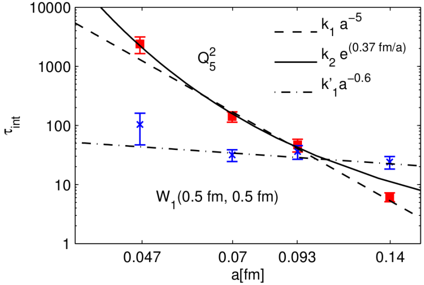
Study of Riemann Manifold Monte Carlo for Lattice QCD

Chulwoo Jung
in collaboration with Norman Christ, Yong-Chull Jang, Peter Boyle, ...

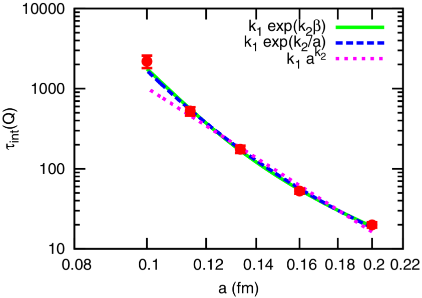
DWQ25, December 13, 2021

Critical Slowing Down in Hybrid Monte Carlo

Autocorrelation increases rapidly as the lattice spacing(a) decreases ($\tau_{int} \sim a^{-(5\sim6)}$ or $exp[a_0/a]$)



Wilson, Q^2, Wilson loop (Schaefer et. al., arXiv:1009.5228)



DBW2, Q (McGlynn & Mawhinney, PhysRevD.90.074502)

Numerical cost to generate the same number of decorrelated configurations increase at least $\sim (1/a)^9$ for the same physical volume!

While different observables may have different autocorrelation time, it is critical to ensure the Markov chains generated from current and future machines are much longer than the observed autocorrelation times.

- This meeting (all Monday):
    - Field Transformations HMC (Trivializing map) (N. Matsumoto)
    - Gauge Covariant NN (A. Tomiya)
    - Riemann Manifold Hybrid Monte Carlo (RMHMC)
    - L2HMC (S. Foreman): ML based approach to optimze HMC integrator for Autocorrelation reduction.
- Gauge fixed HMC: Include Gauge fixing term in the evolution(Y. Zhao, N. Christ, A. Sheta..)
- Many other interesting approaches presented in lattice conference and other meetings.

Riemann Manifold Hybrid Monte Carlo(RMHMC)

Origin: Duane & Pendleton (Phys. Lett. B206, 101–106 (1988))
Introduce auxiliary field π and momenta ϕ.

$$\mathrm{H} = \mathrm{S(U)} + \frac{1}{2}\sum_{\mu}\left[\mathrm{p}_{\mu}{}^{\dagger}\mathrm{M(U)}^{-1}\mathrm{p}_{\mu}\right] + \log|\mathrm{M}| = \mathrm{S(U)} + \frac{1}{2}\sum_{\mu}\left[\mathrm{p}_{\mu}{}^{\dagger}\mathrm{M(U)}^{-1}\mathrm{p}_{\mu} + \pi_{\mu}^{\dagger}\mathrm{M(U)}\pi_{\mu} + \phi_{\mu}^{2}\right]$$

Gauge invariant Laplace operator is used as the acceleration kernel:

$$M(\mathrm{U})\phi_{\nu}(x) = (1-\kappa)\phi_{\nu}(x) - \frac{\kappa}{4d}\nabla^2\phi_{\nu}(x)$$

$$\nabla^2\phi_{\nu}(x) = \sum_{\mu=0}^{d-1}\left[U_{\mu}(x)\phi_{\nu}(x+\mu)U_{\mu}^{\dagger}(x) + U_{\mu}^{\dagger}(x-\mu)\phi_{\nu}(x-\mu)U_{\mu}(x-\mu) - 2\phi_{\nu}(x)\right] \sim p^2$$

