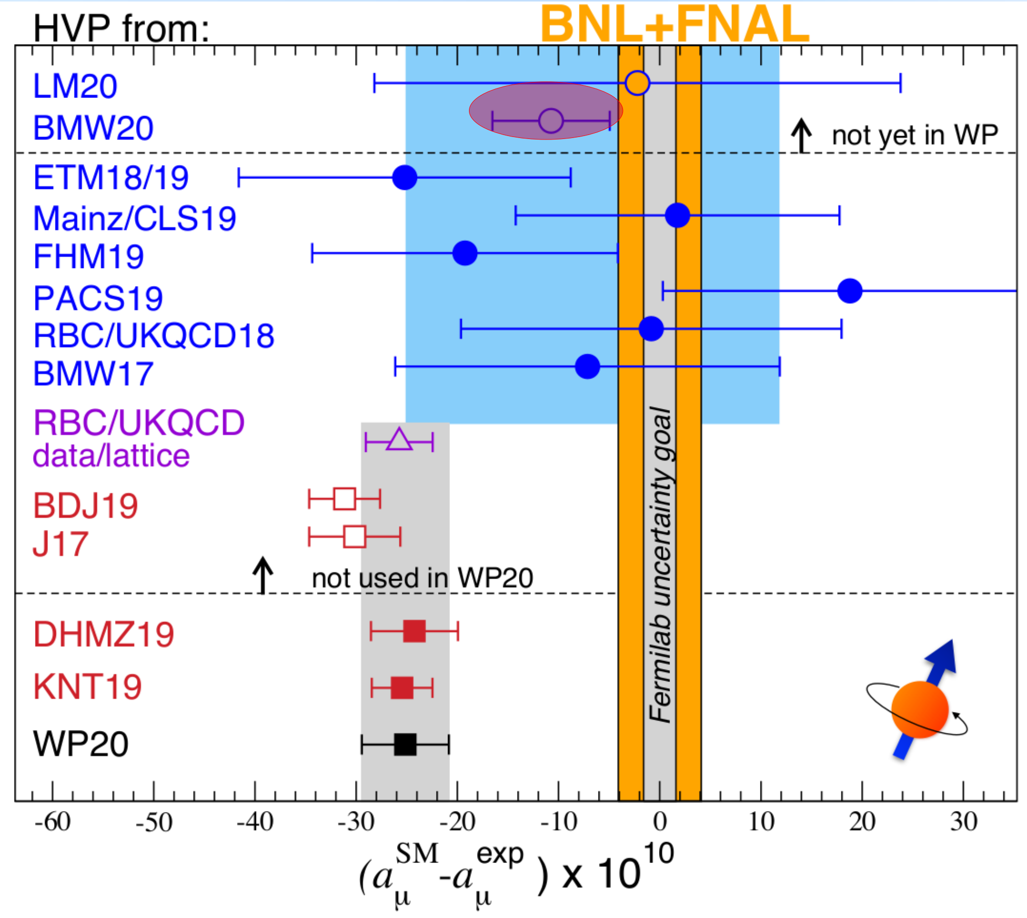
Computing hadronic vacuum polarization on 0.048 fm lattices with GPUs

Gen Wang
Centre de Physique Théorique de Marseille

Budapest-Marseille-Wuppertal collaboration (BMWc)

Lattice 2023, Fermilab, 1 Aug, 2023

Motivation



Colangelo et al., arXiv:2203.15810

- Experimental result is 4.2σ SM predictions (WP Aoyama et al., 2020)
- BMW20 result 707.5(2.3)(5.0) is 2.1σ higher than R-ratio and consistent with experiment value at 1.5σ level

	β	a[fm]	$L \times T$
	3.7000	0.1315	48×64
	3.7500	0.1191	56×96
	3.7753	0.1116	56×84
	3.8400	0.0952	64×96
	3.9200	0.0787	80×128
	4.0126	0.0640	96×144
NEW ⟶	4.1479	0.0483	128×192

