Riemannian Manifold HMC with Fermions

Chulwoo Jung
Brookhaven National Laboratory, USA
in collaboration with Norman Christ, Yong-Chull Jang, Peter Boyle, Guido Cossu ....

Lattice 2023
Fermilab, Aug. 1, 2023

US Exascale machines

Frontier (Oak Ridge)

AMD GPU(MI250X)+ CPU(Epyc),

Aurora (Argonne)

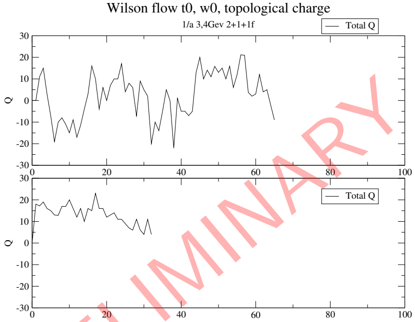
Intel GPU(Ponte Vecchio)+CPU(Sapphire Rapids)

DWF 2+1+1 flavor, physical ensemble program

Parameter $(\beta, m_l, m_s, m_c \cdots)$ tuning done on small volume. Duplicated to create starting lattice for the production run

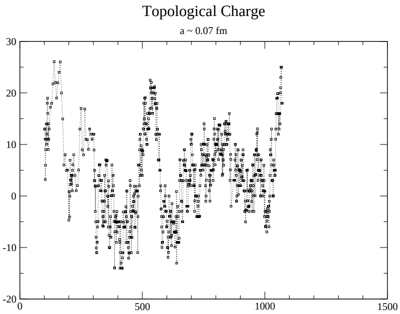
- $96^3 \times 192(L_s = 16), 1/a \sim 3$Gev
- $128^3 \times 288(12), 1/a \sim 4$Gev
- $160^3 \times \sim 384(12), 1/a \sim 5$Gev

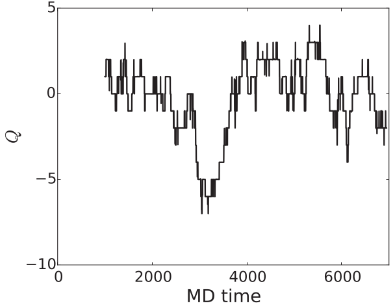
DWF+Wilson, 2+1+1f



Topological charge for $1/a \sim 3,4$Gev ensembles

Critical Slowing Down in Hybrid Monte Carlo



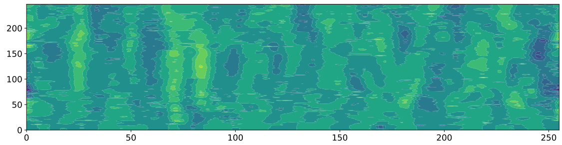


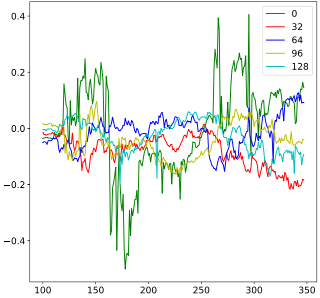
DWF+I, 2+1f $1/a \sim 2.7$Gev $\qquad$ $1/a \sim 3.1$Gev

Autocorrelation increases rapidly as the lattice spacing(a) decreases ($\tau_{int} \sim a^{-(5\sim6)}$ or $exp[a_0/a]$). Numerical cost to generate the same number of decorrelated configurations increase at least $\sim (1/a)^9$ for the same physical volume!

Open boundary in time direction with increased time extent for $1/a > 3$Gev. However..

Topological density per time slice
x axis: T y axis: MD time





x axis: MD time

It is not clear the (anti)instantons created at the open boundaries move fast enough to the interior to be helpful at this lattice spacing.

Riemann Manifold Hybrid Monte Carlo(RMHMC)

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

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

Gauge invariant Laplace operator is used as the acceleration kernel:

