Modèles réduits non-linéaires basés sur le transport optimal pour le calcul de structures électroniques

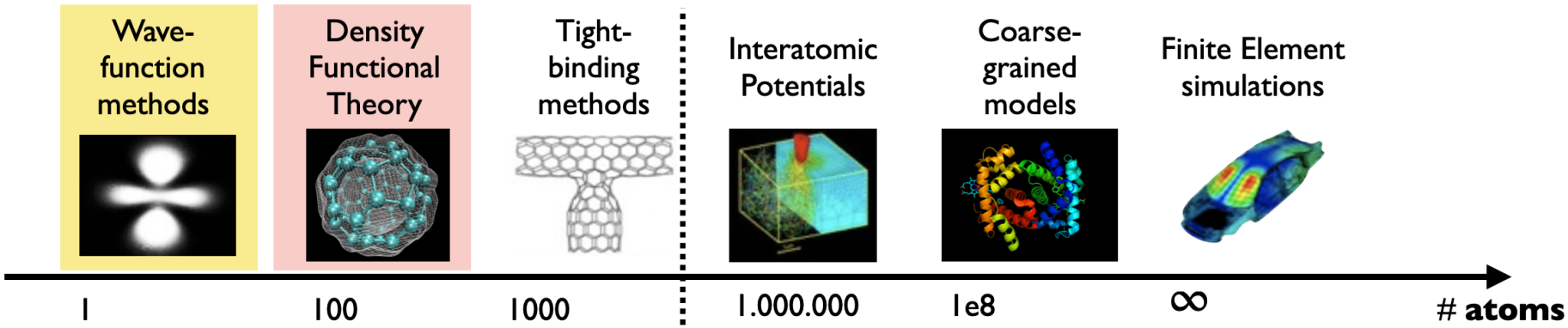
Geneviève Dusson

CNRS & Université Franche-Comté

joint work with Maxime Dalery (UFC), Virginie Ehrlacher (Ecole des Ponts), Gero Friesecke (TUM), Alexei Lozinski (UFC), Etienne Polack (Ecole des Ponts)

LJLL, 26 avril 2023

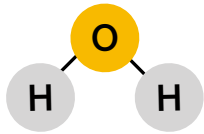
Context : Molecular simulations



Electronic structure calculations

- Modelling of the electrons
- A large number of models : Schrödinger, Hartree–Fock, Density Functional Theory
- Costly but high accuracy

Context : Electronic structure calculations

Water molecule : 

- $M = 3$ nuclei → quantum particles ~~quantum~~ **classical** particles described by (2 hydrogen and 1 oxygen) **positions** and **velocities**
- $N = 10$ electrons → quantum particles described by a **wavefunction**

Born–Oppenheimer approximation :
Proton-to-Electron mass ratio $\frac{m_p}{m_e} \approx 1838$.

Wavefunction $\Psi_{\mathbf{R}}(r_1, r_2, \ldots, r_N)$:

- $|\Psi_{\mathbf{R}}|^2$ probability density
- **Pauli principle** : $\Psi_{\mathbf{R}}(\ldots, r_i, \ldots, r_j, \ldots) = -\Psi_{\mathbf{R}}(\ldots, r_j, \ldots, r_i, \ldots)$

Focus in this presentation on the ground state problem

The ground state problem : Schrödinger equation

Born–Oppenheimer approximation : molecular system described by

- M (classical) nuclei with positions $\mathbf{R} \in \mathbb{R}^{3M}$
- N electrons described by a wave-function (or orbitals in DFT) $\Psi_{\mathbf{R}} : \mathbb{R}^{3N} \to \mathbb{C}$