Nature 593 (2021) no.7857, 51-55

- Omega baryon measurements
- Staggered eigensystem on GPU

Omega baryon operators

- For the Ω baryon, two staggered baryon operators in [1] and an operator which only couples to a single taste [2]

$$\Omega_{\mathrm{VI}}(t) = \sum_{x_k even} \epsilon_{abc} \left[S_1\chi_a S_{12}\chi_b S_{13}\chi_c - S_2\chi_a S_{21}\chi_b S_{23}\chi_c + S_3\chi_a S_{31}\chi_b S_{32}\chi_c \right](x)$$

$$\Omega_{\mathrm{XI}}(t) = \sum_{x_k even} \epsilon_{abc} \left[S_1\chi_a S_2\chi_b S_3\chi_c \right](x)$$

$$\Omega_{\mathrm{Ba}}(t) = [2\delta_{\alpha 1}\delta_{\beta 2}\delta_{\gamma 3} - \delta_{\alpha 3}\delta_{\beta 1}\delta_{\gamma 2} - \delta_{\alpha 2}\delta_{\beta 3}\delta_{\gamma 1} + (\cdots \beta \leftrightarrow \gamma \cdots)]$$

$$\sum_{x_k even} \epsilon_{abc} \left[S_1\chi_{a\alpha} S_{12}\chi_{b\beta} S_{13}\chi_{c\gamma} - S_2\chi_{a\alpha} S_{21}\chi_{b\beta} S_{23}\chi_{c\gamma} + S_3\chi_{a\alpha} S_{31}\chi_{b\beta} S_{32}\chi_{c\gamma} \right](x)$$

- Wuppertal smearing connects 2a lattice spacing

$$\left[\hat{W}v\right]_x = (1-\sigma)v_x + \frac{\sigma}{6}\sum_{\mu=1,2,3}\left(U^{3d}_{\mu,x} U^{3d}_{\mu,x+\mu} v_{x+2\mu} + U_{u,x-\mu} U^{3d,\dagger}_{\mu,x-2\mu} v_{x-2\mu} \right)$$

- Use CPU memory as buffers for smeared propagator
- 128^3 x 196 could fit into 64 V100 GPUs with around 1.6 TFlops per GPU

[1] M. F. L. Golterman and J. Smit, Nucl. Phys. B 255 (1985), 328-340
[2] J. A. Bailey, Phys. Rev. D 75 (2007), 114505 [arXiv:hep-lat/0611023

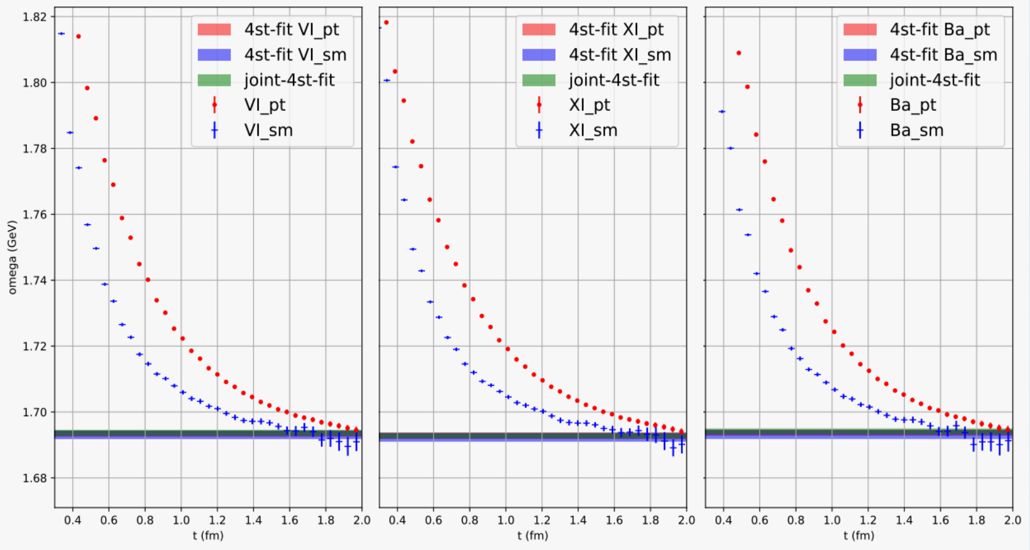
Four-state fits

- A four-state fit function with two positive and two negative parity states:

$$h(t,A,M) = A_0 h_+(M_0,t) + A_1 h_-(M_1,t) + A_2 h_+(M_2,t) + A_3 h_-(M_3,t)$$

$$h_+(M,t) = e^{-Mt} + (-t)^{t-1} e^{-M(T-t)}$$
$$h_-(M,t) = -h_+(M,T-t)$$

s	$\mu_s \cdot 1672$ MeV	$\delta\mu_s$
1	2012 MeV	0.10
2	2250 MeV	0.10
3	2250 MeV	0.15