In the perturbative limit,

$$Frequency(p)^2 \sim \beta\frac{p^2}{(1-\kappa)+\kappa\frac{p^2}{4d}}$$

Accelerated 'mass term' $M(U)$ should be gauge invariant $\rightarrow$ field-dependent metric, non-separable hamiltonian. Implicit integrator is needed to keep it symplectic and reversible. RMHMC (Girolami & Calderhead, 2011)):
Typical integrator algorithms such as leapfrog is non-reversible for non-separable hamiltonian: implicit integrator to maintain reversibility

$$p^{n+\frac{1}{2}} = p^n - \frac{\epsilon}{2}\frac{\delta H}{\delta U}(U^n, p^{n+\frac{1}{2}}), \qquad U^{n+1} = U^n + \frac{\epsilon}{2}\left[\frac{\delta H}{\delta p}(U^n, p^{n+\frac{1}{2}}) + \frac{\delta H}{\delta p}(U^{n+1}, p^{n+\frac{1}{2}})\right]$$

$$p^{n+1} = p^{n+\frac{1}{2}} - \frac{\epsilon}{2}\frac{\delta H}{\delta U}(U^{n+1}, p^{n+\frac{1}{2}})$$

Here we study more generalized function of ∇:

$$M(\mathrm{U})\phi_\nu(x)^{-1} = G[\nabla^2]^2\phi_\nu(x)$$

$$G(x) = \frac{\sum_{i=0}^{n}\beta_i x^i}{\sum_{i=0}^{n}\alpha_i x^i} = G_0 + \sum_{i=0}^{n/2}[a_i' x + b_i'][x^2 + c_i' x + d_i']^{-1}$$

Fourier Acceleration

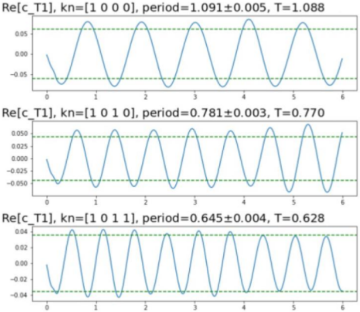
Treat gauge force as independent harmonic oscillators for modes of Laplacian operator

$$H = \sum_{\omega}\left(\frac{1}{2m(\omega)}p^2 + \frac{1}{2}k(\omega)x^2\right) \quad \langle F\rangle \sim -k\langle x\rangle \sim -\sqrt{\frac{k(\omega)}{m(\omega)}} \quad \text{Frequency} =\sim \sqrt{\frac{k(\omega)}{m(\omega)}}$$

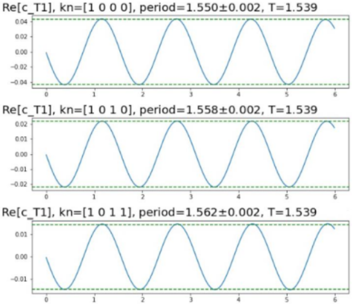
To eliminate critical slowing down, meassure $\langle F\rangle$ with normal HMC((constant m)
tune $m(\omega) \sim k(\omega) \sim \langle F\rangle^2$

Gauge Fixed Fourier Accelerated HMC (Y. Zhao, N. Christ)

- Plot MD time dependence, 4^4, M=3, β=100:



(a) HMC kinetic term

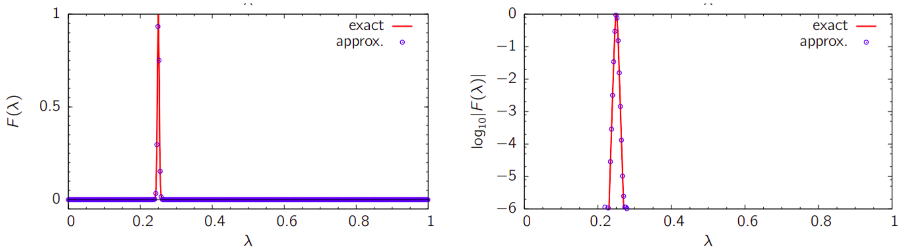


(b) Fourier accelerated kinetic term

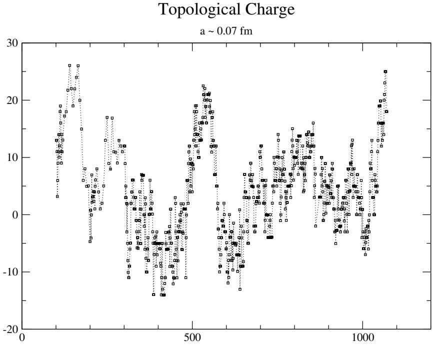
Measurement of $k(\omega)$

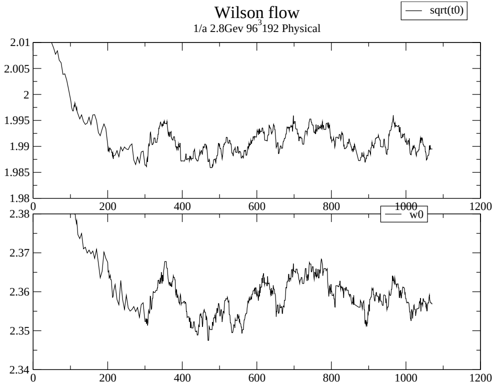
Implemented a sharply peaked bandpass filter with Chebyshef polynomial of Laplace operator.

- Used to estimate mode density ($\langle \xi | B_\lambda(\nabla^2) | \xi \rangle$) and contribution to force ($\langle F | B_\lambda(\nabla^2) | F \rangle$) from different Laplace modes.
- : Response to observable: Apply $B_\lambda(\nabla^2)$ to Initial momenta, measure change in observables (Action S, Wilson Flowed energy $\langle E \rangle, \cdots$) after a very short ($\tau = 10^{-4} \sim 10^{-3}$) trajectory, necessary as the HMC diffuse momemta very quickly.



Observables: Wilson flow smeared energy





Topology tunnelling can be address by open boundary conditions. t_0, w_0 show similar autocorrelation time as topological charge for current finest physcial mass ensembles $1/a \sim 2.8$Gev. More flexible quantity than Q: change in t_0, w_0 or E(t), energy at wilson flow time t, can be studied at small step size

Studied Ensembles

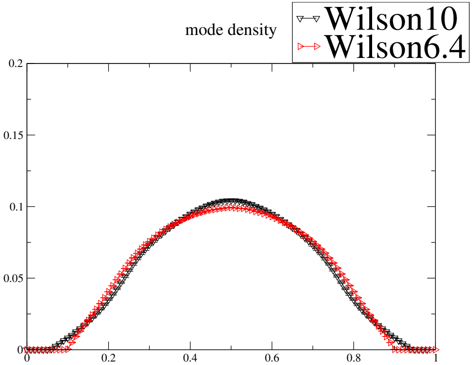
Quenched ensembles

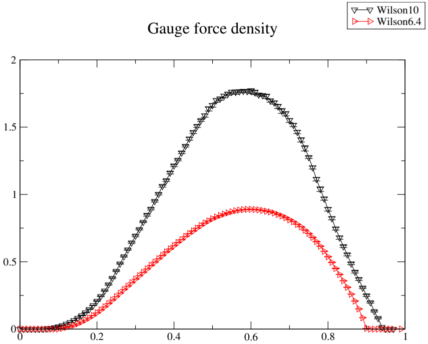
- Quenched $\beta = 10$ 16^4 Wilson gauge action
- Quenched $a^{-1} \sim 4$Gev ($\beta = 6.4$) 32^4 Wilson gauge action
- Quenched $a^{-1} \sim 2$Gev 16^4 DBW2 gauge action

Dynamical (with fermions) ensembles :RBC/UKQCD DWF 2+1 flavor ensembles

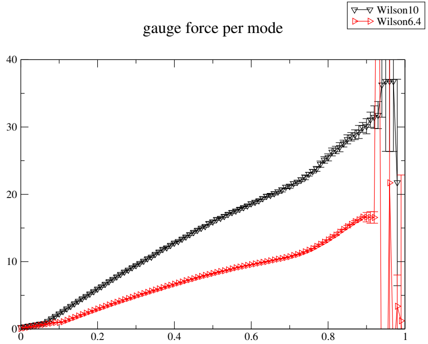
- 48I: $48^3 \times 96, 1/a \sim 2.3$Gev, Physical
- 32I2.8Gev: $32^3 \times 64, 1/a \sim 2.8$Gev, Physical
- 32I3.1Gev: $32^3 \times 64, 1/a \sim 3.1$Gev $m_\pi \sim 300$Mev

