

曙光DCU上的格点QCD软件

Yi-Bo Yang (ITP/CAS)



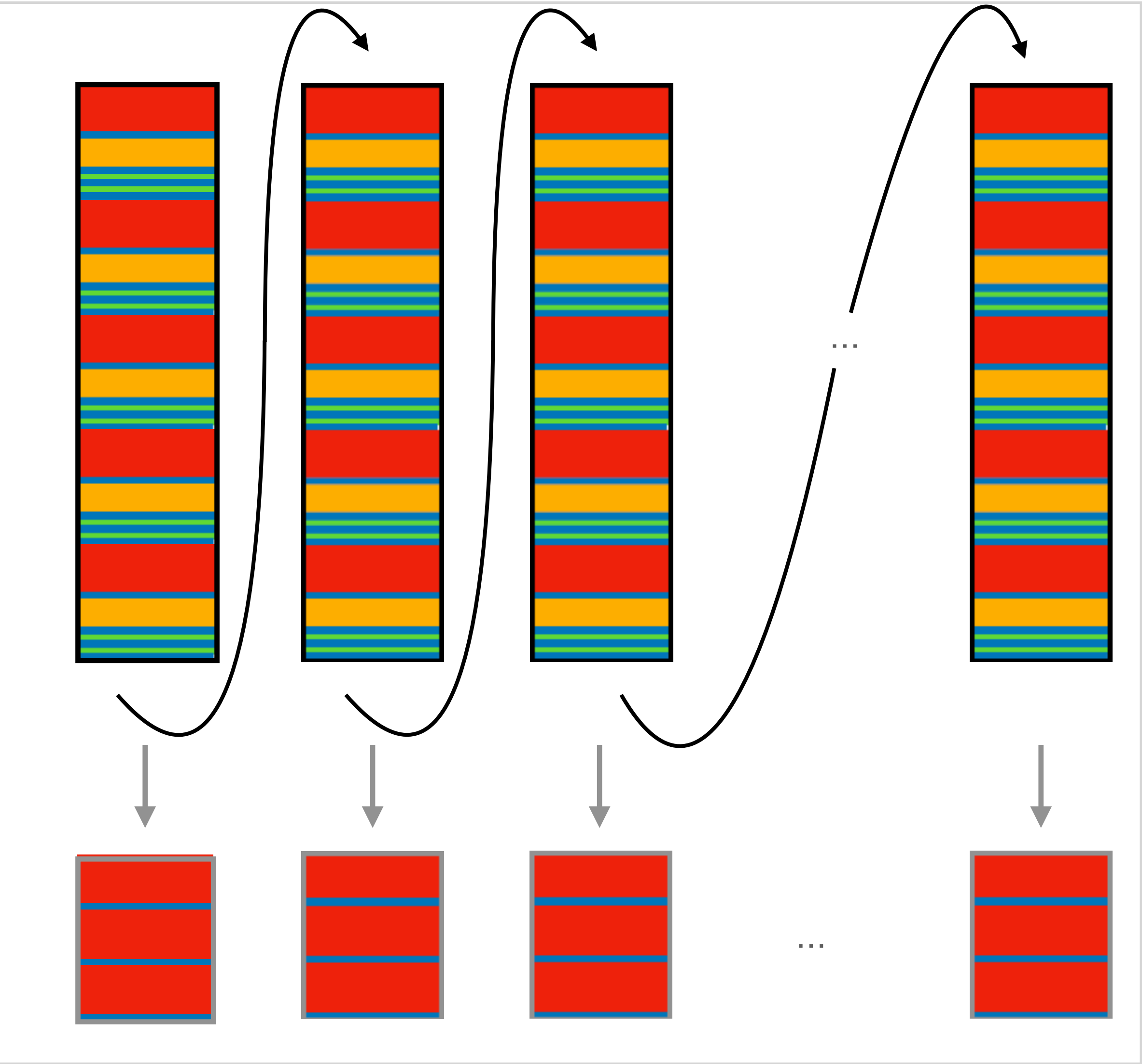
ICTP-AP
International Centre
for Theoretical Physics Asia-Pacific
国际理论物理中心-亚太地区



Peng Sun (IMP/CAS)

格点QCD计算流程与热点

核心计算热点：稀疏线性映射， $\xi = D\eta$



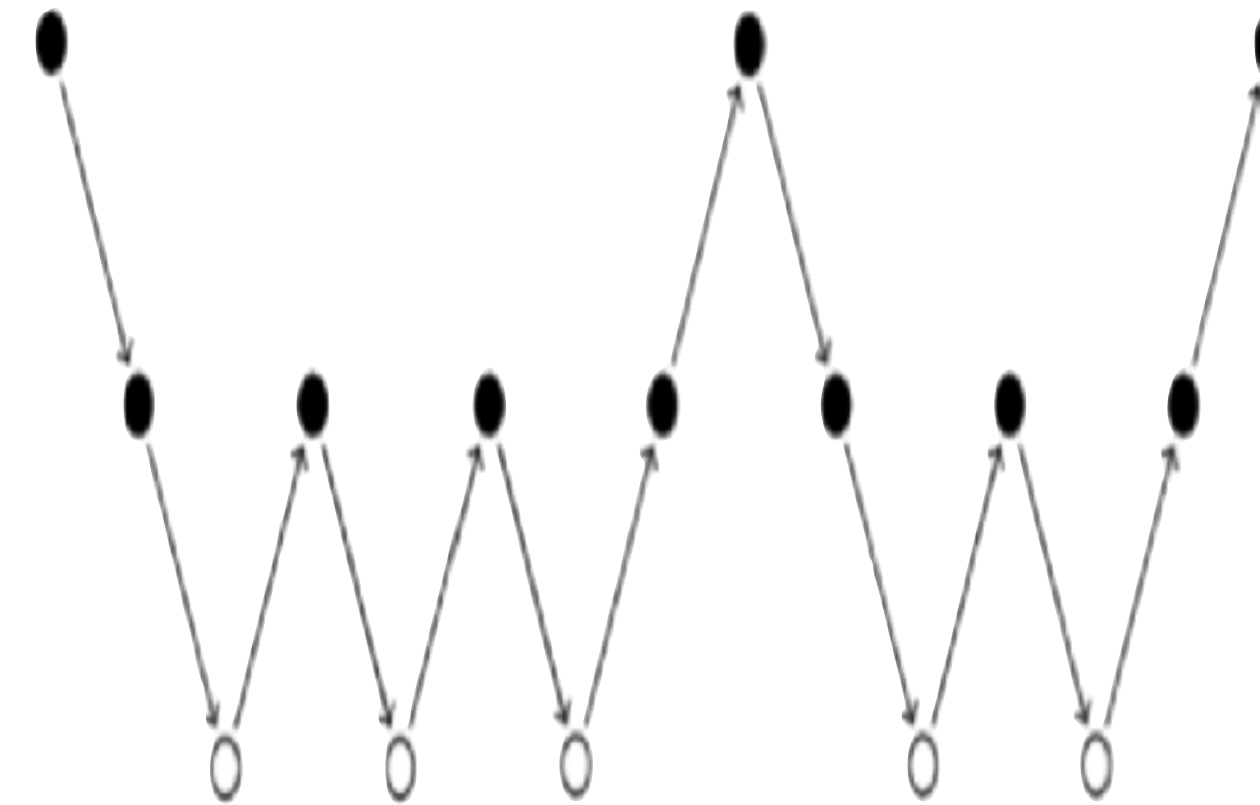
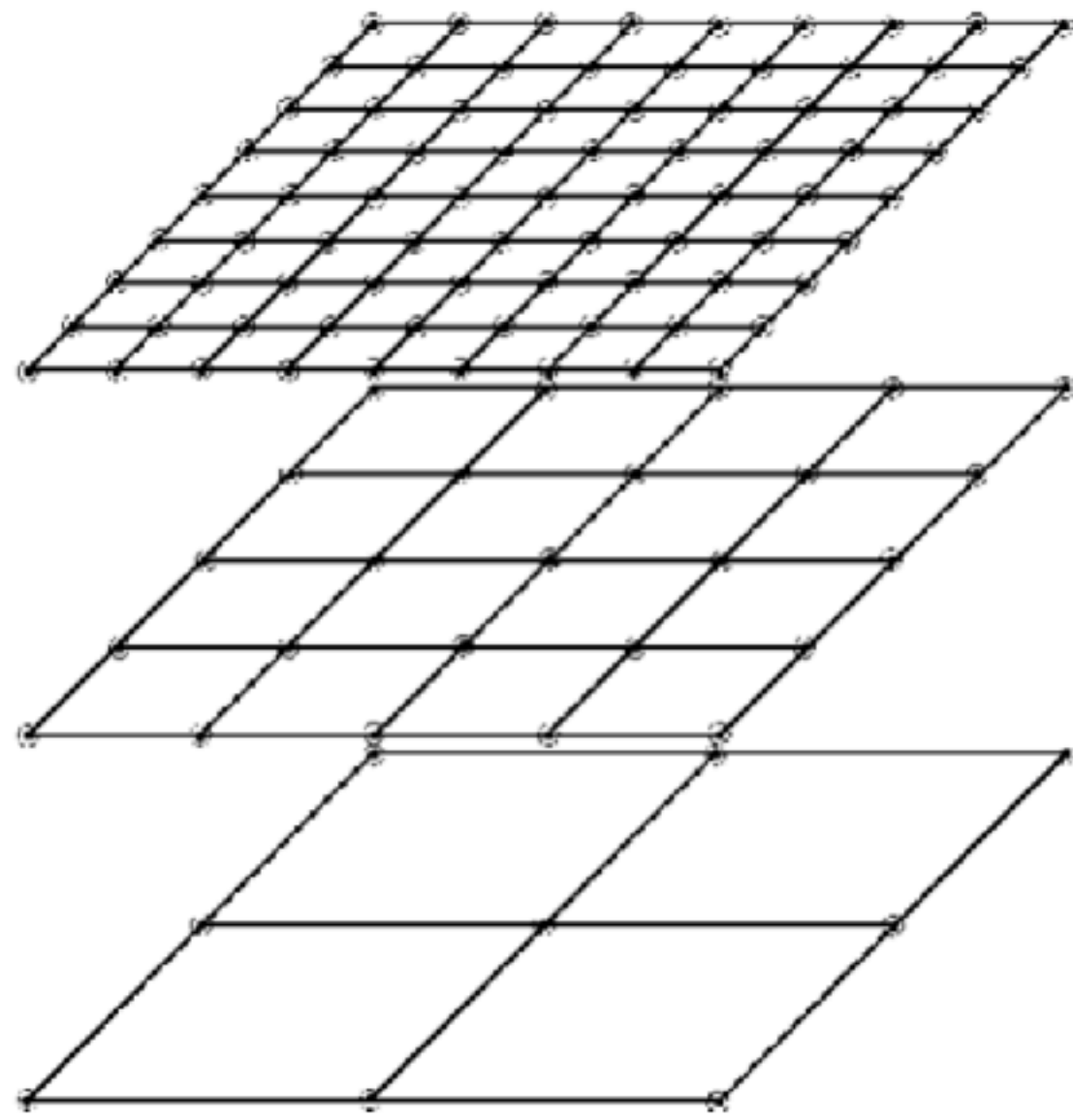
顺序产生重点采样样本(组态)