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

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

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

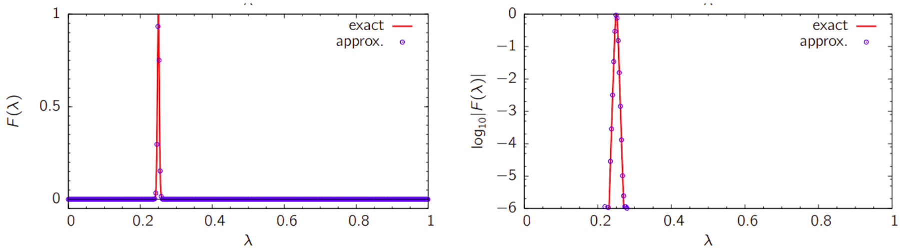
$$p^{n+\frac{1}{2}} = p^n - \frac{\epsilon}{2}\frac{\delta H}{\delta U}(U^n, p^{n+\frac{1}{2}}), \qquad U^{n+1} = U^n + \frac{\epsilon}{2}\left[\frac{\delta H}{\delta p}(U^n, p^{n+\frac{1}{2}}) + \frac{\delta H}{\delta p}(U^{n+1}, p^{n+\frac{1}{2}})\right]$$
$$p^{n+1} = p^{n+\frac{1}{2}} - \frac{\epsilon}{2}\frac{\delta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

Here we study more generalized function of ∇:

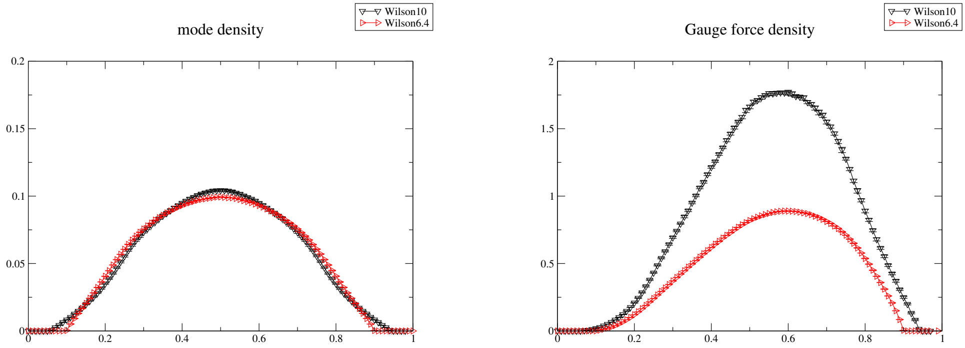
$$M(\mathbf{U})^{-1} = G[\nabla^2]^2$$
$$G(x) = \frac{\sum_{i=0}^{n} \beta_i x^i}{\sum_{i=0}^{n} \alpha_i x^i} = G_0 + \sum_{i=0}^{n/2} [a_i' x + b_i'][x^2 + c_i' x + d_i']^{-1}$$

Tuning strategy

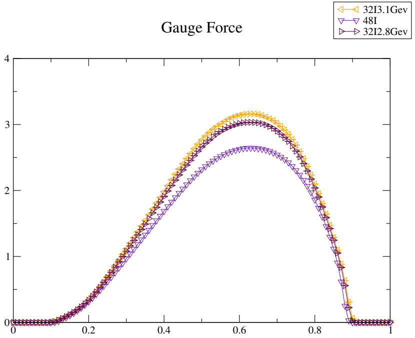
- Employ a sharply peaked bandpass filter with Chebyshef polynomial of Laplace operator, to estimate mode density ($\langle \xi | B_\lambda(\nabla^2) | \xi \rangle$) and contribution to force ($\langle F | B_\lambda(\nabla^2) | F \rangle$) from different Laplace modes.
- : Force ($\frac{\partial S_i}{\partial U}$): amplitude for low- and high- Laplace modes can be extracted during HMC
- Response to observable: Apply $B_\lambda(\nabla^2)$ to Initial momenta, measure change in observables (H, Wilson Flowed energy $\langle E \rangle$, $Q(L/2)$) after a very short ($\tau = 10^{-4} \sim 10^{-3}$) trajectory.