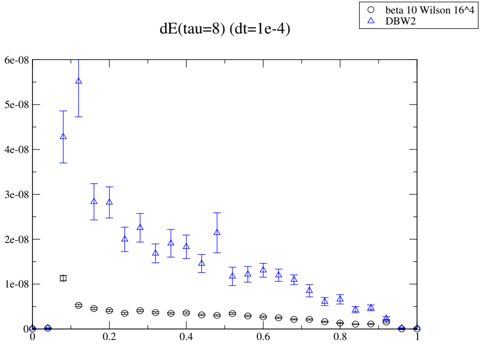
Laplace mode density and Force Density for quenched ensembles



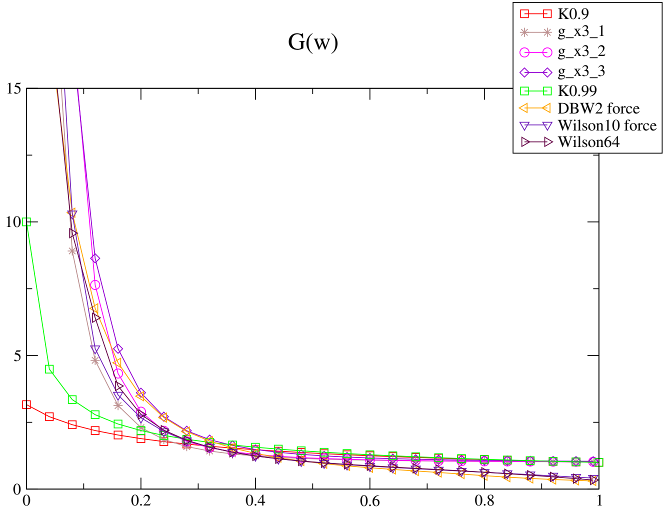


Force density and wilson flow response for quenched emsemble





Despite the small density, low Laplace modes create (relatively) huge change in wilson flowed energy

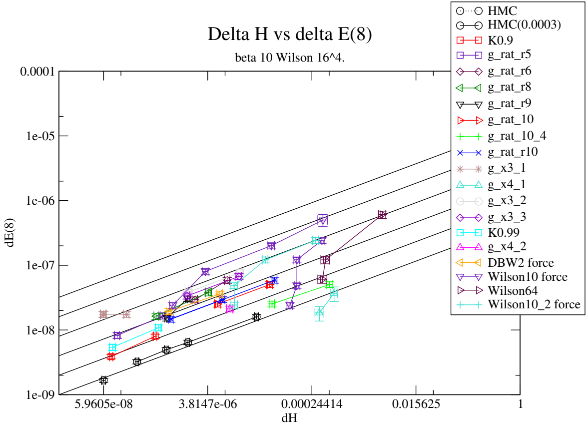


g_x3_[123] $\sim 1 + \frac{c}{(x+b)^3}$

Duane & Pendleton:

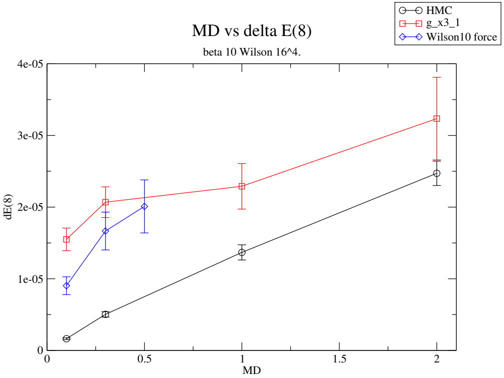
$\sim \sqrt{\frac{1}{(1-\kappa)+\kappa x}}$

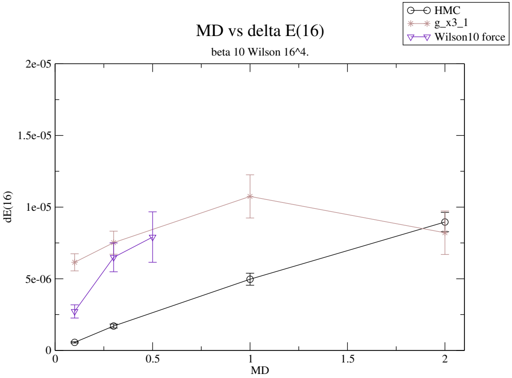
$\Delta H/dt^3$ and $\Delta E(\tau = 8)$ from short HMC for Wilson $\beta = 10$.