对各样本进行分析

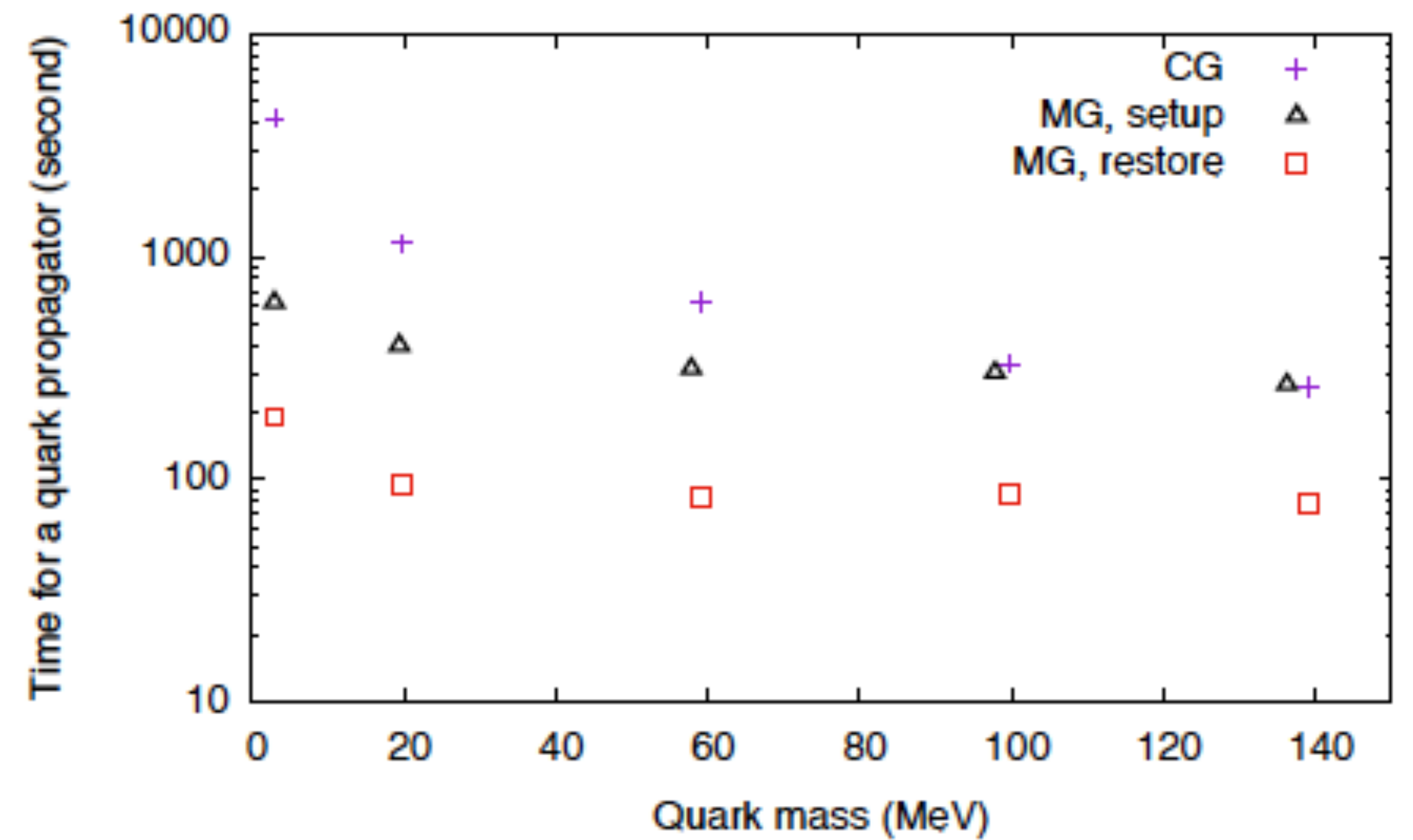
	稀疏线性系统求解, $\xi = D^{-1}\eta = \sum_i c_i D^i \eta + \mathcal{O}(10^{-12})$
	代数操作, $\xi = c_1 \eta_1 + (\eta_2^\dagger \cdot \eta_3) \eta_4$
	稀疏矩阵开平方, $\xi = (D + m_0)^{-1/2} \eta = \sum_i \frac{d_i}{D + m_0 + e_i} \eta + \mathcal{O}(10^{-12})$
	稀疏矩阵求导, $\xi = \frac{\partial D}{\partial U} \eta = D_1 D_2 D_3 \eta$

- 稀疏线性系统求解是主要热点;
- 但加速主要热点后, 矩阵开平方、代数和求导操作将成为新的瓶颈。
- 组态是一切物理分析的基础!

自适应几何多重网格算法



M. A. Clark, A. Strelchenko, M. Cheng, A. Gambhir, and R. Brower, "Accelerating Lattice QCD Multigrid on GPUs Using Fine-Grained Parallelization," International Conference for High Performance Computing, Networking, Storage and Analysis (SC), 2016



- 基于QCD渐近自由的特性，量子涨落效应在短程小，长程大；
- 因而格点QCD天然地适合使用平滑粗粒化的方式使用几何多重网格。
- 在粗粒化格子上低精度求解的结果将在细格子上作为预条件子再次进行迭代以达到更高精度，形成类W形循环。
- 通信量减少也有利于缓解存储墙！

Discretized QCD

Fermion actions

- The naive discretization suffers from the doubling problem:

- $$\mathcal{S}_q^{Naive}(m) = \sum_{x,y} \bar{\psi}(x) D_{Naive}(m; x, y) \psi(y), \quad D_{Naive}(m; x, y) = \frac{1}{2a} \sum_{\mu} \gamma_{\mu} (U_{\mu}(x) \delta_{y, x+a\hat{\mu}} - U_{\mu}^{\dagger}(x - a\hat{\mu}) \delta_{y, x-a\hat{\mu}}) + m \delta_{y,x}$$

- Staggered fermion action:** $\psi^{st}(x) = \gamma_4^{x_4} \gamma_1^{x_1} \gamma_2^{x_2} \gamma_3^{x_3} \psi(x)$, $\{\gamma_1^{st}, \gamma_2^{st}, \gamma_3^{st}, \gamma_4^{st}\} = \{(-1)^{x_4}, (-1)^{x_1+x_4}, (-1)^{x_1+x_2+x_4}, 1\}$;

- Wilson fermion action:** $D + m \rightarrow D + aD^2 + m$

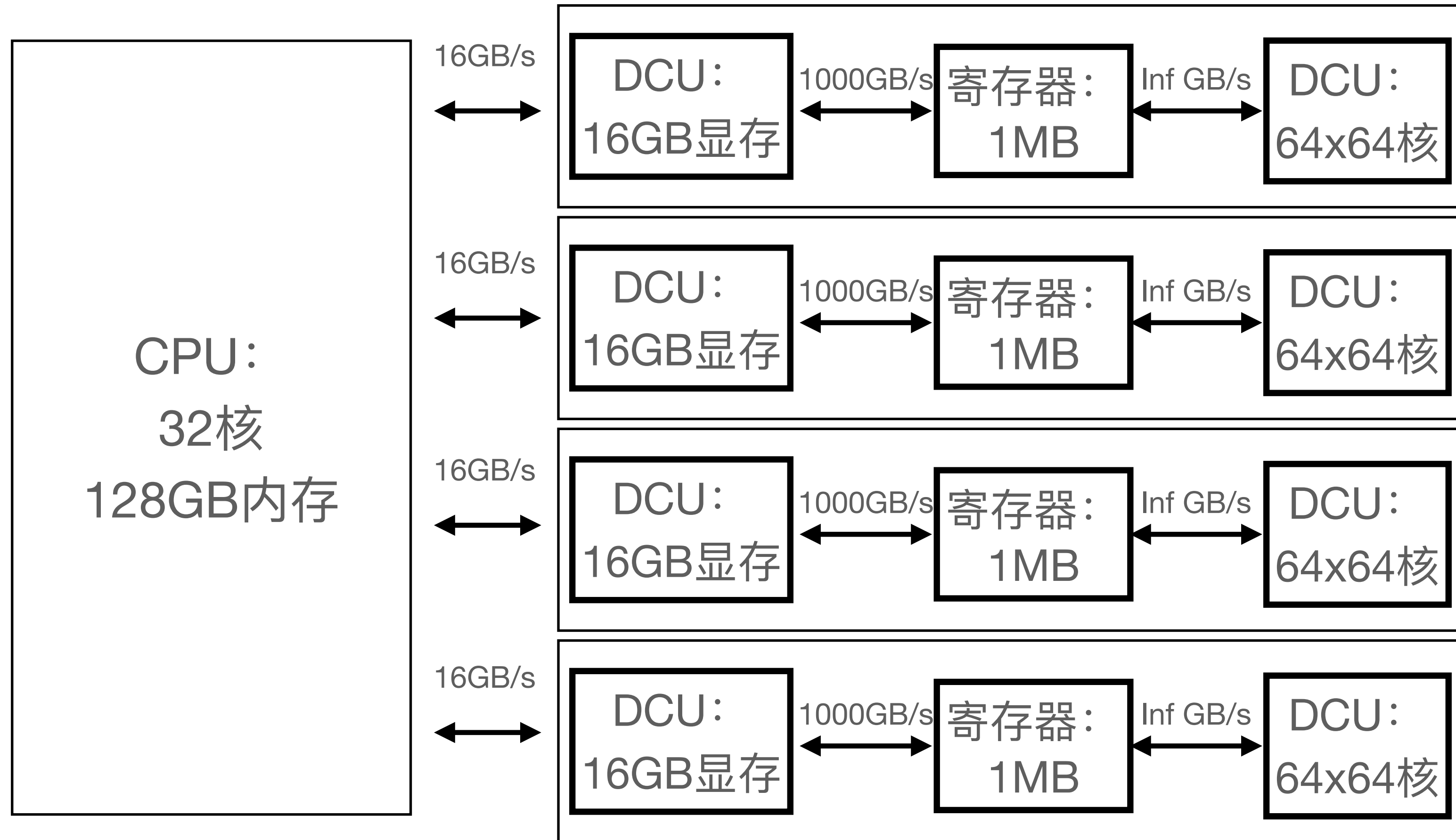
- Clover fermion action:** $D + m \rightarrow D + aD^2 + m + ac_{sw} \sigma_{\mu\nu} F^{\mu\nu}$

