Some Mathematical and Numerical Challenges in First-Principle and Multiscale Modeling of Nano-Systems

Eric CANCES

Ecole des Ponts and INRIA, Paris, France

Banff, August 29 - September 2, 2016

2. Modeling and simulation of infinite aperiodic systems

Molecular simulation is a huge consumer of CPU time.



In some applications, comparison with experiment is not possible.







Total error: $\mathbf{e} = \mathbf{e}_{m} + \mathbf{e}_{d} + \mathbf{e}_{a} + \mathbf{e}_{i} + \mathbf{e}_{c}$

Goal 1: provide error estimates (certification) **Goal 2:** minimize the computation cost to obtain the desired accuracy.



Total error: $\mathbf{e} = \mathbf{e}_{m} + \mathbf{e}_{d} + \mathbf{e}_{a} + \mathbf{e}_{i} + \mathbf{e}_{c}$

Goal 1: provide error estimates (certification) **Goal 2:** minimize the computation cost to obtain the desired accuracy.

$\label{eq:intermediation} \mbox{Implementation error} \ e_i$

• Human error (bugs): manual/automatic code validation.

$\label{eq:intermediation} \mbox{Implementation error } \mathbf{e}_i$

- Human error (bugs): manual/automatic code validation.
- Finite arithmetic errors (single / double / quadruple precision)

$\label{eq:intermediation} \mbox{Implementation error} \ e_i$

- Human error (bugs): manual/automatic code validation.
- Finite arithmetic errors (single / double / quadruple precision):
 - Between Jan. 1982 and Nov. 1983, Vancouver stock exchange dropped from 1,000 to 524 instead of going up to 1,098 due to truncation errors.
 - Patriot missile failure (Feb. 1991) due to round-off errors: storage of 0.1 on 24 bits only in the internal clock (500 m drift after 100 h).
 - Ariane 5 rocket crash (Jun. 1996) due to overflow errors during the conversion of 64-bit floating point numbers into 16-bit signed integers.

$\label{eq:intermediation} \mbox{Implementation error} \ e_i$

- Human error (bugs): manual/automatic code validation.
- Finite arithmetic errors (single / double / quadruple precision):
 - Between Jan. 1982 and Nov. 1983, Vancouver stock exchange dropped from 1,000 to 524 instead of going up to 1,098 due to truncation errors.
 - Patriot missile failure (Feb. 1991) due to round-off errors: storage of 0.1 on 24 bits only in the internal clock (500 m drift after 100 h).
 - Ariane 5 rocket crash (Jun. 1996) due to overflow errors during the conversion of 64-bit floating point numbers into 16-bit signed integers.

Computing error \mathbf{e}_{c}

Resiliency will be one of the toughest challenges in future exascale systems. Memory errors contribute more than 40% of the total hardware-related failures and are projected to increase in future exascale systems. The use of error correction codes (ECC) and checkpointing are two effective approaches to fault tolerance. (Li et al. 2011)

Periodic Gross-Pitaevskii model

- used to simulate Bose-Einstein condensates;
- simplified version of the Kohn-Sham model for condensed matter.

Periodic Gross-Pitaevskii model

- used to simulate Bose-Einstein condensates;
- simplified version of the Kohn-Sham model for condensed matter.

$$I = \inf \left\{ E(v), \ v \in H^1_{\#}(\Omega), \ \int_{\Omega} |v|^2 = 1 \right\}$$

where $\Omega = (0, 2\pi)^d$ (d = 1, 2 or 3) and where

$$E(v) = \int_{\Omega} |\nabla v|^2 + \int_{\Omega} V|v|^2 + \frac{\mu}{2} \int_{\Omega} |v|^4,$$
$$H^1_{\#}(\Omega) = \left\{ v \in L^2_{\text{loc}}(\mathbb{R}^d) \, | \, \nabla v \in (L^2_{\text{loc}}(\mathbb{R}^d))^d, \, v \, 2\pi \mathbb{Z}^d \text{-periodic} \right\},$$

V being a $2\pi\mathbb{Z}^d$ -periodic real-valued continuous function and $\mu > 0$.

Periodic Gross-Pitaevskii model

- used to simulate Bose-Einstein condensates;
- simplified version of the Kohn-Sham model for condensed matter.

$$I = \inf \left\{ E(v), \ v \in H^{1}_{\#}(\Omega), \ \int_{\Omega} |v|^{2} = 1 \right\}$$
(1)

where $\Omega = (0, 2\pi)^d$ (d = 1, 2 or 3) and where

$$E(v) = \int_{\Omega} |\nabla v|^2 + \int_{\Omega} V|v|^2 + \frac{\mu}{2} \int_{\Omega} |v|^4,$$
$$H^1_{\#}(\Omega) = \left\{ v \in L^2_{\text{loc}}(\mathbb{R}^d) \, | \, \nabla v \in (L^2_{\text{loc}}(\mathbb{R}^d))^d, \ v \ 2\pi \mathbb{Z}^d \text{-periodic} \right\},$$

V being a $2\pi\mathbb{Z}^d$ -periodic real-valued continuous function and $\mu > 0$.

