The crucial role of parallel $d = 6$ FFT in a new computational algorithm for electronic structure

D. Foerster

Laboratoire Ondes et Matière d'Aquitaine (LOMA), Université de Bordeaux

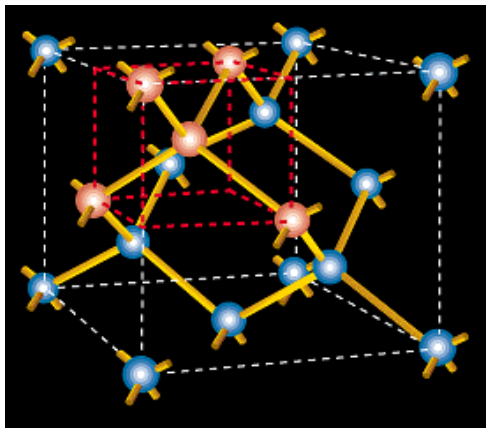
SIAM, CSE21, March 1-5, 2021

Outline

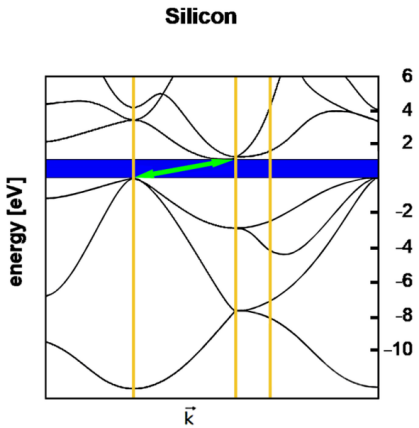
Gaps of semiconductors

Semiconductors (SC) $\rightarrow$ electronics, LED lighting, solar cells ...
Liquid of interacting electrons in crystalline environment.

Electronic wave functions $\psi_p(r) = e^{ip\cdot r}\phi_p(r)$, with $\phi_p(r)$ periodic in **unit cell**
Band energies $E(p)$ depend on **Bloch momenta** $p = (p_1, p_2, p_3)$.



silicon unit cell



bands of silicon

Bands are **filled** $E(p) < 0$ or **empty** $E(p) > 0$.

$$\textbf{band gap} = \min_{p, E(p)>0} E(p) - \max_{p, E(p)\leq 0} E(p)$$

band gap of SC $\rightarrow$ applications.
Example: silicon band gap matches peak in solar spectrum.

GW approximation

Lars Hedin [1] added "self energy" $\Sigma = iGW$ to include Coulomb interaction.

Propagator G, density response χ, screened Coulomb W, self energy Σ

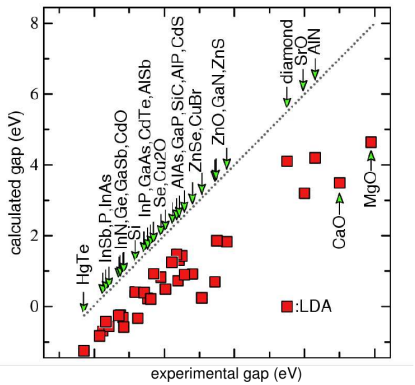
$$\chi_q(r,r',\tau) = -2i \sum_{p_1 - p_2 \equiv q} G^+_{p_1}(r,r',\tau) G^-_{p_2}(r',r,-\tau) \quad (1)$$

$$W_q(g,g',\omega) = V_q^{1/2}(g) \frac{1}{\delta_{gg'} - V_q^{1/2}(g)\chi_q(g,g',\omega)V_q^{1/2}(g')} V_q^{1/2}(g') \quad (2)$$

$$\Sigma_p(r,r',t) = i \sum_{p_1 + p_2 \equiv p} G_{p_1}(r,r',t) W_{p_2}(r',r,t) \quad (3)$$

r, r' points in the unit cell, t =time, p = Bloch momenta.
Gap computed from self energy Σ via Dyson's equation.