- Overlap fermion action:** as a possible solution of: $\gamma_5 D_{GW} + D_{GW} \gamma_5 = \frac{1}{\rho} D_{GW} \gamma_5 D_{GW}$

- Domain wall fermion action:** an approximation of overlap fermion with $\mathcal{O}(10)$ cost of the Wilson/Clover action.



硬件



曙光DCU

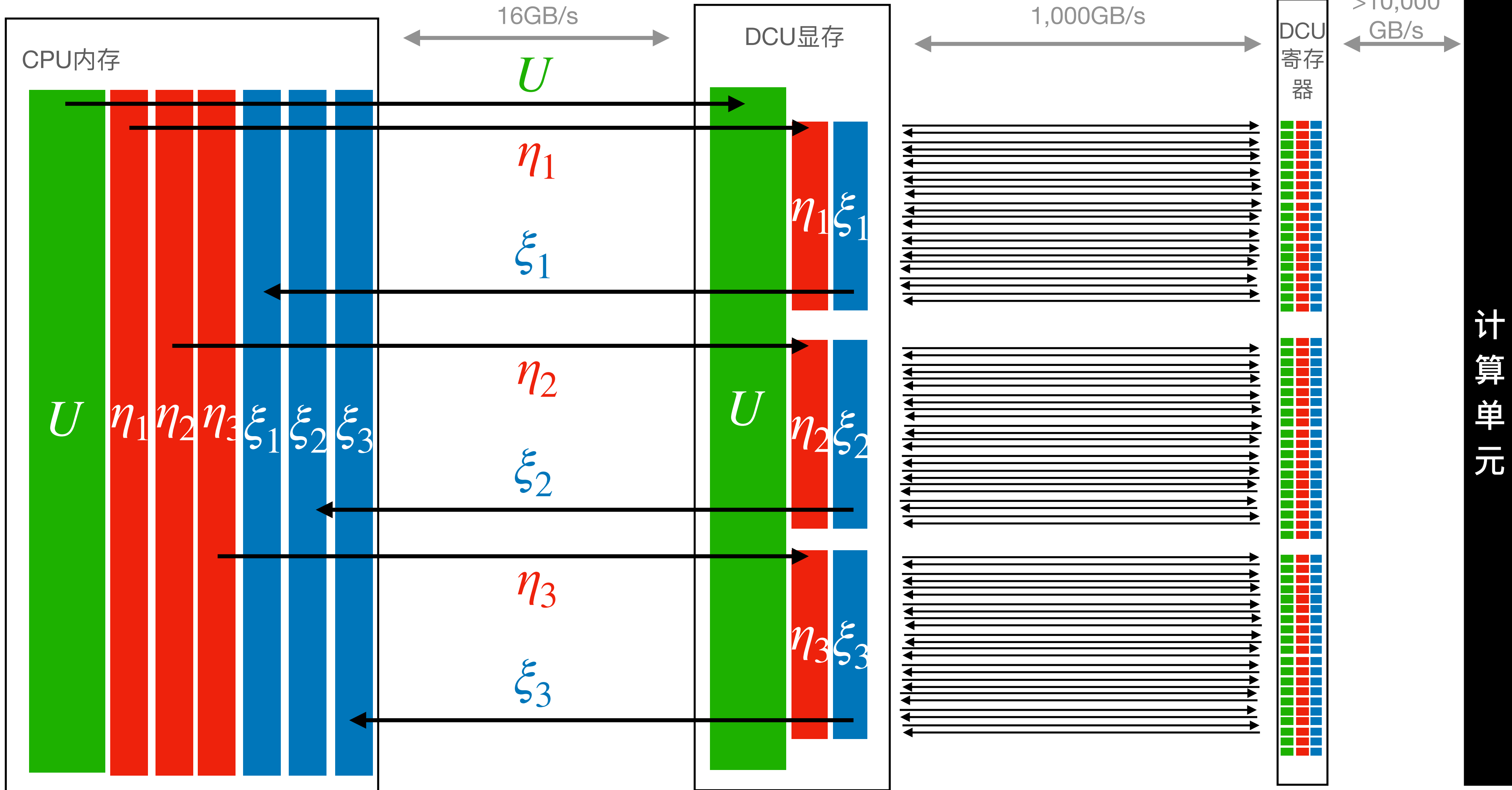
- 实践上可以把大部分数据长期存储在显存上:
- 每个DCU上最大可以存储 48^4 的双精度组态1个, 以及12列传播子8个;
- 或者运行最大 $32^3 \times 12$ 的格子上单精度的Clover-Multigrid solver;
- 结合显存与内存, 可以运行 16^4 的overlap 2-level deflation solver。
- 单节点峰值计算速度约 4-8TFlops (multi rhs or DW).

硬件

稀疏线性系统求解,

$$\xi_j = D(U)^{-1} \eta_j = \sum_i c_i D(U)^i \eta_j + \mathcal{O}(10^{-12})$$

DCU上的计算流程



计算单元

- 寄存器到计算单元带宽非常大，但是容量非常小；
- 显存到寄存器带宽较小，但是容量较大；
- 内存到显存带宽很小，但是容量可以很大。
- GPU/DCU等异构计算中需要尽量充分利用容量有限的寄存器和显存，以降低带宽的瓶颈。

软件 I

QUDA: rocm-devel分支

依托中科先导专项的支持，基于QUDA主线版本移植：

- 2019.03开始移植，2019.06开始调试，2019.11验证multigrid功能正常工作；
- 2020至今在最大3亿格点数的格子上验证移植软件可靠性，并完成多项研究。

Fermion actions	Propagator	Configuration generation	Current domestic users
Clover	Done (Multgrid)	Done (Multgrid)	ITP, IHEP, SJTU, NJNU, SCNU, HNNU
Twisted mass	Done (Multgrid)	N/A	PKU, NWPU
Staggered	Done (Multgrid)	TODO	CCNU, JSU, SCU
Domain wall	TODO	Based on CPS+QUDA	PKU
Overlap	Done (Deflated CG)	N/A	ITP, IHEP, IMP, NJNU, SCNU, HNNU, ZJU

