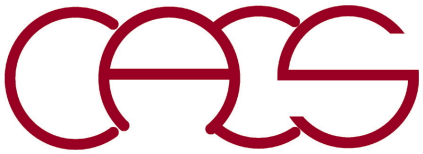
Linear-Scaling Quantum Molecular Dynamics

Aiichiro Nakano

Collaboratory for Advanced Computing & Simulations
Department of Computer Science
Department of Physics & Astronomy
Department of Quantitative & Computational Biology
University of Southern California

Email: anakano@usc.edu



Quantum Molecular Dynamics (QMD)

$$M_I \frac{d^2}{dt^2}\mathbf{R}_I = -\frac{\partial}{\partial \mathbf{R}_I} E[\{\mathbf{R}_I\}, \psi(\mathbf{r}_1 \ldots, \mathbf{r}_N)] \; (I = 1, \ldots, N_{\text{atom}})$$

First molecular dynamics using an empirical interatomic interaction

A. Rahman, *Phys. Rev.* **136**, A405 ('64)

$$\psi(\mathbf{r}_1 \ldots, \mathbf{r}_N) \leftarrow \text{argmin}\, E[\{\mathbf{R}_I\}, \psi(\mathbf{r}_1 \ldots, \mathbf{r}_N)]$$

Density functional theory (DFT)

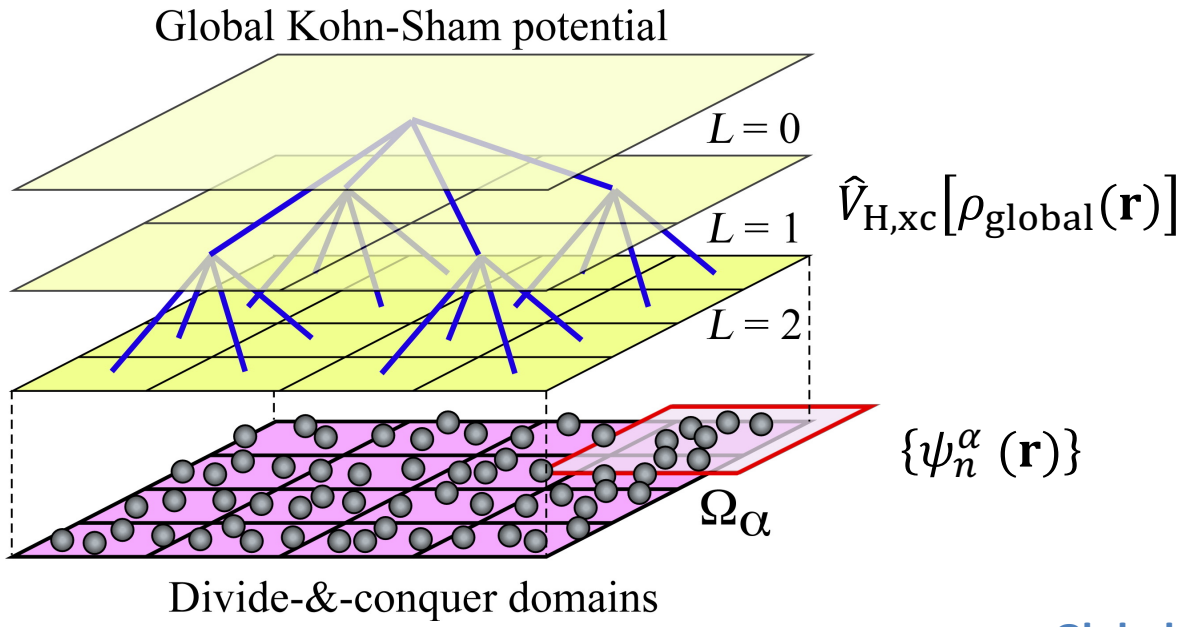
Hohenberg & Kohn, *Phys. Rev.* **136**, B864 ('64)
W. Kohn, *Nobel chemistry prize*, '98

$O(C^N)$	$\rightarrow$	$O(N^3)$
1 N-electron problem		N 1-electron problems
intractable		tractable
$\psi(\mathbf{r}_1 \ldots, \mathbf{r}_N)$		$\{\psi_i(\mathbf{r}) \mid i = 1, \ldots, N\}$

$O(N)$ DFT algorithms

- **Divide-&-conquer DFT** [W. Yang, *Phys. Rev. Lett.* **66**, 1438 ('91); F. Shimojo *et al.*, *Comput. Phys. Commun.* **167**, 151 ('05); *Phys Rev. B* **77**, 085103 ('08); *Appl. Phys. Lett.* **95**, 043114 ('09); *J. Chem. Phys.* **140**, 18A529 ('14)]
- **Quantum nearsightedness principle** [W. Kohn, *Phys. Rev. Lett.* **76**, 3168 ('96); E. Prodan & W. Kohn, *P. Nat. Acad. Sci.* **102**, 11635 ('05)]
- **A recent review** [Bowler & Miyazaki, *Rep. Prog. Phys.* **75**, 036503 ('12)]

Divide-&-Conquer Density Functional Theory



- **Overlapping spatial domains:** $\Omega = \cup_\alpha \Omega_\alpha$
- **Domain Kohn-Sham equations**

Global-local self-consistent field (SCF) iteration

$$\left(-\frac{1}{2}\nabla^2 + \hat{V}_{\mathrm{ion}} + \hat{V}_{\mathrm{H,xc}}[\rho_{\mathrm{global}}(\mathbf{r})]\right)\psi_n^\alpha(\mathbf{r}) = \epsilon_n^\alpha \psi_n^\alpha(\mathbf{r})$$

- **Global & domain electron densities**

$$\rho_{\mathrm{global}}(\mathbf{r}) = \Sigma_\alpha\, p_\alpha(\mathbf{r})\rho_\alpha(\mathbf{r}) \leftarrow \rho_\alpha(\mathbf{r}) = \Sigma_n |\psi_n^\alpha|^2\, \Theta(\mu - \epsilon_n^\alpha)$$

Domain support function

$$\Sigma_\alpha\, p_\alpha(\mathbf{r}) = 1$$

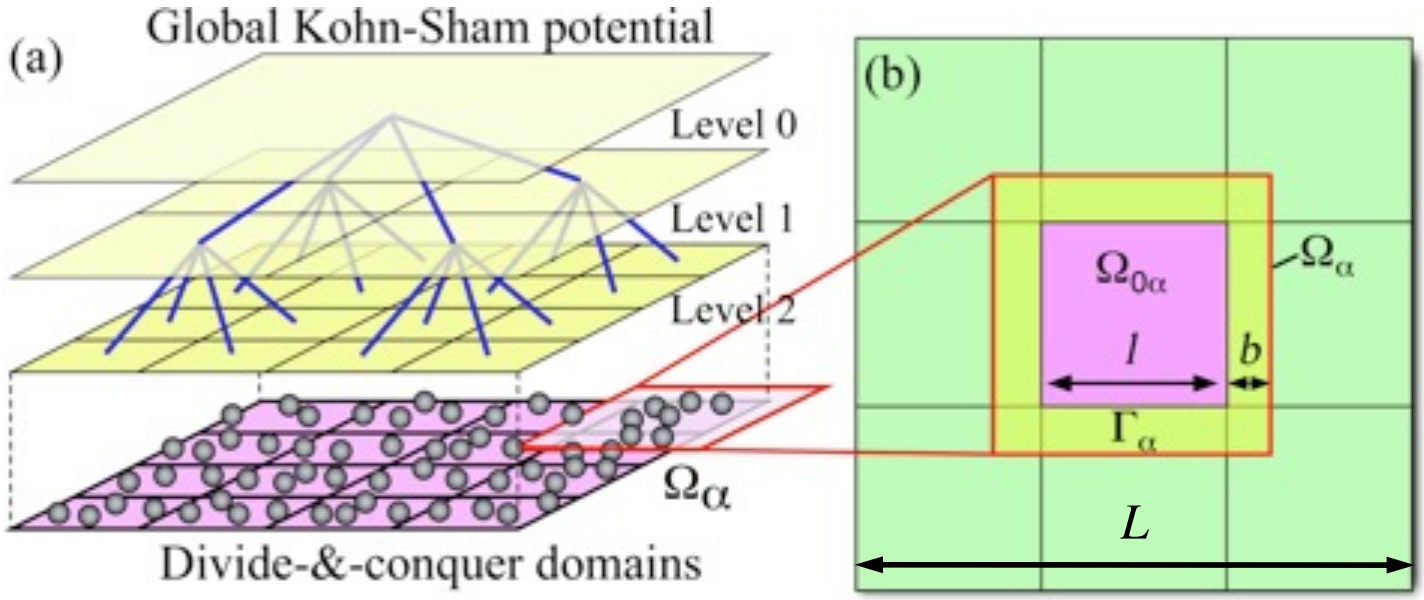
Global chemical potential

$$N = \int d\mathbf{r}\, \rho_{\mathrm{global}}(\mathbf{r})$$

cf. subsystem DFT [W. Mi *et al.*, *Comput. Phys. Commun.* **269**, 108122 ('21)]

Optimization of Divide-&-Conquer DFT

- **Computational parameters of DC-DFT = domain size (l) + buffer thickness (b)**



- **Complexity analysis to optimize the domain size l**

$$l_* = \mathrm{argmin}\left(T_{\mathrm{comp}}(l)\right) = \mathrm{argmin}\left(\left(\frac{L}{l}\right)^3 (l+2b)^{3\nu}\right) = \frac{2b}{\nu-1}$$

Per-domain computational complexity of DFT = $O(n^\nu)$: $\nu = 2$ or 3 ($n <$ or $> 10^3$)

- **Error analysis: Buffer thickness b is dictated by the accuracy requirement**

$$b = \lambda \ln\left(\max\left\{|\Delta\rho_\alpha(\mathbf{r})| \,|\, \mathbf{r} \in \partial\Omega_\alpha\right\}\right)/\varepsilon\langle\rho_\alpha(\mathbf{r})\rangle \qquad |\Delta\rho|e^{-b/\lambda} = \varepsilon\langle\rho\rangle$$