Resulting gaps correct for non magnetic semiconductors:



gaps via GW (green) and LDA (red) vs experiment for some semiconductors

Complexity and memory bottleneck of GW

N^3 points (r_1, r_2, r_3) in unit cell.
For n_{atoms} atoms in unit cell $N^3 \simeq 500 * n_{atoms}$ empirically sufficient.
Products of correlators cheap, with $O(N^6) = O(n_{atoms}^2)$ operations

$$\Sigma_p(r, r', t) = i \sum_{p_1 + p_2 \equiv p} G_{p_1}(r, r', t) W_{p_2}(r', r, t)$$

Construction of G_p and the inversion in W_q cost $O(n_{atoms}^3)$ operations.

Complexity of GW :

1. $O(n_{atoms}^3)$ *if correlations fit into memory*
2. $O(n_{atoms}^4)$ *or worse otherwise*

Difference between a day and a year for $n_{atoms} = 100..1000$.

Memory bottle neck. Complex correlation function $F(r, r')$ in the unit cell.

$$\begin{aligned} \text{needed memory} &= 16 \cdot N^6 \text{ bytes} \\ \text{memory per core} &= 2 \cdot 10^9 \text{ bytes} \end{aligned}$$

Only $N \leq 23$ modes/direction fit into a core - too few for interesting semi conductors.

Reduce storage by fft + Nyquist

Consider $d = 3 + 3$ Fourier transform $f(g, g')$ of complex correlator $F(r, r')$

$$f(g, g') = \sum_{r,r'} \exp i(gr - g'r')F(r, r') \tag{4}$$

Nyquist $|g_i|, |g'_i| \leq \frac{N}{4}$ reduces storage by $2^6 = 64$.

With Nyquist, gap still converges towards correct limit [5].
Probable reason: Coulomb interaction is long range.

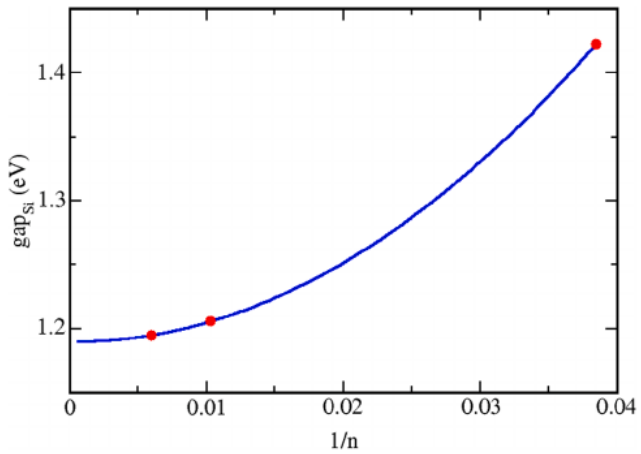


Fig. 1. Computed gap (in eV) of silicon (1.422 eV, 1.21 eV, and 1.20 eV corresponding to 26, 97, and 168 bands with red dots representing the raw data) versus the inverse of the number of bands.

Convergence, with Nyquist's condition, in Si of the band gap, as N tends to infinity

Details of the reduction of storage

Compute density correlator χ from Fourier image

$G^{\pm}_{\mathbf{p}}(\mathbf{g},\mathbf{g}',\tau)$ for $|g_i|, |g'_i| \leq \frac{N}{4}$

$$
\begin{aligned}
 & G^{\pm}_{\mathbf{p}}(g,g',\tau) \overset{Fourier}{\Rightarrow} G^{\pm}_{\mathbf{p}}(r,r',\tau) \\
\chi_q(r,r',\tau) \quad = \quad & -2i \sum_{p_1 - p_2 = q} G^{+}_{p_1}(r,r',\tau) G^{-}_{p_2}(r',r,-\tau) \\
 & \chi_q(r,r',\tau) \overset{Inverse\ Fourier}{\Rightarrow} \chi_q(g,g',\tau)
\end{aligned}
$$

$d = 6$ fft costs only $O(n^2_{atoms}) log(n_{atoms})$ operations.

