mathsf{pv}})^{-1}$$

- BUT cannot write $r(\mathcal{M}^{DW})$ else nested inversion
- Therefore write action as

$$S_f = \bar{\phi}\left(M_{\mathsf{pv}}^\dagger M_{\mathsf{pv}}\right)^{1/2}\phi + \bar{\chi}\left(M_{\mathsf{f}}^\dagger M_{\mathsf{f}}\right)^{-1/2}\chi$$

$\Rightarrow$ 2 fermion fields to simulate 1 flavour contribution

- Naive additional cost small, but more noisy (Dawson)
- Monitor eigenvalues
- R algorithm - include bosonic contribution through negative flavour number

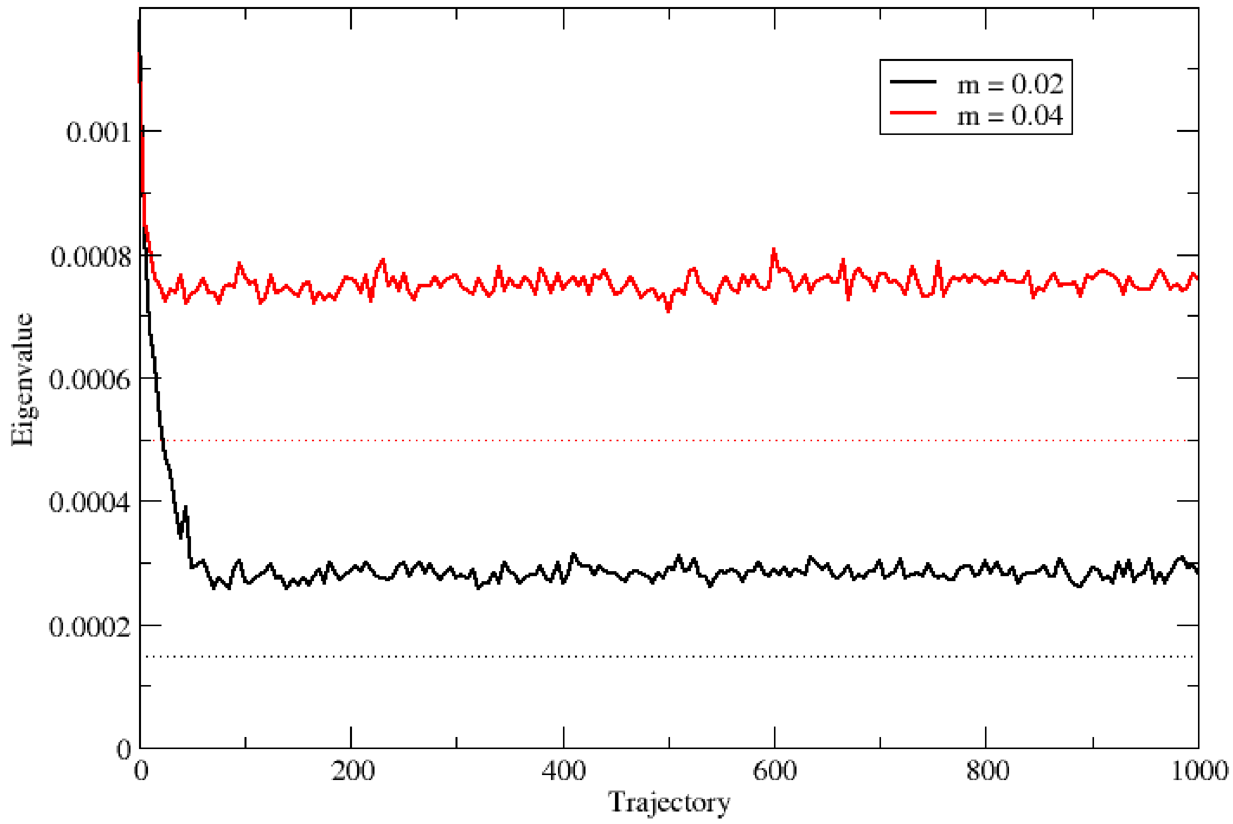
2+1 Domain Wall Fermions

- Lattices generated on QCDOC
- RHMC: use nroots for light pair $(1+1+1)$
- $V = 16^3.32.8$, DBW2 $\beta = 0.72$
- $(n_{\rm l}^{\rm md} = 10, n_{\rm l}^{\rm mc} = 15)$, $(n_{\rm s}^{\rm md} = 9, n_{\rm s}^{\rm mc} = 14)$