Decay length — $\rho_\alpha(\mathbf{r}) - \rho_{\mathrm{global}}(\mathbf{r})$ — Error tolerance

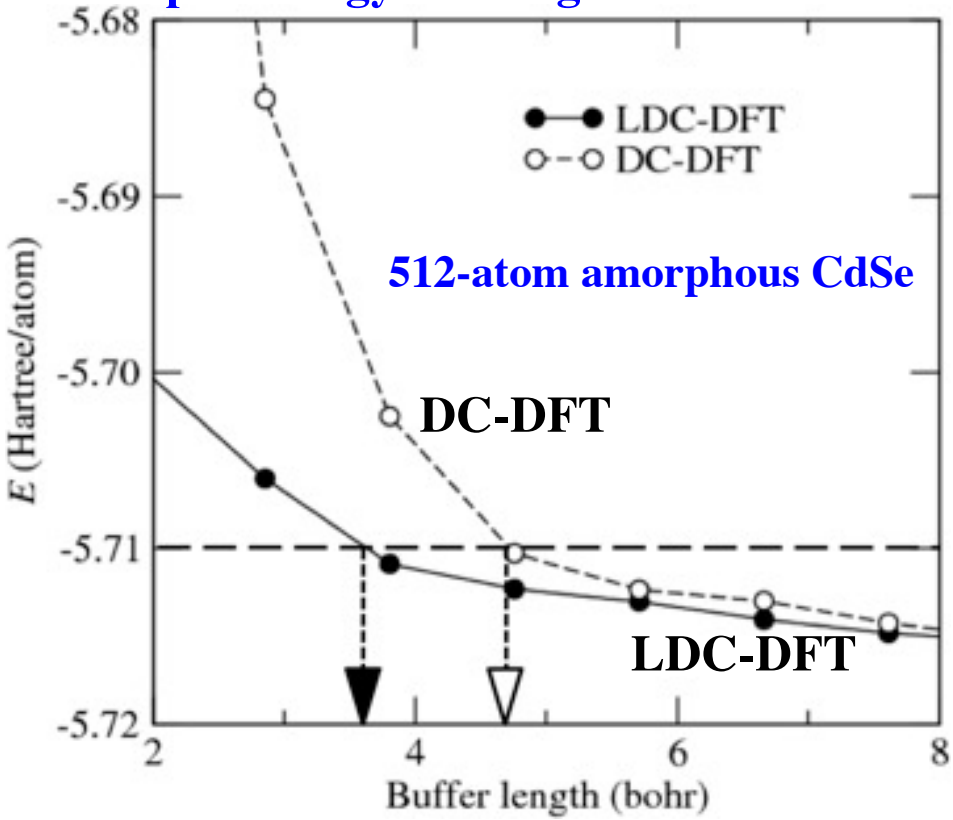
cf. quantum nearsightedness [Kohn, *Phys. Rev. Lett.* **76**, 3168 ('96); Prodan & Kohn, *P. Nat. Acad. Sci.* **102**, 11635 ('05)]

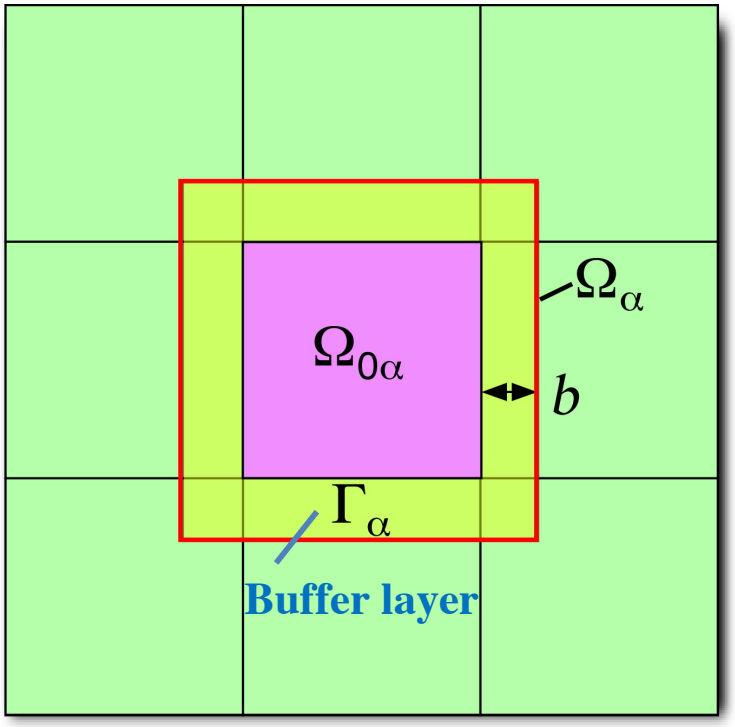
Lean Divide-&-Conquer (LDC) DFT

- **Density-adaptive boundary potential to reduce the *O*(*N*) prefactor** local approximation

$$v_\alpha^{\mathrm{bc}}(\mathbf{r}) = \int d\mathbf{r}' \frac{\partial v(\mathbf{r})}{\partial \rho(\mathbf{r}')} \left(\rho_\alpha(\mathbf{r}') - \rho_{\mathrm{global}}(\mathbf{r}') \right) \cong \frac{\rho_\alpha(\mathbf{r}) - \rho_{\mathrm{global}}(\mathbf{r})}{\xi}$$

- **More rapid energy convergence of LDC-DFT compared with nonadaptive DC-DFT**



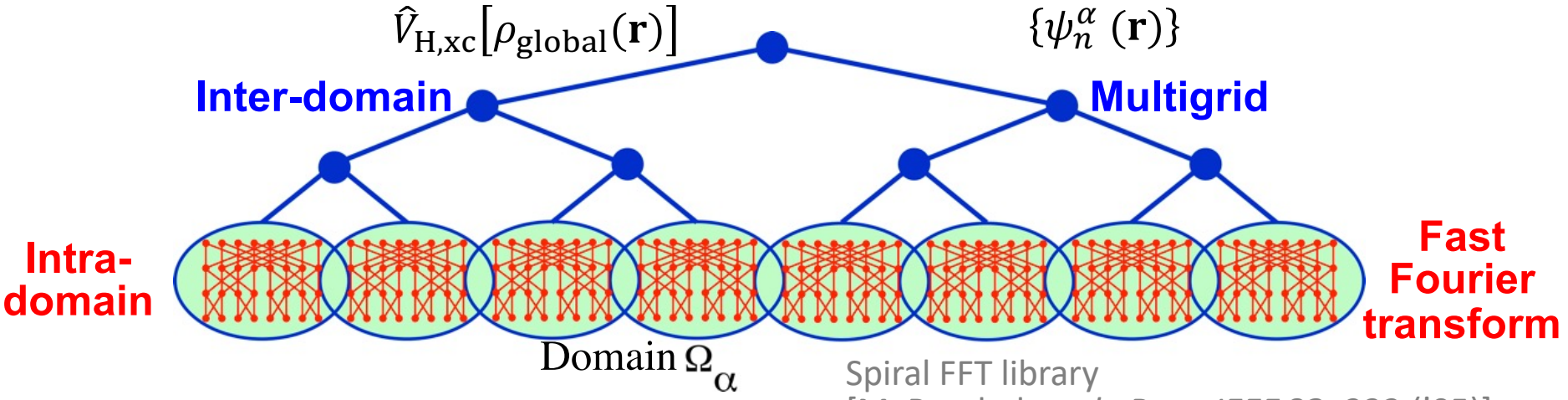


- **Factor 2.03 (for ν = 2) ~ 2.89 (for ν = 3) reduction of the computational cost with an error tolerance of 5×10^{-3} a.u. (per-domain complexity: n^ν)**

F. Shimojo *et al.*, *J. Chem. Phys.* **140**, 18A529 ('14);
Phys. Rev. B **77**, 085103 ('08); *Comput. Phys. Commun.* **167**, 151 ('05)

Hierarchical Computing

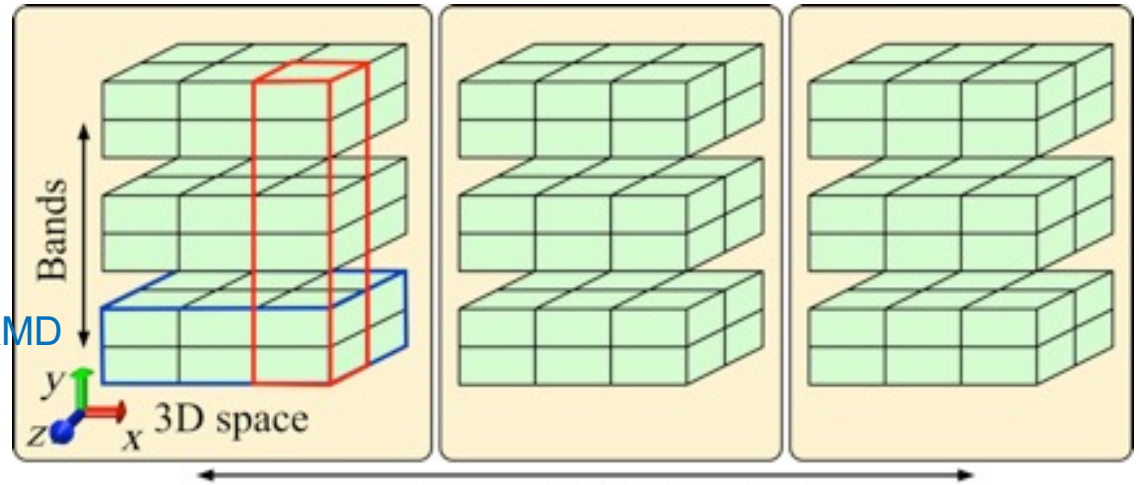
- **Globally scalable (real-space multigrid) + locally fast (plane wave) electronic solver**



Spiral FFT library
[M. Puschel *et al.*, *Proc. IEEE* **93**, 232 ('05)]

cf. globally- sparse-yet-locally-dense eigensolver [J. H. Lam *et al.*, *Nature Commun*. **15**, 3479 ('24)]