- (1) has exactly two minimizers u (with u > 0 in Ω) and -u;
- $\exists ! \lambda \in \mathbb{R}$ such that (λ, u) satisfies the nonlinear Schrödinger equation

$$-\Delta u + Vu + \mu u^3 = \lambda u, \qquad \|u\|_{L^2_{\#}} = 1$$

• λ is the lowest eigenvalue of the self-adjoint operator $-\Delta + V + \mu u^2$.

Approximation space: for all $\mathcal{N} \in \mathbb{N}$, $X_{\mathcal{N}} = \left\{ v_{\mathcal{N}} = \sum_{|\mathbf{k}| \leq \mathcal{N}} \widehat{v}_{\mathbf{k}} e_{\mathbf{k}}, v_{\mathcal{N}} \text{ real valued} \right\}$ where $e_{\mathbf{k}}(\mathbf{r}) = \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{(2\pi)^{d/2}}, \quad \mathbf{k} \in \mathbb{Z}^{d}$.

Approximation space: for all $\mathcal{N} \in \mathbb{N}$,

$$X_{\mathcal{N}} = \left\{ v_{\mathcal{N}} = \sum_{\substack{|\mathbf{k}| \leq \mathcal{N} \\ |\mathbf{k}| \leq \mathcal{N}}} \widehat{v}_{\mathbf{k}} e_{\mathbf{k}}, \ v_{\mathcal{N}} \text{ real valued} \right\} \ \subset H^{s}_{\#}(\Omega), \quad \forall s \in \mathbb{R},$$

where $e_{\mathbf{k}}(\mathbf{r}) = \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{(2\pi)^{d/2}}, \quad \mathbf{k} \in \mathbb{Z}^{d}.$

Periodic Sobolev spaces: for all $s \in \mathbb{R}$,

$$H^s_{\#}(\Omega) := \left\{ v = \sum_{\mathbf{k} \in \mathbb{Z}^d} \widehat{v}_{\mathbf{k}} e_{\mathbf{k}}, \ v \text{ real valued}, \ \|v\|^2_{H^s_{\#}} := \sum_{\mathbf{k} \in \mathbb{Z}^d} \left(1 + |\mathbf{k}|^2\right)^s |\widehat{v}_{\mathbf{k}}|^2 < \infty \right\}.$$

)

Approximation space: for all $\mathcal{N} \in \mathbb{N}$,

$$X_{\mathcal{N}} = \left\{ v_{\mathcal{N}} = \sum_{\substack{|\mathbf{k}| \leq \mathcal{N} \\ \#}} \widehat{v}_{\mathbf{k}} e_{\mathbf{k}}, \ v_{\mathcal{N}} \text{ real valued} \right\} \ \subset H^{s}_{\#}(\Omega), \quad \forall s \in \mathbb{R},$$

where $e_{\mathbf{k}}(\mathbf{r}) = \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{(2\pi)^{d/2}}, \quad \mathbf{k} \in \mathbb{Z}^{d}.$

Periodic Sobolev spaces: for all $s \in \mathbb{R}$,

$$H^s_{\#}(\Omega) := \left\{ v = \sum_{\mathbf{k} \in \mathbb{Z}^d} \widehat{v}_{\mathbf{k}} e_{\mathbf{k}}, \ v \text{ real valued}, \ \|v\|^2_{H^s_{\#}} := \sum_{\mathbf{k} \in \mathbb{Z}^d} \left(1 + |\mathbf{k}|^2\right)^s |\widehat{v}_{\mathbf{k}}|^2 < \infty \right\}.$$

1

Truncation operator $\Pi_{\mathcal{N}}$: for all $\mathcal{N} \in \mathbb{N}$, $\Pi_{\mathcal{N}}$ is defined on $\mathcal{S}'_{\#}(\Omega)$ by

$$\Pi_{\mathcal{N}}\left(\sum_{\mathbf{k}\in\mathbb{Z}^d}\widehat{v}_{\mathbf{k}}e_{\mathbf{k}}\right) = \sum_{|\mathbf{k}|\leq\mathcal{N}}\widehat{v}_{\mathbf{k}}e_{\mathbf{k}}.$$

 $\Pi_{\mathcal{N}}$ is the orthogonal projector from $H^s_{\#}(\Omega)$ onto $X_{\mathcal{N}}$ for all $s \in \mathbb{R}$.

Galerkin approximation

Find a minimizer u_N to the finite-dimensional variational problem

$$I_{\mathcal{N}} = \inf \left\{ E(v_{\mathcal{N}}), \ v_{\mathcal{N}} \in X_{\mathcal{N}}, \ \int_{\Omega} |v_{\mathcal{N}}|^2 = 1 \right\} \qquad (u_{\mathcal{N}}, 1)_{L^2_{\#}} \ge 0.$$

$$u_{\mathcal{N}} = u_{\mathcal{N}}$$

$$X_{\mathcal{N}}$$

Galerkin approximation

Find a minimizer u_N to the finite-dimensional variational problem



If $V \in H^{\sigma}_{\#}(\Omega)$ for some $\sigma \geq 0$, then $u \in H^{\sigma+2}_{\#}(\Omega)$ and

$$\forall s < \sigma + 2, \quad \|u - \Pi_{\mathcal{N}} u\|_{H^{s}_{\#}} = \left(\sum_{|\mathbf{k}| > \mathcal{N}} \left(1 + |\mathbf{k}|^{2}\right)^{s} |\widehat{u}_{\mathbf{k}}|^{2}\right)^{1/2} \le \frac{\|u\|_{H^{\sigma+2}_{\#}}}{\mathcal{N}^{\sigma+2-s}}.$$

