Grand Challenge program

Large Scale applications on TSUBAME 2.5

TOKYO TECH
Pursuing Excellence

Petascale Solver for Semidefinite Programming Problems

- SDP (semidefinite programming) is one of the most central problems in mathematical optimization.
- Many applications: structural optimization, combinatorial optimization, quantum chemistry, sensor network location, etc.

Primal-Dual Formulation

Primal

$$\text{minimize} \quad \sum_{k=1}^{m} c_k x_k$$

$$\text{subject to} \quad X = \sum_{k=1}^{m} F_k x_k - F_0, \quad X \succeq O.$$

$c,\ x \in R^n$

$F_0, \ldots, F_m \in \mathbb{S}^n$

$$F_k = \begin{pmatrix} F_k^1 & O & O & O \\ O & F_k^2 & O & O \\ O & O & \ddots & O \\ O & O & O & F_k^h \end{pmatrix}$$

where $F_k^1 \in \mathbb{S}^{n_1}, F_k^2 \in \mathbb{S}^{n_2}, \ldots, F_k^h \in \mathbb{S}^{n_h}$. $\sum_{\ell=1}^{h} n_\ell = n$

$X \in \mathbb{S}^n$ is a positive semidefinite matrix.

Dual

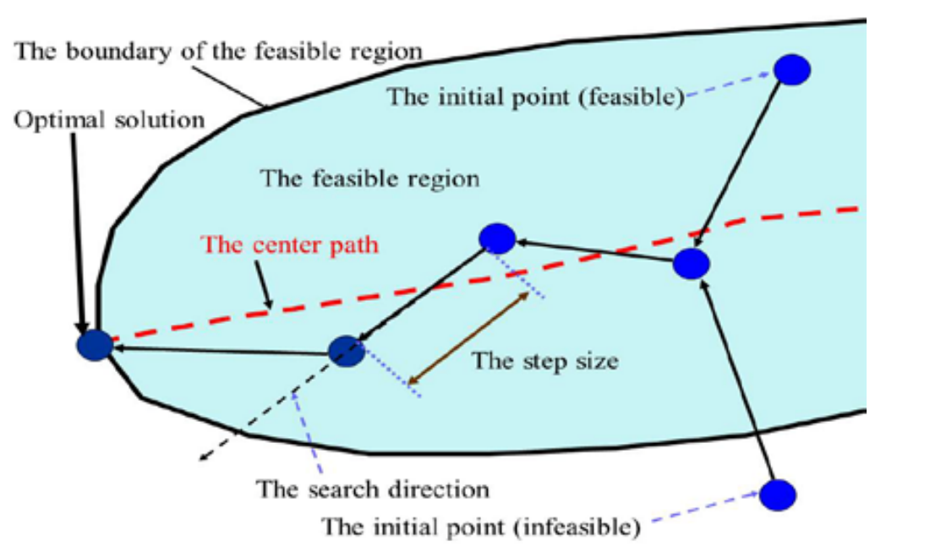
$$\text{maximize} \quad F_0 \bullet Y$$

$$\text{subject to} \quad F_k \bullet Y = c_k \ (k = 1, \ldots, m), \quad Y \succeq O.$$

$$U \bullet V = \sum_{i=1}^{n}\sum_{j=1}^{n} U_{ij} V_{ij}.$$

Algorithmic Framework of PDIPM

Step 0 Choose an initial point (x, X, Y) with $X \succ O, Y \succ O$
Step 1 Compute Search Direction (dx, dX, dY)
Step 2 Compute Step Length α_P, α_D $X + \alpha_P dX \succ O, \quad Y + \alpha_D dY \succ O$
Step 3 Update $(x, X, Y) \leftarrow (x + \alpha_P dx,\ X + \alpha_P dX,\ Y + \alpha_D dY)$
Step 4 Goto Step 1 if (x, X, Y) is not close to optimal



Two major bottleneck parts (80~90% of execution time)

- **ELEMENTS** : Computation of the SCM
  - ✓ Memory Access-intensive
  - ✓ Time-complexity: $O(mn^3 + m^2n^2)$

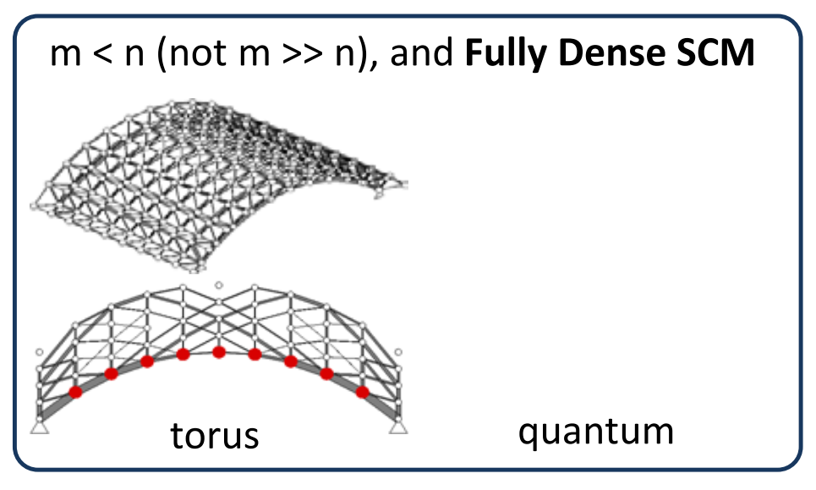
or

- **CHOLESKY** : Cholesky factorizations of the SCM
  - ✓ Compute-intensive
  - ✓ Time-complexity: $O(m^3)$