- **Hierarchical band (*i.e.*, Kohn-Sham orbital) + space + domain (BSD) decomposition**



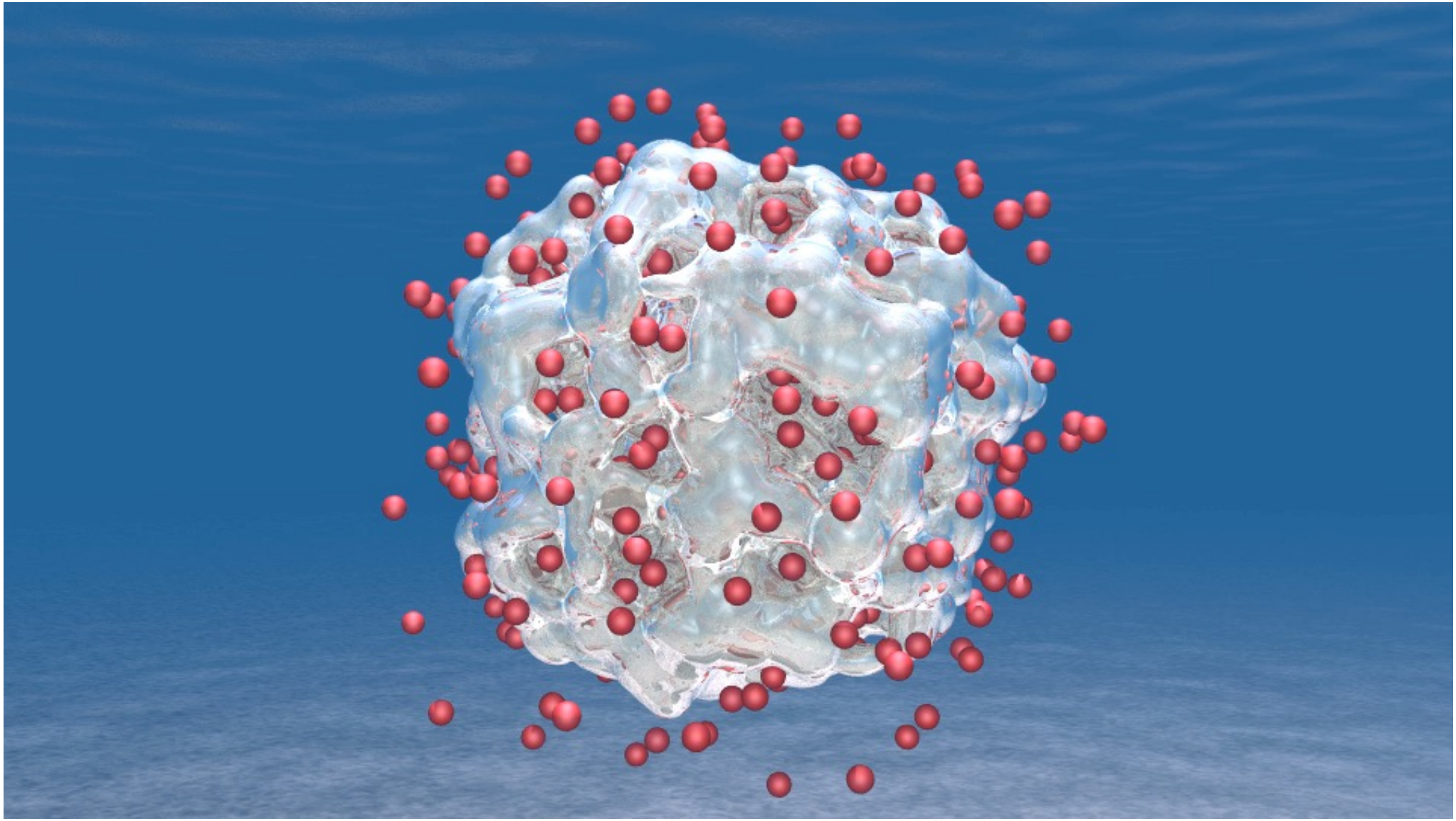
QM-nodes in QXMD

MD-nodes in QXMD software

H_2 Production from Water Using LiAl Particles

16,661-atom QMD simulation of $Li_{441}Al_{441}$ in water on 786,432 IBM Blue Gene/Q cores

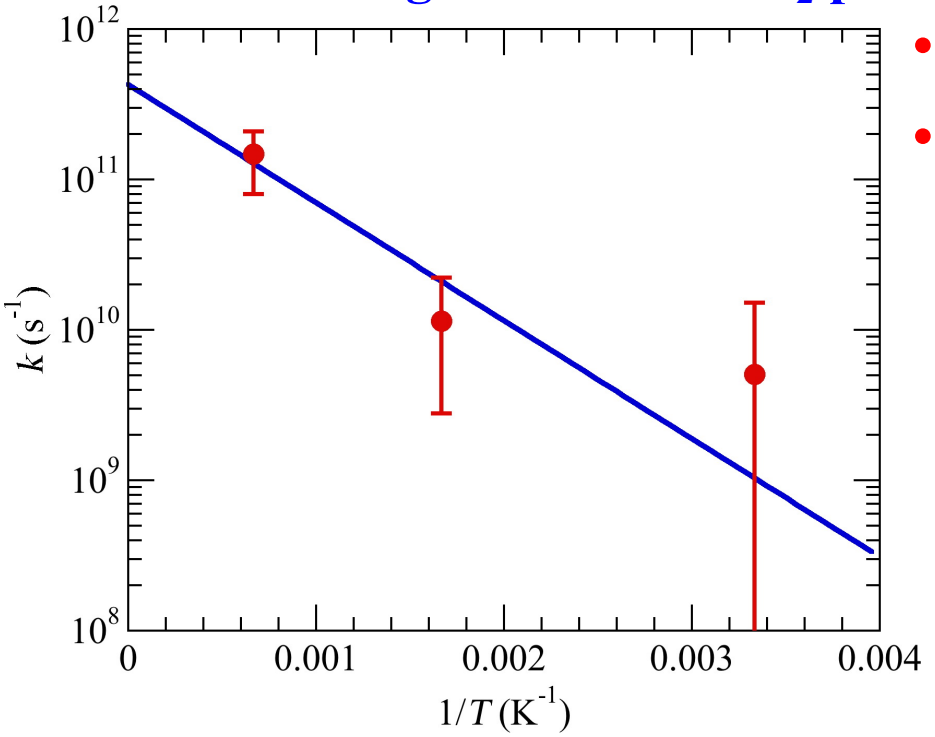
K. Shimamura *et al.*, *Nano Lett.* **14**, 4090 ('14)



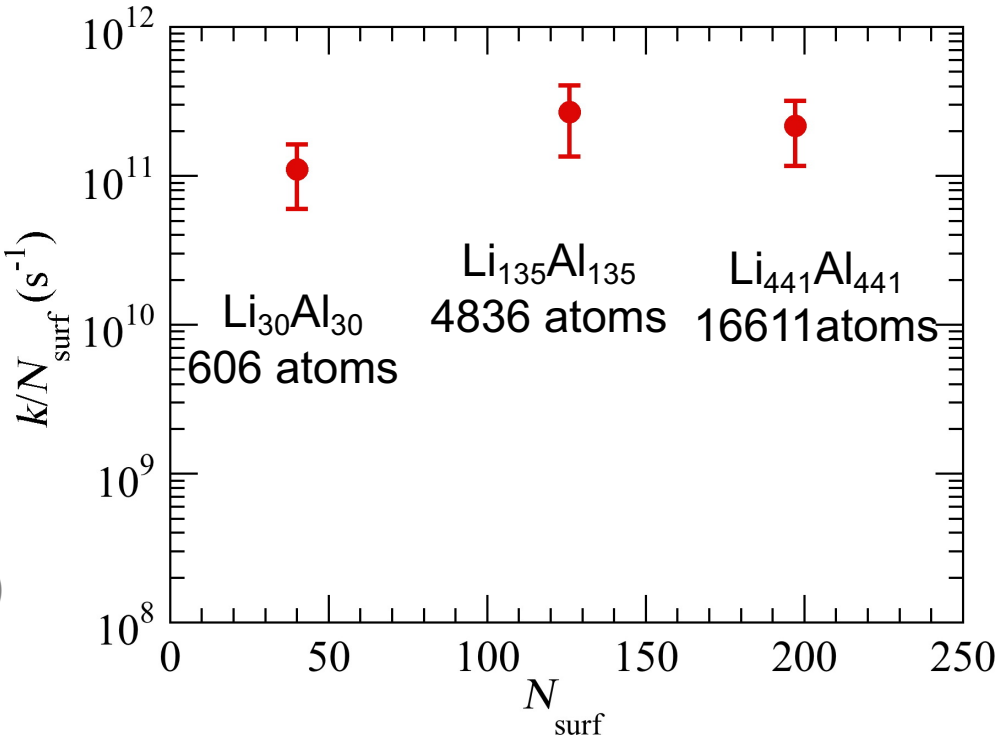
21,140 time steps (129,208 self-consistent-field iterations)

Rapid & Scalable H_2 Production

- **Orders-of-magnitude faster H_2 production from water than with pure Al**



- **Activation barrier = 0.068 eV**
- **Reaction rate = 1.04×10^{9} (s^{-1}) per LiAl pair at 300 K**



K. Shimamura *et al.*, *Nano Lett.* **14**, 4090 ('14)
K. Nomura *et al.*, *IEEE/ACM SC14* ('14)

- **Reaction rate does not decrease for larger particles → industrial scalability**

***cf.* discontinuous Galerkin DFT** [Lin *et al.*, *J. Comput. Phys.* **231**, 2140 ('12)]

See notes on (1) origin of DCDFT, (2) parallel DCDFT, (3) DCDFT data structures, (4) DCDFT algorithm, (5) DC forces, & (6) lean DCDFT

Fermi Operator

- **Fermi operator**

$$F(\hat{H}) = \frac{2}{\exp\left(\dfrac{\hat{H}-\mu}{k_{\mathrm{B}}T}\right)+1}$$