Theorem (EC, Chakir, Maday '10)

There exists $0 < c \le C < \infty$ such that for all $\mathcal{N} \in \mathbb{N}$,

$$\|u - \Pi_{\mathcal{N}} u\|_{H^{1}_{\#}} \leq \|u - u_{\mathcal{N}}\|_{H^{1}_{\#}} \leq C \|u - \Pi_{\mathcal{N}} u\|_{H^{1}_{\#}} \xrightarrow{} 0$$
$$c\|u - u_{\mathcal{N}}\|_{H^{1}_{\#}}^{2} \leq I_{N} - I = E(u_{\mathcal{N}}) - E(u) \leq C \|u - u_{\mathcal{N}}\|_{H^{1}_{\#}}^{2}.$$

Theorem (EC, Chakir, Maday '10)

There exists $0 < c \le C < \infty$ such that for all $\mathcal{N} \in \mathbb{N}$,

$$\|u - \Pi_{\mathcal{N}} u\|_{H^{1}_{\#}} \leq \|u - u_{\mathcal{N}}\|_{H^{1}_{\#}} \leq C \|u - \Pi_{\mathcal{N}} u\|_{H^{1}_{\#}} \xrightarrow{} 0$$
$$c\|u - u_{\mathcal{N}}\|_{H^{1}_{\#}}^{2} \leq I_{N} - I = E(u_{\mathcal{N}}) - E(u) \leq C \|u - u_{\mathcal{N}}\|_{H^{1}_{\#}}^{2}.$$

Assume that $V \in H^{\sigma}_{\#}(\Omega)$ for some $\sigma > d/2$. Then

- $(u_{\mathcal{N}})_{\mathcal{N}\in\mathbb{N}}$ converges to u in $H^{\sigma+2}_{\#}(\Omega)$;
- there exists positive constants C and C_s such that

$$\forall -\sigma \leq s < \sigma + 2, \quad \|u - u_{\mathcal{N}}\|_{H^s_{\#}} \leq \frac{C_s}{\mathcal{N}^{\sigma+2-s}}, \quad |\lambda - \lambda_{\mathcal{N}}| \leq \frac{C}{\mathcal{N}^{2(\sigma+1)}}.$$

Remark: these *a priori* **convergence rates are optimal.**

Theorem (EC, Chakir, Maday '10)

There exists $0 < c \le C < \infty$ such that for all $\mathcal{N} \in \mathbb{N}$,

$$\|u - \Pi_{\mathcal{N}} u\|_{H^{1}_{\#}} \leq \|u - u_{\mathcal{N}}\|_{H^{1}_{\#}} \leq C \|u - \Pi_{\mathcal{N}} u\|_{H^{1}_{\#}} \xrightarrow{} 0$$
$$c\|u - u_{\mathcal{N}}\|_{H^{1}_{\#}}^{2} \leq I_{N} - I = E(u_{\mathcal{N}}) - E(u) \leq C \|u - u_{\mathcal{N}}\|_{H^{1}_{\#}}^{2}.$$

Assume that $V \in H^{\sigma}_{\#}(\Omega)$ for some $\sigma > d/2$. Then

- $(u_{\mathcal{N}})_{\mathcal{N}\in\mathbb{N}}$ converges to u in $H^{\sigma+2}_{\#}(\Omega)$;
- there exists positive constants C and C_s such that

$$\forall -\sigma \leq s < \sigma + 2, \quad \|u - u_{\mathcal{N}}\|_{H^s_{\#}} \leq \frac{C_s}{\mathcal{N}^{\sigma+2-s}}, \quad |\lambda - \lambda_{\mathcal{N}}| \leq \frac{C}{\mathcal{N}^{2(\sigma+1)}}.$$

Remark: these *a priori* **convergence rates are optimal. They are useful**

• to estimate the convergence rate of various quantities of interest;

Ex
$$(d = 3, \sigma = 5/2)$$
: $0 \le E(u_N) - E(u) \le \frac{C}{N^7}, \qquad |u(0) - u_N(0)| \le \frac{C}{N^3}.$

Theorem (EC, Chakir, Maday '10)

There exists $0 < c \le C < \infty$ such that for all $\mathcal{N} \in \mathbb{N}$,

$$\|u - \Pi_{\mathcal{N}} u\|_{H^{1}_{\#}} \leq \|u - u_{\mathcal{N}}\|_{H^{1}_{\#}} \leq C \|u - \Pi_{\mathcal{N}} u\|_{H^{1}_{\#}} \xrightarrow{} 0$$
$$c\|u - u_{\mathcal{N}}\|_{H^{1}_{\#}}^{2} \leq I_{N} - I = E(u_{\mathcal{N}}) - E(u) \leq C \|u - u_{\mathcal{N}}\|_{H^{1}_{\#}}^{2}.$$

Assume that $V \in H^{\sigma}_{\#}(\Omega)$ for some $\sigma > d/2$. Then

- $(u_{\mathcal{N}})_{\mathcal{N}\in\mathbb{N}}$ converges to u in $H^{\sigma+2}_{\#}(\Omega)$;
- there exists positive constants C and C_s such that

$$\forall -\sigma \leq s < \sigma + 2, \quad \|u - u_{\mathcal{N}}\|_{H^s_{\#}} \leq \frac{C_s}{\mathcal{N}^{\sigma + 2 - s}}, \quad |\lambda - \lambda_{\mathcal{N}}| \leq \frac{C}{\mathcal{N}^{2(\sigma + 1)}}.$$

Remark: these *a priori* **convergence rates are optimal. They are useful**

- to estimate the convergence rate of various quantities of interest;
- to design more efficient numerical methods.