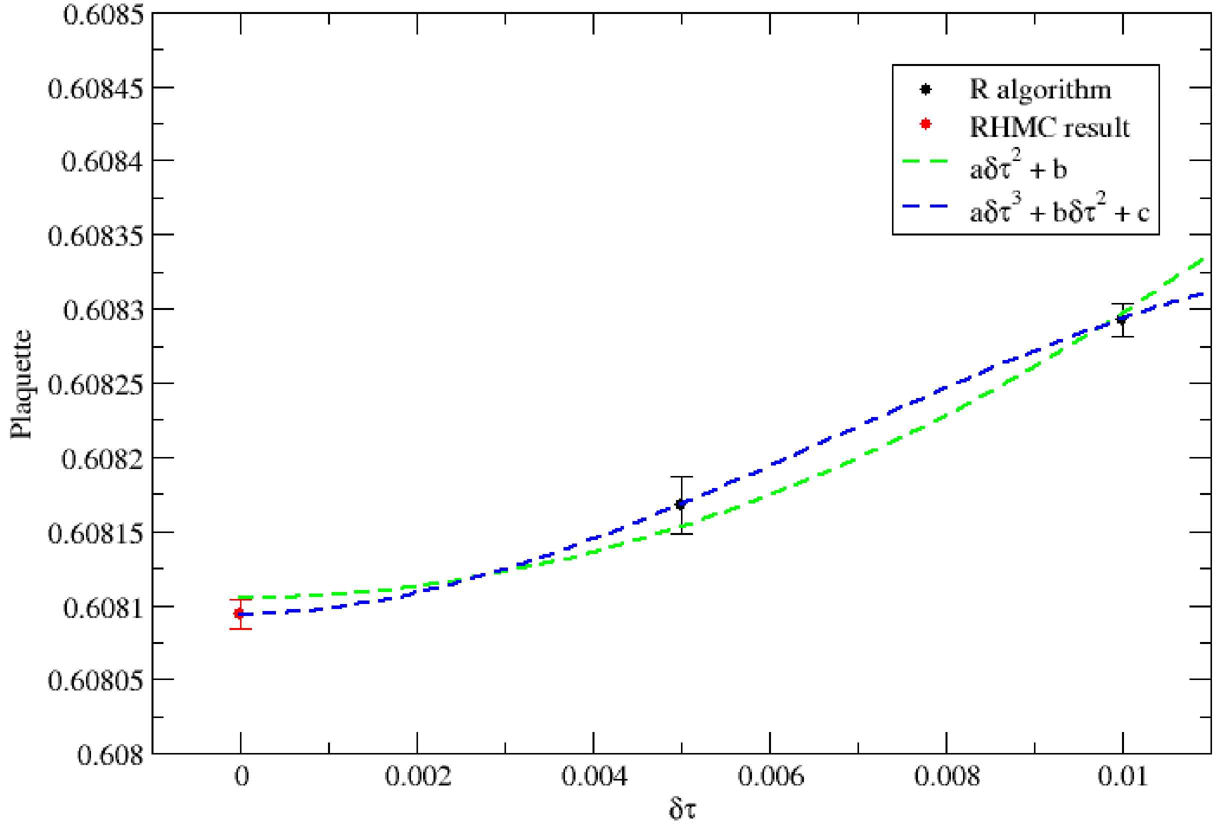
Alg	$m_{\rm l}$	$m_{\rm s}$	$\delta\tau$	$N_{\rm mv}$	A
R	0.04	0.04	0.01	13449(4)	-
R	0.02	0.04	0.01	32782(12)	-
R	0.02	0.04	0.005	65345(15)	-
RHMC	0.04	0.04	0.02	20300(3)	65.5(6)%
RHMC	0.02	0.04	0.0185	36265(8)	69.3(4)%
RHMC	0.01	0.04	0.02	45343(24)	64.5(6)%