Possibility for factor of $8 \sim 10$ gain?

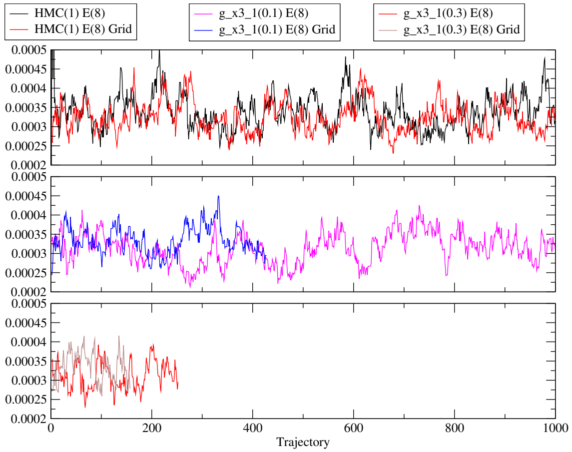
Trajectory Length vs. Wilson flowed energy change

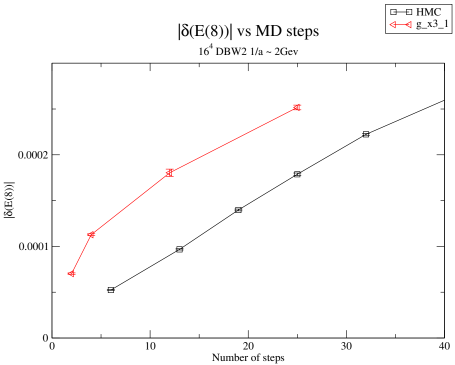


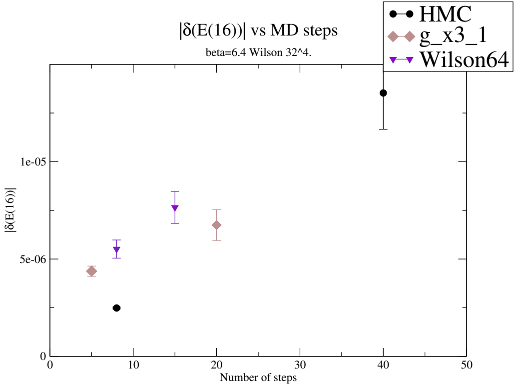


A significant gain suggested for short trajectory length.

Wilson flow energy evolution for Wilson $\beta = 10$.