n : matrix size
m: # of constraints

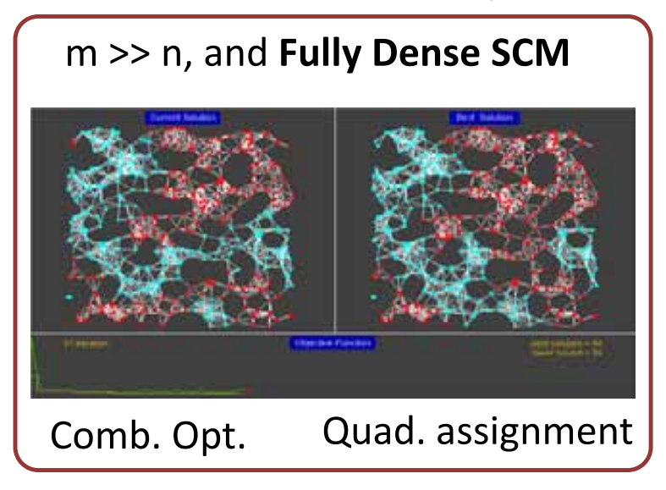
ELEMENTS-bound SDP problems

m < n (not m >> n), and **Fully Dense SCM**



CHOLESKY-bound SDP problems

m >> n, and **Fully Dense SCM**

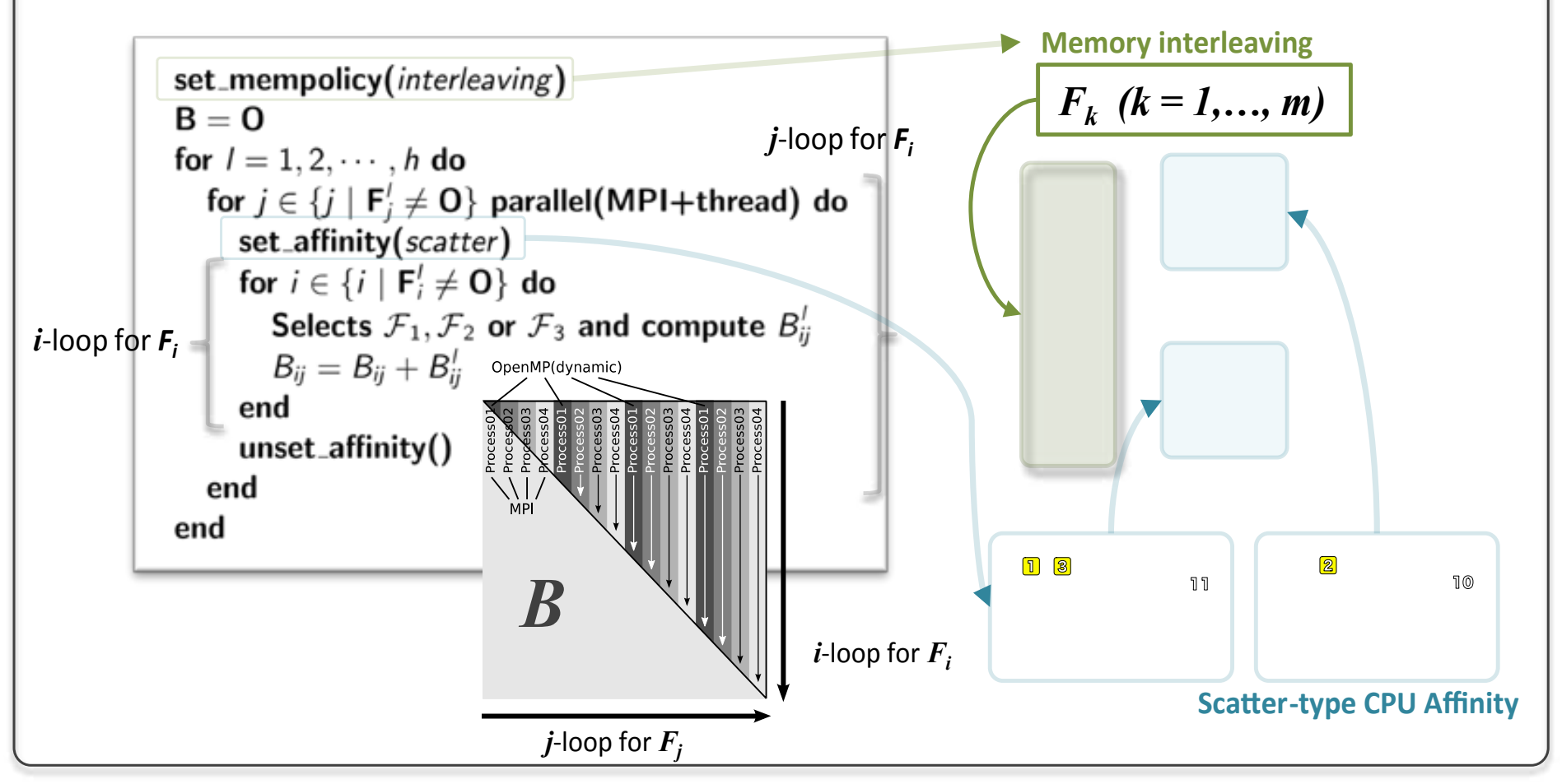


Scalable Implementation of ELEMENTS

ELMENTS for large-scale SDP problems generally requires significant computational resources in terms of CPU cores and memory bandwidth.

Our algorithm

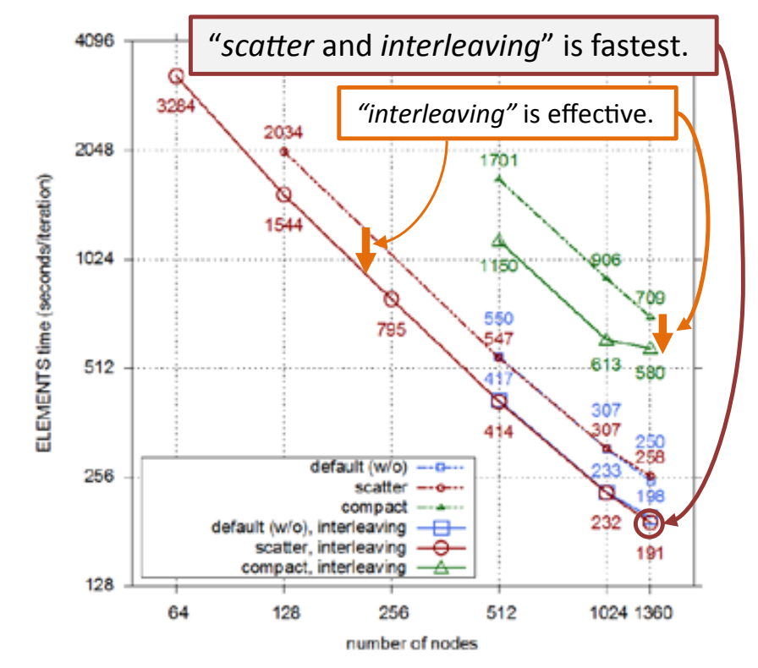
Efficient Hybrid (MPI-OpenMP) parallel Computation of SCM B