Energy minimization

$$\inf_{\substack{\Psi \in L^2_{as}(\mathbb{R}^{3N}) \\ \|\Psi\|_{L^2}=1}} \langle \Psi_{\mathbf{R}}, H_{\mathbf{R}} \Psi_{\mathbf{R}} \rangle,$$

where $H_{\mathbf{R}}$ is the Hamiltonian of the problem, parametrized by the positions of the nuclei, typically

$$H_{\mathbf{R}} = -\frac{1}{2} \sum_{i=1}^{N} \Delta_{r_i} + \sum_{i=1}^{N} V_{\mathbf{R}}^{ne}(r_i) + \sum_{1 \le i < j \le N} \frac{1}{|r_i - r_j|}, \quad V_{\mathbf{R}}^{ne}(r_i) = \sum_{k=1}^{M} \frac{1}{|R_k - r_i|}$$

Eigenvalue problem

$$H_{\mathbf{R}} \Psi_{\mathbf{R}} = E_{\mathbf{R}} \Psi_{\mathbf{R}}.$$

Aim

A problem parametrized by the nuclei positions.
Goal : approximate

$$\mathcal{M} := \{\Psi_{\mathbf{R}} \text{ for } \mathbf{R} \in \mathcal{R}\}, \quad \mathcal{R} \text{ being the set of configurations}$$

Used for different purposes :

- Ab initio molecular dynamics
- Geometry optimization
- Building databases to construct interatomic potentials

Questions :

- Understand the structure of this manifold
- How to efficiently approximate the elements on this manifold ?
- More precisely, efficiently approximate all solutions (for varying positions) from the computation of only **a few** solutions

We will work on a toy problem but keep in mind that we want to deal with high-dimensional problems.

Toy problem

One-dimensional, one electron

Energy minimization :

$$\min_{\substack{\Psi \in H^1(\mathbb{R}) \\ \|\Psi\|_{L^2(\mathbb{R})}=1}} \frac{1}{2}\int_{\mathbb{R}} |\Psi'|^2 - \sum_{m=1}^{M} z_m \Psi(R_m)^2$$

Eigenvalue problem :