CPU time reduction by post-treatment



Solving the pb in a fine approximation space X_N is costly (~ $\mathcal{KN}^d \ln \mathcal{N}$ op.).

CPU time reduction by post-treatment



Solving the pb in a fine approximation space X_N is costly (~ $\mathcal{KN}^d \ln \mathcal{N}$ op.). Alternative:

1. solve the pb on a coarser approximation space X_n , $n \ll \mathcal{N}$ (~ $\mathcal{K}n^d \ln n$ op.);

CPU time reduction by post-treatment



Solving the pb in a fine approximation space X_N is costly (~ $\mathcal{KN}^d \ln \mathcal{N}$ op.). Alternative:

- **1.** solve the pb on a coarser approximation space X_n , $n \ll \mathcal{N}$ (~ $\mathcal{K}n^d \ln n$ op.);
- **2.** construct from u_n an approx. $u_{n,\mathcal{N}}$ of u in $X_{\mathcal{N}}$ (~ $\kappa \mathcal{N}^d \ln \mathcal{N}$ op., $\kappa \ll \mathcal{K}$).

CPU time reduction by post-treatment: two grid methods (introduced by Xu and Zhou in 2000 for solving nonlinear elliptic pbs)

CPU time reduction by post-treatment: two grid methods (introduced by Xu and Zhou in 2000 for solving nonlinear elliptic pbs)

- **1.** solve the *nonlinear* eigenvalue problem in the coarse approx. space X_n ;
- 2. freeze the self-consistent potential $V + \mu u_n^2$ and/or the eigenvalue λ_n ;
- **3.** solve the *linear* eigenpb $(-\Delta + V + \mu u_n^2)\widetilde{u} = \widetilde{\lambda}\widetilde{u}$ in the fine approx. space X_N .

CPU time reduction by post-treatment: two grid methods (introduced by Xu and Zhou in 2000 for solving nonlinear elliptic pbs)

solve the *nonlinear* eigenvalue problem in the coarse approx. space X_n;
 freeze the self-consistent potential V + μu_n² and/or the eigenvalue λ_n;
 solve the *linear* eigenpb (-Δ + V + μu_n²)ũ = λũ in the fine approx. space X_N.

Theorem (EC, Chakir, He, Maday '14)

Assume that $V \in H^{\sigma}_{\#}(\Omega)$ for some $\sigma > d/2$. There exists $C \in \mathbb{R}_+$ s.t.

$$\|u_{n,\mathcal{N}} - u\|_{H^{1}_{\#}} \leq C\left(n^{-\sigma-3} + \mathcal{N}^{-\sigma-1}\right),$$

$$0 \leq E(u_{n,\mathcal{N}}) - E(u) \leq C\left(n^{-\sigma-3} + \mathcal{N}^{-\sigma-1}\right)^{2}.$$

 \longrightarrow For $n \sim N^{\frac{\sigma+1}{\sigma+3}}$, significant gain in CPU time with no loss of accuracy.

Self-consistent algorithm (simple fixed point - not optimal!)

$$\begin{aligned} & (-\Delta v_k + V v_k + \mu v_{k-1}^2 v_k = \lambda_k v_k, \quad v_k \in H^1_{\#}(\Omega), \quad \|v_k\|_{L^2} = 1, \\ & \lambda_k \text{ lowest eigenvalue of } -\Delta + V + \mu v_{k-1}^2. \end{aligned}$$

Self-consistent algorithm (simple fixed point - not optimal!)

$$\begin{pmatrix} -\Delta v_k + V v_k + \mu v_{k-1}^2 v_k = \lambda_k v_k, & v_k \in H^1_{\#}(\Omega), & \|v_k\|_{L^2} = 1, \\ \lambda_k \text{ lowest eigenvalue of } -\Delta + V + \mu v_{k-1}^2. \end{cases}$$

 \rightarrow convergence (μ small) or oscillation between two states (μ large).

Analysis in EC - Le Bris '00, Levitt '12 (for the Hartree-Fock model).