软件 I

Chroma/CPS+QUDA

US Lattice Quantum Chromodynamics

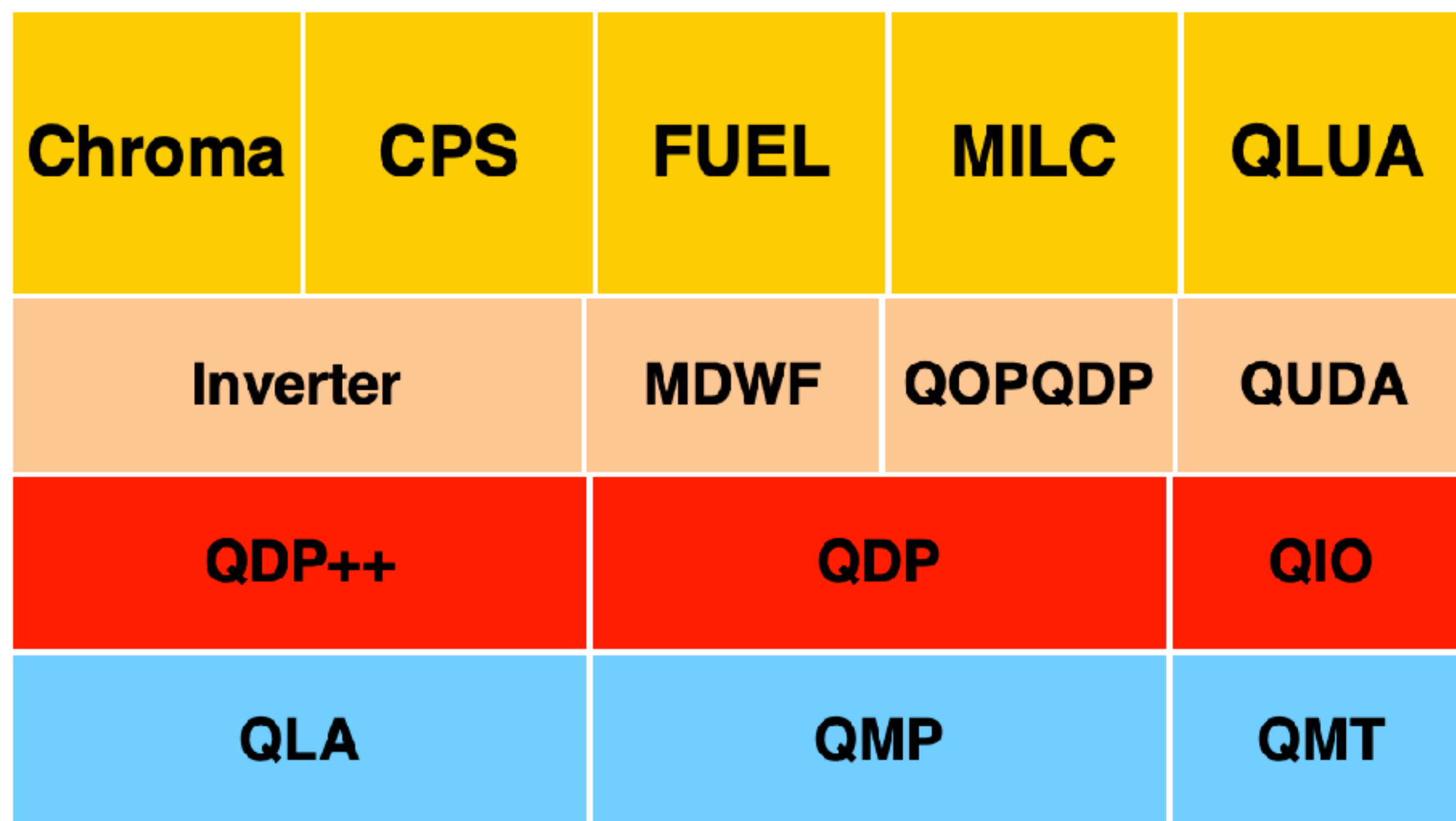


Figure 1: The SciDAC Layers and the software module architecture.

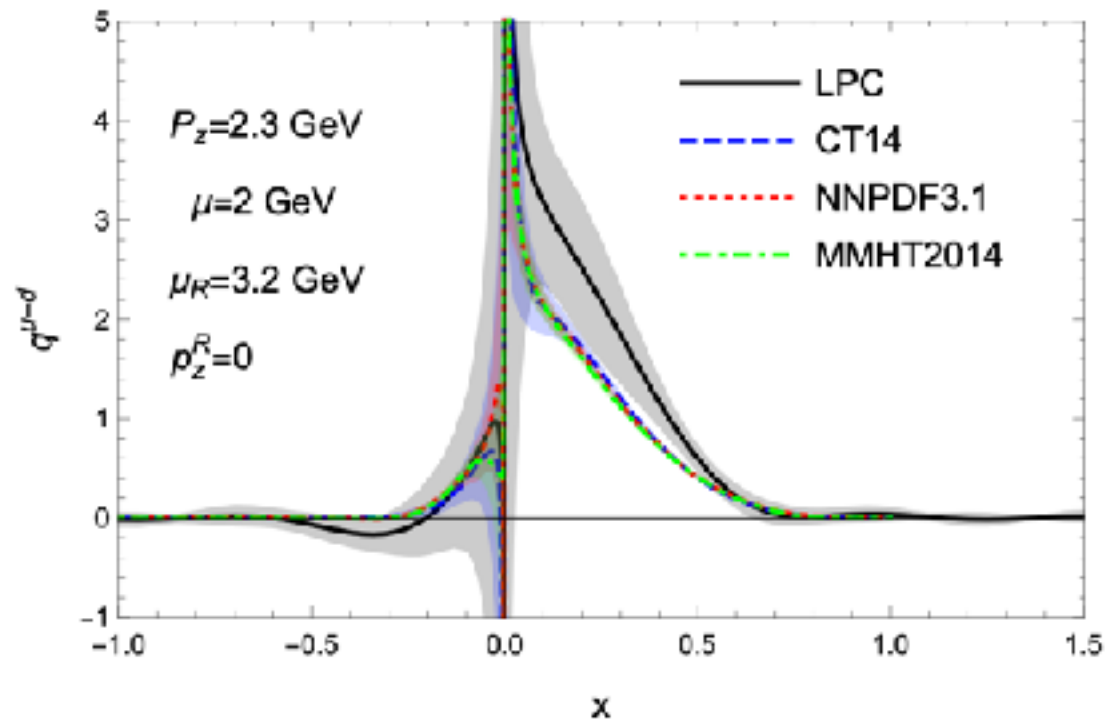
<http://usqcd-software.github.io>

- QUDA作为算法加速模块，可以嵌入Chroma、CPS和MILC等前端软件包，实现完整的格点QCD计算；
- 对于较为便宜的Clover作用量，还需要QDP-JIT等GPU加速模块才能充分地GPU加速；
- 对于更昂贵的DWF和Overlap作用量，QUDA的加速已经基本可以满足需要。