- No mass dependance on acceptance rate

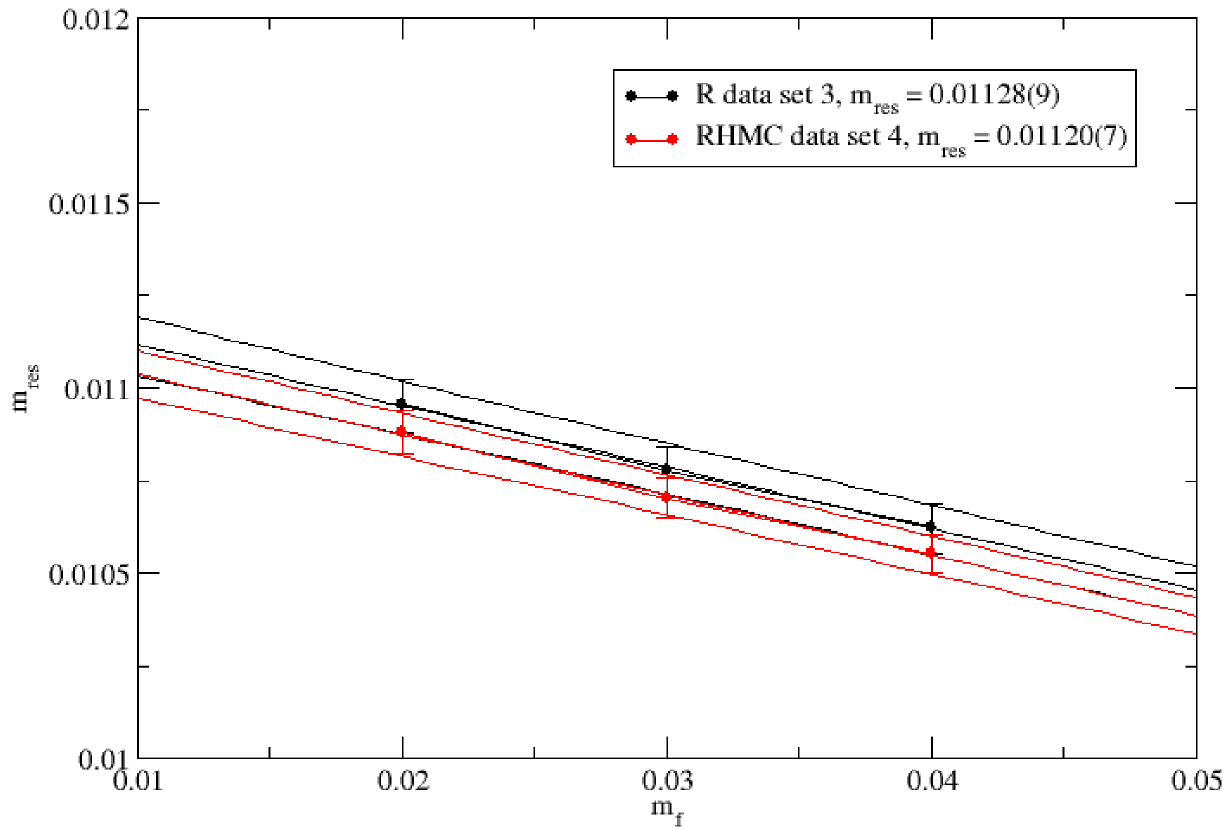
Lowest Eigen-value Behaviour



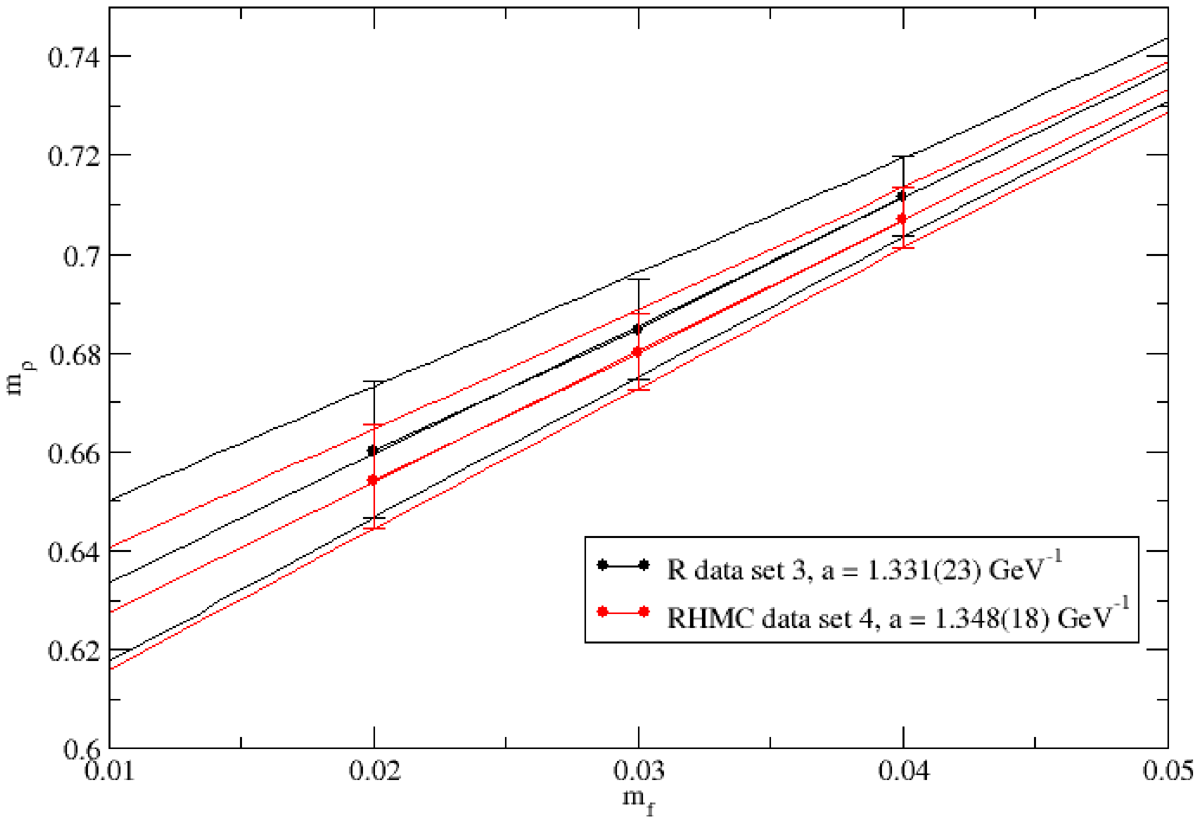
Plaquette $\delta\tau$ dependance $m_l = 0.02$