A posteriori error estimators and error balancing: theory

$$\begin{pmatrix} -\Delta \boldsymbol{v}_{\boldsymbol{k}} + \Pi_{n_{k}} \left(V + \mu \boldsymbol{v}_{\boldsymbol{k}-1}^{2} \right) \Pi_{n_{k}} \boldsymbol{v}_{\boldsymbol{k}} = \boldsymbol{\lambda}_{\boldsymbol{k}} \ \boldsymbol{v}_{\boldsymbol{k}}, \quad \boldsymbol{v}_{\boldsymbol{k}} \in X_{n_{k}}, \quad \|\boldsymbol{v}_{\boldsymbol{k}}\|_{L^{2}} = 1, \\ \lambda_{k} = \lambda_{v_{k-1}, n_{k}}, \quad \text{where } \lambda_{v, n} \text{ is the lowest eigenvalue of } -\Delta + \Pi_{n} \left(V + \mu \boldsymbol{v}^{2} \right) \Pi_{n}.$$

A posteriori error estimators and error balancing: theory

$$\left\{ \begin{array}{l} -\Delta \boldsymbol{v}_{\boldsymbol{k}} + \Pi_{n_{\boldsymbol{k}}} \left(V + \mu \boldsymbol{v}_{\boldsymbol{k}-1}^{2} \right) \Pi_{n_{\boldsymbol{k}}} \boldsymbol{v}_{\boldsymbol{k}} = \boldsymbol{\lambda}_{\boldsymbol{k}} \; \boldsymbol{v}_{\boldsymbol{k}}, \qquad \boldsymbol{v}_{\boldsymbol{k}} \in X_{n_{\boldsymbol{k}}}, \quad \|\boldsymbol{v}_{\boldsymbol{k}}\|_{L^{2}} = 1, \\ \lambda_{\boldsymbol{k}} = \lambda_{v_{\boldsymbol{k}-1},n_{\boldsymbol{k}}}, \quad \text{where } \lambda_{v,n} \text{ is the lowest eigenvalue of } -\Delta + \Pi_{n} \left(V + \mu \boldsymbol{v}^{2} \right) \Pi_{n}. \end{array} \right.$$

Error estimator

$$J(v_k) = E(v_k) - E(u) + \frac{1}{2} ||u - v_k||_{L^2}^2.$$

A posteriori error estimators and error balancing: theory

$$\left\{ \begin{array}{l} -\Delta \boldsymbol{v}_{\boldsymbol{k}} + \Pi_{n_{k}} \left(V + \mu \boldsymbol{v}_{\boldsymbol{k}-1}^{2} \right) \Pi_{n_{k}} \boldsymbol{v}_{\boldsymbol{k}} = \boldsymbol{\lambda}_{\boldsymbol{k}} \; \boldsymbol{v}_{\boldsymbol{k}}, \qquad \boldsymbol{v}_{\boldsymbol{k}} \in X_{n_{k}}, \quad \|\boldsymbol{v}_{\boldsymbol{k}}\|_{L^{2}} = 1, \\ \lambda_{k} = \lambda_{v_{k-1}, n_{k}}, \quad \text{where } \lambda_{v, n} \text{ is the lowest eigenvalue of } -\Delta + \Pi_{n} \left(V + \mu \boldsymbol{v}^{2} \right) \Pi_{n}. \end{array} \right\}$$

Error estimator

$$J(v_k) = E(v_k) - E(u) + \frac{1}{2} ||u - v_k||_{L^2}^2.$$

Theorem (EC, Dusson, Maday, Stamm, Vohralik '14).

$$0 \le J(v_k) \le \eta_{\mathrm{d},k} + \eta_{\mathrm{a},k},$$

$$\eta_{\mathrm{d},k} = \frac{1}{2} \left(\boldsymbol{\lambda}_{\boldsymbol{u}_{\boldsymbol{k}},\boldsymbol{n}_{\boldsymbol{k}}} - \boldsymbol{\lambda}_{\boldsymbol{v}_{\boldsymbol{k}},\boldsymbol{\infty}} \right) \ge 0, \quad \eta_{\mathrm{a},k} = \frac{1}{2} \left(\mu \int_{\Omega} (\boldsymbol{v}_{\boldsymbol{k}}^2 - \boldsymbol{v}_{\boldsymbol{k}-1}^2) \boldsymbol{v}_{\boldsymbol{k}}^2 + \boldsymbol{\lambda}_{\boldsymbol{k}} - \boldsymbol{\lambda}_{\boldsymbol{v}_{\boldsymbol{k}},\boldsymbol{n}_{\boldsymbol{k}}} \right) \ge 0.$$

A posteriori error estimators and error balancing: theory

$$\left\{ \begin{array}{l} -\Delta \boldsymbol{v}_{\boldsymbol{k}} + \Pi_{n_{k}} \left(V + \mu \boldsymbol{v}_{\boldsymbol{k}-1}^{2} \right) \Pi_{n_{k}} \boldsymbol{v}_{\boldsymbol{k}} = \boldsymbol{\lambda}_{\boldsymbol{k}} \; \boldsymbol{v}_{\boldsymbol{k}}, \qquad \boldsymbol{v}_{\boldsymbol{k}} \in X_{n_{k}}, \quad \|\boldsymbol{v}_{\boldsymbol{k}}\|_{L^{2}} = 1, \\ \lambda_{k} = \lambda_{v_{k-1}, n_{k}}, \quad \text{where } \lambda_{v, n} \text{ is the lowest eigenvalue of } -\Delta + \Pi_{n} \left(V + \mu \boldsymbol{v}^{2} \right) \Pi_{n}. \end{array} \right\}$$

Error estimator

$$J(v_k) = E(v_k) - E(u) + \frac{1}{2} ||u - v_k||_{L^2}^2.$$

Theorem (EC, Dusson, Maday, Stamm, Vohralik '14).

