Toward sharp error analysis of extended Langrangian molecular dynamics for polarizable force field simulation

Michael Lindsey, Courant Institute
(Joint work with Dong An, Lin Lin)

Ki-Net Young Researchers Workshop
UMD, October 2019
Background

- In molecular dynamics (MD), interested in evolution of state vector $\mathbf{r} = (r^{(1)}, \ldots, r^{(N)}) \in \mathbb{R}^d$ of atomic positions $r^{(i)} \in \mathbb{R}^3$ (so $d = 3N$).
- Often a latent variable $\mathbf{x} = \Phi(\mathbf{r}) \in \mathbb{R}^{d'}$ needs to be computed as auxiliary step to perform the evolution, i.e., dynamics given by

$$\ddot{\mathbf{r}}(t) = G(\mathbf{r}(t), \mathbf{x}(t)), \text{ where } \mathbf{x}(t) = \Phi(\mathbf{r}(t)).$$

- $\Phi$ can be complicated, e.g., in *ab initio* MD simulations, one must solve a quantum many-body problem at each time step to compute the forces!
  - In practice, within DFT approximation, must solve nonlinear eigenvalue problem at each time step.
- Spoiler: we will study case $\mathbf{x}(t) = A(r)^{-1} \mathbf{b}(r)$, to be motivated later. Must solve linear system at each time step.
- Maintain general perspective for the time being.
More structure: dynamics of \((r, x)\) are Hamiltonian with respect to

\[
H(r, x, p, \phi) = \frac{1}{2}|p|^2 + U(r) + Q(r, x),
\]

where \(\phi\) is the momentum variable for \(x\), viewed as having zero mass. (Mass of all particles taken to be 1 for simplicity.)

What I really mean is that the auxiliary variable \(x\) has \(\varepsilon \to 0\) mass, so consider

\[
H_\varepsilon(r, x, p, \phi) = \frac{1}{2}|p|^2 + \frac{\varepsilon}{2}|\phi|^2 + U(r) + Q(r, x).
\]

This yields dynamics:

\[
\ddot{r}_\varepsilon(t) = -\frac{\partial U}{\partial r}(r_\varepsilon(t)) - \frac{\partial Q}{\partial r}(r_\varepsilon(t), x_\varepsilon(t))
\]

\[
\varepsilon^2 \ddot{x}_\varepsilon(t) = -\frac{\partial Q}{\partial x}(r_\varepsilon(t), x_\varepsilon(t)).
\]
Take limit:

\[ \ddot{r}(t) = -\frac{\partial U}{\partial r}(r(t)) - \frac{\partial Q}{\partial r}(r(t), x(t)) \]

\[ 0 = -\frac{\partial Q}{\partial x}(r(t), x(t)). \]

Map \( x = \Phi(r) \) implicitly defined by solving \( \frac{\partial Q}{\partial x}(r, x) = 0 \) for given \( r \).

Don’t think of this as a ‘derivation’ of our model...will define specific model shortly.

But it is a way to understand the structure, and the practical method that we consider we will reverse these steps.
Model

- Classical molecular dynamics with polarizable force field.
- Potential energy landscape $U(r)$ includes:
  - ‘nonbonded terms’ (e.g., permanent electrostatic and van der Waals interactions).
  - ‘bonded valence terms’ (bond-stretching, angle-bending, etc.).
- $x = (x^{(1)}, \cdots, x^{(N)}) \in \mathbb{R}^d$ is the vector of induced dipoles associated to each atom (so $d = d' = 3N$).
- Contribution of induced dipole given by

$$Q(r, x) = \frac{1}{2} x^\top A(r) x - b(r)^\top x,$$

where

- $b^{(i)}(r)$ indicates the electric field felt by the $i$-th atom due to the permanent electrostatics of the system in configuration $r$.
- $A(r) \succeq 0$ specifies the interaction of the induced dipoles with themselves and each other.
  - $A^{(i,j)}(r) = \delta_{ij} [\alpha^{(i)}]^{-1} I_3 + T^{(i,j)}(r)$, where $\alpha^{(i)}$ is the atomic polarizability of the $i$-th atom, $T^{(i,j)}(r) \in \mathbb{R}^{3 \times 3}$ is the induced dipole interaction matrix.
Model summary