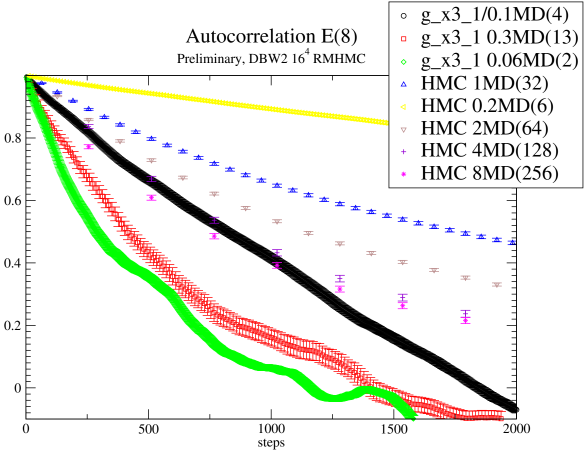




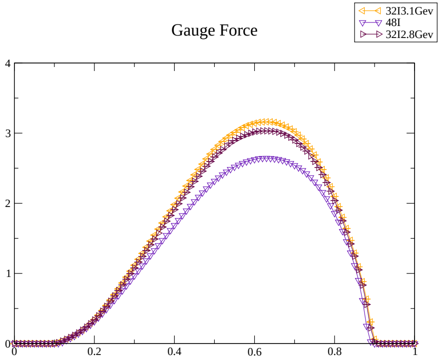


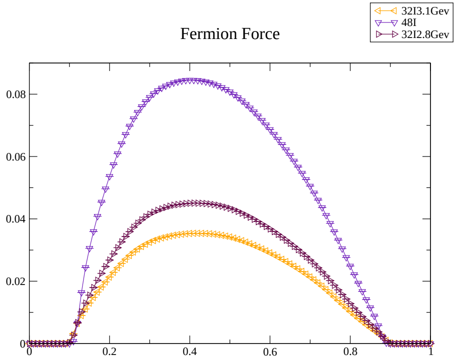
Similar gain on short trajectory length compared to HMC. Less gain for longer trajectory.

Autocorrelation of E(8) for $1/a \sim$ 2Gev DBW2 ensemble (Preliminary)

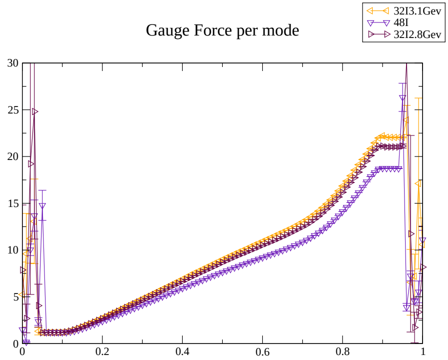


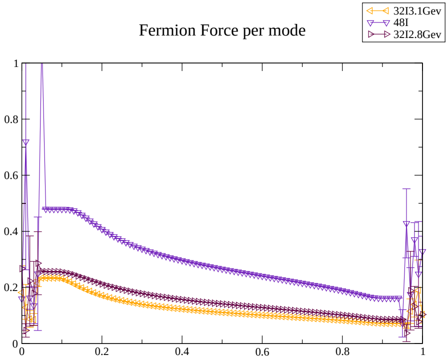
Analysis of HMC forces with Laplace modes for dynamical ensembles



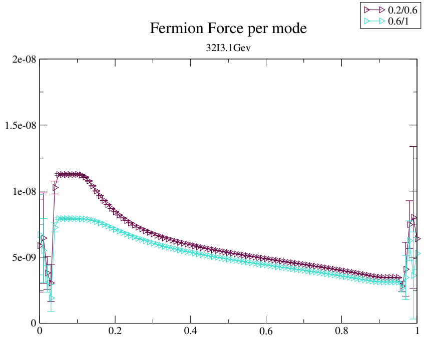


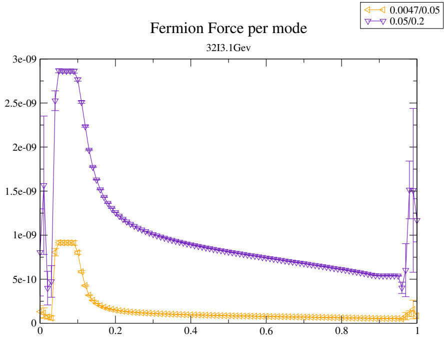
Gauge and Fermion force density per mode for dynamical esembles





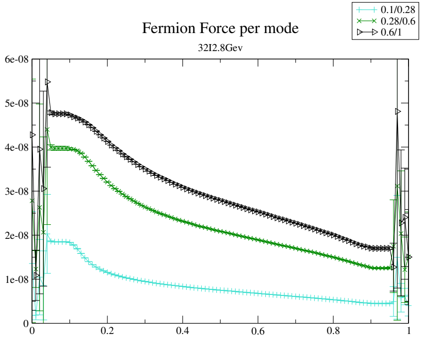
Fermion force density per mode for 32I3.1Gev esembles