- **Projection to the occupied subspace**

$$\left|\psi_{\mathrm{proj}}\right\rangle = F(\hat{H})\left|\psi\right\rangle$$

- **The expectation value of any operator *A* is obtained by**

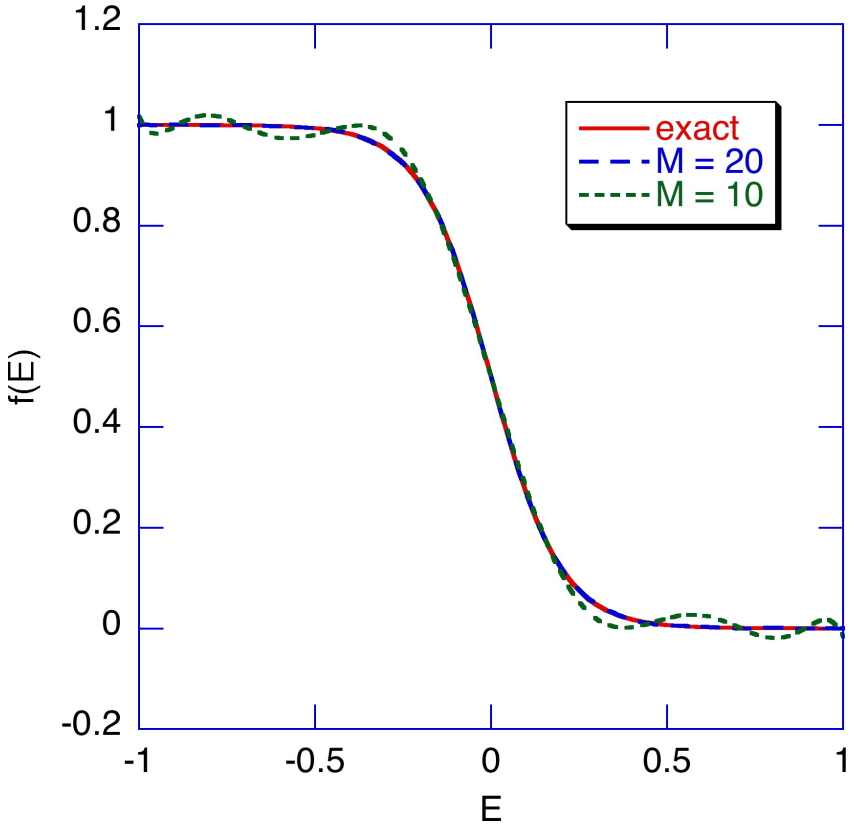
$$\left\langle \hat{A} \right\rangle = \mathrm{tr}\left[\hat{A}\hat{F}\right]$$

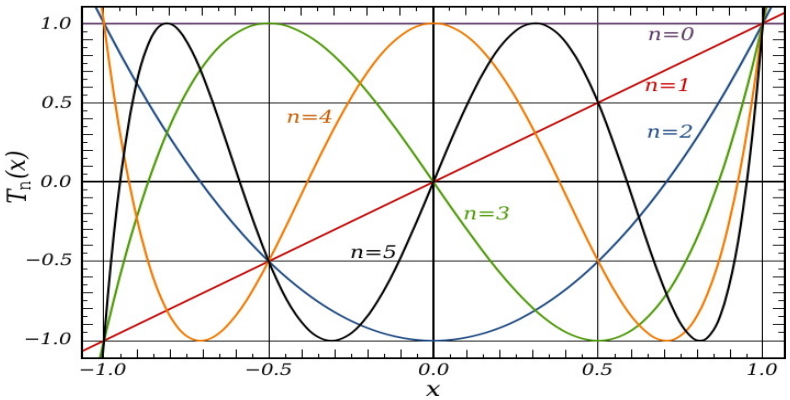
- **Widely used in *O*(*N*) electronic structure calculations (*N* = number of electrons) through its sparse representation**

S. Goedecker, *Rev. Mod. Phys*. **71**, 1085 ('99);
K. Tsuruta *et al*., *Phil. Mag. Lett*. **81**, 357 ('01)

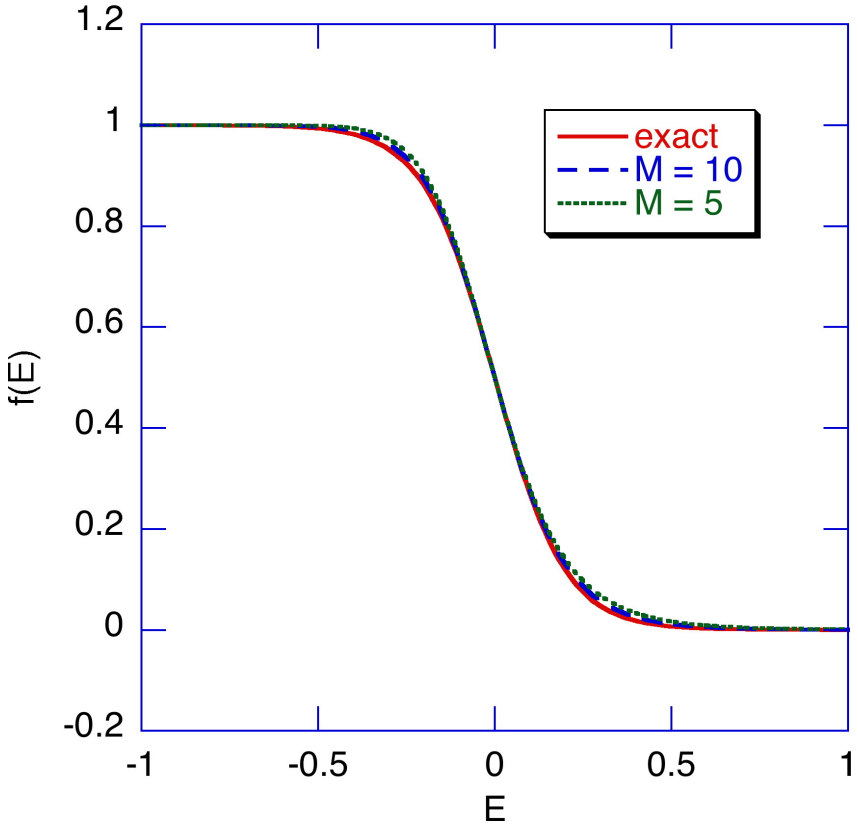
Fermi-Operator Approximations

Chebyshev polynomial





Rational



$$F(\hat{H}) \cong \sum_{\nu=1}^{M} \frac{R_\nu}{\hat{H} - z_\nu}$$

$$\left(\hat{H} - z_\nu\right)\left|\psi_{\text{out}}^{\nu}\right\rangle = R_\nu \left|\psi_{\text{in}}\right\rangle$$

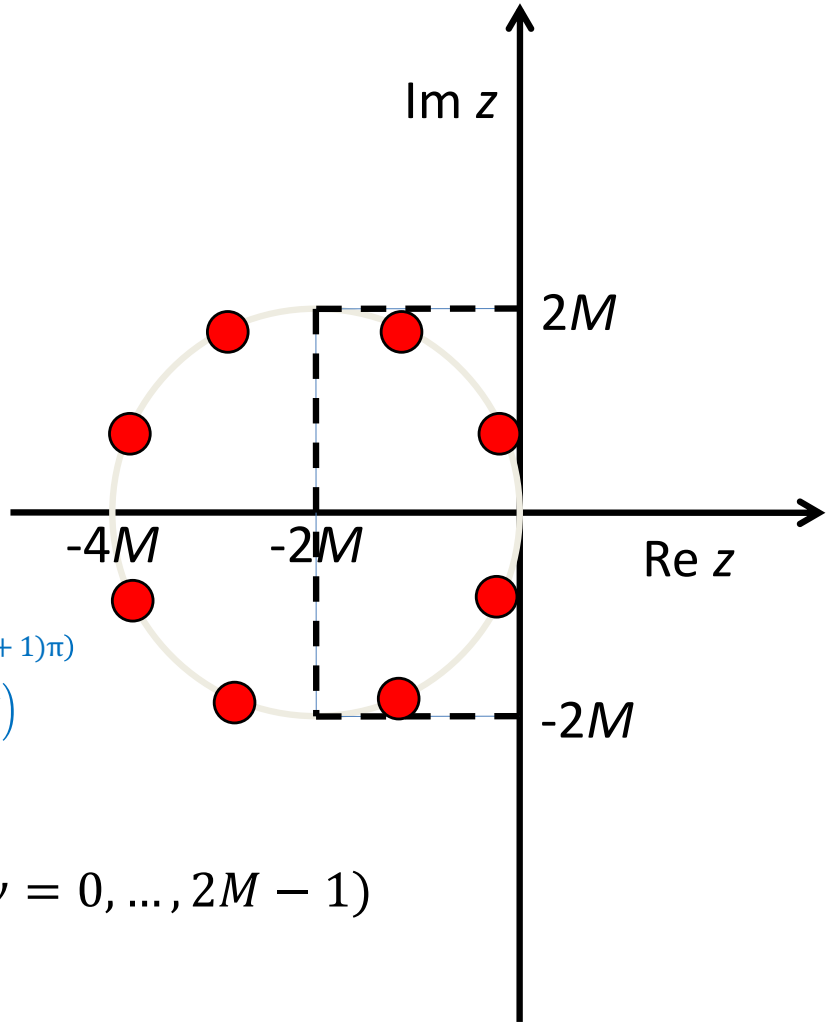
Rational Fermi-Operator Expansion

$$f(z) = \frac{1}{\exp(z)+1}$$

$\left(1+\frac{z}{n}\right)^n \xrightarrow{n\to\infty} \exp(z)$

$$\cong \frac{1}{\left(1+\frac{z}{2M}\right)^{2M}+1}$$

$$\cong \sum_{\nu=0}^{2M-1} \frac{R_\nu}{z-z_\nu}$$



$\left(1+\frac{z_\nu}{2M}\right)^{2M} = -1 = \exp(i(2\nu+1)\pi)$

$\therefore 1+\frac{z_\nu}{2M} = \exp\left(i\pi\frac{2\nu+1}{2M}\right)$

Poles

$$z_\nu = 2M\left(\exp\left(i\frac{(2\nu+1)\pi}{2M}\right)-1\right)$$

$$R_\nu = -\exp\left(i\frac{(2\nu+1)\pi}{2M}\right)$$

Residues

$(\nu = 0, \ldots, 2M-1)$

Expand the denominator in $\Delta = z - z^\nu$ & keep the linear term

D. M. C. Nicholson *et al.*, *Phys. Rev. B* **50**, 14686 ('94);
A. P. Horsfield *et al.*, *Phys. Rev. B* **53**, 12694 ('96);
L. Lin *et al.*, *J. Phys. Condes. Matter* **25**, 1295501 ('13)