▶ Plugging in $Q$ to our earlier formulas, the (‘massless’) dynamics are given by differential-algebraic system

$$
\ddot{r}(t) = G(r(t), x(t))
$$

$$
0 = b(r(t)) - A(r(t))x(t),
$$

where

$$
G(r, x) = - \frac{\partial U}{\partial r}(r) - \frac{1}{2} x^\top \frac{\partial A}{\partial r}(r)x - \frac{\partial b^\top}{\partial r}(r)x.
$$

▶ Computational bottleneck: apparently have to invert $A(r)$ at each time-step in numerical integration.

▶ Aside: our main analysis will work for quite general $G$ (and even $d' \neq d$), but...

▶ We do use Hamiltonian structure to guarantee some useful global-in-time *a priori* bounds.

▶ For numerical discretization, Hamiltonian structure exploited by symplectic integrators for long-time stability.
Extended Lagrangian scheme

- **Extended Lagrangian** approach solves the $\varepsilon > 0$ (‘massive’) dynamics from before:
  \[
  \ddot{r}_\varepsilon(t) = G(r_\varepsilon(t), x_\varepsilon(t))
  \]
  \[
  \varepsilon\ddot{x}_\varepsilon(t) = b(r_\varepsilon(t)) - A(r_\varepsilon(t))x_\varepsilon(t).
  \]

- In this setting, called the ‘inertial extended Lagrangian SCF-free’ or ‘iEL/0-SCF’ method [Albaugh, Niklasson, Head-Gordon 2017].
  - SCF = self-consistent field iteration.
  - 0-SCF means here that you don’t have solve linear systems at each step.

- Can consider extended Lagrangian molecular dynamics (XLMD) for more general $Q$ (e.g., XL Born-Oppenheimer MD (XL-BOMD) for *ab initio* simulations [Niklasson et al 2006; Niklasson 2008; Niklasson, Cawkwell 2012]).

- Auxiliary dynamics oscillatory on time scale $\sim \sqrt{\varepsilon}$, hence must take time steps at least this small.

- With careful choice of $\varepsilon$, iEL/0-SCF (and related methods) can outperform discretizations of the original dynamics in terms of efficiency and long-time stability while still maintaining the accuracy for the atomic trajectories.
Assumptions

- We consider a fixed time interval $[0, T]$ (independent of $\varepsilon$).
- Let $(r_\star, p_\star)$, where $p_\star := \dot{r}_\star$, be the solution of the massless equations, given fixed initial conditions $r_\star(0), p_\star(0)$.
- Let $(r_\varepsilon, p_\varepsilon, x_\varepsilon, \dot{x}_\varepsilon)$ be the solution of the $\varepsilon$-massive equation, given initial conditions $r_\varepsilon(0) = r_\star(0), p_\varepsilon(0) = p_\star(0)$, as well as $x_\varepsilon(0), \dot{x}_\varepsilon(0)$ arbitrary (for now) other than being bounded independent of $\varepsilon$.
- Technical assumptions:
  1. $A : \mathbb{R}^d \to \mathbb{S}_{++}^d$ is a $C^3$ map, and there exists $C > 0$ such that $A(r) \succeq C^{-1}$ for all $r \in \mathbb{R}^d$.
  2. $b : \mathbb{R}^d \to \mathbb{R}^d$ is a bounded $C^3$ map.
  3. $U$ is bounded below, and $F := -\frac{\partial U}{\partial r} : \mathbb{R}^d \to \mathbb{R}^d$ is a locally Lipschitz $C^0$ map.