with *Automatic configuration* for CPU Affinity and memory interleaving



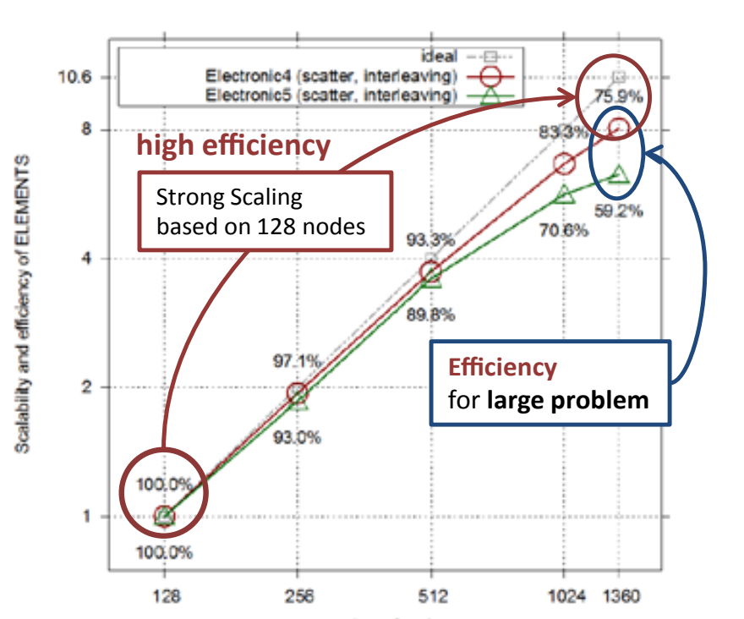
```
set_mempolicy(interleaving)
B = O
for l = 1, 2, ..., h do
  for j ∈ {j | F_j^l ≠ O} parallel(MPI+thread) do
    set_affinity(scatter)
    for i ∈ {i | F_i^l ≠ O} do
      Selects 𝓕1, 𝓕2 or 𝓕3 and compute B_ij^l
      B_ij = B_ij + B_ij^l
    end
    unset_affinity()
  end
end
```

Performance of ELEMENTS for quantum on CPUs of TSUBAME 2.0

Electronic5 (SCM size: 47.7k x 47.7k)
CPU time varied with affinity/interleaving setting