$O(N)$ Fermi Operator Expansion

- **Truncated expansion of Fermi-operator by Chebyshev polynomial $\{T_p\}$**

$$F(\hat{H}) \cong \sum_{p=0}^{P} c_p T_p(\hat{H})$$

- ***O*(*N*) algorithm**

prepare a basis set of size $O(N)$
(let the size be N for simplicity)

for $l = 1, N$

let an N-dimensional unit vector be $|e_l\rangle = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} l$

recursively construct the l^{th} column of matrix T_p, $\left|t_l^p\right\rangle$, keeping only $O(1)$ off-diagonal elements (*cf.* quantum nearsightedness)

$$\begin{cases} \left|t_l^0\right\rangle = |e_l\rangle \\ \left|t_l^1\right\rangle = \hat{H}|e_l\rangle \\ \left|t_l^{p+1}\right\rangle = 2\hat{H}\left|t_l^p\right\rangle - \left|t_l^{p-1}\right\rangle \end{cases}$$

build a sparse representation of the l^{th} column of F as

$$|f_l\rangle \cong \sum_{p=0}^{P} c_p \left|t_l^p\right\rangle$$

See note on Fermi-operator expansion

Local Orbital Minimization

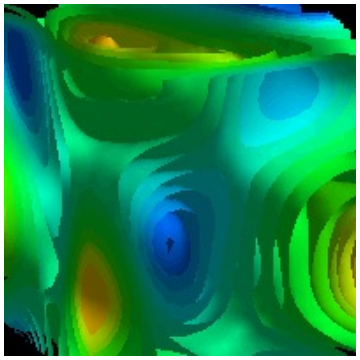
O(*N*) DFT algorithm

- **Asymptotic decay of density matrix:**

$$\rho(\mathbf{r},\mathbf{r}') \equiv \sum_{n=1}^{N_{\mathrm{el}}} \psi_n(\mathbf{r})\psi_n^*(\mathbf{r}')$$

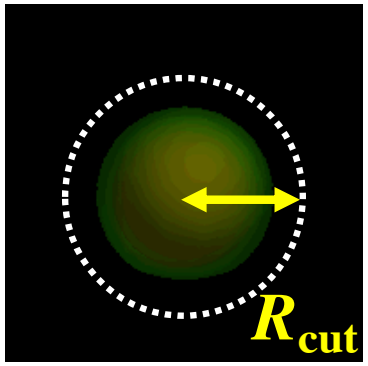
Nearsightedness! $\propto \exp(-C\,|\mathbf{r}-\mathbf{r}'|)$

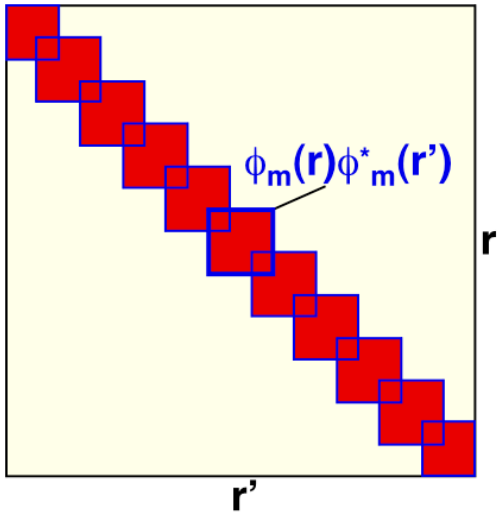
- **Localized functions:**



$\psi_n(\mathbf{r})$ $\longrightarrow$ $\phi_m(\mathbf{r})$

$$\phi_m(\mathbf{r}) = \sum_n \psi_n(\mathbf{r})U_{nm}$$





- **Unconstrained minimization:**

$$\tilde{E}[\{\phi_n\}] = \sum_{m=1}^{N_{\mathrm{wf}}}\sum_{n=1}^{N_{\mathrm{wf}}} \int d^3r\,\phi_m^*(\mathbf{r})(H-\eta I)\phi_n(\mathbf{r})\left(2\delta_{nm} - \int d^3r\,\phi_n^*(\mathbf{r})\phi_m(\mathbf{r})\right) + \eta N_{\mathrm{el}}$$

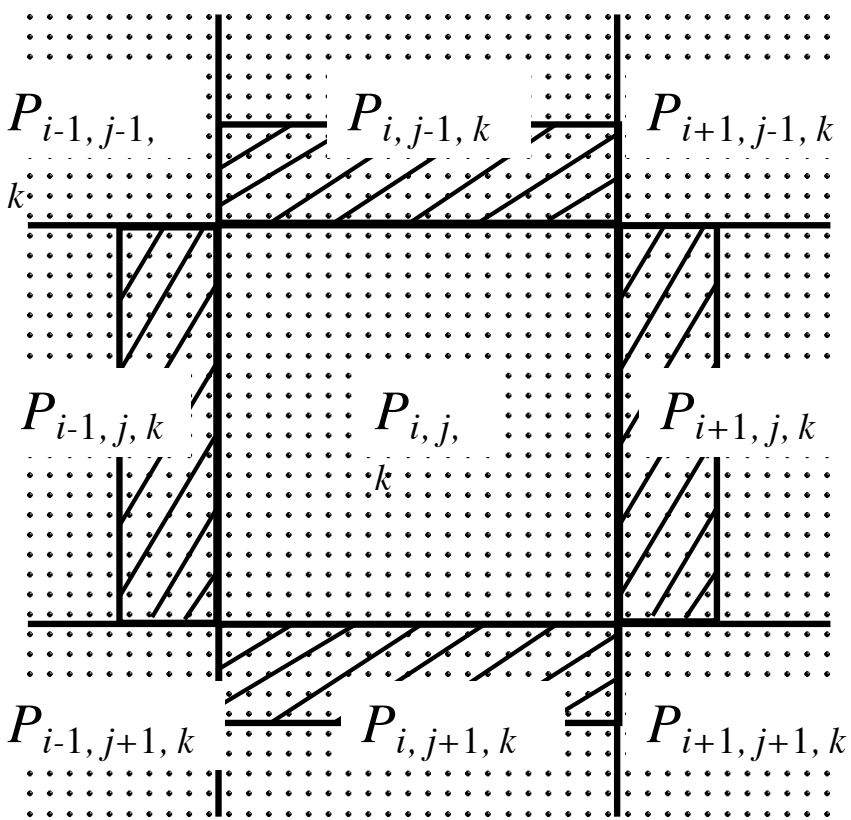
***cf.* *O*(*N*³) QR or Cholesky decomposition for orthogonalization**

See notes on (1) nonorthogonal orbitals & (2) local orbital minimization

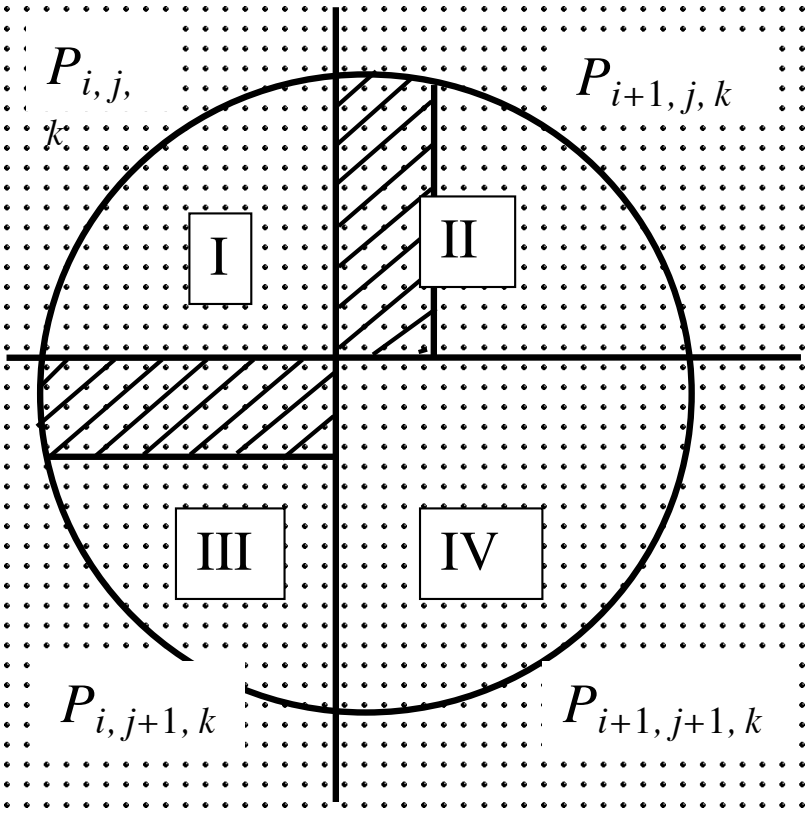
Kim, Mauri & Galli, *Phys. Rev. B* **52**, 1640 ('95); Ordejon *et al.*, *ibid*. **51**, 1456 ('95)
Shimojo *et al.*, *Comput. Phys. Commun*. **140**, 303 ('01)

Analysis of Parallel DFT Algorithms

Spatial decomposition



DFT



Linear-scaling (LS) DFT

	Computation	Communication
DFT	$O(N^3/P)$	$O(N(N/P)^{2/3} + N^2 \log P)$
DC-DFT	$O(N/P)$	$O((N/P)^{2/3})$