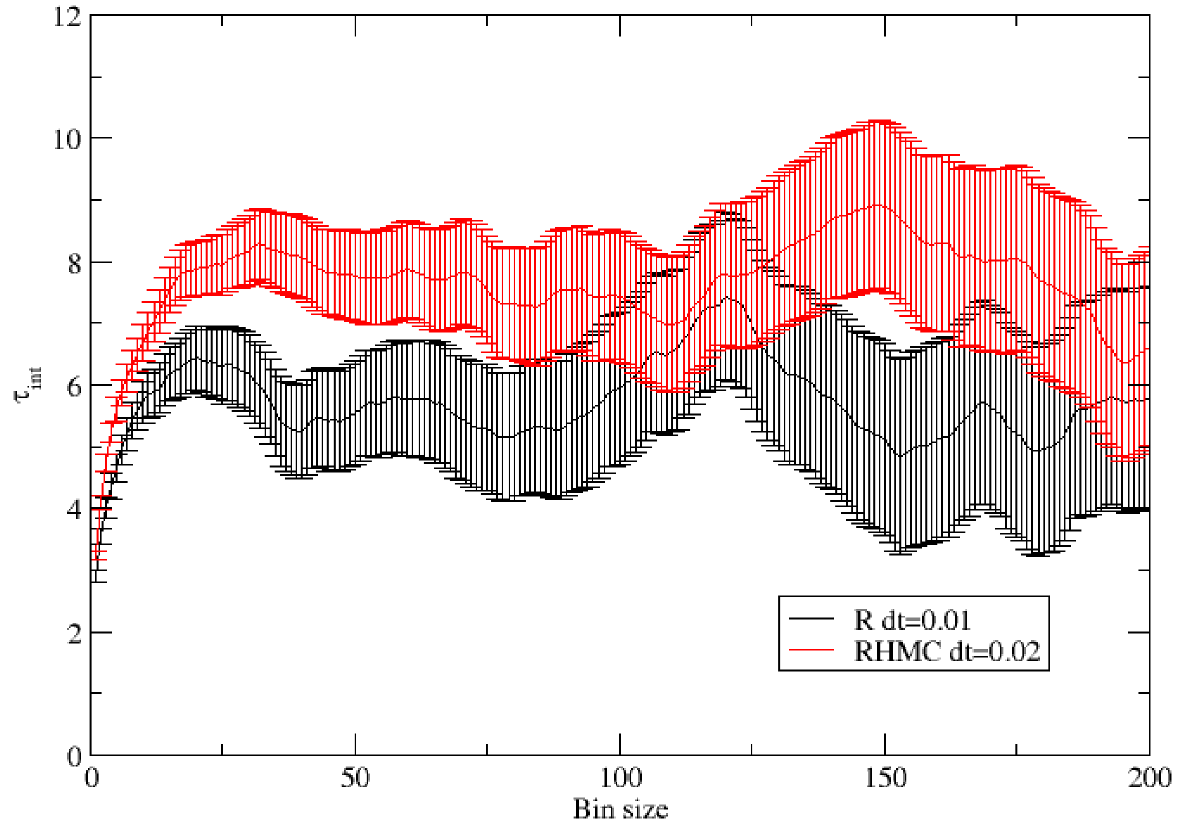
$m_{\rm res}$, $m_{\rm l} = 0.02$



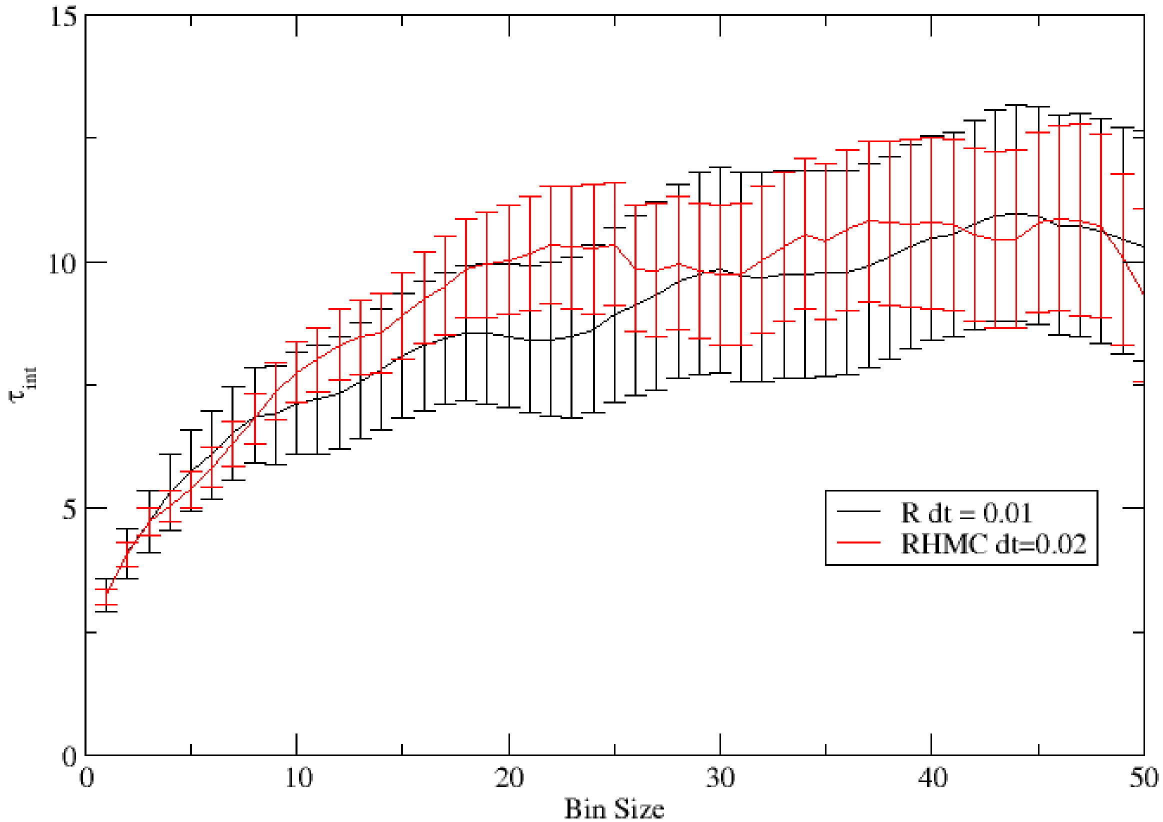
m_ρ, $m_l = 0.02$


0.74
0.72
0.7
0.68
0.66
0.64
0.62
0.6
0.01
0.02
0.03
0.04
0.05
m_ρ
m_f
R data set 3, a = 1.331(23) GeV^-1
RHMC data set 4, a = 1.348(18) GeV^-1