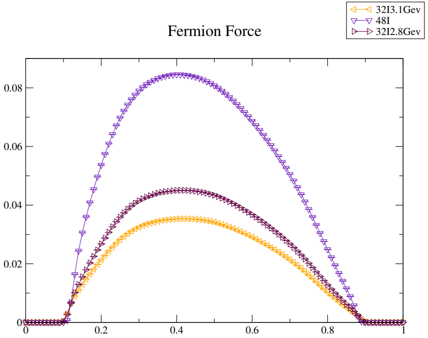


Laplace mode density and Force Density for quenched ensembles

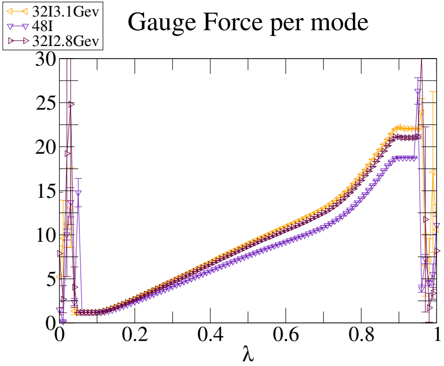


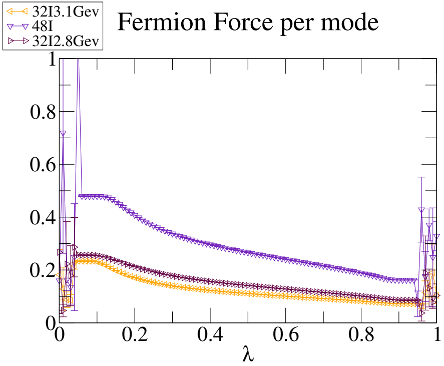
Analysis of HMC forces with Laplace modes for dynamical ensembles



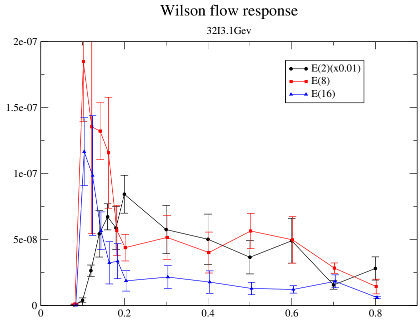


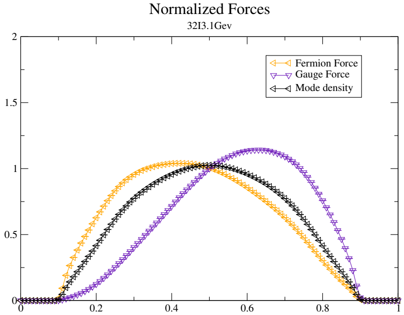
Gauge and Fermion force density per mode for dynamical esembles





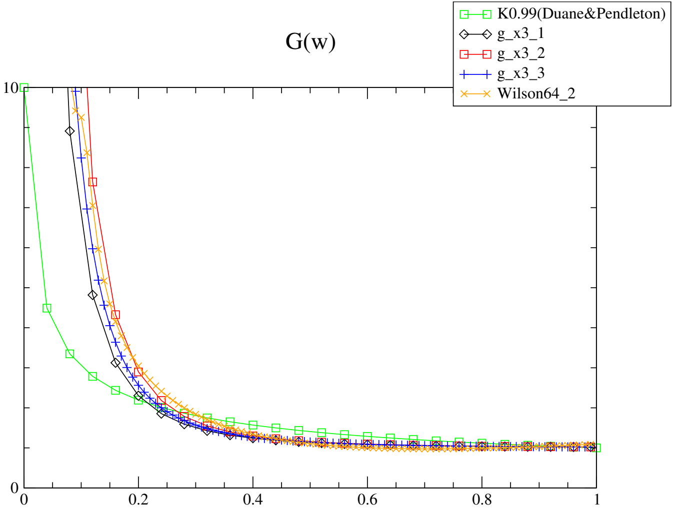
Force density and wilson flow response for Laplace modes for 32I3.1Gev ensemble