Shimojo *et al.*, *Comput. Phys. Commun.* **140**, 303 ('01)

Zoo of $O(N)$ DFT Algorithms (Small Subset)

- **Density-matrix minimization (DMM)** [Li *et al.*, *Phys Rev. B* **47**, 10891 ('93); Nunes & Vanderbilt, *ibid*. **50**, 17611 ('94); Hernandez *et al.*, *ibid*. **53**, 7147 ('96)]
  See notes on (1) idempotency, (2) orthogonal DMM, (3) nonorthogonal DMM & (4) real-space DMM
- **Filter diagonalization** [Wall & Neuhauser, *J. Chem. Phys*. **102**, 8011 ('95)]
  See note on filter diagonalization *cf.* **Koshiba's "egg of dream":** 小柴昌俊–夢の卵を孵す
- **Green's function (GF) approaches** [Horsfield *et al.*, *Phys. Rev. B* **53**, 12694 ('96); Hoshi *et al.*, *Proc. SCALA16* ('16)]
  See (1) slide & notes 1 & 2 on Lanczos tridiagonalization, and (2) note on Pade *via* Lanczos
  cf. Bethe lattice & Laughlin *et al.*, *Phys. Rev. B* **20**, 5228 ('79)
- **Block tridiagonal divide-&-conquer (DC)** [Gansterer *et al.*, *ACM T. Math. Software* **28**, 45 ('02)]
  See note on block tridiagonal DC
- **Embedded-cluster boundary condition (BC): Reduce $O(N)$ prefactor of DCDFT?**
  See notes on (1) embedded-cluster BC, (2) orbital BC, (3) multiple scattering & (4) surface GF
  cf. **Dynamical mean-field theory (DMFT)** [Georges *et al.*, *Rev. Mod. Phys*. **68**, 13 ('96); Kotliar *et al.*, *ibid*. **78**, 865 ('06)]

Stochastic DFT

- **Project onto a vector space spanned by a set of *I* stochastic wave functions** $\kappa(r)$

  $\hat{I} = \sum_{\kappa=1}^{I} |\kappa\rangle\langle\kappa|$

- **Electron density is obtained by projecting onto the occupied subspace spanned by**

  $|\xi\rangle = \theta(\mu - \hat{H})|\kappa\rangle$

  where the step function θ is approximated by Chebyshev polynomials with μ & $\hat{H}$ being the chemical potential & Kohn-Sham Hamiltonian

- **With sparse representation of $\hat{H}$, the algorithmic complexity can be sublinear, since the order of Chebyshev expansion to achieve a prescribed accuracy decreases as a function of the number of electrons (self-averaging)**

Baer *et al*., *Phys. Rev. Lett*. **111**, 106402 ('13); *Annu. Rev. Phys. Chem*. **73**, 255 ('22)

- **Sublinear stochastic algorithm also applies to TDDFT**

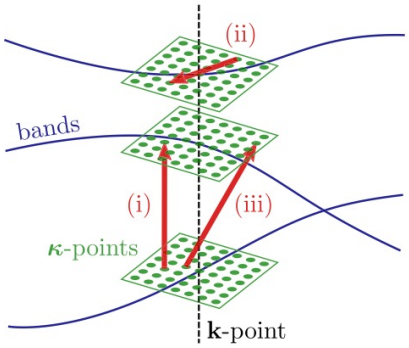
Vlcek *et al*., *J. Chem. Phys*. **150**, 184118 ('19)

- **Key idea: Projection onto randomized vector space**

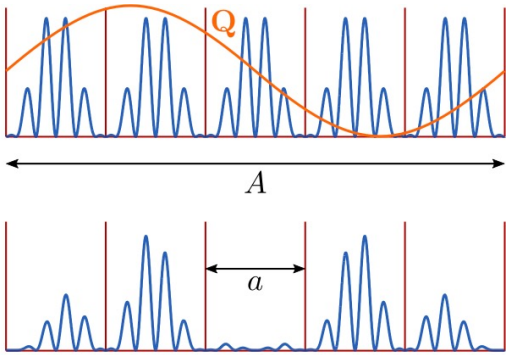
cf. randomized linear algebra [Murray *et al*., *arXiv*: 2302.11474 ('23)]

UltraQ DFT: Recombine in Reciprocal Space

- **Additional sum over a finer grid around each k-point in the reciprocal space describes physics at ultra-long length scales (generalized Block theorem)**



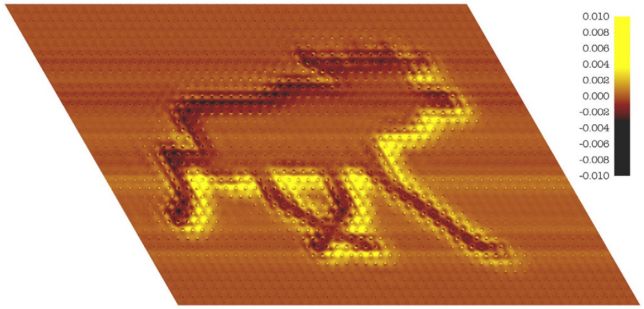
(a)



(b)

Ultrafine Q modulation

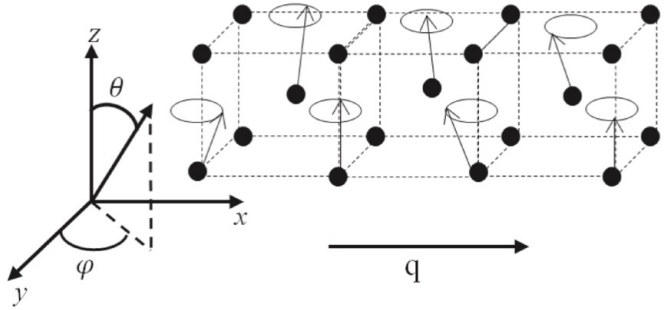
SCF electron density
48×36 LiF unit cells

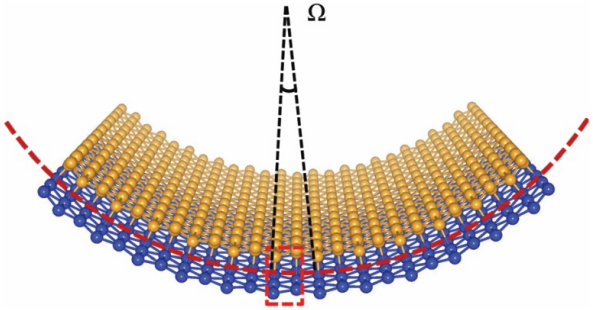


Müller *et al.*, *Phys. Rev. Lett.* **125**, 256402 ('20)

- **Generalized Block theorem constrains the Hamiltonian invariant under slow spatial modulation (*e.g.*, spiral magnet & bended 2D sheet) *via* Block-like k-point sampling**

Prayitno *et al.*, *JPSJ* **87**, 114709 ('18); Shi *et al.*, *PCCP* **22**, 11567 ('20)



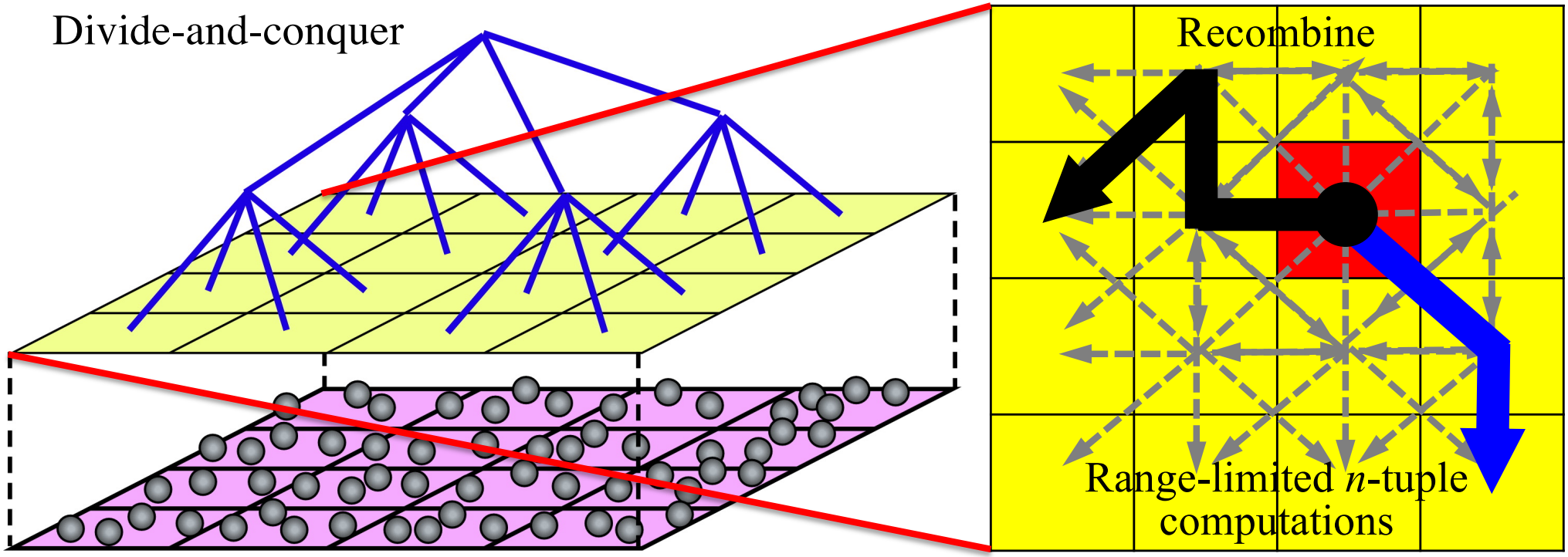


- **Wannier interpolation: Sparse k-point sampling → construct smooth Wannier functions in real space (*cf.* tight-binding Hamiltonian) → transform back to dense reciprocal space** [Wannier-Berri code: Tsirkin, *npjCM* **7**, 33 ('21)]