软件I应用

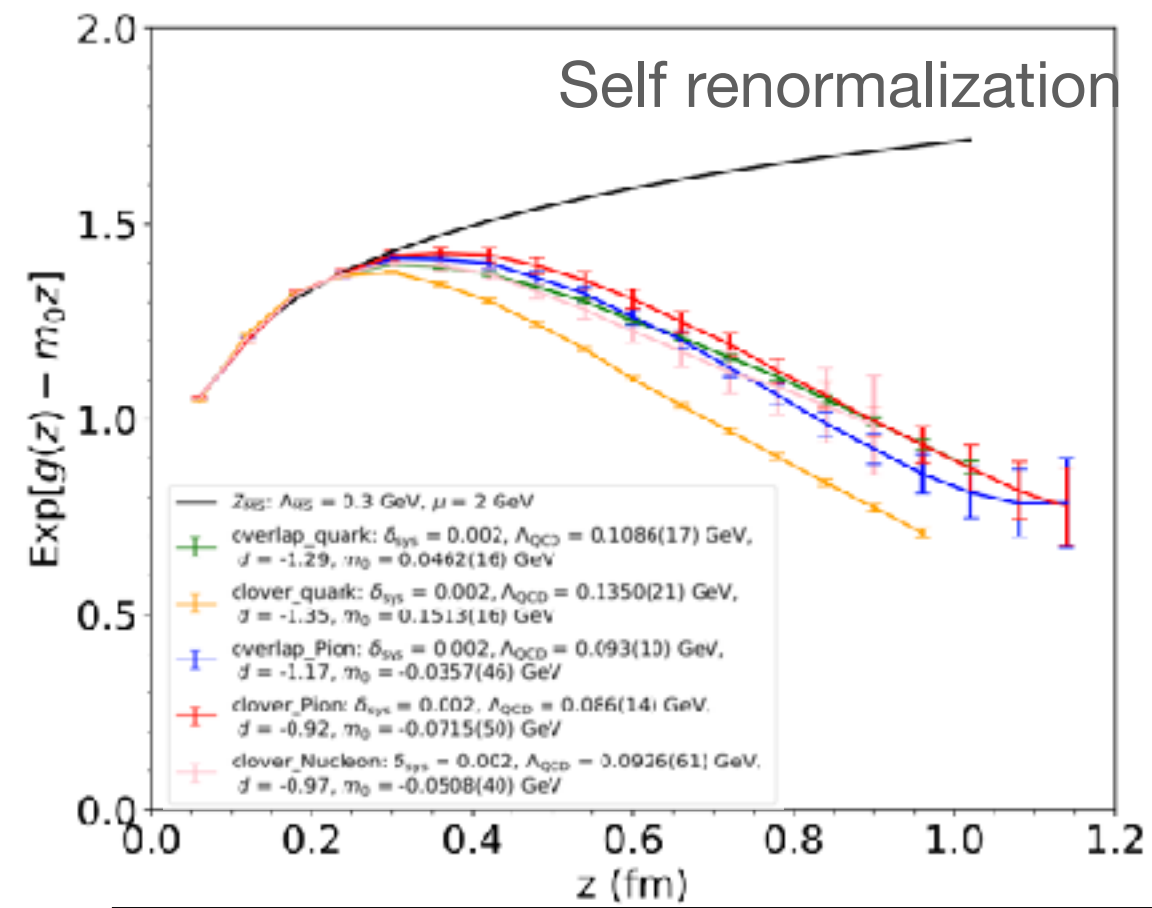


QUADA: rocm-devel分支的应用

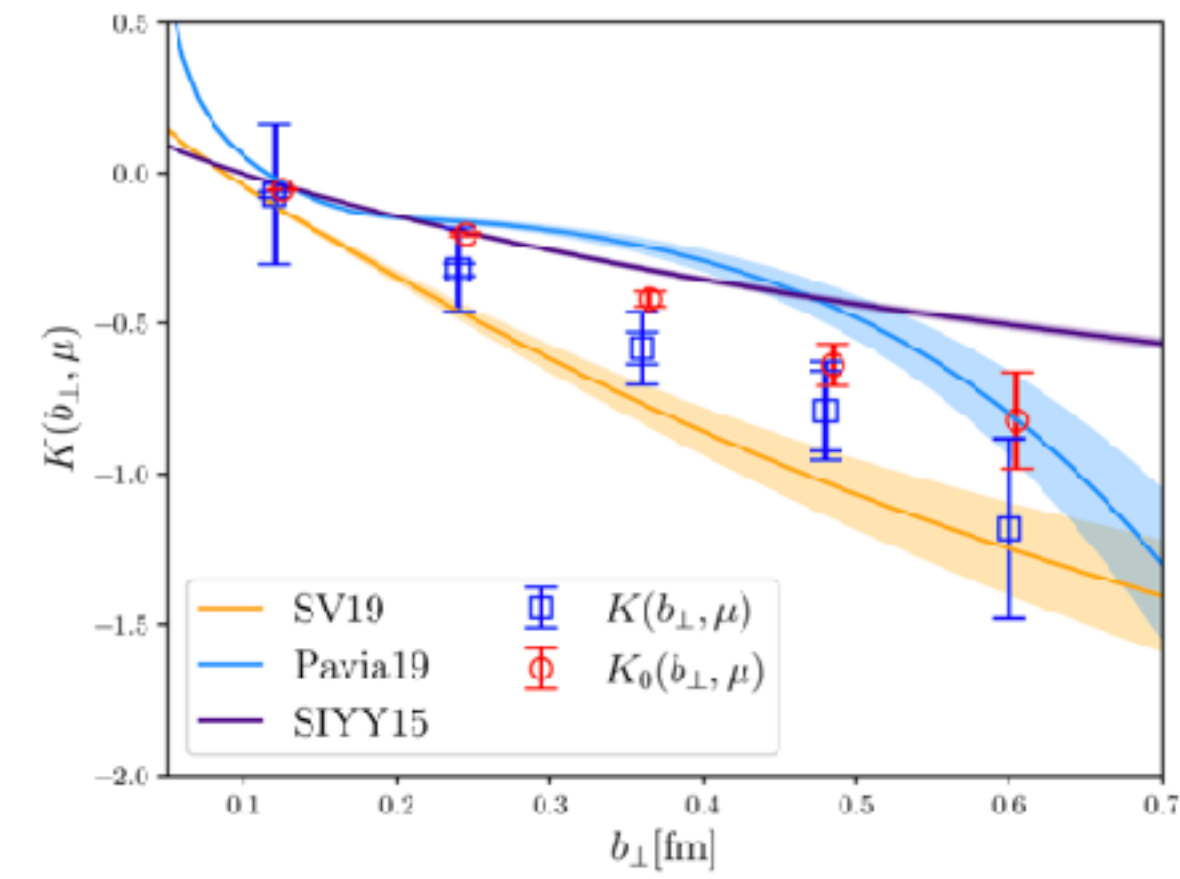


Unpolarized nucleon PDF

Y. Liu, et.al., LPC, PRD101(2020)034020

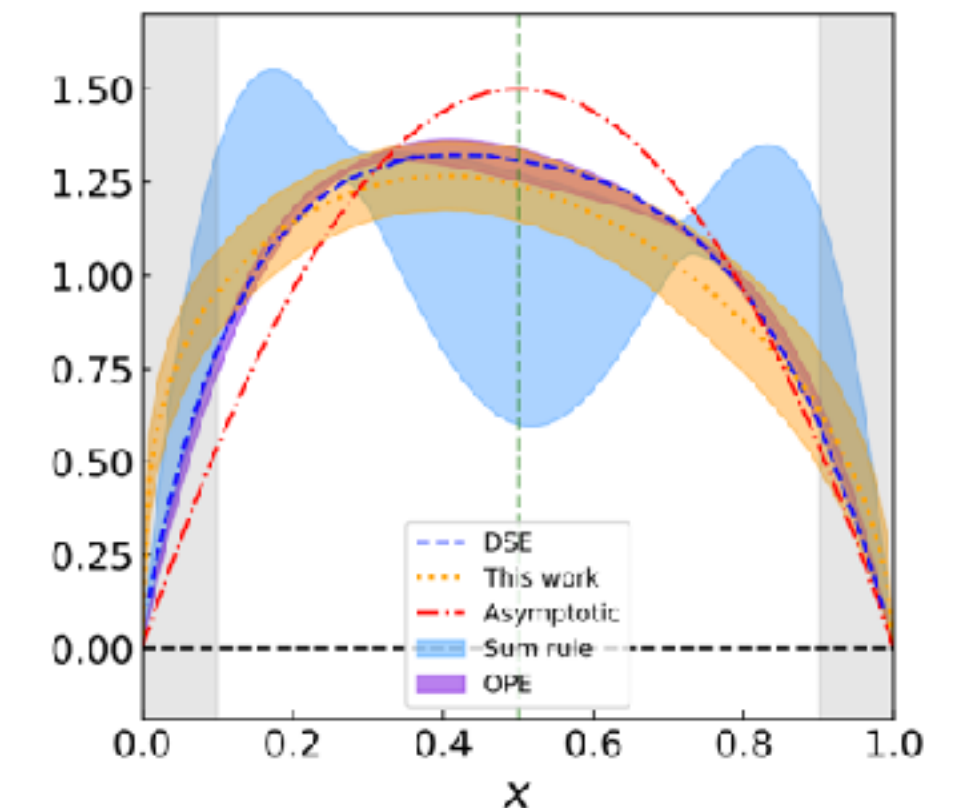


Y.-K. Huo. et.al., LPC, NPB969(2021)115443



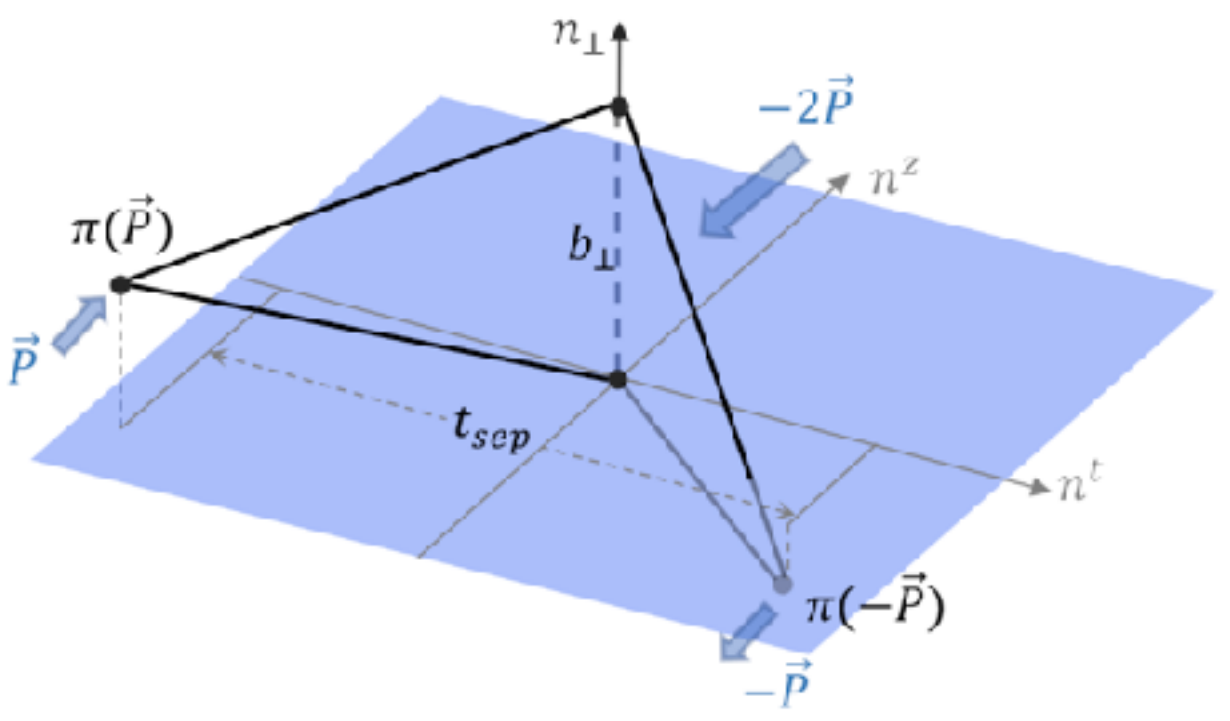
CS-Kernel with 1-loop matching

M. Chu. et.al., LPC, PRD106(2022)034509



J. Hua, et.al., LPC, PRL129(2022)132001

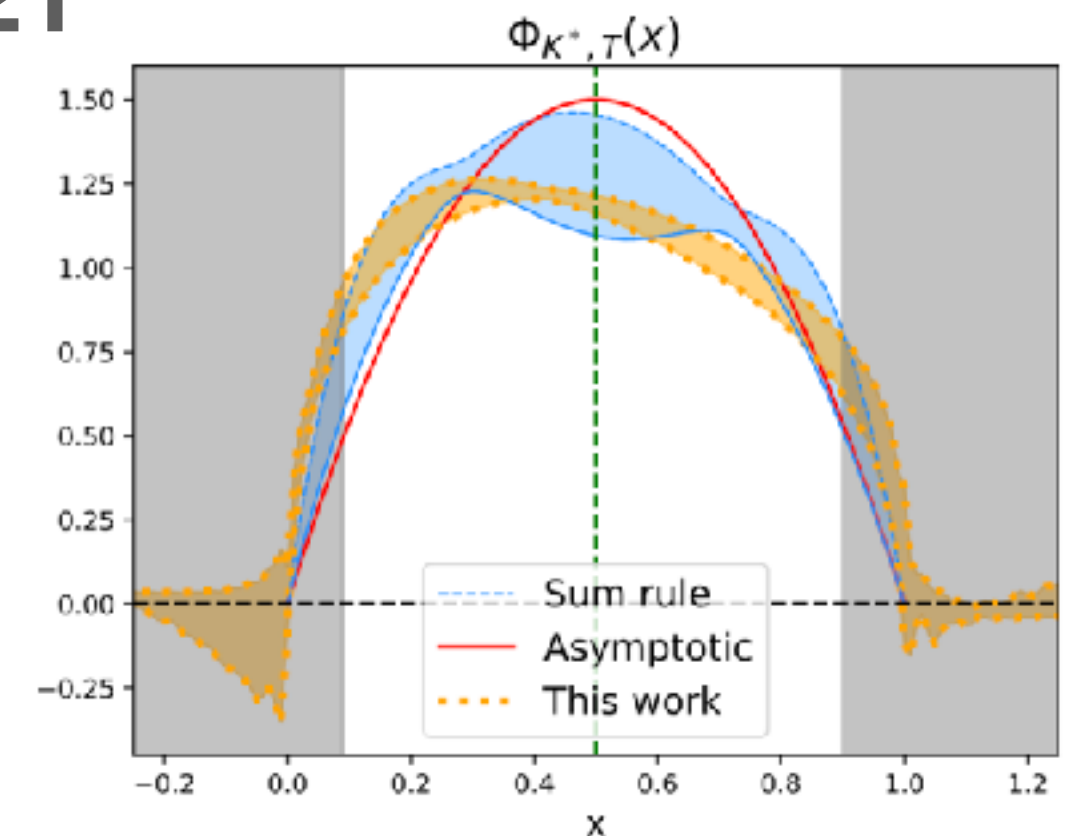
2020



TMD Soft function

Q.-A. Zhang. et.al., LPC, PRL125(2020)192001

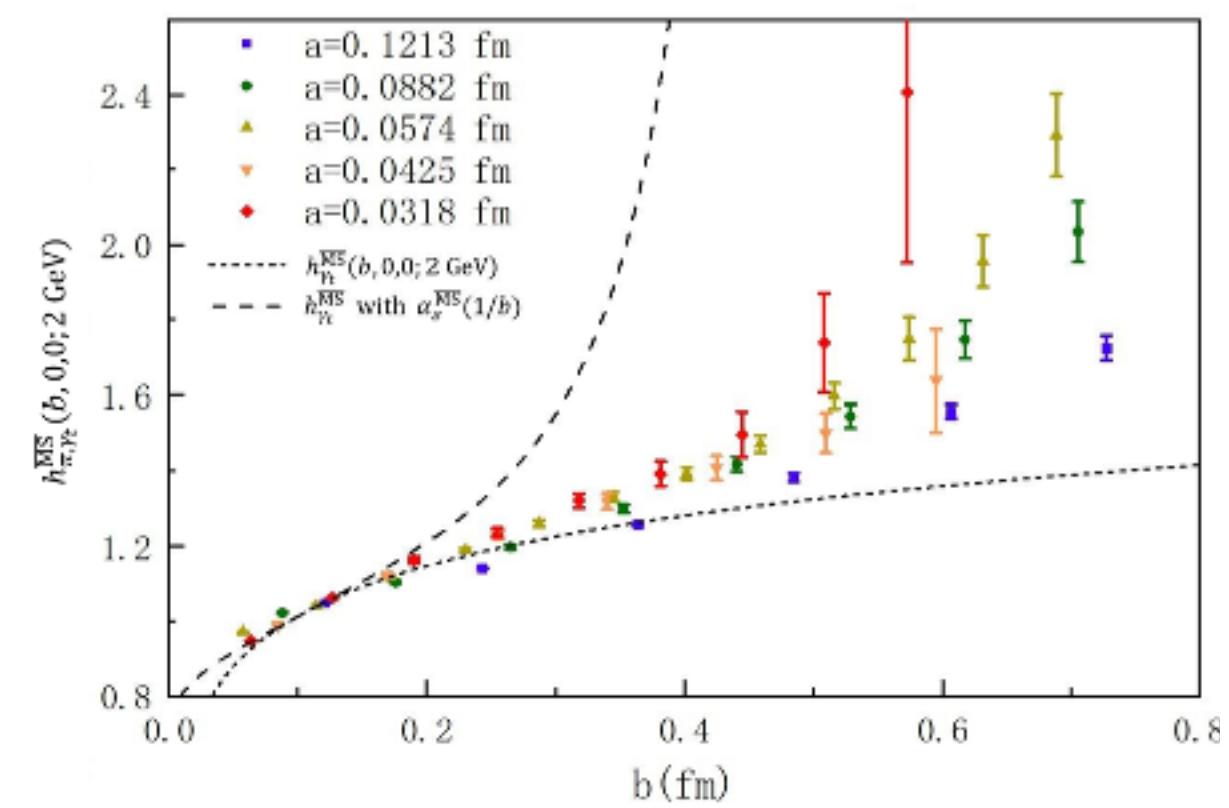
2021



Vector meson DA