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

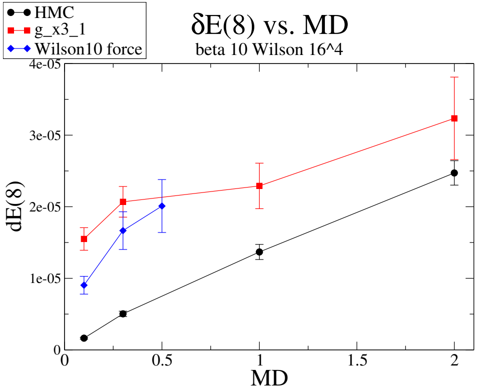
Choice of RMHMC mass kernel

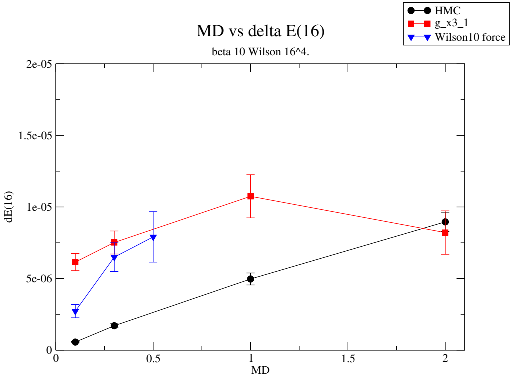


g_x3_[123] $\sim 1 + \frac{c}{(x+b)^3}$

Duane & Pendleton: $\sim \sqrt{\frac{1}{(1-\kappa)+\kappa x}}$

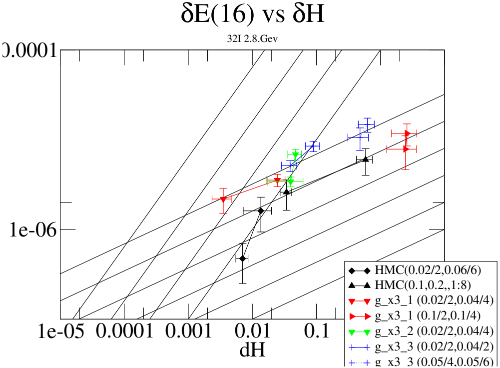
Quenched $\beta = 10$ Wilson ensemble

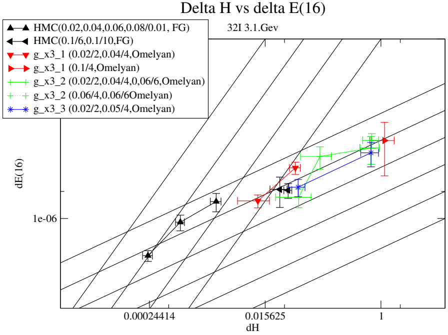




A significant gain shown for short trajectory length.

Single step RMHMC 2+1-flavor ensembles



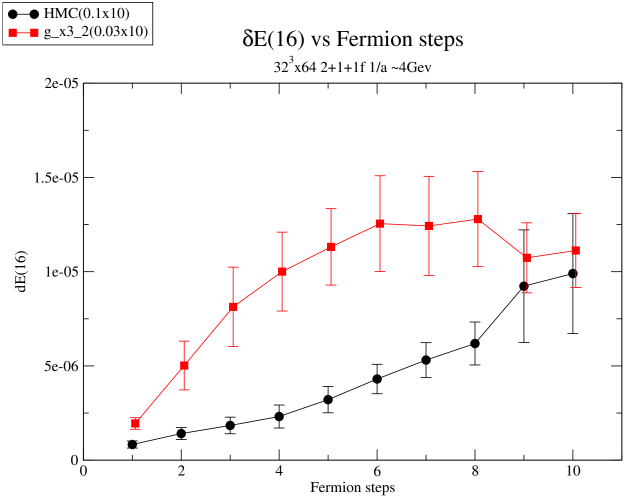


$1/a \sim$2.8Gev, near physical

$1/a \sim$3.1Gev, heavier than physical ($m_\pi \sim$ 300Mev)

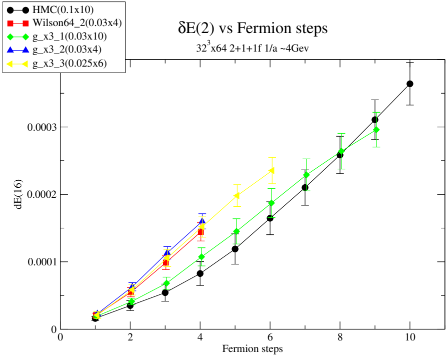
Factor of $\sim$ 2 on physical mass ensemble.

RMHMC on 2+1+1 flavor $1/a \sim$ 4Gev ensemble, full trajectory

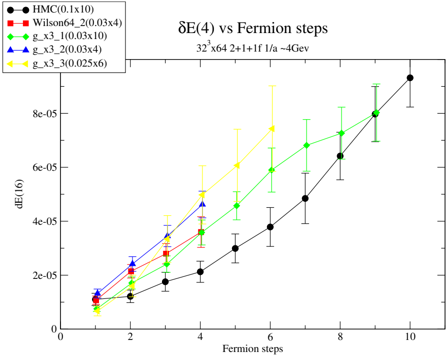


ECP Highlights, 2023 Q1

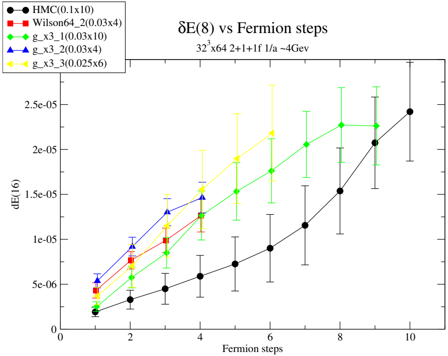
E(2)