Related Topics: Research = Directed Random Walk

- **Fragment molecular orbital (FMO) method: fragment dimer & beyond**
  Kitaura *et al.*, *Chem. Phys. Lett.* **312**, 319 ('99); Tanaka *et al.*, *ibid*. **556**, 272 ('13)
- **Linearly scaling 3D fragment (LS3DF) method ~ real-space FMO**
  Wang *et al.*, *Proc. Supercomputing, SC08* ('08)
- **Density matrix renormalization group (DMRG)**
  White, *Phys. Rev. B* **48**, 10345 ('93)
  See notes on (1) singular value decomposition (SVD) & (2) DMRG
- **Tensor networks: modern DMRG — combine with message passing (belief propagation or Bethe lattice) in 3D?**
  Chan & Sharma, *Annu. Rev. Phys. Chem.* **62**, 465 ('11)

Next: Divide-Conquer-Recombine (DCR)



M. Kunaseth *et al.*, *ACM/IEEE SC13* ('13)

Globally-informed local DC solutions are used in the recombine phase as compact bases to synthesize global properties in broad applications

F. Shimojo *et al.*, *J. Chem. Phys.* **140**, 18A529 ('14)
K. Nomura *et al.*, *IEEE/ACM SC14* ('14)
T. Razakh *et al.*, *IEEE PDESC* ('24)

Divide-Conquer-Recombine Applications

- **High-order inter-molecular-fragment correlation**
  Tanaka *et al.*, *Chem. Phys. Lett.* **556**, 272 ('13)
- **Global frontier orbitals (HOMO & LUMO) *via* linear combination of domain orbitals (LCDO)**
  Yamada *et al.*, *Phys. Rev. B* **95**, 045106 ('17)
  It's projection!
  Note the projection (P & Q = 1–P) — generalized Langevin approach?
  Mori, *Prog. Theor. Phys.* **33**, 423 ('65); Kinjo & Hyodo, *Mol. Sim.* **33**, 417 ('07)
- **Dielectric response: Simplified inter-layer coupling in layered materials; equation-of-motion decoupling for density response function, *cf.*** linear-response time-dependent DFT in nonadiabatic QMD lecture
  Andersen *et al.*, *Nano Lett.* **15**, 4616 ('15)
- **Electron dynamics: Local plane-wave time-dependent KS equations glued together by Maxwell's equations**
  Sato & Yabana *et al.*, *J. Adv. Sim. Sci. Eng.* **1**, 98 ('14)
- **Exciton dynamics: Local nonadiabatic QMD simulations recombined into a global exciton-kinetics graph**
  Mou *et al.*, *Appl. Phys. Lett.* **102**, 173301 ('13)

Divide-Conquer-Recombine Approach (1)

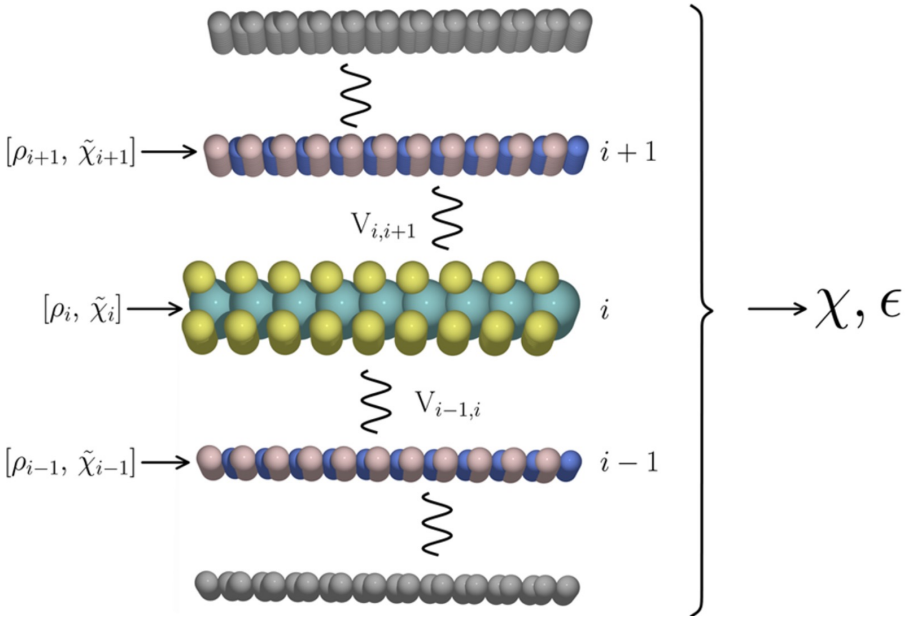
- **Dielectric response: Simplified inter-layer coupling in layered materials; equation-of-motion decoupling for density response function, *cf.*** linear-response time-dependent DFT in nonadiabatic QMD lecture
- **Density response function**

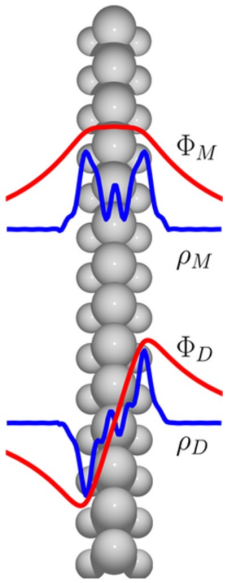
$$n_{\mathrm{ind}}(\mathbf{r},\omega) = \int dr' \chi(\mathbf{r},\mathbf{r}',\omega) V_{ext}(\mathbf{r}',\omega)$$

- **Multiscale Dyson equations**

$$\tilde{\boldsymbol{\chi}} = \boldsymbol{\chi}^{\mathbf{0}} + \boldsymbol{\chi}^{\mathbf{0}}\tilde{\mathbf{V}}\tilde{\boldsymbol{\chi}} \qquad \text{Intra-layer (accurate)}$$

$$\boldsymbol{\chi} = \tilde{\boldsymbol{\chi}} + \tilde{\boldsymbol{\chi}}\mathbf{V}^{\mathbf{I}}\boldsymbol{\chi} \qquad \text{Inter-layer (simplified, 2 DOF)}$$



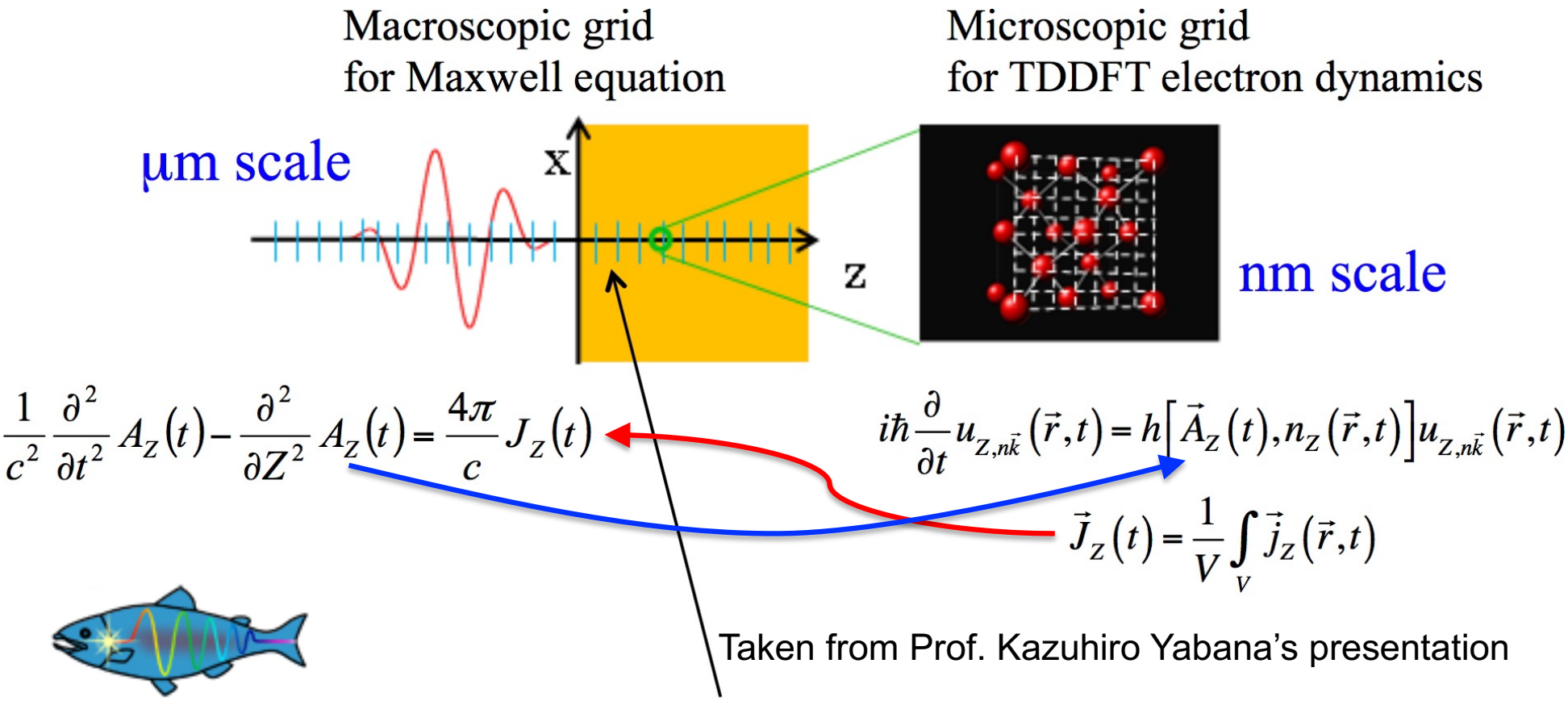


Andersen *et al.*, *Nano Lett.* **15**, 4616 ('15)

See Dyson equations in a nutshell

Divide-Conquer-Recombine Approach (2)

- **Electron dynamics — Maxwell + TDDFT: Local plane-wave time-dependent KS equations glued together by global Maxwell's equations**