Proposition (a priori bounds)

Under assumptions (1), (2), (3), there exists $C$ such that both $|x_\varepsilon| \leq C$ and $\left| \frac{d^k r_\star}{dt^k} \right|, \left| \frac{d^k r_\varepsilon}{dt^k} \right| \leq C$ for $k = 0, 1, 2$ and on the entire time interval $[0, T]$. 
Goals and results

- Want to study the convergence of \((r_\varepsilon, p_\varepsilon)\) to \((r_\star, p_\star)\) as \(\varepsilon \to 0\), as well as the convergence \(x_\varepsilon \to x_\star\), where \(x_\star(t) := A(r_\star(t))^{-1}b(r_\star(t))\).

- What about the initial condition for \(x_\varepsilon, \dot{x}_\varepsilon\) again? Three cases:
  - **Compatible**: choose \(x_\varepsilon(0) = x_\star(0) = A(r_\star(0))^{-1}b(r_\star(0))\), but can take \(\dot{x}_\varepsilon(0)\) arbitrary.
  - **Optimally compatible**: compatible but also take \(\dot{x}_\varepsilon(0) = \dot{x}_\star(0)\) (can get explicit formula).
  - **Incompatible**: otherwise.

- Mostly care about accuracy of \(r_\varepsilon\). Numerical experiments say:
  - **Compatible**: \(r_\varepsilon - r_\star = O(\varepsilon), \ x_\varepsilon - x_\star = O(\sqrt{\varepsilon})\).
  - **Optimally compatible**: \(r_\varepsilon - r_\star = O(\varepsilon), \ x_\varepsilon - x_\star = O(\varepsilon)\).
  - **Incompatible**: no convergence of anything.

Theorem

*For compatible initial condition, it holds that \(r_\varepsilon - r_\star, x_\varepsilon - x_\star = O(\sqrt{\varepsilon})\). In dimension \(d = 1\), all the sharp rates mentioned above hold.*

Remark

*Sharp analysis in \(d = 1\) should extend, but it’s at least a bit more difficult. (It is only the dimension of the auxiliary variable \(x\) that matters.)*
Numerical example

\[ U(r) = r_1^2 + r_2^2 = |r|^2, \quad A(r) = \begin{pmatrix} 2 + |r|^2 & |r|^2 \\ |r|^2 & 1 + |r|^2 \end{pmatrix} \]

\[ b(r) = (\sin(r_1 + r_2), \cos(r_1 - 2r_2))^\top \]

**Figure:** Example \((d = 2)\) of DAE (massless) and SPE (massive, \(\varepsilon = 0.001\)) dynamics with compatible initial condition.
Related work

- To our knowledge, there are no convergence results for such extended Lagrangian schemes, except in setting (linear-response regime) where dynamics are linear in both \( r \) and \( x \) and the system can be directly diagonalized [L. Lin, J. Lu, S. Shao 2014].

- An interesting related work [D. An, T. Head-Gordon, L. Lin, J. Lu 2019] studies the related stochastic dynamics:

\[
\ddot{r}_\varepsilon(t) = G(r_\varepsilon(t), x_\varepsilon(t)) \\
\varepsilon \ddot{x}(t) = b(r_\varepsilon(t)) - A(r_\varepsilon(t))x_\varepsilon(t) - \sqrt{\varepsilon \gamma} \dot{x}_\varepsilon + \sqrt{2\gamma/\beta \varepsilon^{1/4}} \dot{W}(t),
\]

where \( \beta \) is an inverse temperature, \( \gamma \) is a damping parameter, and \( W \) is a Brownian motion.

- Choosing \( \beta^{-1} = O(\varepsilon^{1/2}) \) yields \( O(\sqrt{\varepsilon}) \) error for both \( r_\varepsilon, x_\varepsilon \).

- These results are sharp in this setting, but the setting itself (even without noise) loses the \( O(\varepsilon) \) error.
Residual dynamics

- Recall the dynamics
  \[
  \ddot{r}_\varepsilon(t) = G(r_\varepsilon(t), x_\varepsilon(t))
  \]
  \[
  \varepsilon \ddot{x}_\varepsilon(t) = b(r_\varepsilon(t)) - A(r_\varepsilon(t))x_\varepsilon(t).
  \]

- Think of path \( r_\varepsilon \) as being fixed, then study the behavior of \( x_\varepsilon \).
  - Governed by linear inhomogeneous ODE with time-dependent coefficients.