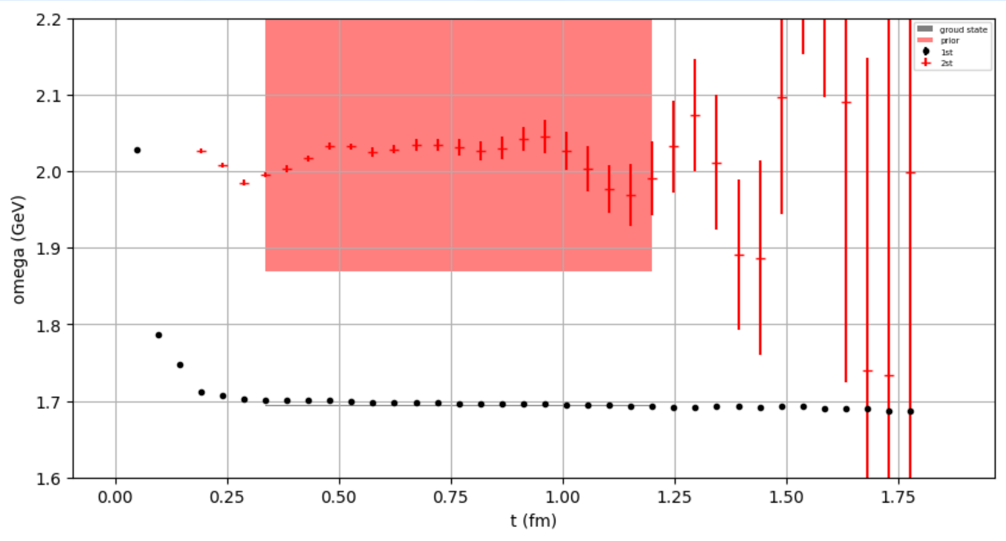
- Priors
  - s=1 motivated from Belle collaboration [1]
  - s = 2, 3, from the quark model [2]
- Fits to smear to smear/point propagators
- Joint fits of above two propagators
- The results are consistent across all fits and operators

[1] Yelton, J. et al. Phys. Rev. Lett. 121, 052003. arXiv:1805.09384 [hep-ex] (2018)
[2] Capstick, S. & Isgur, Phys. Rev. D34. [AIP Conf. Proc.132,267(1985)], 2809 (1986)

GEVP method

Construct matrix [1] from folded propagator C(t) :

$$H(t) = \begin{pmatrix} C_{t+0} & C_{t+1} & C_{t+2} & C_{t+3} \\ C_{t+1} & C_{t+2} & C_{t+3} & C_{t+4} \\ C_{t+2} & C_{t+3} & C_{t+4} & C_{t+5} \\ C_{t+3} & C_{t+4} & C_{t+5} & C_{t+6} \end{pmatrix}$$



[1] C. Aubin and K. Orginos, AIP Conf. Proc. 1374 (2011) no.1, 621-624

Solve GEVP at two time slices:

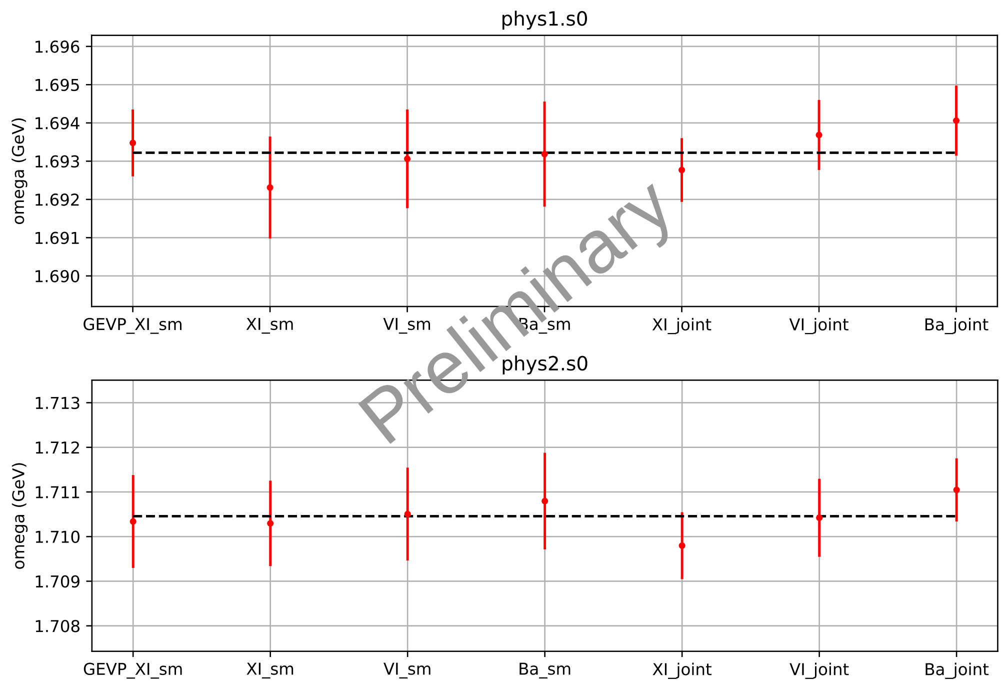
$$H(t_a)v(t_a,t_b) = \lambda(t_a,t_b)H(t_b)v(t_a,t_b)$$