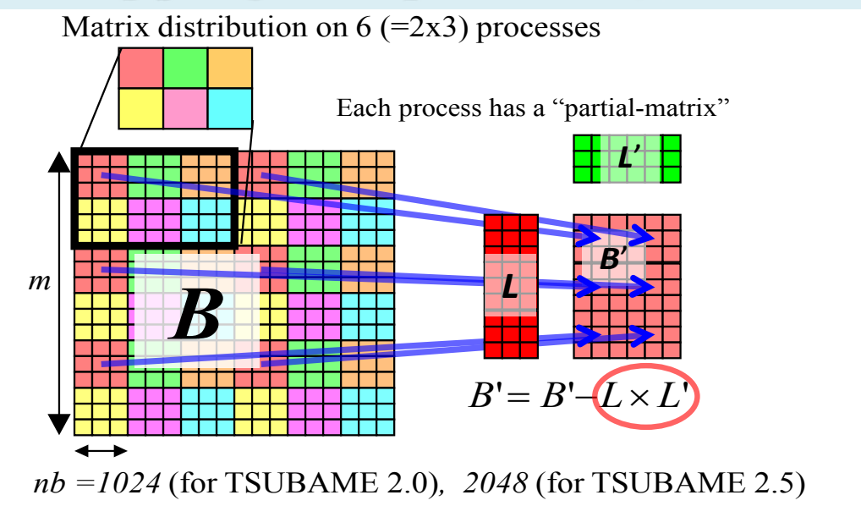
Electronic4 (SCM size: 76.6k x 76.6k)
Scalability of "*scatter and interleaving*"



Scalable Implementation of CHOLESKY

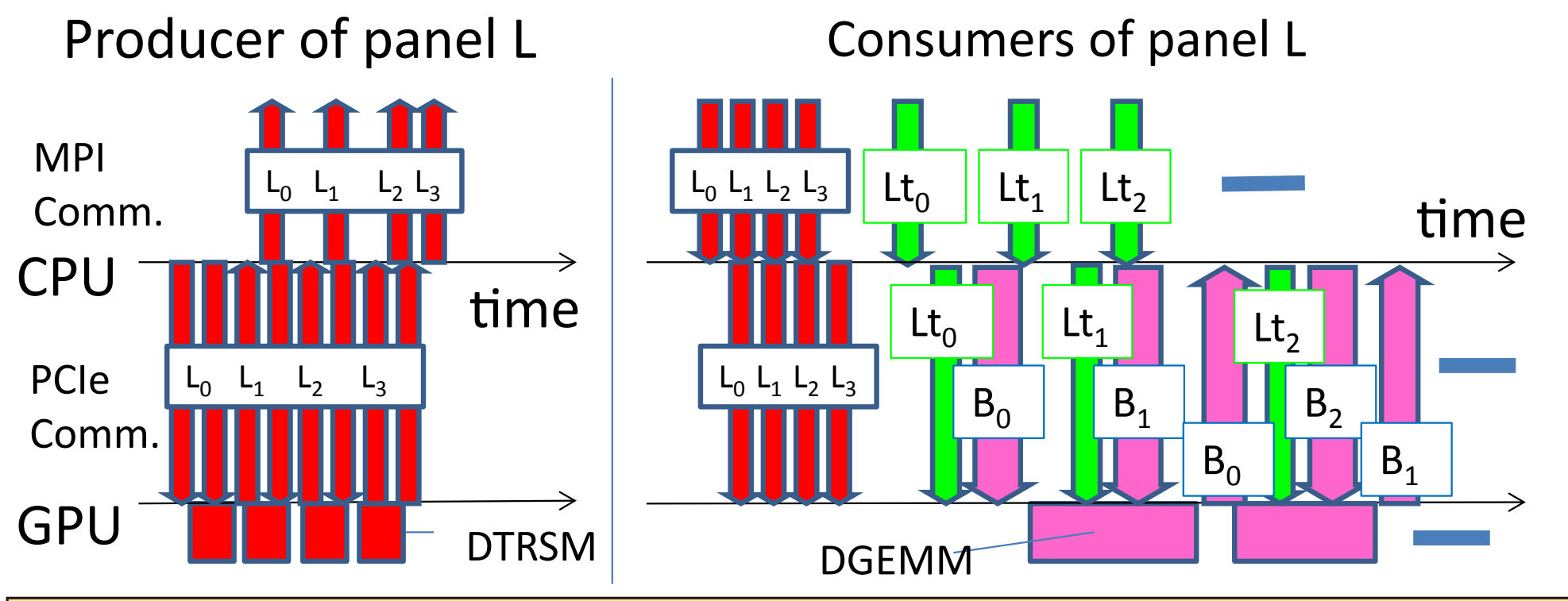
For problems with m >> n, high performance **CHOLESKY** is implemented for GPU supercomputers.
Key for petaflops is **overlapping computation, PCI-Express communication and MPI communication.**

Data Distribution:
The dense matrix B (m x m) is distributed in 2D Block cylic distribution