Plaquette τ_{int}, $m_l = 0.02$



Pseudo-Scalar Correlator τ_{int}, $t = 13$, $m_l = 0.02$



Finished/Current RHMC Runs

- $V = 16^3.32.8$

Action	β	$\frac{m_{ud}}{m_s}$	am_{ud}	N_{traj}	am_{res}
DBW2	0.72	1.0	0.04	3395	0.0106(1)
DBW2	0.72	0.5	0.02	6000	
DBW2	0.72	0.25	0.01	6000	
DBW2	0.764	1.0	0.04	1615	0.00546(7)
DBW2	0.764	0.5	0.02	1800	
DBW2	0.78	1.0	0.04	1620	0.00437(6)
DBW2	0.78	0.5	0.02	1505	
Iwasaki	2.13	1.0	0.04	2380	0.0104(2)
Iwasaki	2.13	0.5	0.02	2450	
Iwasaki	2.2	1.0	0.04	4565	0.0065(1)
Iwasaki	2.2	0.5	0.02	3175	

Future RHMC Runs

- $V = 24^3.64.16$

Action	β	$\frac{m_{ud}}{m_s}$	am_{ud}	N_{traj}
Iwasaki	2.13	0.75	0.04	5000+
Iwasaki	2.13	0.5	0.02	5000+
Iwasaki	2.13	0.25	0.02	5000+

- 12+ month run using standard RHMC!
- Need to go faster.....

Omelyan Integrator

- Omelyan integrator given by

$$\hat{U}_{\mathrm{QPQPQ}}(\delta\tau) = e^{\lambda\delta\tau Q} e^{\frac{\delta\tau}{2}P} e^{(1-2\lambda)\delta\tau Q} e^{\frac{\delta\tau}{2}P} e^{\lambda\delta\tau Q}$$

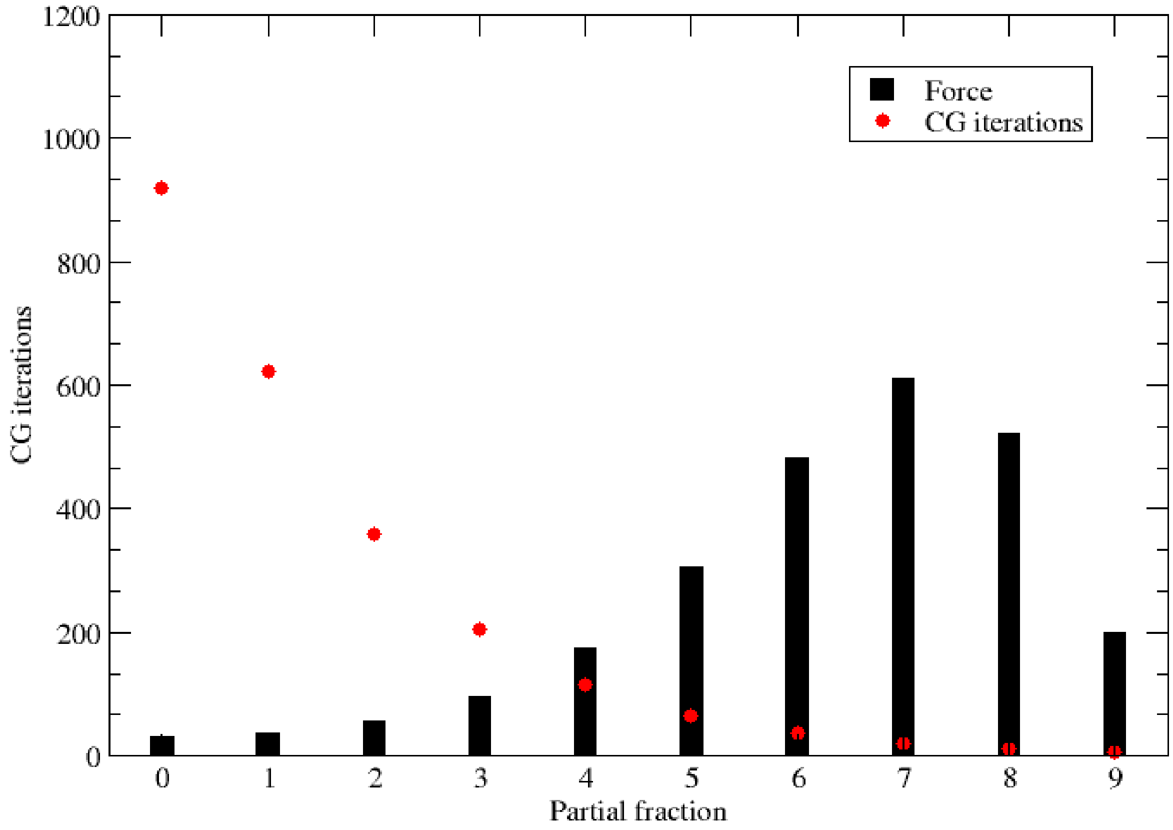
  ($\lambda \approx 0.1932$)
- $\approx$ 50% improvement in efficiency (de Forcrand and Takaishi)
- Use multiple timescales

$$\hat{U}_1(\delta\tau) = e^{\lambda_1\delta\tau Q} e^{\frac{\delta\tau}{2}P_1} e^{(1-2\lambda_1)\delta\tau Q} e^{\frac{\delta\tau}{2}P_1} e^{\lambda_1\delta\tau Q}$$