$$0 \le J(v_k) \le \eta_{\mathrm{d},k} + \eta_{\mathrm{a},k},$$

$$\eta_{\mathrm{d},k} = \frac{1}{2} \left(\boldsymbol{\lambda}_{\boldsymbol{u}_{\boldsymbol{k}},\boldsymbol{n}_{\boldsymbol{k}}} - \boldsymbol{\lambda}_{\boldsymbol{v}_{\boldsymbol{k}},\boldsymbol{\infty}} \right) \ge 0, \quad \eta_{\mathrm{a},k} = \frac{1}{2} \left(\mu \int_{\Omega} (\boldsymbol{v}_{\boldsymbol{k}}^2 - \boldsymbol{v}_{\boldsymbol{k}-1}^2) \boldsymbol{v}_{\boldsymbol{k}}^2 + \boldsymbol{\lambda}_{\boldsymbol{k}} - \boldsymbol{\lambda}_{\boldsymbol{v}_{\boldsymbol{k}},\boldsymbol{n}_{\boldsymbol{k}}} \right) \ge 0.$$

 λ_{v_k,n_k} and $\lambda_{v_k,\infty}$ can be estimated accurately at a low computational cost by means of Rayleigh-Schrödinger perturbation method.

A posteriori error estimators and error balancing: in practice,

• the inequalities

$$E(v_k) - E(u) \le J(v_k) \le \eta_{\mathrm{d},k} + \eta_{\mathrm{a},k}$$

are "almost" equalities

(observed in numerical simulation, justified by a priori error estimates);

• $\eta_{d,k}$ and $\eta_{a,k}$ provide relatively cheap and sharp estimators of e_d and e_a resp. at iteration k if the quantity of interest is the energy.

A posteriori error estimators and error balancing: in practice,

• the inequalities

$$E(v_k) - E(u) \le J(v_k) \le \eta_{\mathrm{d},k} + \eta_{\mathrm{a},k}$$

are "almost" equalities

(observed in numerical simulation, justified by *a priori* error estimates);

• $\eta_{d,k}$ and $\eta_{a,k}$ provide relatively cheap and sharp estimators of e_d and e_a resp. at iteration k if the quantity of interest is the energy.

Numerical strategy:

- refine if $\eta_{\mathrm{d},k} \gg \eta_{\mathrm{a},k}$,
- iterate otherwise,

until $\eta_{\mathrm{d},k} + \eta_{\mathrm{a},k} \leq \varepsilon$ (desired accuracy).

Extension to Kohn-Sham models:

- Optimal *a priori* error estimators for Kohn-Sham are available (EC, Chakir, Maday M2AN '12)
- Numerical quadrature errors can be estimated (EC, Chakir, Maday, J. Sci. Comput. '10)
- A posteriori error estimators and computational load reduction strategies are under development: Lin, Yang et al. (Berkeley), Schneider et al. (Berlin), Zhou et al. (Beijing), our (now delocalized) team (EC, Dusson, Maday, Stamm, Vohralík, J. Comp. Phys. '16)
- k-point sampling on the Brillouin zone for metals (EC, Ehrlacher, Gontier, Levitt, Lombardi, in prep.)
- Construction of optimized pseudopotentials (EC, Mourad CMS '16, Blanc, EC, Dupuy, in prep)
- Error estimators on energy differences (with G. Dusson and Y. Maday).

Model error e_m: can certified *a posteriori* error estimators be computed?



Model error e_m: can certified *a posteriori* error estimators be computed?



• Wavefunction methods: yes, at least in principle

Model error e_m: can certified *a posteriori* error estimators be computed?



- Wavefunction methods: yes, at least in principle
- Density functional theory: probably not

Model error e_m : can certified *a posteriori* error estimators be computed?



- Wavefunction methods: yes, at least in principle
- Density functional theory: probably not
- Quantum Monte Carlo methods: maybe

Educated choice of the exchange-correlation potential in DFT

$$\begin{cases} \mathbf{ground state density } \rho_0^{\{\mathbf{R}_k\}}(\mathbf{r}) = \rho_0(\mathbf{r}) = \sum_{i=1}^N |\phi_i(\mathbf{r})|^2 \\ -\frac{1}{2} \Delta \phi_i + V_{\rho_0}^{\mathrm{KS}} \phi_i = \lambda_i \phi_i, \qquad \lambda_1 < \lambda_2 \le \lambda_3 \le \cdots \\ \int_{\mathbb{R}^3} \phi_i \phi_j = \delta_{ij} \\ V_{\rho_0}^{\mathrm{KS}} = V_{\rho_0}^{\mathrm{H}} + v_{\rho_0}^{\mathrm{xc}} \qquad v_{\rho_0}^{\mathrm{xc}} : \text{ exchange-correlation potential} \\ -\Delta V_{\rho_0}^{\mathrm{H}} = 4\pi \left(\rho_0 - \sum_{k=1}^M z_k \delta_{\mathbf{R}_k} \right) \end{cases}$$