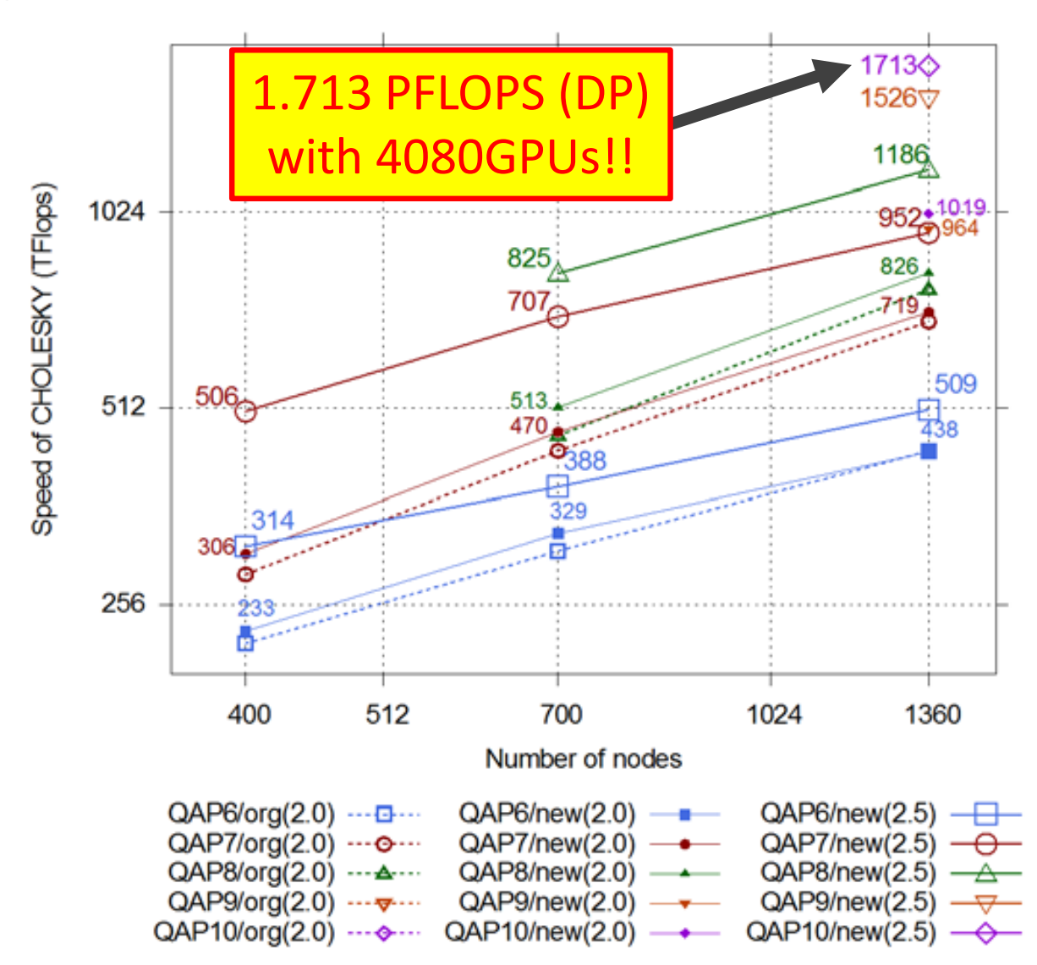


nb =*1024* (for TSUBAME 2.0), *2048* (for TSUBAME 2.5)

Our algorithm with overlapping



Performance of CHOLESKY for QAP on GPUs of TSUBAME 2.0/2.5



SDPARA can solve the largest SDP problem in 1.7PFLOPS !
- DNN relaxation problem for QAPLIB with 2.3million constraints

Project leader: Katsuki Fujisawa (Chuo Univ)