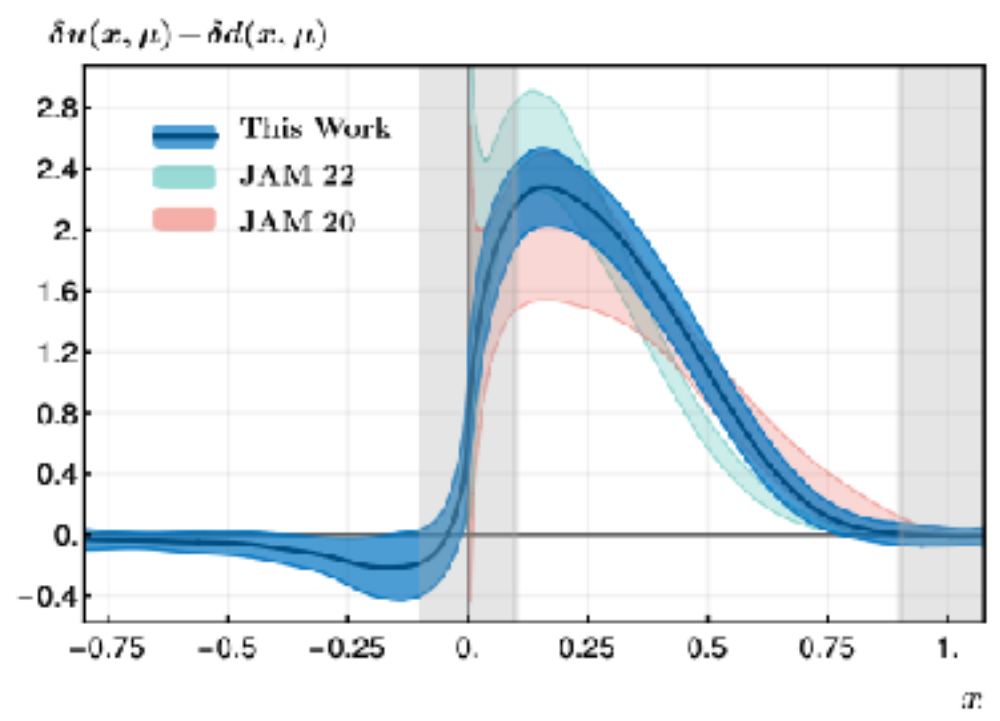
J. Hua. et.al., LPC, PRL127(2021)062002

2022



Renormalization of quasi-TMD

K. Zhang. et.al., LPC, PRL129(2022)082002



Transversity PDF

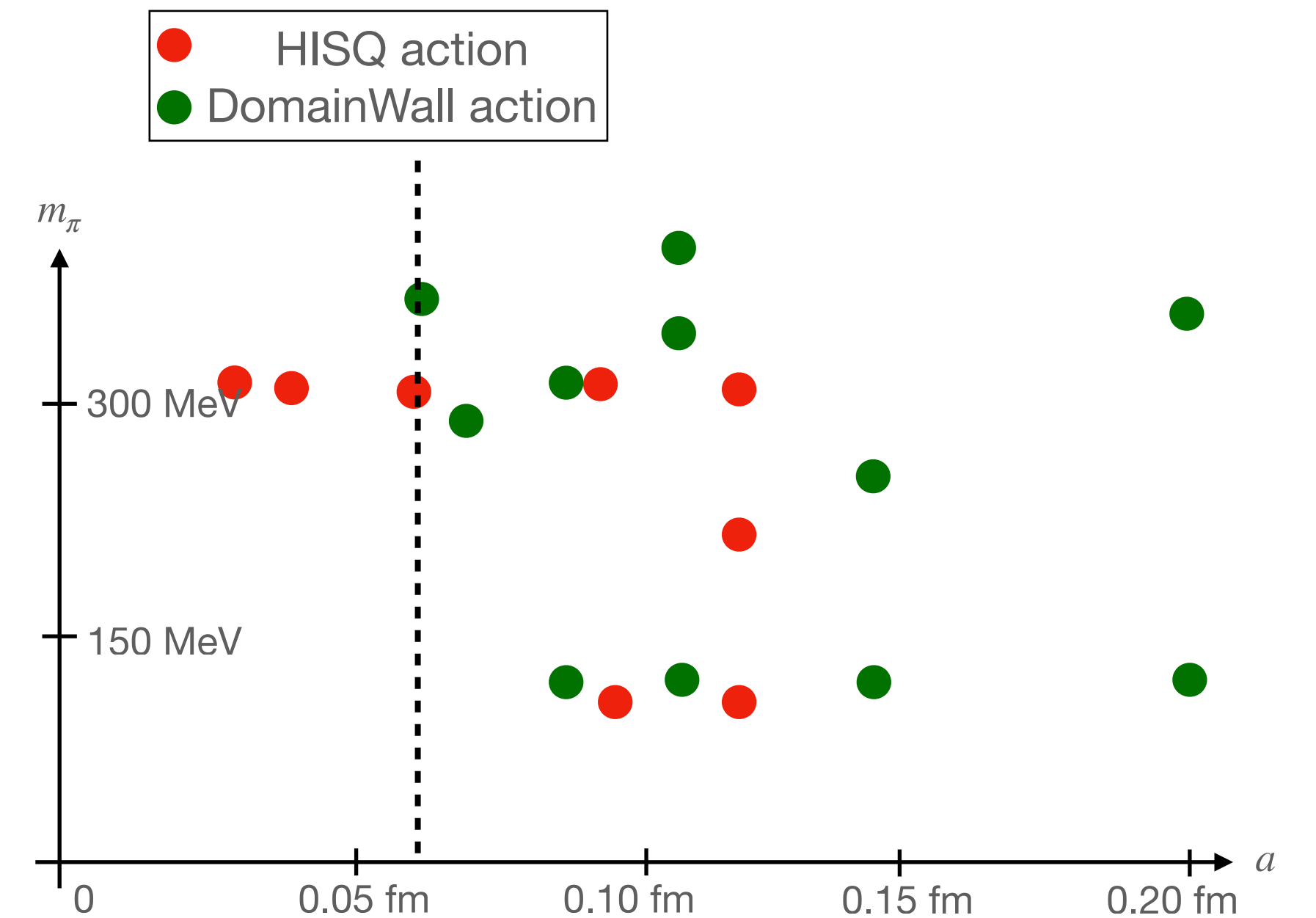
F. Yao. et.al., LPC, 2208.08008

软件2

- 相比QUDA, GWU-code专注于Overlap离散化方案的计算, 代码较为简单。
- 目前已经基于DCU版本的GWU-code产生大量数据。

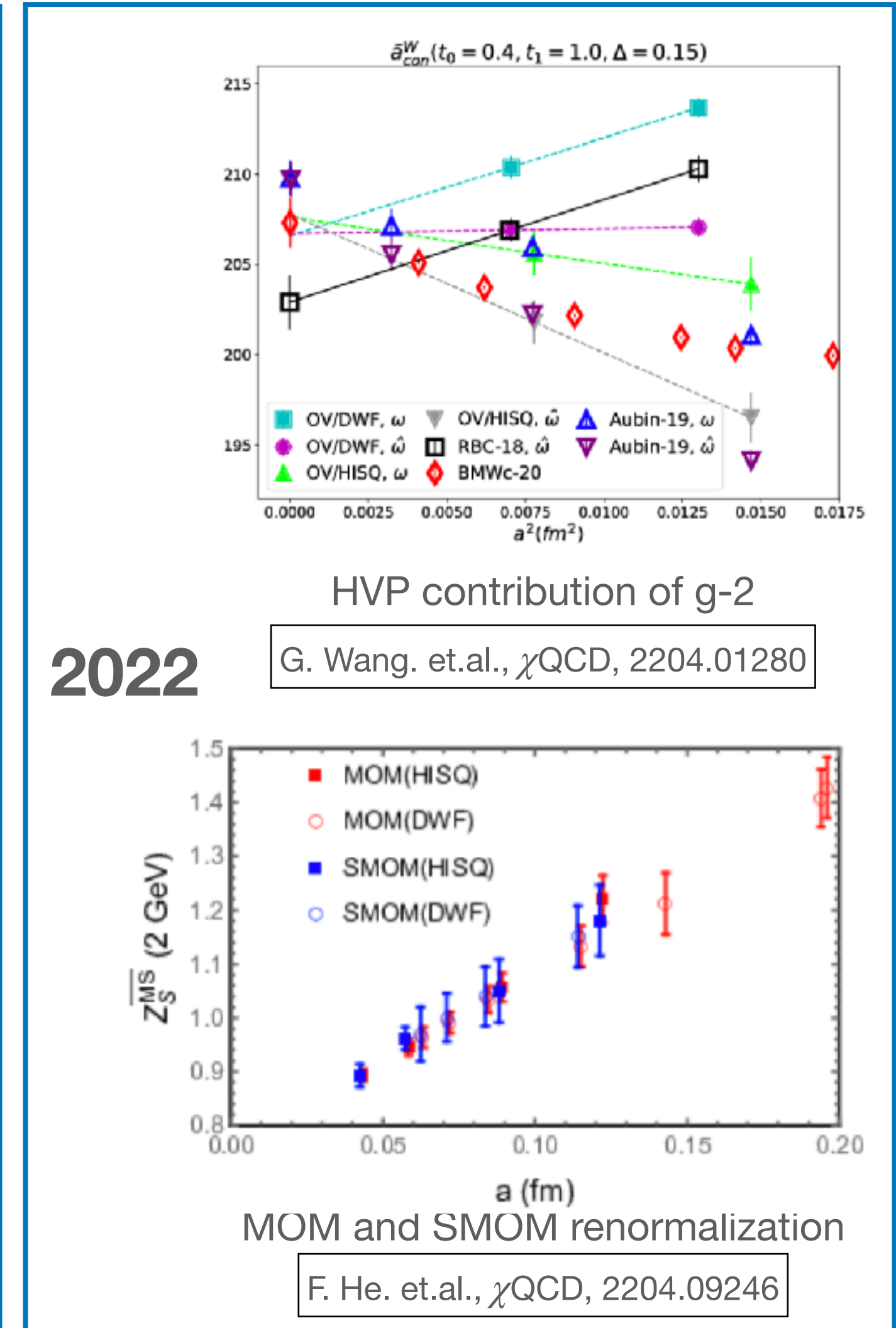
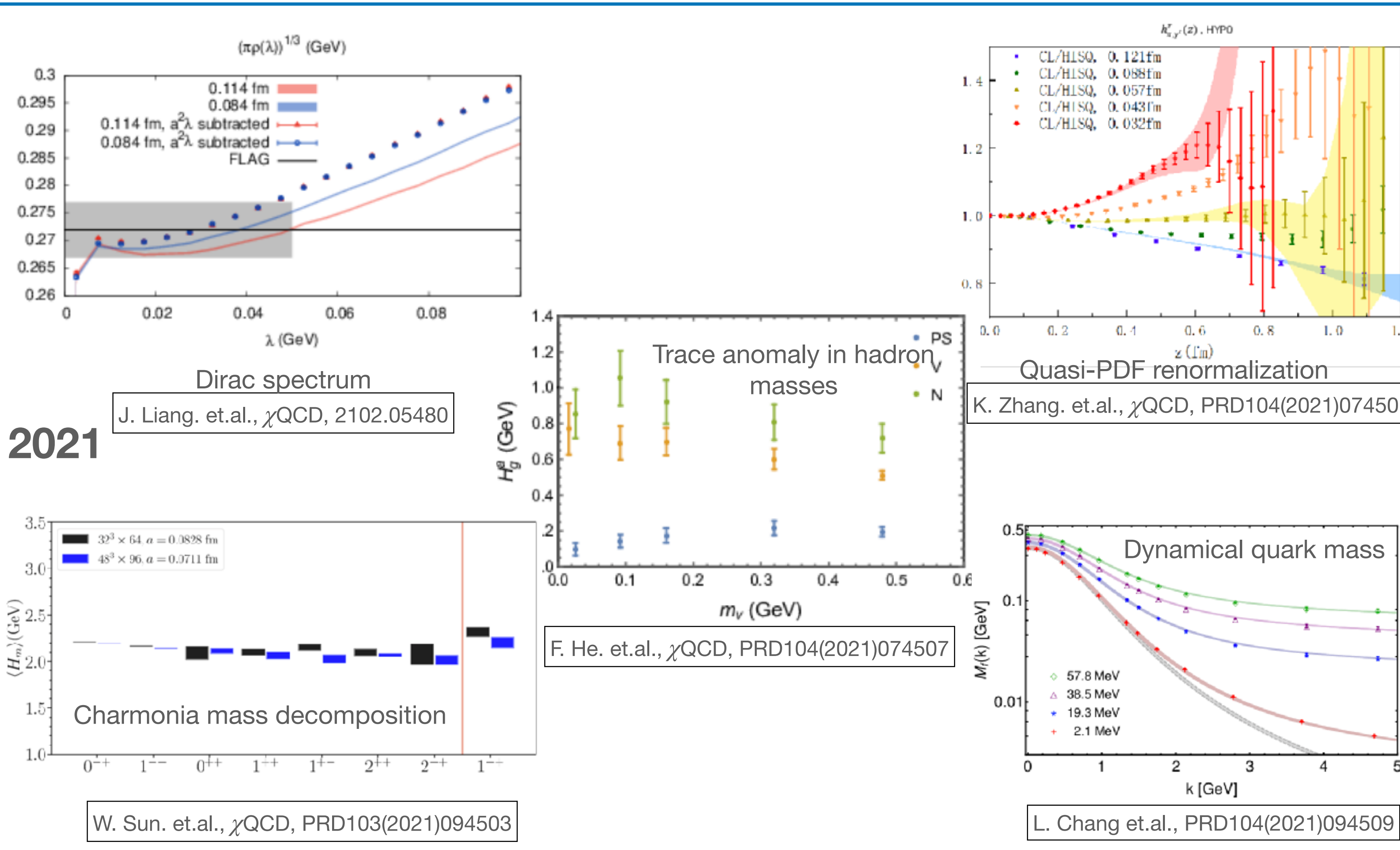
GWU-code

	24 ³ x6 4	32 ³ x6 4	48 ³ x 96	24 ³ x6 4	64 ³ x1 28	32 ³ x6 4	48 ³ x9 6	32 ³ x6 4	24 ³ x6 4	48 ³ x6 4	32 ³ x9 6	64 ³ x9 6	48 ³ x1 44	64 ³ x1 92	96 ³ x2 88	
lattice spacing (fm)	0.197	0.143	0.114	0.1105	0.837	0.0828	0.0711	0.0627	0.1213	0.1213	0.0882	0.0882	0.0574	0.0425	0.0318	
Pion mass (GeV)	140	140	0.14	0.33	0.14	0.30	0.28	0.38	0.31	0.13	0.31	0.13	0.31	0.31	0.31	
Obtained at China	58	35	313	523	313	624	185	241	93	29	122	25	87	54	17	
Data size in China (TB)	1	4	464	12	1467	34	51	13	1	23	10	96	36	58	97	
	overlap@DWF								overlap@HISQ							



软件2应用

GWU-code的应用



已完成工作

核子与轻核内部结构的第一性原理计算

课题编号: ghfund202107011598

课题负责人: 孙鹏

性能优化测试:

计算时间 (ms)	1*dcu	1*cpu	1*dcu/ 1*cpu
$8^3 * 8$	38.9	1042.2	2679.17%

计算时间 (ms)	4*dcu	4*cpu	4*dcu/ 4*cpu	8*dcu	8*cpu	8*v100	8*dcu/ 8*cpu	8*dcu/ 8*v100
$24^3 * 64$	643.68	81038.8	12589.9%					