- Want to think of \( x_\varepsilon \) as perturbation of \( A(r_\varepsilon(t))^{-1}b(r_\varepsilon(t)) \), so consider instead dynamics for the residual \( y_\varepsilon := x_\varepsilon - A(r_\varepsilon(t))^{-1}b(r_\varepsilon(t)) \):
  \[
  \varepsilon \ddot{y}_\varepsilon = -A(r_\varepsilon(t))^{-1}y_\varepsilon(t) + \varepsilon \psi_\varepsilon(t),
  \]
  where
  \[
  \psi_\varepsilon := -\frac{d^2}{dt^2} \left[ A(r_\varepsilon)^{-1}b(r_\varepsilon) \right].
  \]

- \textit{A priori} bounds guarantee that \(|\psi_\varepsilon| \leq C\) uniformly in \( \varepsilon, t \).

- Initial conditions for \( y_\varepsilon \) and \( \dot{y}_\varepsilon \) are given by
  \[
  y_\varepsilon(0) = 0, \quad \dot{y}_\varepsilon(0) = z_0 := \dot{x}_\varepsilon(0) - \dot{x}_\star(0).
  \]
Warm-up

- As a warm-up, consider the constant-coefficient, homogeneous version of the residual dynamics:

\[
\begin{aligned}
\epsilon \ddot{y}_\varepsilon(t) &= -Ay_\varepsilon(t) \\
y_\varepsilon(0) &= 0, \quad \dot{y}_\varepsilon(0) = z_0,
\end{aligned}
\]

where $A$ is constant.

- Then we have the explicit solution

\[
y_\varepsilon(t) = \sqrt{\epsilon} \sin \left( \sqrt{A} \frac{t}{\sqrt{\epsilon}} \right) A^{-1/2} z_0.
\]

- $|y_\varepsilon| \leq C\sqrt{\epsilon}$ for arbitrary $z_0$, and moreover $\dot{y}_\varepsilon$ is $O(1)$ but oscillatory on time scale $\sim \sqrt{\epsilon}$.

- If $z_0 = 0$, then in fact we have $y_\varepsilon \equiv 0$, which is of course much stronger than real case (will get $|y_\varepsilon| \leq C\varepsilon$).
Homogeneous system

- First study homogeneous equation

\[ \varepsilon \ddot{\tilde{y}}_\varepsilon = -A(r_\varepsilon)\tilde{y}_\varepsilon, \quad (1a) \]
\[ \tilde{y}_\varepsilon(s) = 0, \quad \dot{\tilde{y}}_\varepsilon(s) = \xi_0, \quad (1b) \]

where the starting time \( s < T \) and initial value \( \xi_0 \) is arbitrary.

- For future reference, we define the flow map for the homogeneous system (1) by

\[ \Phi^{s,t}_{\varepsilon}(0, \xi_0) = \left( \begin{array}{c} \tilde{y}_\varepsilon(t) \\ \dot{\tilde{y}}_\varepsilon(t) \end{array} \right) \]

for \( t \geq s \), where \( \tilde{y}_\varepsilon \) is the solution of (1).
Homogeneous system

- Define
  \[ K_\varepsilon(t) = A(\varepsilon(t))^{1/2}. \]

- Let
  \[ U_{\varepsilon,+}^s(t) = T e^{i\varepsilon^{-1/2} \int_s^t K_\varepsilon(t') dt'}, \]
  where \( T e^{[\cdot]} \) denotes is ‘the time-ordered exponential.’