δE(2) vs Fermion steps
32^3x64 2+1+1f 1/a ~4Gev
HMC(0.1x10)
Wilson64_2(0.03x4)
g_x3_1(0.03x10)
g_x3_2(0.03x4)
g_x3_3(0.025x6)
dE(16)
0.0003
0.0002
0.0001
0
0
2
4
6
8
10
Fermion steps

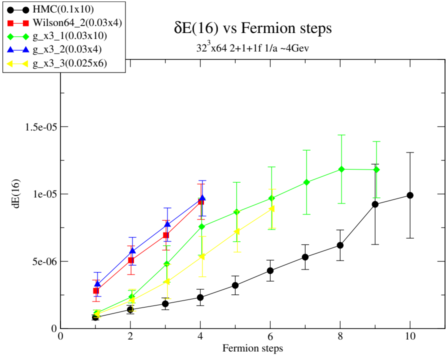
E(4)


δE(4) vs Fermion steps
32^3x64 2+1+1f 1/a ~4Gev
HMC(0.1x10)
Wilson64_2(0.03x4)
g_x3_1(0.03x10)
g_x3_2(0.03x4)
g_x3_3(0.025x6)
dE(16)
8e-05
6e-05
4e-05
2e-05
0
0
2
4
6
8
10
Fermion steps

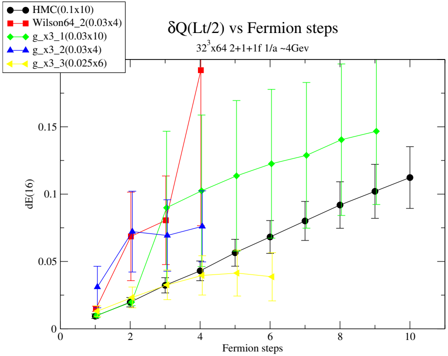
E(8)


δE(8) vs Fermion steps
32^3x64 2+1+1f 1/a ~4Gev
HMC(0.1x10)
Wilson64_2(0.03x4)
g_x3_1(0.03x10)
g_x3_2(0.03x4)
g_x3_3(0.025x6)
dE(16)
Fermion steps
0
5e-06
1e-05
1.5e-05
2e-05
2.5e-05
0
2
4
6
8
10

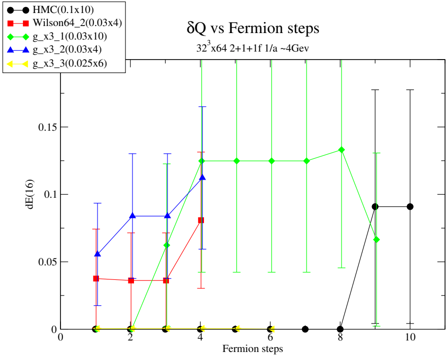
E(16)


δE(16) vs Fermion steps
32^3x64 2+1+1f 1/a ~4Gev
HMC(0.1x10)
Wilson64_2(0.03x4)
g_x3_1(0.03x10)
g_x3_2(0.03x4)
g_x3_3(0.025x6)
dE(16)
Fermion steps
0
5e-06
1e-05
1.5e-05
0
2
4
6
8
10

$Q(L_t/2)$


HMC(0.1x10)
Wilson64_2(0.03x4)
g_x3_1(0.03x10)
g_x3_2(0.03x4)
g_x3_3(0.025x6)
δQ(Lt/2) vs Fermion steps
32^3x64 2+1+1f 1/a ~4Gev
dE(16)
Fermion steps
0
0.05
0.1
0.15
0
2
4
6
8
10

Q

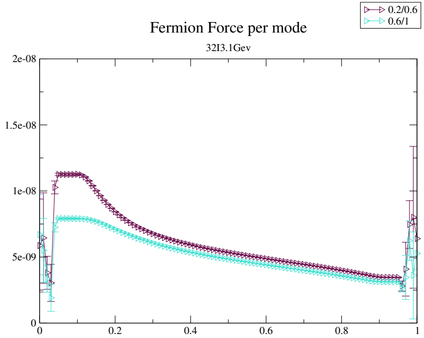

HMC(0.1x10)
Wilson64_2(0.03x4)
g_x3_1(0.03x10)
g_x3_2(0.03x4)
g_x3_3(0.025x6)
δQ vs Fermion steps
32^3x64 2+1+1f 1/a ~4Gev
dE(16)
Fermion steps
0
0.05
0.1
0.15
0
2
4
6
8
10