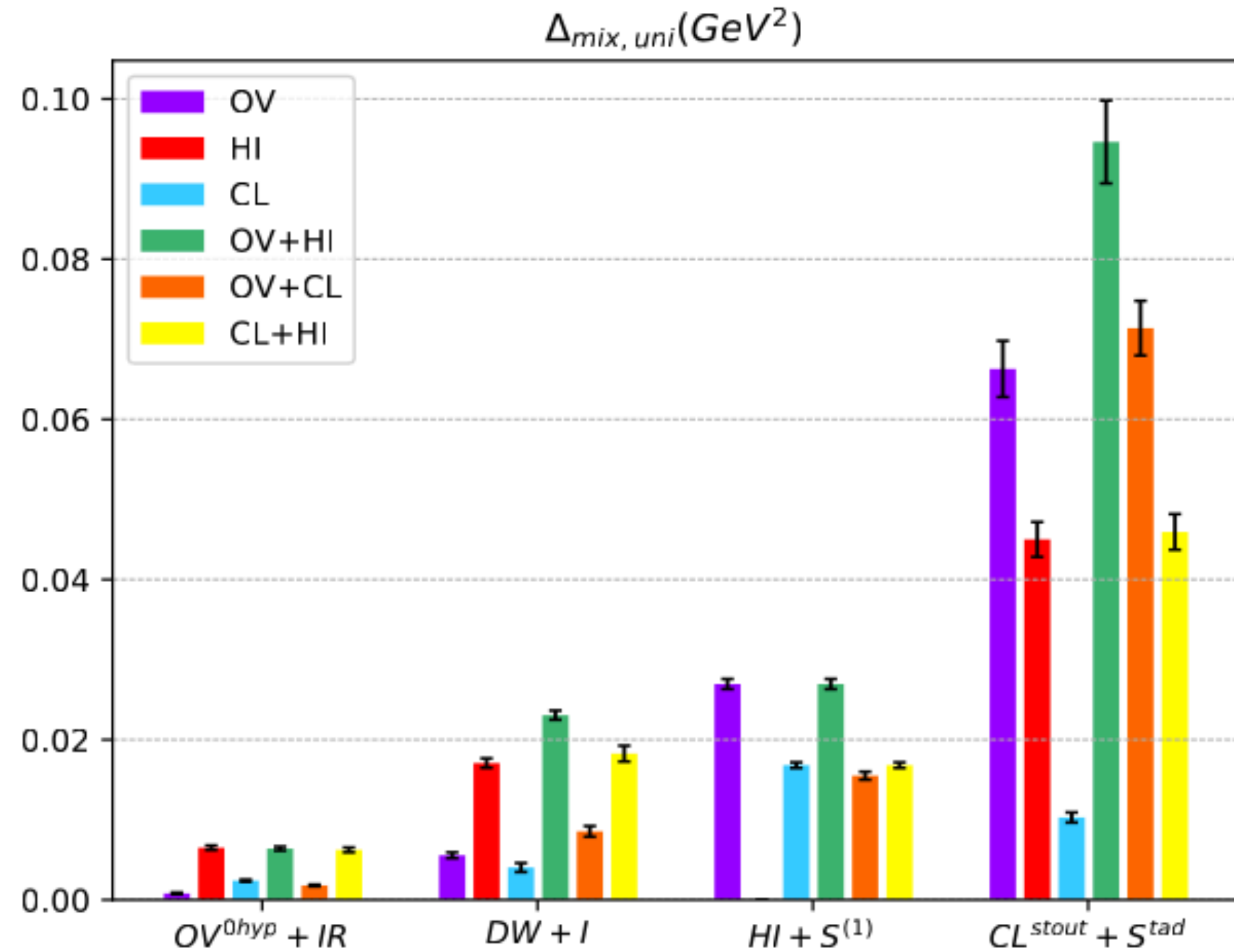
$24^3 * 64$				538.153	34667.1	457.56	6429.9%	84.86%
-------------	--	--	--	---------	---------	--------	---------	--------

- 跨节点多显卡性能表现优异, 8块DCU表现性能与8块V100表现性能相比达到82.9%。

- 单卡比单核心加速比超过860倍, 单卡比单CPU加速26.9倍, 实现整个模式性能单卡超单颗处理器。

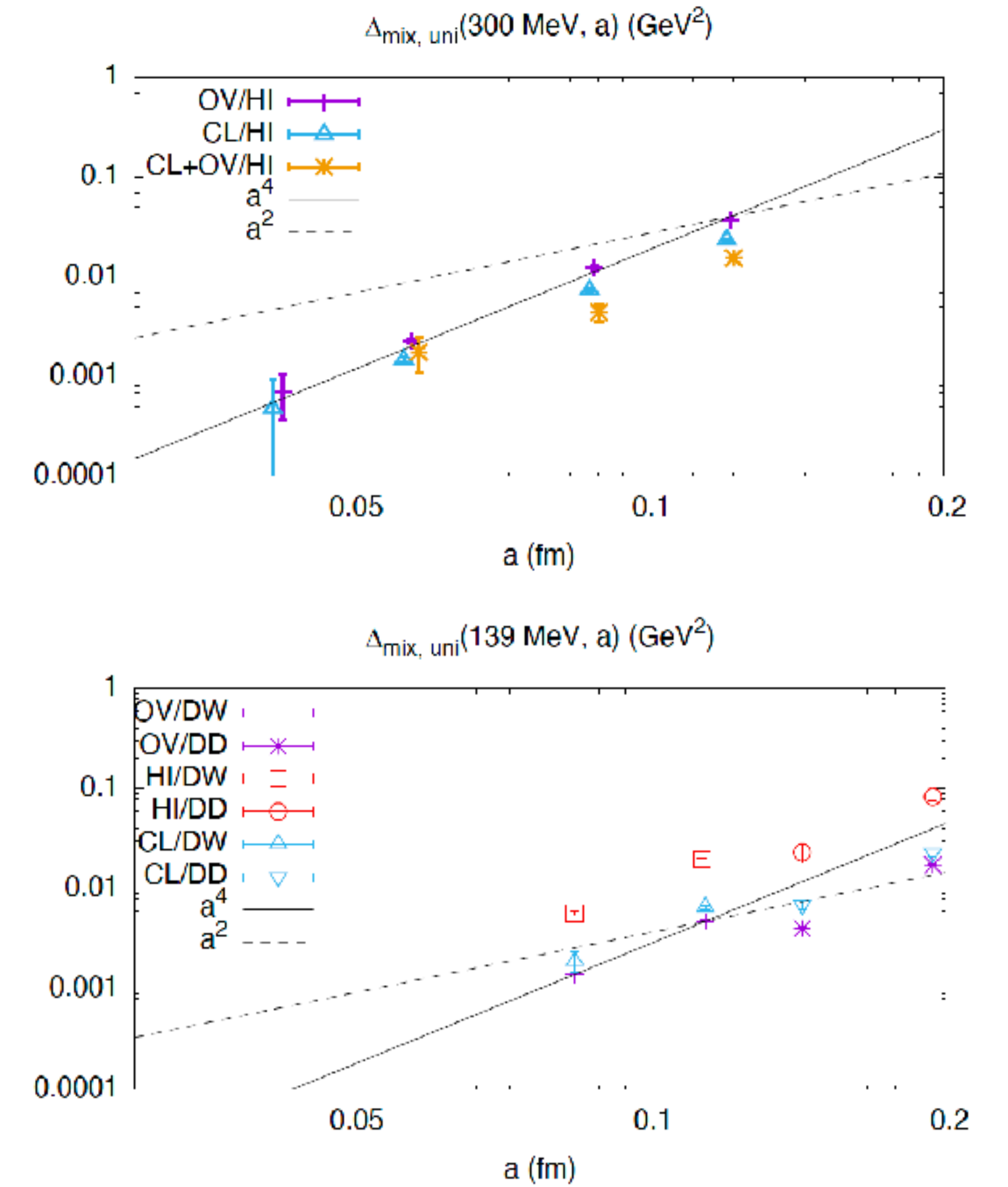
- 经OpenChiral与中科可控联合测试, OpenChiral软件与DCU深度计算处理器兼容性良好, 可以顺利安装、配置, 并稳定运行。

软件3应用



- Our study shows that it is an $\mathcal{O}(a^4)$ effect in the lattice spacing range we studied, **once the fermion action in the ensemble has the chiral symmetry**, while would be $\mathcal{O}(a^2)$ in the other cases.
- The HISQ action would be the cheapest choice to build a large lattice at small lattice spacing.

OpenChiral的应用



D. Zhao, et.al, χ QCD, 2207.14132

Based on the calculation on SunRising-1 computing environment

软件部分

代码现状

	Dslash	Solver	HMC
曙光	<ol style="list-style-type: none">1. QUDA: 目前已移植, 支持HISQ(对反周期边界条件与精度切换的支持还不完善), CLV, TM+C, DW(没有chroma接口, 只支持CPS), OV(有代码但是没有整理完毕并用于大规模实际计算);2. GWU code: 目前已移植, 支持Clover和OV, 存在Dslash长时间迭代随机出错的问题;3. 光合基金项目: 原生代码, 单DCU性能达到V100的80%, 多DCU兼容性还需要进一步调试。	<ol style="list-style-type: none">1. QUDA: 提供CG, GCR, BicgSTAB; 对于CLV/TM+C提供完整的Multigrid支持; 对于HISQ支持Multigrid但是效率比CG低; OV的solver可在Chroma中实现;2. GWU code: 完整的OV 2-level deflation solver支持;3. 光合基金项目: 提供CG, BicgSTAB和EO prec。	<ol style="list-style-type: none">1. QUDA: 目前不提供solver以外的功能, HMC所需函数需要通过QDP-JIT加速(QDP-JIT有HIP版但是未测试)。2. CPS: 只支持DW组态产生, 但是大部分参数组合(格距/质量等)需要调试; 不支持multigrid。相比V100版本效率大约在1/4-1/2(依赖于硬件驱动版本)
申威	<ol style="list-style-type: none">1. SWLQCD: 太湖原生代码, 每个核组固定计算16^4的格子的CLV, 不支持EO prec, 不能调整格子大小; 修改后兼容26010p。2. Qsunway: SWLQCD的Chroma接口	<ol style="list-style-type: none">1. SWLQCD: 提供CG和GCR	<ol style="list-style-type: none">1. SWLQCD: 除solver外目前仅支持主核; strange quark rHMC的从核代码调试中。
天河3	N/A	N/A	N/A

Summary

- 在曙光/申威/天河三家国产架构超算中，由于曙光的DCU架构下与计算核心具有高速连接的内存最大，从而是目前格点QCD单节点计算性能最高的架构，软件研发/移植也最为容易。
- 在美国全面禁售不低于A100性能的所有Nvidia GPU的前提下，曙光DCU架构目前是在中国开展格点QCD研究不可或缺的，我们也依托曙光DCU开展了一系列研究工作（发表PRL3篇，PRD5篇，CPC1篇，NPB1篇）。
- 为了充分利用国产超算开展格点QCD研究，尽快完成对申威与天河架构的适配至关重要。