Solving the Schrödinger Equations of Molecules

Solving the Schrödinger Equations of Molecules

Schrödinger equation

$$H\psi = E\psi$$

Free Complement (FC) theory

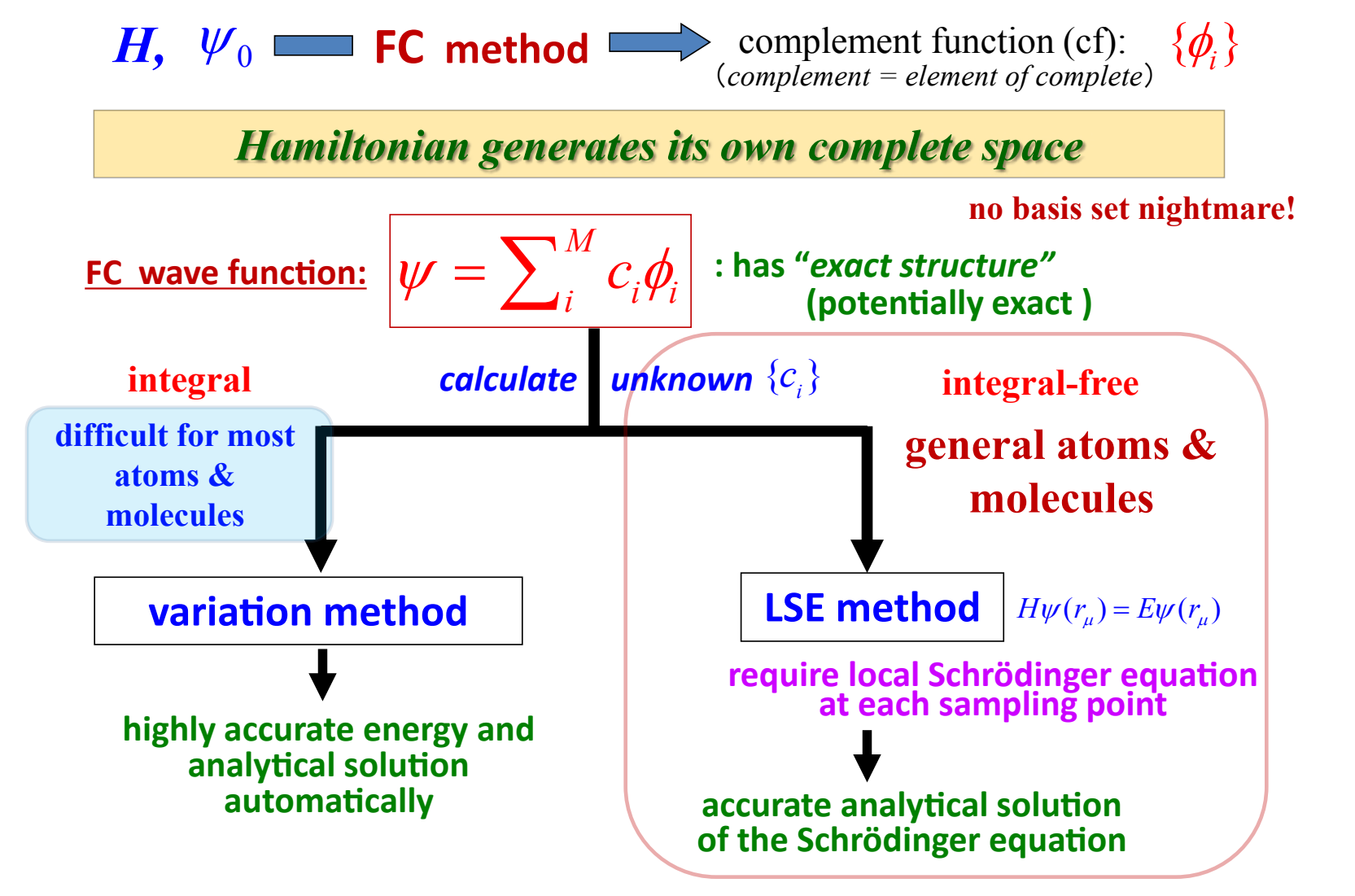
Pauli principle

$$P\psi = (-)^P \psi$$

Must be satisfied for electrons
Very time consuming process

Free complement (FC) method for solving the Schrödinger equation

Review: H. Nakatsuji, *Acc. Chem. Res.* 45, 1480 (2012)



$H,\ \psi_0$ → FC method → complement function (cf): $\{\phi_i\}$
(*complement = element of complete*)

Hamiltonian generates its own complete space

no basis set nightmare!

FC wave function: $\psi = \sum_i^M c_i \phi_i$: has "*exact structure*" (potentially exact)

calculate unknown $\{c_i\}$

- integral: difficult for most atoms & molecules → **variation method** → highly accurate energy and analytical solution automatically
- integral-free: general atoms & molecules → **LSE method** $H\psi(r_\mu) = E\psi(r_\mu)$ → require local Schrödinger equation at each sampling point → accurate analytical solution of the Schrödinger equation

FC method of generating complement functions

When we use Slater-type valence-bond (VB) function as ψ_0, the exact molecular wave function is expressed as

$$\psi = \sum_I c_I \hat{A} \phi_I$$

where the complement functions (cf's) ϕ_I is written in Hartree product form as

$$\phi_I(1,2..N) = \prod_i^{N_e} \left(\exp(-\alpha_i r_{iA_i}) x_{iA_i}^{k_{x,iA_i}} y_{iA_i}^{k_{y,iA_i}} z_{iA_i}^{k_{z,iA_i}} \left|\vec{r}_{iA_i}\right|^{k_{r,iA_i}} \times \prod_{A(\neq A_i)}^{N_A} r_{iA}^{k_{iA}} \times \prod_{j(>i)}^{N_e} r_{ij}^{m_{ij}} \right) \cdot \sigma_I$$

Pauli principle: Anti-symmetrization

1) Nk algorithm

H. Nakashima, H. Nakatsuji, *J. Chem. Phys.* 139, 044112 (2013).

- determinant based $N^{3\text{-}4}$ algorithm
- not suitable for complex wave functions

2) iExg algorithm

- *anti-symmetrization theory for molecules.*
- *natural order-N theory.*

These theories would be helpful for doing chemistry in Schrödinger and Dirac accuracy.