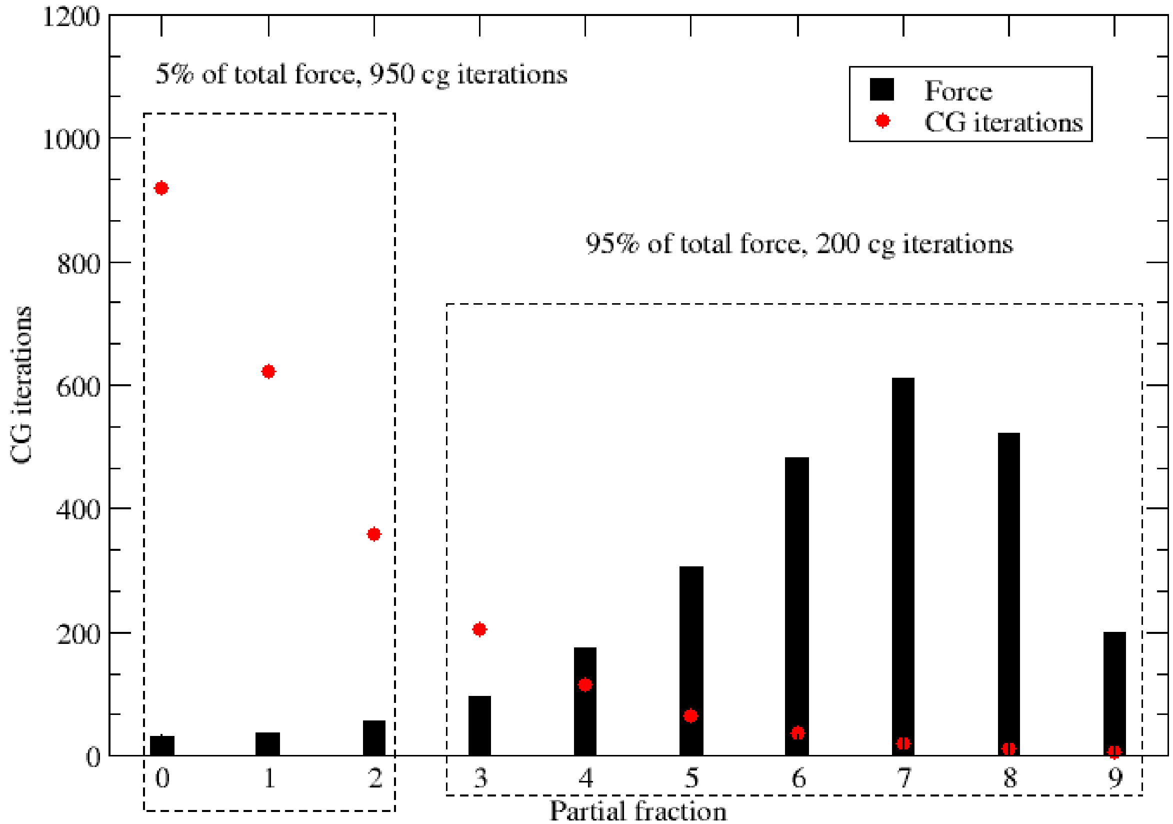
$$\hat{U}_2(\delta\tau) = \left[\hat{U}_1(\lambda_2\frac{\delta\tau}{m})\right]^m e^{\frac{\delta\tau}{2}P_2} \left[\hat{U}_1\left((1-2\lambda_2)\frac{\delta\tau}{m}\right)\right]^m e^{\frac{\delta\tau}{2}P_2} \left[\hat{U}_1(\lambda_2\frac{\delta\tau}{m})\right]^m$$

- $V = 16^3 32 \times 8$, $\beta = 2.13$, $m = \frac{m_s}{2}$
  - Leapfrog: $\delta\tau = 0.02$ $\langle A \rangle = 63.6\%$
  - Omelyan: $\delta\tau = 0.04$ $\langle A \rangle = 88.8\%$
- λ_1, λ_2 independently tunable

DWF Forces ($N_f = 1, \beta = 2.13,\ m = 0.01,\ V = 24^3.64.16$)



DWF Forces ($N_f = 1, \beta = 2.13,\ m = 0.01,\ V = 24^3.64.16$)



Multiple fermion timescales

- Split partial fractions into multiple timescales
- Tune parameters..........