Without Nyquist, $O(n^3_{atoms})$ GW needs special computers with large memory [4].

With Nyquist, $O(n^3_{atoms})$ GW on a PC.

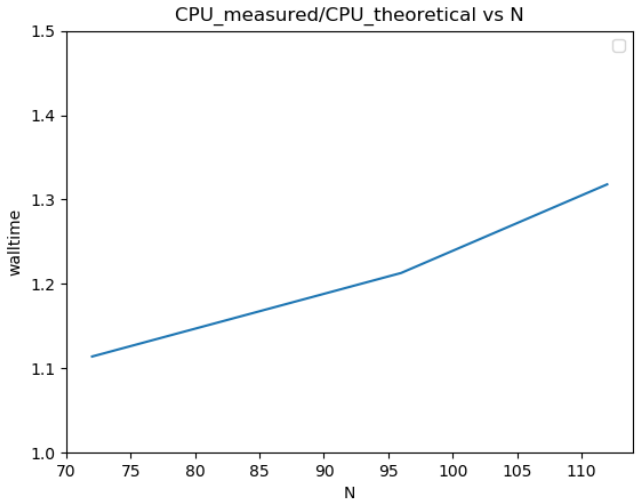
KAUST-Oslo fft in $d = 6$ for $d = 3$ correlators

Distributed correlations need a **distributed** fft in $d = 6$ dimensions

$$F(r, r') = \sum_{g,g'} \exp -i(gr - g'r')f(g, g') \tag{5}$$

Fortunately, the KAUST-Oslo algorithm [6] works in any dimension.

Dr Aseeri of KAUST has found no bandwidth bottleneck up to $N = 112$ and 4096 cores.



wall time, measured/predicted vs N, by Dr Aseeri at Shaheen/KAUST

$N = 112 \rightarrow n_{atoms} = 2800$ - sufficient for organic semiconductors.

Conclusions

- Band gaps of semiconductors involve $d = 3$ correlations $F(r, r')$ -
- If correlators fit into computer memory, an $O(n^3_{atoms})$ algorithm applies -
- Nyquist's condition on the Fourier image of correlations removes memory bottleneck, but correlations must be spread over many cores -
- Use of the KAUST-Oslo fft algorithm in $d = 6$ dimensions should allow $O(n^3_{atoms})$ prediction of bandgaps of semiconductors.

Thanks
I am indebted to Dr. Dalcin of KAUST University for explanations of the KAUST-Oslo algorithm and advice on parallel computing.

Dr Aseeri and Professor Keyes, both also from KAUST University, respectively suggested this contribution and helped with its funding.

Dr Gueddida of Nancy University contributed on form and content of this presentation.


L. Hedin, "New method for calculating the one-particle Green's function with application to the electron-gas problem", Physical. Review **139A**, 796 (1965).


D. Golze, M. Dvorak and P. Rinke, "The GW Compendium: A Practical Guide to Theoretical Photoemission Spectroscopy", Frontiers of Chemistry **7**, 377 (2019)


M. Rieger, L. Steinbeck, I. White, H. Rojas, and R. Godby, "The GW space-time method for the self-energy of large systems", Computational Physics Communications **117**, 211 (1999).


P. Liu, M. Kaltak, J. Klimes, and G. Kresse, "Cubic scaling GW: towards fast quasiparticle calculations", Physical Review **B94**, 165109 (2016).


D. Foerster and S. Gueddida "A low resources space time approach to the GW approximation", Computational Materials Science **187**, 110078 (2021)


L. Dalcin, M. Mortensen and D. E. Keyes, "Fast parallel multidimensional FFT using advanced MPI", Journal of Parallel and Distributed Computing **128**, 137 (2019)