Parallel algorithm of the FC-LSE method

Step 1. Function generation according to the Hamiltonian

$$\psi_{n+1} = \left[1 + C_n g\left(H - E_n\right)\right]\psi_n$$

Step 2. FC wave function

$$\psi_{n+1} = \sum_i^{M_n} c_i^{(n)} \phi_i^{(n)}$$

Analytical evaluation (Small computational cost)

Step 3. LSE equation Most time consuming step

3. 1. $\mathbf{AC} = \mathbf{BCE} \qquad \mathbf{A}_{\mu i} = H\phi_i(\mathbf{r}_\mu),\ \mathbf{B}_{\mu i} = \phi_i(\mathbf{r}_\mu)$

- Distribute sampling points — call MPI_Send(r_μ) call MPI_Recv(r_μ)
- Evaluate $H\phi$ and ϕ at sampling point (10^6-10^8 points)

$$\phi(1,2,\ldots,N_e) = A\left[f_{12}(1,2)\cdots\cdots f_{r-1,r}(r-1,r)\cdot o_1(1)\cdots\cdots o_{N_e}(N_e) \right]$$

Anti-symmetrization: Time consuming

- Even task is distributed to each processor (No communication)
- Peta-flops super parallel machine: Only 1-100 sampling points for single processor
- Cost: $M_n \times N_s \times O(N_e^{3\text{-}5})$

M_n: No. of complement functions
N_s: No. of sampling points
N_e: No. of electrons

3. 2. $\mathbf{HC} = \mathbf{SCE} \qquad \mathbf{H} = \mathbf{B}^\dagger\mathbf{A},\ \mathbf{S} = \mathbf{B}^\dagger\mathbf{B}$

$$H_{ij}^{(Local)} = \sum_\mu^{(Local)} \phi_i(\mathbf{r}_\mu) H \phi_j(\mathbf{r}_\mu)$$

$$S_{ij}^{(Local)} = \sum_\mu^{(Local)} \phi_i(\mathbf{r}_\mu) \phi_j(\mathbf{r}_\mu)$$

BLAS3 library

- Make local H,S matrices
- Cost: $M_n \times M_n \times N_s$

3. 3. Collect H,S matrices and make total H,S matrices

$$H_{ij}^{(All)} = \sum_{n_{proc}} H_{ij}^{(Local)}$$

$$S_{ij}^{(All)} = \sum_{n_{proc}} S_{ij}^{(Local)}$$

call MPI_Reduce(H,MPI_SUM)
call MPI_Reduce(S,MPI_SUM)

- Communication: $M_n \times M_n$

Step 4. Diagonalization (Eigenvalue problem)

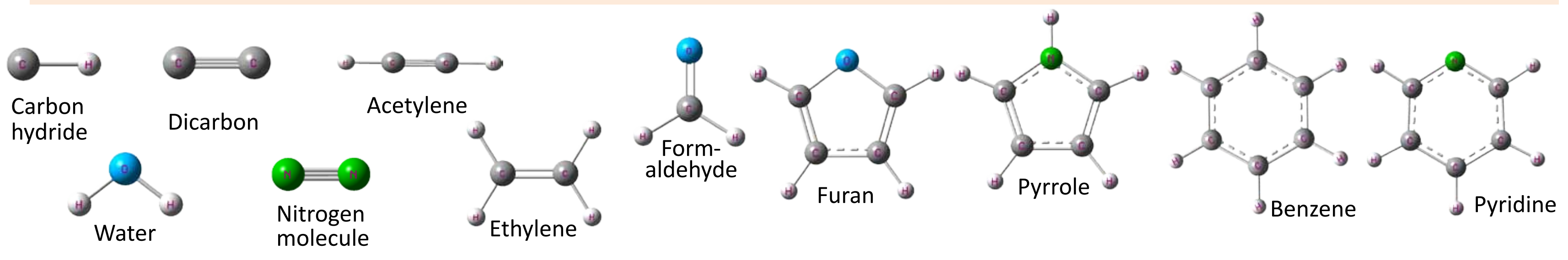
Use parallelized numerical library

Step 5. Physical observables

Energy: $\langle E_L \rangle = \left[\sum_{i,j}^{M_n} C_i C_j \mathbf{H}_{ij}\right] / \left[\sum_{i,j}^{M_n} C_i C_j \mathbf{S}_{ij}\right]$ H-square error: $\sigma^2 = \langle E_L^2 \rangle - \langle E_L \rangle^2$

Small computational cost

Now, organic chemistry can be done in Schrödinger accuracy !!



Small organic molecules: ~20 electrons

Order=2

Molecule	No. of Elec.	M_n	Energy (a.u.) FC-LSE	Energy (a.u.) Exact energy (experiment)	$\Delta E = E_{FC\text{-}LSE} - E_{exact}$ (kcal/mol)
Carbon Hydride(CH)	7	1503	-38.480 41	-38.479 0	-0.884
Water(H_2O)	10	2075	-76.456 78	-76.457 8	0.671
Dicarbon(C_2)	12	1976	-75.923 69	-75.926 5	-0.438
Dinitrogen(N_2)	14	1121	-109.537 10	-109.542 7	3.512
Acetylene (C_2H_2)	14	1709	-77.333 31	-77.335 7	1.49
Ethylene (C_2H_4)	16	2628	-78.577 95	-78.587 4	5.93
Formaldehyde (H_2CO)	16	4083	-114.517 81	-114.508 0	-6.15

The solutions in Schrödinger (Chemical) accuracy could be obtained.