Integ'or	poles	Steps/ traj	Light	Heavy	Gauge	Accept	Cost	Time/ 10traj	Auth (evol)	hdw
Ome'n	ud(3L7H) s(2L7H)	6	1/6 (0.167)	1/(6×5) 0.033	1/(6×5×5) 0.0067	71%	5.0	3.58	cmm	n
		6	1/6 (0.167)	1/(6×5) 0.033	1/(6×5×10) 0.0033	67%	10.4	7	cmm (4)	n
		6	1/6 (0.167)	1/(6×10) 0.017	1/(6×10×5) 0.0033	76%	15.1	11.5	cmm (5)	n
		6	1/6 (0.167)	1/(6×3) 0.055	1/(6×3×8) 0.0069	67%	6.7	4.5	dja (1)	n
		6	1/6 (0.167)	1/(6×5) 0.033	1/(6×5×3) 0.011	69%	8.7	6	dja (3)	n
		6	1/6 (0.167)	1/(6×2) 0.083	1/(6×2×8) 0.010	62%	5.3	3.3	dja (11)	n
		6	1/6 (0.167)	1/(6×3) 0.056	1/(6×3×3) 0.019	68%	5.9	4	dja (10)	n
	ud(2L8H) s(2L7H)	6	1/6 (0.167)	1/(6×5) 0.033	1/(6×5×5) 0.0067	78%	10.0	7.8	bjp (6)	n
	ud(4L6H) s(2L7H)	6	1/6 (0.167)	1/(6×5) 0.033	1/(6×5×5) 0.0067	64%	7.9	5.1	bjp (7)	n
	ud(3L7H) s(3L7H)	6	1/6 (0.167)	1/(6×3) 0.056	1/(6×3×3) 0.019	57%	7.0	4	dja (15)	n
	ud(4L6H) s(4L6H)	6	1/6 (0.167)	1/(6×3) 0.056	1/(6×3×3) 0.019	46%	6.9	3.17	dja (16)	n
	ud(4L6H) s(3L7H)	6	1/6 (0.167)	1/(6×3) 0.056	1/(6×3×3) 0.019	59%	4.3	2.64	dja (18)	y
	ud(4L6H) s(4L6H)	6	1/6 (0.167)	1/(6×3) 0.056	1/(6×3×3) 0.019	43%	6.0	2.6	dja (30)	y
	ud(4L6H) s(3L7H)	5	1/5 (0.2)	1/(5×5) 0.04	1/(5×5×5) 0.008				dja (32)	y
Leap	ud(3L7H) s(2L7H)	50	1/50 (0.02)	1/50 (0.02)	1/200 (0.005)	77%	5.6	4.3	bjp (8)	y
		12	2/25 (0.08)	2/(25×4) 0.02	2/(25×4×4) 0.005	66%	3.6	2.4	bjp (9)	y
		12	2/25 (0.08)	2/(25×2) 0.04	2/(25×2×8) 0.005	6%	29	1.73	bjp (12)	y
		12	2/25 (0.08)	2/(25×4) 0.02	2/(25×4×2) 0.01	54%	4.5	2.43	bjp (13)	y
		12	2/25 (0.08)	2/(25×3) 0.027	2/(25×3×5) 0.0053	32%	6.4	2.05	bjp (17)	y
		10	1/10 (0.10)	1/(10×5) 0.02	1/(10×5×4) 0.005	33%	8.2	2.45	bjp (20)	y
		11	9/100 (0.09)	9/(100×4) 0.0225	9/(100×4×4) 0.0056	27%	8.5	2.3	bjp (21)	y
	ud(4L6H) s(3L7H)	12	2/25 (0.08)	2/(25×4) 0.02	2/(25×4×4) 0.005	28%	7.2	2.01	bjp (22)	y
	ud(4L6H) s(3L7H)	16	(0.06)	0.06/(3) 0.02	0.06/(3×4) 0.005	52%	4.12	2.13	bjp (23)	y
	ud(4L6H) s(3L7H)	14	(0.07)	0.07/(4) 0.0175	0.07/(4×4) 0.0044	42%	5.28	2.22	bjp (24)	y
	ud(4L6H) s(3L7H)	18	(0.055)	0.055/(3) 0.0183	0.055/(3×4) 0.0046	62%	3.86	2.39	bjp (27)	y
L-O-L	ud(3L7H) s(2L7H)	12	2/25 (0.08)	2/(25×2) (0.04)	2/(25×2×8) 0.005	69%	3.5	2.4	pab (14)	y
		11	9/100 (0.09)	9/(100×2) 0.045	9/(100×2×9) 0.005	58%	4.1	2.4	bjp (19)	y
	ud(4L6H) s(3L7H)	12	8/100 (0.08)	12/(25×2) 0.04	12/(15×2×8) 0.005	43%	4.7	2.03	bjp (25)	y
	ud(4L6H) s(3L7H)	14	7/100 (0.07)	0.07/(2) 0.035	0.07/(2×8) 0.0044	67%	3.45	2.31	bjp (26)	y
	ud(4L6H) s(3L7H)	14	7/100 (0.07)	0.07/(1) 0.07	0.07/(1×16) 0.0044	50%	3.58	1.79	bjp (28)	y
	ud(4L6H) s(3L7H)	16	6/100 (0.06)	0.06/(1) 0.06	0.06/(1×16) 0.0038	67%	3.145	2.12	bjp (29)	y