Extract mass from latter time slices:

$$C(t) = v^\dagger(t_a,t_b)H(t)v(t_a,t_b)$$

- Group state can be observed at very early time slice
- GEVP excited state agrees very well with prior in the 4-state fit

Omega mass

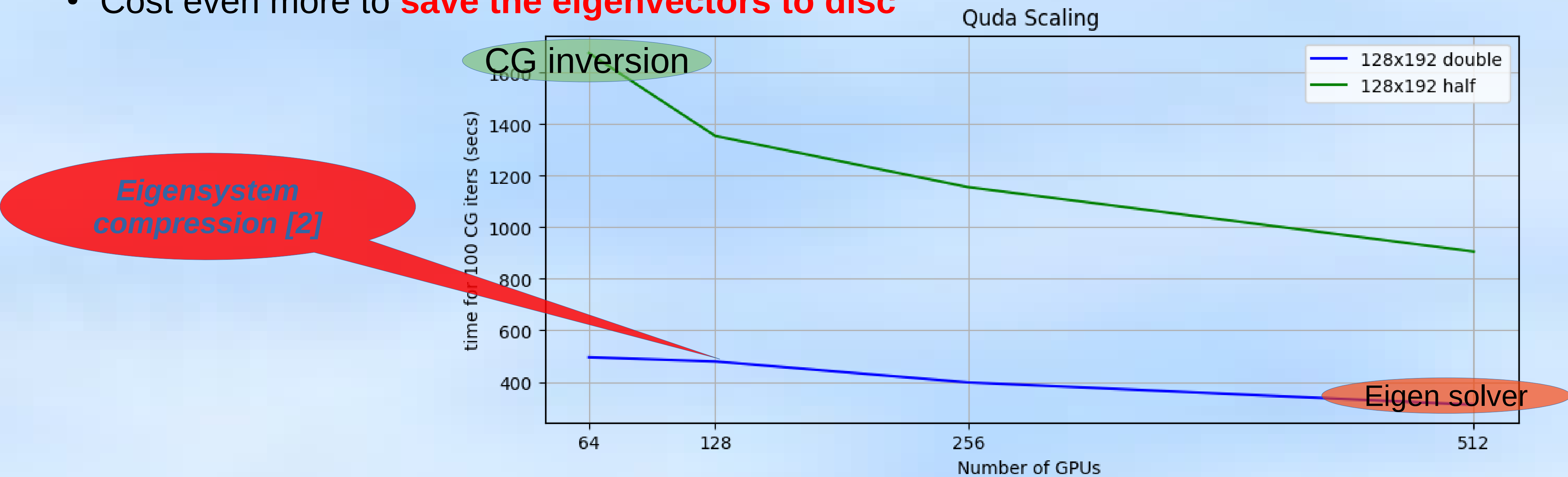


- Results on phys1 and phys2 (different quark masses) with a=0.048 fm
  - GEVP fits
  - Four-state fits to smear-smear
  - Joint fit with smear-point/smear propagator
- All the fits are consistent with each other with noise/signal ratio < 0.1%

- Omega baryon measurements
- Staggered eigensystem on GPU

Staggered-fermion performance with QUDA

- Staggered even-even eigensystem as preconditioning for light quark CG inversion
  - Only 64 GPUs needed
- Need 512 GPUs to solve the eigensystem
  - All double precision eigenvectors on GPU
  - Cost even more to **save the eigenvectors to disc**



[1] M. A. Clark, R. Babich, K. Barros, R. Brower, C. Rebbi and et. al. "https://github.com/lattice/quda"
[2] M.A. Clark, Chulwoo Jung, C. Lehner, EPJ Web Conf. 175 (2018) 14023, arXiv:1710.06884