- This just means that \( U_{\varepsilon,+}^s(t) \) is defined to be the solution of
  \[ \dot{U}_{\varepsilon,+}^s(t) = i\varepsilon^{-1/2} K_\varepsilon(t) U_{\varepsilon,+}^s(t), \quad U_{\varepsilon,+}^s(s) = I_d. \]

- Likewise, let
  \[ U_{\varepsilon,-}^s(t) = T e^{-i\varepsilon^{-1/2} \int_s^t K_\varepsilon(t') dt'}. \]

- By construction \( U_{\varepsilon,+}^s \) and \( U_{\varepsilon,-}^s \) are unitary matrices for all \( t \).
Homogeneous system

Lemma

(i) $\Phi_{s,t}^{\varepsilon}(0, \xi_0)$ can be written in the form

$$
\Phi_{s,t}^{\varepsilon}(0, \xi_0) = \left( U_{s,+}^{\varepsilon}(t) c_{s,+}^{\varepsilon}(t) + U_{s,-}^{\varepsilon}(t) c_{s,-}^{\varepsilon}(t) \right) \varepsilon^{-1/2} K_{s}^{\varepsilon}(t) \left[ U_{s,+}^{\varepsilon}(t) c_{s,+}^{\varepsilon}(t) - U_{s,-}^{\varepsilon}(t) c_{s,-}^{\varepsilon}(t) \right].
$$

(ii) $c_{s,+}^{\varepsilon}(t)$ and $c_{s,-}^{\varepsilon}(t)$ allow the estimate

$$
|c_{s,+}^{\varepsilon}(t)|, |c_{s,-}^{\varepsilon}(t)| \leq C \varepsilon^{1/2} |\xi_0|,
$$

where $C$ is independent of $\varepsilon$ and $\xi_0$.

(iii) $\Phi$ allows the estimate

$$
\Phi_{s,t}^{\varepsilon}(0, \xi_0) = \left( O(\varepsilon^{1/2} |\xi_0|) \right).
$$
Now we return to the residual system

\[ \ddot{y}_\varepsilon = -A(r_\varepsilon(t))^{-1} y_\varepsilon(t) + \varepsilon \psi_\varepsilon(t), \quad y_\varepsilon(0) = 0, \quad \dot{y}_\varepsilon(0) = z_0. \]

Introduce the variable \( z_\varepsilon := \dot{y}_\varepsilon \) to obtain:

\[
\begin{pmatrix}
\dot{y}_\varepsilon \\
\dot{z}_\varepsilon
\end{pmatrix} =
\begin{pmatrix}
z_\varepsilon \\
-\varepsilon^{-1} Ay_\varepsilon
\end{pmatrix} + \begin{pmatrix} 0 \\ \psi_\varepsilon(t) \end{pmatrix}.
\]

Then by Duhamel's principle,

\[
\begin{pmatrix}
y_\varepsilon(t) \\
z_\varepsilon(t)
\end{pmatrix} = \Phi_{\varepsilon}^{0,t}(0, z_0) + \int_0^t \Phi_{\varepsilon}^{s,t}(0, \psi_\varepsilon(s)) \, ds.
\]

Then it directly follows that

\[ |y_\varepsilon| \leq C \varepsilon^{1/2}, \quad |\dot{y}_\varepsilon| \leq C. \]
Completing the proof of the coarse estimate

- Recall dynamics:
  \[ \dot{r}_\varepsilon = p_\varepsilon, \quad \dot{p}_\varepsilon = G(r_\varepsilon, x_\varepsilon). \]

- We just proved \( y_\varepsilon = x_\varepsilon - A(r_\varepsilon)^{-1}b(r_\varepsilon) = O(\varepsilon^{1/2}) \), so by assumptions / a priori bounds

  \[ \dot{r}_\varepsilon = p_\varepsilon, \quad \dot{p}_\varepsilon = G(r_\varepsilon, A(r_\varepsilon)^{-1}b(r_\varepsilon)) + e_\varepsilon, \]

  where \( e_\varepsilon = O(\varepsilon^{1/2}) \).

- Hence \((r_\varepsilon, p_\varepsilon)\) almost satisfy an \( O(\varepsilon^{1/2}) \)-perturbation of the equations defining \((r_\star, p_\star)\).