- Factor 3-4 speedup

Heavy Mode Cancellation

- Rewrite one flavour dwf determinant as

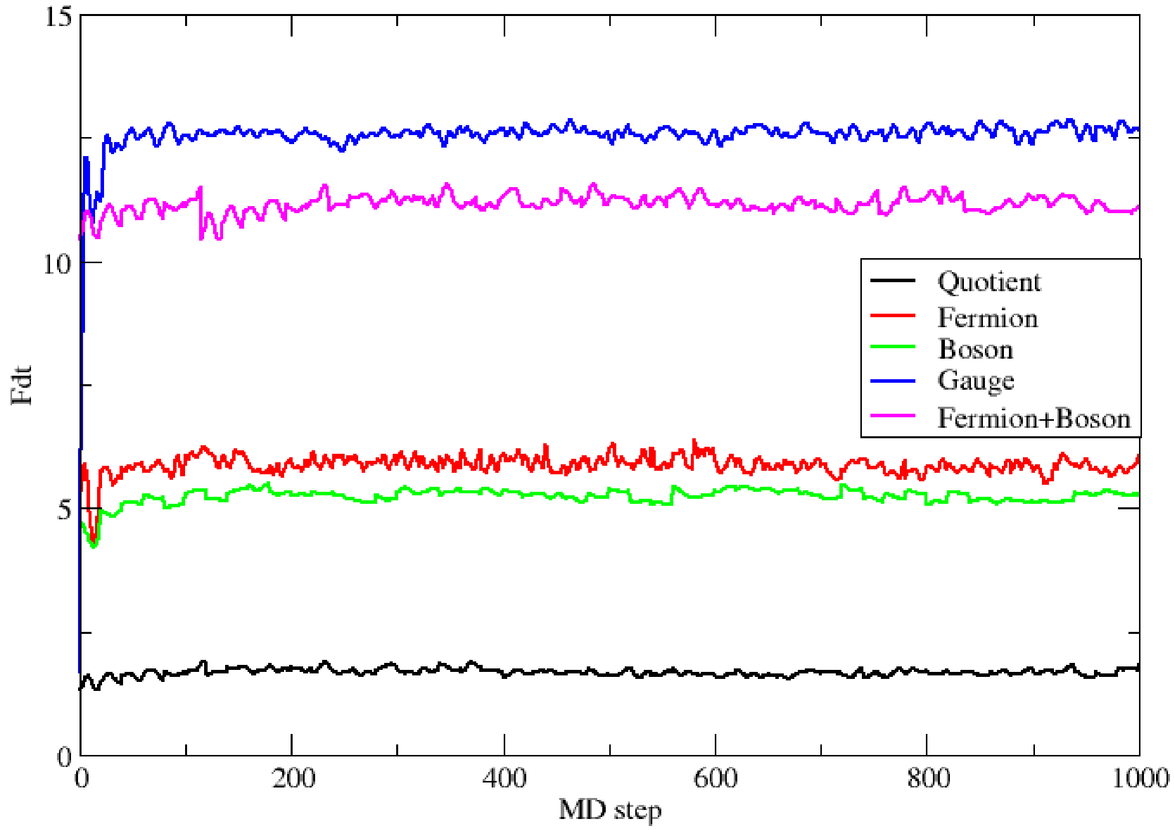
$$\left(\frac{\det M_{\mathrm{f}}^{\dagger}M_{\mathrm{f}}}{\det M_{\mathrm{pv}}^{\dagger}M_{\mathrm{pv}}}\right)^{1/2} = \det\left[(M_{\mathrm{pv}}^{\dagger}M_{\mathrm{pv}})^{-1/4}(M_{\mathrm{f}}^{\dagger}M_{\mathrm{f}})^{1/2}(M_{\mathrm{pv}}^{\dagger}M_{\mathrm{pv}})^{-1/4}\right]$$

- Pseudofermion action is

$$\begin{aligned} S_{\mathrm{f}} &= \bar{\phi}(M_{\mathrm{pv}}^{\dagger}M_{\mathrm{pv}})^{1/4}(M_{\mathrm{f}}^{\dagger}M_{\mathrm{f}})^{-1/2}(M_{\mathrm{pv}}^{\dagger}M_{\mathrm{pv}})^{1/4}\phi \\ &= \bar{\phi}\, r_1\, r_2^2\, r_1\, \phi \end{aligned}$$

- Extra $M_{\mathrm{pv}}^{\dagger}M_{\mathrm{pv}}$ inversion required $\Rightarrow$ negligible cost

Heavy Mode Cancellation ($N_f = 1$, $\beta = 2.13$, $m = 0.02$, $V = 4^5$)



Heavy Mode Cancellation

- Large decrease in force magnitude
- $3-4\times$ fermion stepsize increase
- Can't combine with multi-timescale partial fractions
- Sliding CG solver tolerance, $r = 10^{-4} - 10^{-6}$
- Should gain more than timescale splitting

Conclusions and outlook

- RHMC is a fast and exact algorithm
- Many recent improvements made
- Benchmark dwf calculation $V = 24^3.64.16$, $\beta = 2.13$, $m_s = 0.04$, $m_l = 0.01, 0.02, 0.03$, 12 months $\rightarrow$ 3 months
- Multiple timescale RHMC applicable to all fermion formulations
- Compare against "Domain Decomposition" (Lüscher) and "Multiple timescale mass preconditioning" (Urbach et al)