Educated choice of the exchange-correlation potential in DFT

$$\begin{cases} \mathbf{ground state density } \rho_0^{\{\mathbf{R}_k\}}(\mathbf{r}) = \rho_0(\mathbf{r}) = \sum_{i=1}^N |\phi_i(\mathbf{r})|^2 \\ -\frac{1}{2} \Delta \phi_i + V_{\rho_0}^{\mathbf{KS}} \phi_i = \lambda_i \phi_i, \qquad \lambda_1 < \lambda_2 \le \lambda_3 \le \cdots \\ \int_{\mathbb{R}^3} \phi_i \phi_j = \delta_{ij} & \text{Perspective on Density Functional Theory} \\ V_{\rho_0}^{\mathbf{KS}} = V_{\rho_0}^{\mathbf{H}} + v_{\rho_0}^{\mathbf{xc}} \\ -\Delta V_{\rho_0}^{\mathbf{H}} = 4\pi \left(\rho_0 - \sum_{k=1}^M z_k \delta_{\mathbf{R}_k}\right) & \text{Kleron Burks} \end{cases}$$

on

Educated choice of the exchange-correlation potential in DFT

$$\begin{aligned} \mathbf{ground state density } \rho_0^{\{\mathbf{R}_k\}}(\mathbf{r}) &= \rho_0(\mathbf{r}) = \sum_{i=1}^N |\phi_i(\mathbf{r})|^2 \\ -\frac{1}{2} \Delta \phi_i + V_{\rho_0}^{\mathrm{KS}} \phi_i &= \lambda_i \phi_i, \qquad \lambda_1 < \lambda_2 \leq \lambda_3 \leq \cdots \\ \int_{\mathbb{R}^3} \phi_i \phi_j &= \delta_{ij} \\ V_{\rho_0}^{\mathrm{KS}} &= V_{\rho_0}^{\mathrm{H}} + v_{\rho_0}^{\mathrm{xc}} \\ -\Delta V_{\rho_0}^{\mathrm{H}} &= 4\pi \left(\rho_0 - \sum_{k=1}^M z_k \delta_{\mathbf{R}_k} \right) \end{aligned}$$

citations of the B3LYP paper (Google Scholar, Aug. 2016): 66,338

Educated choice of the exchange-correlation potential in DFT

$$\begin{cases} \mathbf{ground state density } \rho_0^{\{\mathbf{R}_k\}}(\mathbf{r}) = \rho_0(\mathbf{r}) = \sum_{i=1}^N |\phi_i(\mathbf{r})|^2 \\ -\frac{1}{2} \Delta \phi_i + V_{\rho_0}^{\mathrm{KS}} \phi_i = \lambda_i \phi_i, \qquad \lambda_1 < \lambda_2 \le \lambda_3 \le \cdots \\ \int_{\mathbb{R}^3} \phi_i \phi_j = \delta_{ij} & \text{Perspective on Density Functional Theory} \\ V_{\rho_0}^{\mathrm{KS}} = V_{\rho_0}^{\mathrm{H}} + v_{\rho_0}^{\mathrm{xc}} \\ -\Delta V_{\rho_0}^{\mathrm{H}} = 4\pi \left(\rho_0 - \sum_{k=1}^M z_k \delta_{\mathbf{R}_k} \right) & \text{Kieron Burke} \end{cases}$$

Can machine learning help?

Machine learning in molecular simulation: model selection



Machine learning in molecular simulation: direct computation of QOI



DFT calculations are used to compute the QOI in training and validation sets (Ceder et al '10, Rupp, Tkatchenko, Müller, von Lilienfeld '12, Burke et al. '12, Csányi et al. '13, Mallat et al. '14, ..., IPAM program '16)

Homogeneous vs heterogeneous aperiodic systems









Homogenous aperiodic systems

- general theory of one-particle linear models developed in the '90s by Bellissard and co-workers
 - \rightarrow application to incommensurate bilayer systems (talk by P. Cazeaux)
- Kohn-Sham models for homogeneous aperiodic systems are not fully understood from a mathematical point of view (the problem originates from long-range Coulomb interactions)
 - → EC, Lahbabi, Lewin, J. Pure App. Math. '13







Very large or infinite inhomogenous system

• focused models: QM/PCM, QM/MM, QM/MM/PCM, QCM, ...

QM/PCM (quantum mechanics/polarizable continuum models)



Born ('20), Kirkwood ('34), Onsager ('36)

Rivail et al. ('76), Tomasi et al. ('81), Klamt et al. ('93), Cramer and Truhlar ('94), ...

see also the talk by S. Gusarov on reference interaction site models (RISM)

Very large or infinite inhomogenous system

• focused models: QM/PCM, QM/MM, QM/MM/PCM, QCM, ...

QM/MM (quantum mechanics/molecular mechanics)



Warshel and Levitt ('76), Singh and Kollman ('86), Field, Bash and Karplus ('90)

Morokuma et al. (ONIOM method, '96), ...

Note that the ONIOM method also allows one to carry out QM/QM (high-accuracy/low-accuracy) quantum mechanical calculations, e.g. MP4/HF.

Very large or infinite inhomogenous system

• focused models: QM/PCM, QM/MM, QM/MM/PCM, QCM, ...

QM/MM/PCM



QM/MM/PCM simulation (Mennucci et al.)