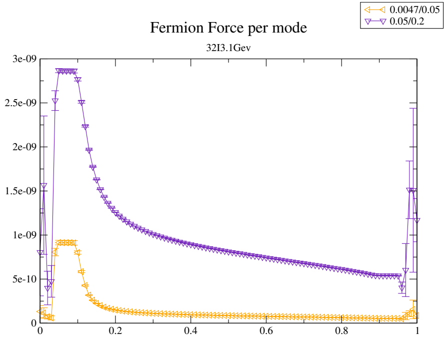
Conclusion & Discussion

- Efficacy of finer lattice spacing/larger lattice volume evolution critically depends on controlling critical slowing down in HMC.
- Low modes of Laplace operator strongly correlates with long range mode probed by wilson flowed energy. Fourier Acceleration with Laplace operator can be tuned to accelerate slow-moving low momentum modes measured by smeared energy significantly faster than normal HMC.
- RMHMC on one of the 2+1+1 flavor ($1/a \sim 4$Gev) ensembles shows factor of ~ 2 improvement in untis in terms of fermion force evaluation.
- What is the nature of the turnover/saturation?
- Can we use it now? Not yet translating to the reduction in walltime per trajectory, due to time spent on Laplace operator evaluation .
  The need for taking small stepsize for MD,especially for gauge ($\sim 2 \times 10^{-3}$) + iterative solve for implicit solver increase the iteration count significantly.
- Room for algorithmic improvement: Algorithms for communicaiton avoidance, bandwidth reduction, etc being investigated.
- We only tried Laplace operator so far. Other choices (Hessian...) being explored.

Thank you!

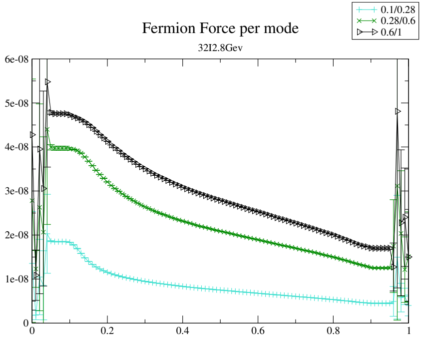
Fermion force density per mode for 32I3.1Gev esembles

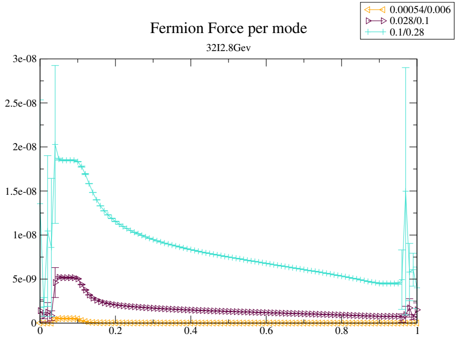




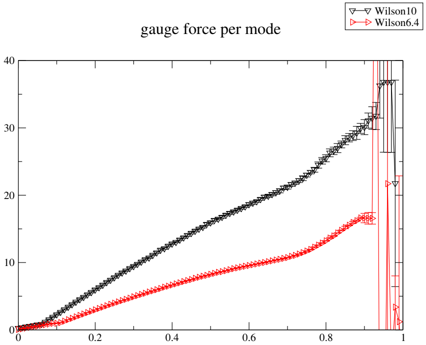
Force for small Hasenbusch mass relatively larger for small Laplace mode, but the force overall small

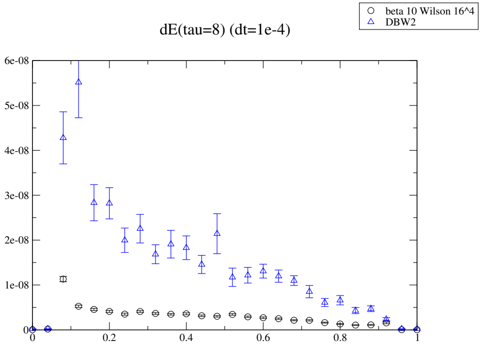
Fermion force density per mode for 32I2.8Gev esembles





Force density and wilson flow response for quenched emsemble





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