Taken from Prof. Kazuhiro Yabana's presentation

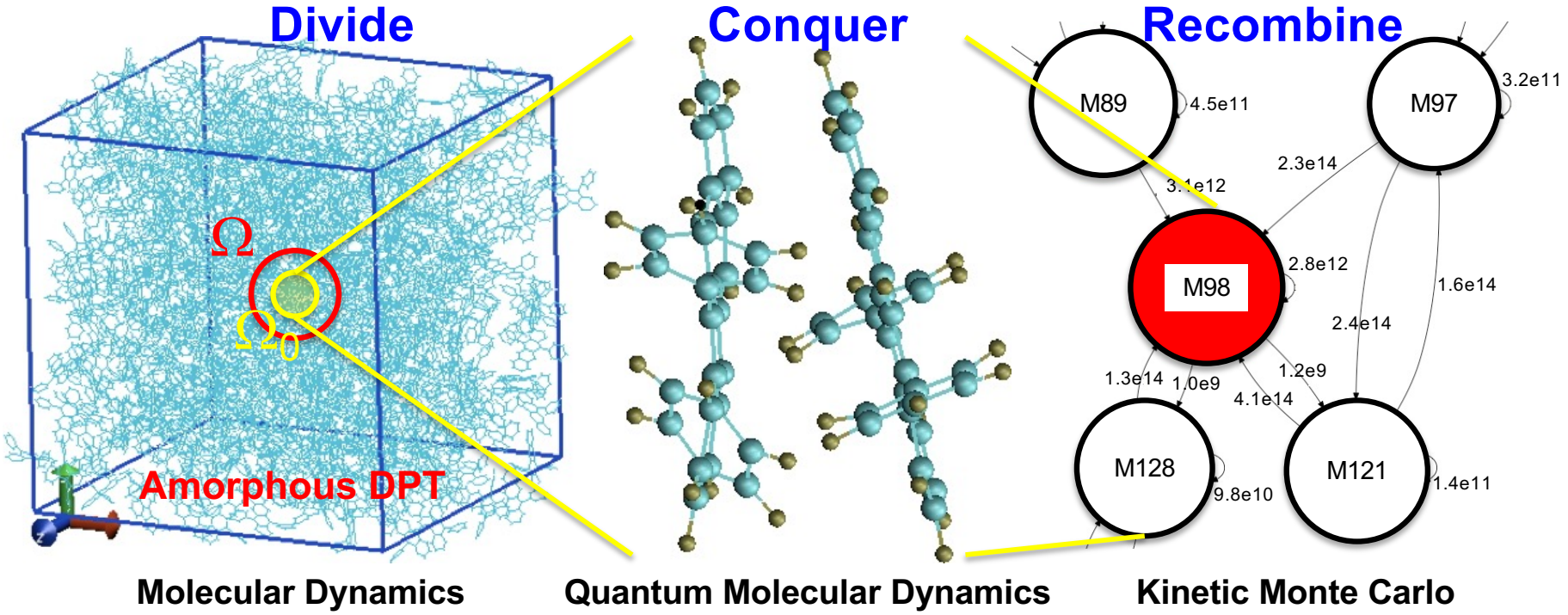
SALMON

http://salmon-tddft.jp

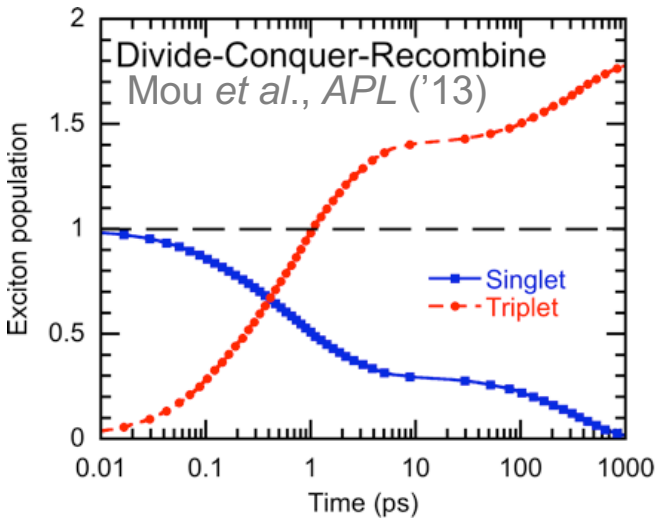
At each macroscopic grid point,
we solve real-time electron dynamics in parallel

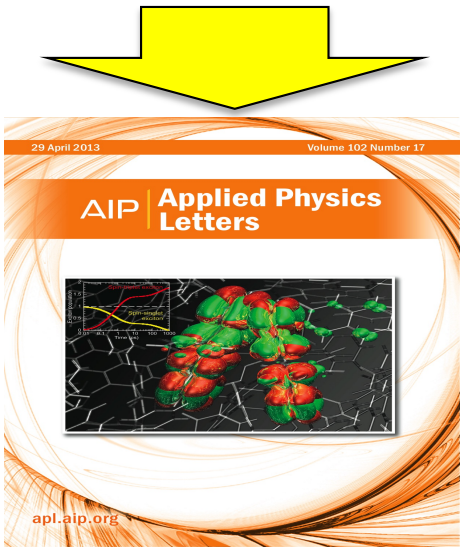
Sato & Yabana *et al.*, *J. Adv. Sim. Sci. Eng.* **1**, 98 ('14)

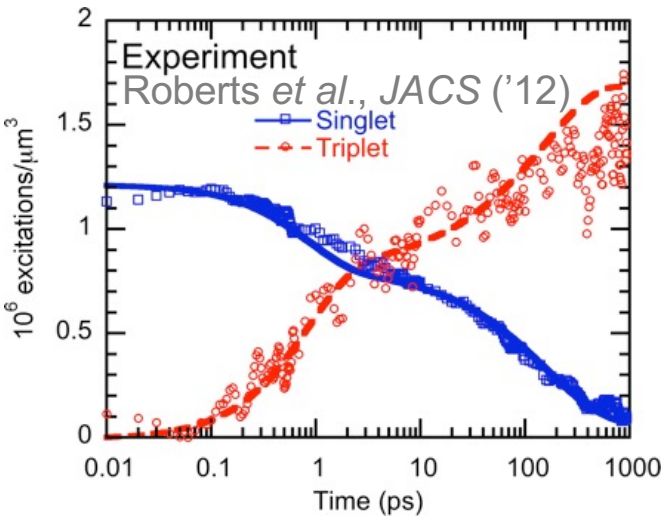
Divide-Conquer-Recombine Approach (3)



Global kinetic graph: Experimental Length & Time Scales

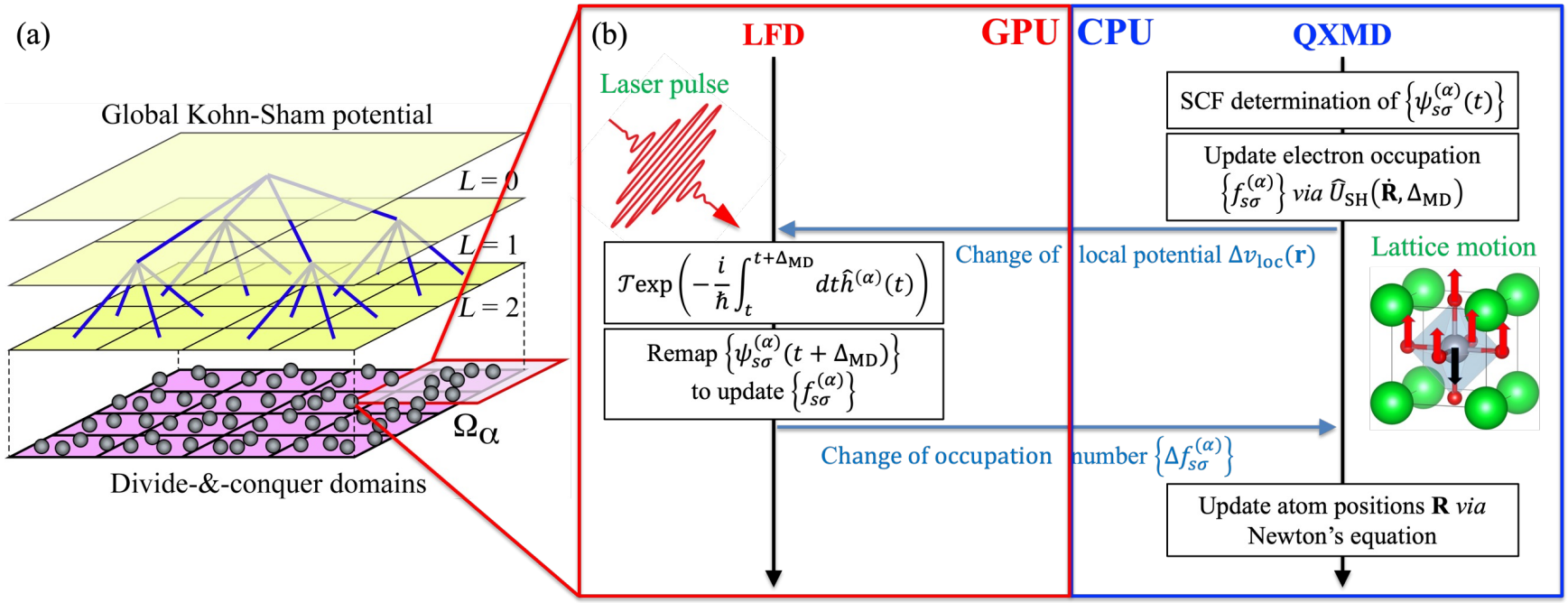






Divide-Conquer-Recombine Approach (4)

Divide-&-conquer Maxwell-Ehrenfest-surface hopping (DC-MESH) method



- **Incorporate multi-physics at appropriate scales & levels of approximation: (1) long-range Hartree coupling of divide-&-conquer (DC) domains; (2) short-range exchange-correlation (*e.g.*, nonlocal xc functional) & light-matter interaction within each domain**
- **See** DC-MESH lecture [Linker *et al.*, *Science Adv.* **8**, eabk2625 (2022); Razakh *et al.*, *PDSEC* (IEEE, '24)]