- This is enough to imply that \( r_\varepsilon = r_\star + O(\varepsilon^{1/2}) \), \( p_\varepsilon = p_\star + O(\varepsilon^{1/2}) \).
  - Several ways to see this, e.g., go back to Picard iteration.
Sharp estimates in 1D

- We focus on the case \( d = 1 \).
- Actually we will have to bootstrap from our coarse estimate!
- Retain definitions from above.
- Since \( d = 1 \), we denote \( k_\varepsilon = K_\varepsilon \) to emphasize that this is a scalar quantity.
- Since \( d = 1 \), time-ordered exponential takes simple form

\[
U_{\varepsilon, \pm}(t) = e^{\pm i (\kappa_\varepsilon(t) - \kappa_\varepsilon(s))/\sqrt{\varepsilon}}, \quad \kappa_\varepsilon(t) = \int_0^t k_\varepsilon(s) \, ds.
\]

- By assumption \( k \geq C^{-1} \), so \( \kappa_\varepsilon \) is strictly increasing and inverse mapping \( \kappa_\varepsilon^{-1} \) is well-defined.
- *A priori* bounds imply \(|\dot{\kappa}|, |\ddot{\kappa}| \leq C\).
Sharp homogeneous oscillator asymptotics

Lemma

With the flow map $\Phi^{t,s}$ defined as before,

$$
\Phi^{s,t}(0, \xi_0) = \left( \begin{array}{c}
\varepsilon^{1/2} k_\varepsilon(t)^{-1/2} k_\varepsilon(s)^{-1/2} \sin \left( \frac{\kappa_\varepsilon(t) - \kappa_\varepsilon(s)}{\sqrt{\varepsilon}} \right) \xi_0 \\
\kappa_\varepsilon(t)^{1/2} k_\varepsilon(s)^{-1/2} \cos \left( \frac{\kappa_\varepsilon(t) - \kappa_\varepsilon(s)}{\sqrt{\varepsilon}} \right) \xi_0
\end{array} \right) + \left( \begin{array}{c}
O(\varepsilon) \\
O(\varepsilon^{1/2})
\end{array} \right)
$$

- Compare: proof of asymptotics of Hermite polynomials from Tao’s book on RMT.
- In earlier proof, obtained the system for $c_\pm$ (omit $\varepsilon$ from notation):

$$
\dot{c}_+ = -\frac{1}{2} U_+^{-1} K^{-1} \dot{K} U_+ c_+ + \frac{1}{2} U_+^{-1} K^{-1} \dot{K} U_- c_- \\
\dot{c}_- = \frac{1}{2} U_-^{-1} K^{-1} \dot{K} U_+ c_+ - \frac{1}{2} U_-^{-1} K^{-1} \dot{K} U_- c_-.
$$

- Since $d = 1$, we can now commute operators to obtain

$$
\dot{c}_+ = -\frac{\dot{k}}{2k} c_+ + \frac{\dot{k}}{2k} e^{-2i\kappa/\sqrt{\varepsilon}} c_-, \quad \dot{c}_- = -\frac{\dot{k}}{2k} c_- + \frac{\dot{k}}{2k} e^{2i\kappa/\sqrt{\varepsilon}} c_+.
$$
Sharp inhomogeneous asymptotics

Lemma

Let \( y_\varepsilon \) be the solution to the residual system \( \varepsilon \ddot{y}_\varepsilon = -Ay_\varepsilon + \psi_\varepsilon \), with \( y_\varepsilon(0) = 0, \dot{y}_\varepsilon(0) = z_0 \). Then for \( t \in [0, T] \),

\[
\begin{pmatrix}
  y_\varepsilon(t) \\
  \dot{y}_\varepsilon(t)
\end{pmatrix} = \begin{pmatrix}
  \varepsilon^{1/2} k(t)^{-1/2} k(0)^{-1/2} \sin \left( \frac{\kappa(t)}{\sqrt{\varepsilon}} \right) z_0 \\
  k(t)^{1/2} k(0)^{-1/2} \cos \left( \frac{\kappa(t)}{\sqrt{\varepsilon}} \right) z_0
\end{pmatrix} + \begin{pmatrix}
  O(\varepsilon) \\
  O(\varepsilon^{1/2})
\end{pmatrix}.
\]