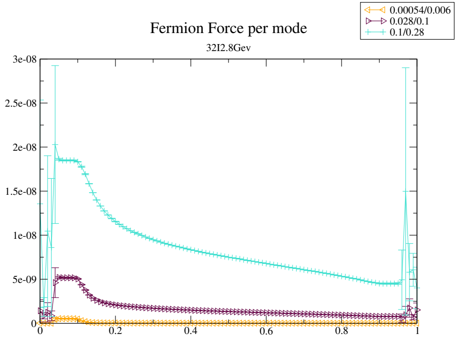




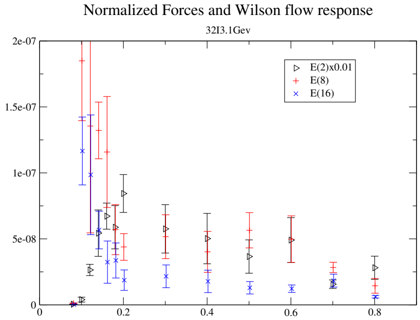
Force for small Hasenbusch mass relatively larger for small Laplace mode, but the force overall small

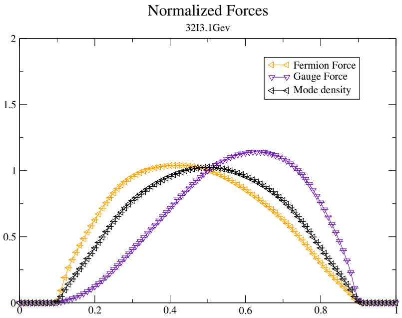
Fermion force density per mode for 32I2.8Gev esembles

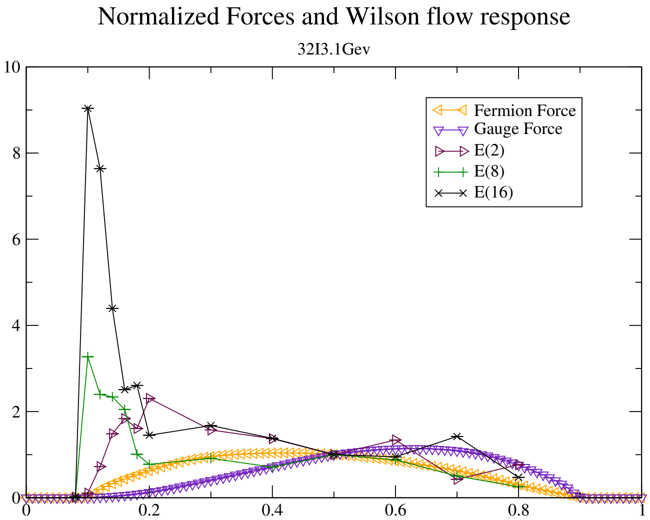




Force density and wilson flow response for Laplace modes for 32I3.1Gev ensemble






Normalized Forces and Wilson flow response
32I3.1Gev
10
8
6
4
2
0
0
0.2
0.4
0.6
0.8
1
Fermion Force
Gauge Force
E(2)
E(8)
E(16)

Putting it all together: Tuning for dynamical simulation

- Start with single Fermion step evolution, with various Sexton-Weingarten ratio
- Turned out leapfrog was really suboptimal: a scan of parameters with implicit leapfrog suggested fermion step has to be as small as typical gauge steps. Dramatic change with implicit Omelyan integrator.
- Does the increasing behavior of fermion force per mode affect the S-W ratio? Initial results suggests it may be, at least for some kernels. Can we adjust the shape of the kernel to improve?
- While we have been mostly ignoring numerical cost for each RMHMC step as our aim is to use RMHMC for dynamical simulation, the Laplace operator is very network bound for small volume, more careful optimization/tuning may be necessary.

Status & Discussion

- Fourier Acceleration with Laplace operator can be tuned to accelerate slow-moving low momentum modes significantly faster than normal HMC.
- Low modes of Laplace operator strongly correlates with long range mode probed by wilson flowed energy. $\sim \times 10$ acceleration possible?
- Fermions?: Force density from Fermion action is relatively larger for low Laplace modes compared to gauge force. Still relatively small for low Laplace modes due to low density of modes. Potential for cost savings if we can make each trajectory length small without sacrificing decorrelation of low modes.
- Do we understand the saturation/stagnation?
- $1/a \sim$ 4Gev Wilson, $1/a \sim$ 2Gev DBW2, and $1/a \sim (3.1, 2.8)$Gev 2+1f run ongoing.
- Parameter space is large. There may be room for improvement in $F[\nabla^2]$: Machine learning, etc.
- Different choice of operators, other than Laplace, e.g. Hessian.

Thank you!