Fully GPU-Accelerated Distributed Full Configuration Interaction Beyond Trillion-Determinant Scale

Hong Gao*†, Satoshi Imamura*, Akihiko Kasagi*, Yasuaki Ito†, Koji Nakano†
* Fujitsu Limited, † Hiroshima University

Background

Full configuration interaction (FCI) provides an **exact benchmark** for quantum chemistry methods (VQE, QPE, etc)

Large-scale FCI becomes essential as VQE and QPE scale to **larger systems**

Problem

Exponential growth in memory and computational cost

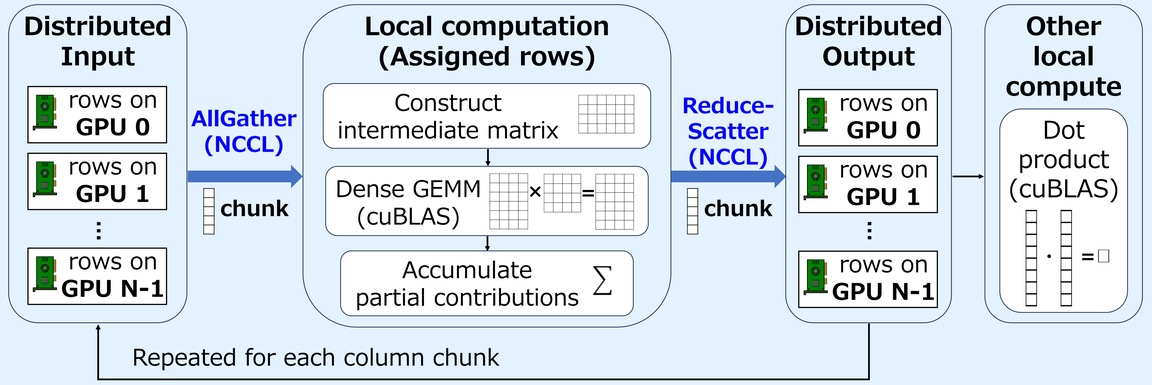
- Limited to **small systems**
- **Long** execution time

Our target: 10^{12} Determinants
in practical time

Data volume: > 210 TB
(> 2600× H100 GPU memory)

Proposed

Fully GPU-accelerated distributed FCI workflow
with chunk-based determinant partitioning



GPU Memory: performance-critical data

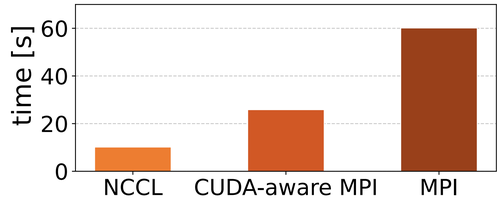
Async Data Transfer

CPU Memory: large auxiliary vectors offloaded

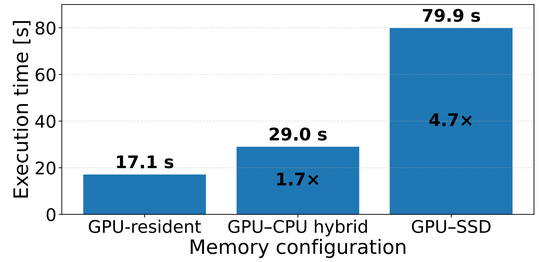
CPU 0 | CPU 1 | CPU 2 | ... | CPU K-1

■ Optimization (CH_3CN, 10^9 determinants, with 16 GPUs)

- GPU-native communication (NCCL)
  - Comparison with other communication libraries



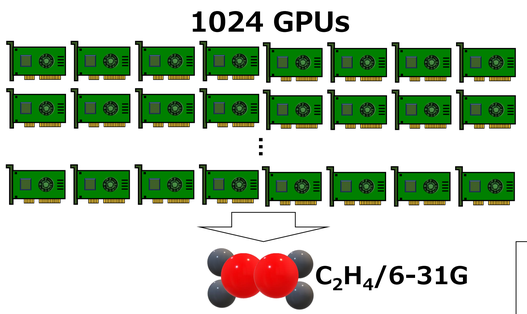
- Hybrid GPU–CPU memory scheme
  - Comparison of memory configurations



Test Platform: ABCI-Q system H [1]
- 4×H100/node, InfiniBand NDR, 1 TB host memory

Results

■ Largest uncompressed FCI



2.4×10^{12} determinants
2 iterations (3 hours)

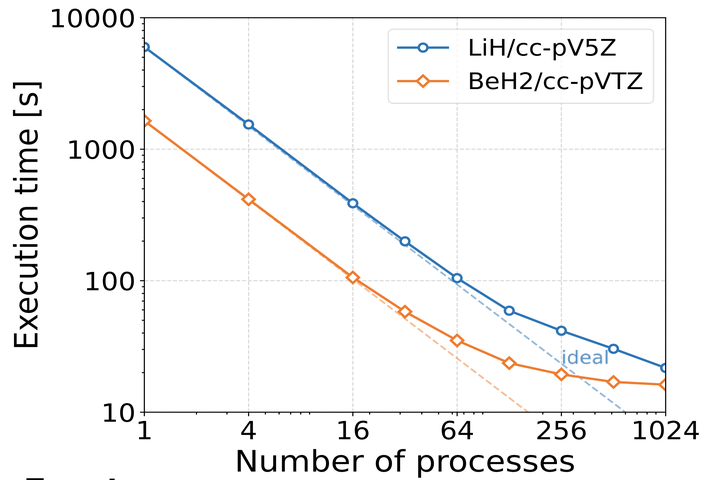
1000× scale of prior GPU-accelerated FCI [2]
(1 iteration)

■ Speedup over prior work [3]

mol	$N_{determinants}$	$N_{devices}$	t_{CPU}[3]	t_{GPU}	Speedup
LiH	1.3×10^{7}	4	29.3 m	30s	**58×**
CH4	7.7×10^{10}	32	16.1h	18.1m	**53×**
C3H8	**1.3×10^{12}**	**512**	**113.6h**	**8.6h**	**13×**

■ Strong scalability

- Continued time reduction to 1024 processes
- **Near-ideal strong scalability** to 64 processes



■ Exactness

- Energy agrees with prior works [3][4]
- Validated on 107 molecule-basis-set combinations

Conclusion

First fully GPU-accelerated distributed FCI for over 2 trillion determinants in practical time

- Provides large-scale exact benchmarks
- Paves the way for emerging quantum chemistry methods and quantum computing

Acknowledgement

This work was supported by computational resources of the Quantum-AI Hybrid Computing Infrastructure (ABCI-Q), awarded under the "ABCI-Q Grand Challenge" Program of the National Institute of Advanced Industrial Science and Technology (AIST).

Reference

[1] G-QuAT, "Quantum-AI Technology," https://unit.aist.go.jp/g-quat/G-QuAT_0617_ol_PDF%E7%94%A8r.pdf, 2025, [Online; accessed August 2026].
[2] B. S. Fales and T. J. Martinez, "Efficient treatment of large active spaces through multi-GPU parallel implementation of direct configuration interaction," Journal of Chemical Theory and Computation, vol. 16, no. 3,pp. 1586–1596, 2020.
[3] H. Gao, S. Imamura, A. Kasagi, and E. Yoshida, "Distributed implementation of full configuration interaction for one trillion determinants," Journal of Chemical Theory and Computation, vol. 20, no. 3, pp. 1185–1192, 2024.
[4] Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; Wouters, S. PySCF: the Python-based simulations of chemistry framework. Wiley Interdisciplinary Reviews: Computational Molecular Science 2017, 8, e1340.