- Reformulate inhomogeneous as the first-order system

\[
\begin{pmatrix}
  \dot{y}_\varepsilon(t) \\
  \dot{z}_\varepsilon(t)
\end{pmatrix} = \begin{pmatrix}
  z_\varepsilon(t) \\
  -\varepsilon^{-1} A(r_\varepsilon(t)) y_\varepsilon(t)
\end{pmatrix} + \begin{pmatrix}
  0 \\
  \psi_\varepsilon(t)
\end{pmatrix}.
\]

- Duhamel’s principle yields

\[
\begin{pmatrix}
  y_\varepsilon(t) \\
  z_\varepsilon(t)
\end{pmatrix} = \Phi_{\varepsilon,0}^t (0, z_0) + \int_0^t \Phi_{\varepsilon,s}^t (0, \psi_\varepsilon(s)) \, ds.
\]

- Want to show integral term is \( O(\varepsilon) \) (oscillatory integral argument).
Completion of proof of sharp estimate

- Recall the $\varepsilon$-massive dynamics

  \[
  \ddot{r}_\varepsilon = G(r_\varepsilon, x_\varepsilon) \\
  \varepsilon \ddot{x}_\varepsilon = b(r_\varepsilon) - A(r_\varepsilon)x_\varepsilon
  \]

  and the massless dynamics

  \[
  \ddot{r}_\star = G \left( r_\star, A(r_\star)^{-1}b(r_\star) \right).
  \]

- Already know that $|x_\varepsilon - A(r_\varepsilon)^{-1}b(r_\varepsilon)| \leq C\varepsilon^{1/2}$.

- Also know that $|r_\varepsilon - r_\star| \leq C\varepsilon^{1/2}$ as well by bootstrapping from the coarse estimate.

- Combining previous two lines, get $|x_\varepsilon - A(r_\star)^{-1}b(r_\star)| \leq C\varepsilon^{1/2}$.

- Hence can Taylor-expand the ODE and retain only linear terms, with $O(\varepsilon)$ error.

- Defining new variable $\theta_\varepsilon := r_\varepsilon - r_\star$, which measures the error in the $r$ variable, can obtain:

  \[
  \ddot{\theta}_\varepsilon = \Upsilon \theta_\varepsilon + \Gamma y_\varepsilon + O(\varepsilon), \quad \theta_\varepsilon(0) = 0, \quad \dot{\theta}_\varepsilon(0) = 0.
  \]
Completion of proof of sharp estimate

- We view the ODE for $\theta_\varepsilon$ as a forced modification of the homogeneous ODE

$$\ddot{\theta}_\varepsilon = \Upsilon \tilde{\theta}_\varepsilon.$$

- Let $\Theta^s(t) \in \mathbb{R}^{2 \times 2}$ be the flow map for this homogeneous ODE.
- Then by Duhamel’s principle and oscillatory integral argument (forcing is of magnitude $\sim \sqrt{\varepsilon}$ and oscillation $\sim \sqrt{\varepsilon}$ show forcing contributes only $O(\varepsilon)$.
- In optimally compatible case, forcing is of magnitude $O(\varepsilon)$ to begin with.
Conclusions and further directions

- We have proved first convergence results for iEL/0-SCF method.
- Can get sharp asymptotics in case $d = 1$. Qualitatively, the analysis seems to capture what is happening for $d \geq 1$.
- Nonetheless, how to extend the analysis?
  - There should probably be an extra $O(\varepsilon)$ non-oscillatory drift term appearing in the homogeneous oscillator dynamics, depending only on commutators that are equal to zero in $d = 1$.
- Can one extend to more general $Q(r, x)$? Schemes for *ab initio* MD?
- Any practical insights to be gained from the analysis?