Dynamical simulation of very large systems made possible by recent advances in numerical methods (EC, Lagardère, Lipparini, Maday, Mennucci, Piquemal, Scalmani, Stamm, ... '13-'16)

Very large or infinite inhomogenous system

• focused models: QM/PCM, QM/MM, QM/MM/PCM, QCM, ...

QCM (quasi-continuum model, Ortiz, Phillips and Tadmor '96)



Simulation of an AFM tip (Whalen, Shi, Tadmor)

Further developments and applications: Bernstein, Bhattacharya, Carter, Csanyi, Gavini, Kaxiras, Kermode, Knapp, Miller, Molinari, Perez, Rodney, Voter, ... Mathematical analysis by Blanc, Dobson, E, Le Bris, Legoll, Lu, Luskin, Ortner, Plechac, Shapeev, Suli, ...

- focused models: QM/PCM, QM/MM, QM/MM/PCM, QCM, ...
- (Schur complement-like) self-energy / Green's function methods

Schur complement method: if the square matrix D is invertible,

$$\underbrace{\begin{pmatrix} A & B \\ C & D \end{pmatrix}}_{M} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} \qquad \Leftrightarrow \qquad \begin{cases} \left(A - BD^{-1}C\right)u = f - BD^{-1}g \\ v = D^{-1}\left(g - Cu\right) \end{cases}$$

Terminology coined by Haynsworth in 1968.

Decomposition used by Schur in his determinant formula ('17)

$$\det(M) = \det(D) \det(A - BD^{-1}C)$$

and already used by Laplace (1812) and Sylvester (1851) in earlier works.

- focused models: QM/PCM, QM/MM, QM/MM/PCM, QCM, ...
- (Schur complement-like) self-energy / Green's function methods

Self-energy formalism: for all $E \notin \sigma(H_{\rm e})$,

$$\begin{pmatrix} H_{\rm s} & H_{\rm es}^{\dagger} \\ H_{\rm es} & H_{\rm e} \end{pmatrix} \begin{pmatrix} \psi_{\rm s} \\ \psi_{\rm e} \end{pmatrix} = E \begin{pmatrix} \psi_{\rm s} \\ \psi_{\rm e} \end{pmatrix} \qquad \Leftrightarrow \qquad \begin{cases} (H_{\rm s} + \Sigma(E)) \,\psi_{\rm s} = E \psi_{\rm s} \\ \psi_{\rm e} = (E - H_{\rm e})^{-1} H_{\rm es} \psi_{\rm s} \end{cases}$$

where $\Sigma(E) = H_{es}^{\dagger}(E - H_{e})^{-1}H_{es}$ is called the self-energy operator.

Somehow related to Feshbach-Schur perturbation method.

- focused models: QM/PCM, QM/MM, QM/MM/PCM, QCM, ...
- (Schur complement-like) self-energy / Green's function methods

Non-equilibrium Green's function (NEGF) for coherent electron transport



$$\begin{pmatrix} z - H_{\rm l} & -H_{\rm ls} & 0\\ -H_{\rm ls}^{\dagger} & z - H_{\rm s} & -H_{\rm rs}^{\dagger}\\ 0 & -H_{\rm rs} & z - H_{r} \end{pmatrix} \begin{pmatrix} G_{\rm l}(z) & G_{\rm ls}(z) & G_{\rm lr}(z)\\ G_{\rm sl}(z) & G_{\rm s}(z) & G_{\rm sr}(z)\\ G_{\rm rl}(z) & G_{\rm rs}(z) & G_{\rm r}(z) \end{pmatrix} = \begin{pmatrix} I & 0 & 0\\ 0 & I & 0\\ 0 & 0 & I \end{pmatrix}$$

- focused models: QM/PCM, QM/MM, QM/MM/PCM, QCM, ...
- (Schur complement-like) self-energy / Green's function methods

Non-equilibrium Green's function (NEGF) for coherent electron transport



2.3 nm

 $G_{\rm s}(z) = (z - (H_{\rm s} + \Sigma_{\rm l}(z) + \Sigma_{\rm r}(z)))^{-1}.$

with

 $\Sigma_{\rm l}(z) = H_{ls}^{\dagger}(z - H_{\rm l})^{-1}H_{\rm ls}$ and $\Sigma_{\rm r}(z) = H_{rs}^{\dagger}(z - H_{\rm r})^{-1}H_{\rm rs}.$

- focused models: QM/PCM, QM/MM, QM/MM/PCM, QCM, ...
- (Schur complement-like) self-energy / Green's function methods
- superimposed multiscale models:

 $\hat{\mathbf{x}}$



- concurrent coupling (Kaxiras et al. '99)
- equation-free method (Kevrekidis et al. '03)
- HMM: heterogeneous multiscale method (E and Engquist '03)
- coarse-grained models (talk by N. Zabaras)



Conclusion

Conclusion

How can mathematicians help?

- analysis of the models
 - rule out or fix models with bad properties,
 - establish rigorous connections between models (e.g. at different scales),
 - construct suitable approximations/discretizations with controlled accuracy,
- contribution to the design a new generation of codes
 - generating numerical results supplemented with error bars,
 - optimizing the available computational resources,
 - adapted to massively parallel and hybrid architectures.
- knowledge transfer
 - between pure mathematics and applied fields,
 - between different applied fields.