Eigensystem compression

1) Compute N basis vectors with lowest eigen values

$$\vec{v}_i^N \to V \times s \times c$$

2) For a given block (4x4x4x4), create a locally orthogonal basis with the N basis vectors

$$\vec{v}_i^N \to L_B \times \vec{B} \text{ with } L_B = V \times s \times c/B \qquad \vec{u} = \left[\sum_{n=0}^{N} a_n^b \vec{B}_n^b\right]_{b \in L_B} + \vec{X}_u$$

3) Compress ratio

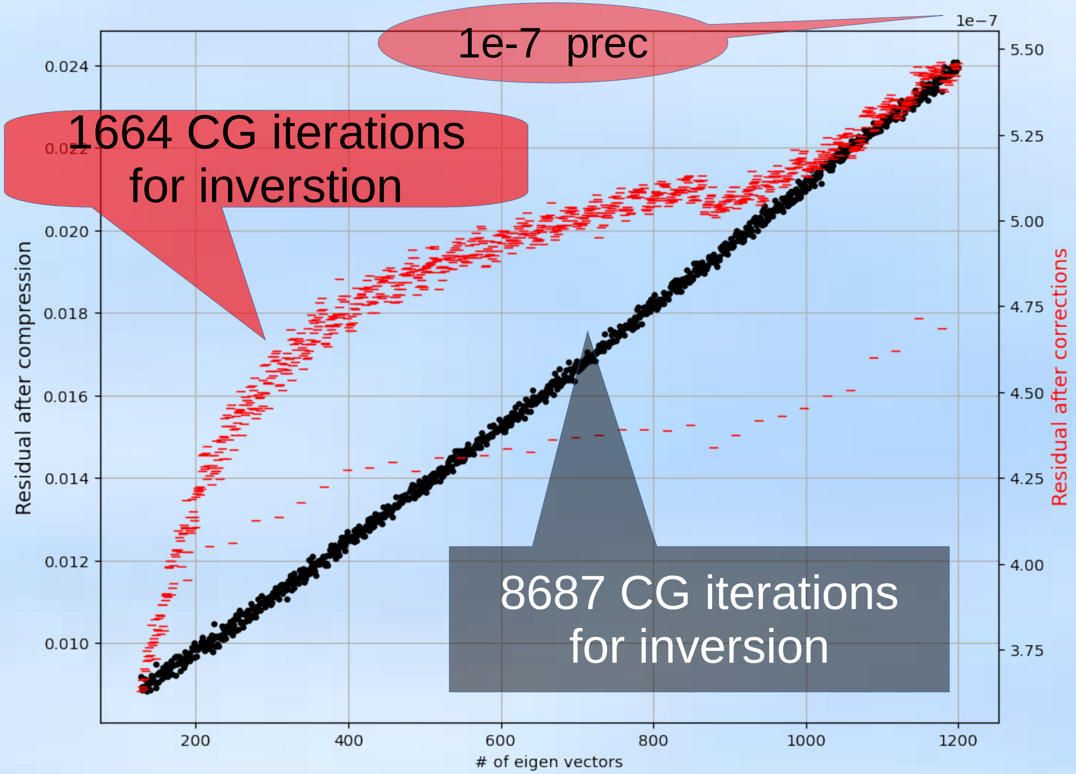
$$\text{Ratio} = \frac{4\times4\times4\times4\times s\times c}{N}$$

4) Solve the K basis vectors on the coarse Grid to obtain the full set of eigenvectors

5) Compute the eigenvalues of the K vectors

[1] M.A. Clark, Chulwoo Jung, C. Lehner, EPJ Web Conf. 175 (2018) 14023, arXiv:1710.06884

Staggered eigensystem compression



- Compute N + K =128 + 1072 full eigen vectors
- Use N=128 as the basis to compress K = 1072 vectors
- Calculate the residual of reconstructed K's

$$r \equiv ||\vec{u}_d - (\vec{u}_d^\dagger \vec{u}')\vec{u}'||^{-1/2},\ \vec{u}_d = D_{eo}D_{oe}\vec{u}'$$

$$\vec{u}' = \vec{u} - \vec{X}_u$$

- Two step of CG corrections

$$D_{eo}D_{oe}\vec{u}'' = \vec{u}',\ D_{eo}D_{oe}\vec{u}''' = \vec{u}''$$

- Calculate the residual of u'''

Conclusions

- Omega mass on 0.048 fm lattice shows consistent results across different fit methods
- Eigensystem compression on staggered fermion is promising for light quark measurements on GPUs

Thank You