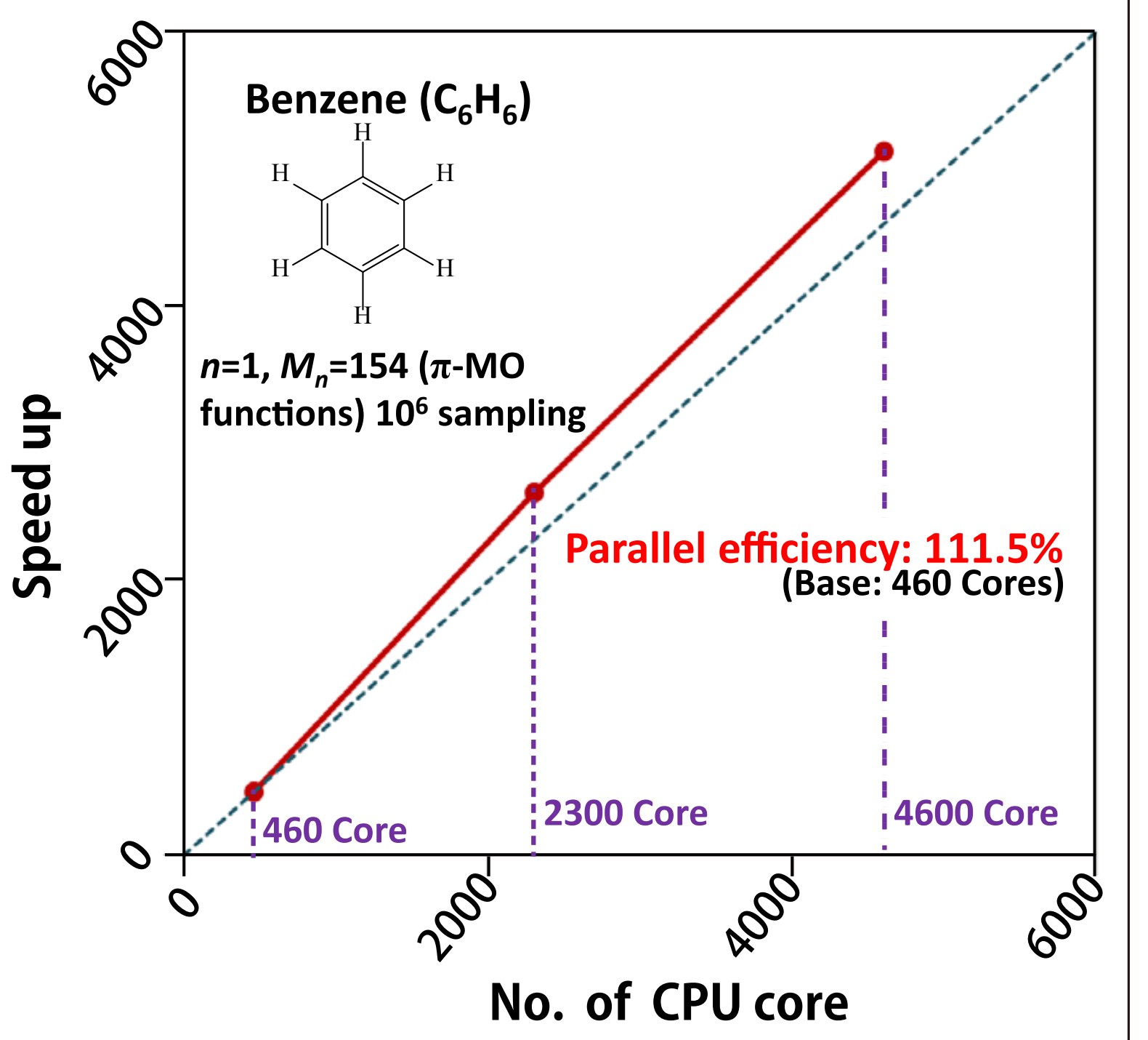
Medium size molecules: ~40 electrons

Order=1: test calculation

Molecule	No. of Elec.	M_n	Energy (a.u.) FC-LSE	Energy (a.u.) Estimated exact from experiment	$\Delta E = E_{FC\text{-}LSE} - E_{exact}$ (a.u.)
Furan (C_4H_4O)	36	161	-229.860 1	-230.027	0.167
Pyrrole (C_4H_5N)	36	174	-209.974 3	-210.173	0.199
Benzene (C_6H_6)	42	398	-232.409 3	-232.248	-0.161
Pyridine (C_5H_5N)	42	386	-247.704 1	(-248.290)	0.586

Order=2 calculations are necessary for more accurate results.

Parallel efficiency on TSUBAME2



http://www.gsic.titech.ac.jp/sc13