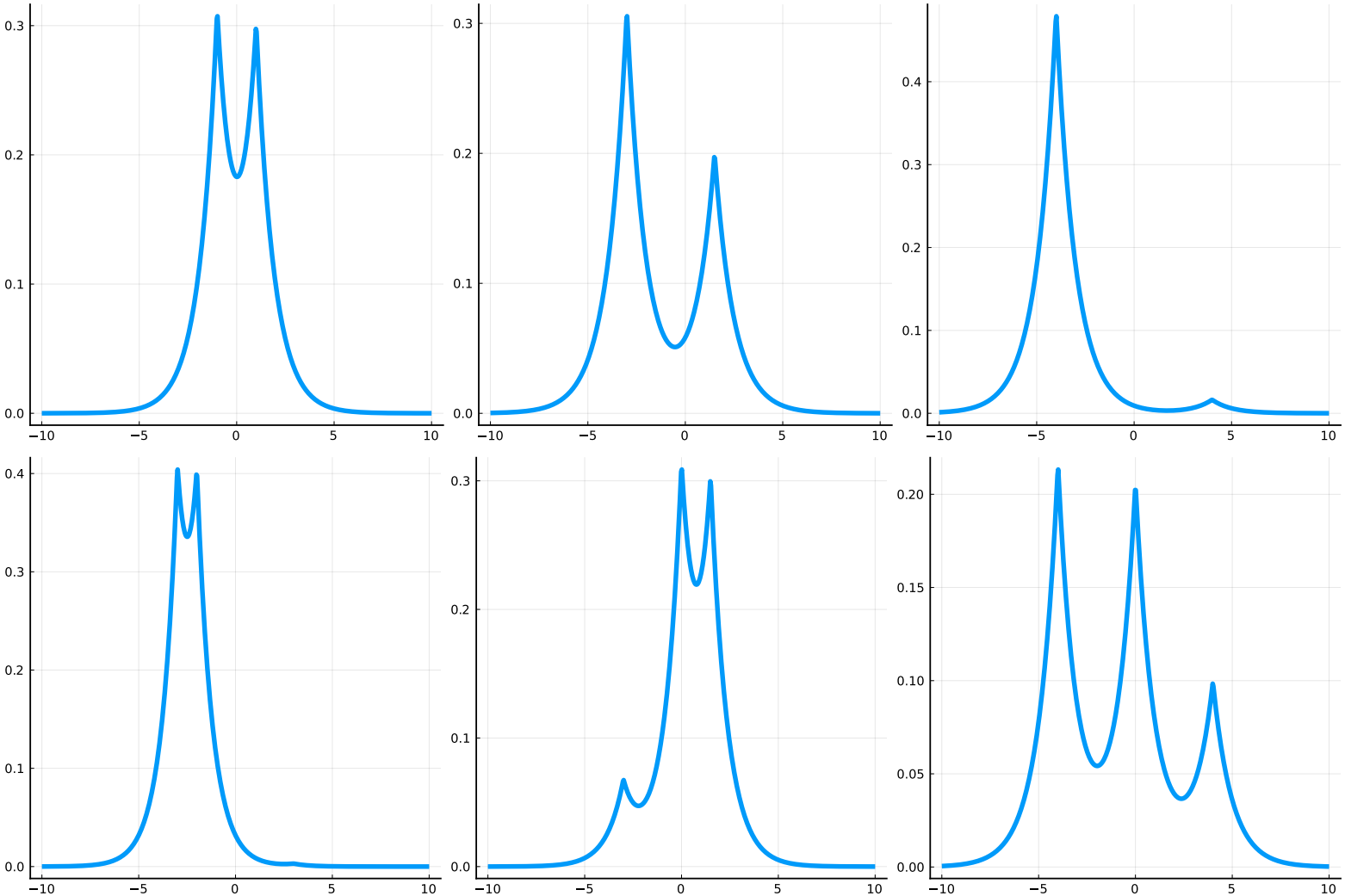
$$\begin{cases} -\frac{1}{2}\Psi_{\mathbf{R}}'' + \left(-\sum_{m=1}^{M} z_m \delta_{R_m}\right)\Psi_{\mathbf{R}} &= E_{\mathbf{R}}\Psi_{\mathbf{R}} \\ \|\Psi_{\mathbf{R}}\|_{L^2(\mathbb{R})} &= 1. \end{cases}$$

- Dirac potential
- Similar regularity as in the 3D case with Coulomb
- Analytic solutions

$$\Psi_{\mathbf{R}} = \sum_{m=1}^{M} \pi_m^{\mathbf{R}} e^{-\zeta_{\mathbf{R}}|x-R_m|},$$

for some positive weights $\boldsymbol{\pi}^{\mathbf{R}} = \left(\pi_m^{\mathbf{R}}\right)_{m=1}^{M} \in (\mathbf{R}_+)^M$ and $\zeta_{\mathbf{R}} > 0$.

Plots of a few solutions



Aim : efficiently approximate all solutions (for varying positions) from the computation of only **a few** solutions

Outline

Linear reduced-order modelling

Optimal transport : a few properties

Practical strategy and numerical results for a 1D toy problem

An example in 2D : Approximating the pair-density

Linear reduced basis method

Problem : Parametrized PDE with parameters $\mathbf{R} \in \mathcal{R}$'s.

Needs to be solved for **many** parameters $\mathbf{R}$.

Offline part : Select accurate solutions for a few **wisely chosen** parameters $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_K \in \mathcal{R}$

- Generate a training set of snapshots for parameters $\mathbf{R} \in \mathcal{R}_{\text{train}}$
- Select "good" snapshots with a greedy algorithm
  - Select one parameter $\mathbf{R}_1 \in \mathcal{R}$
  - at each iteration $K \geq 2$, select the snapshot that is worse approximated in the basis of the previously selected snapshots $\Psi_{\mathbf{R}_1}, \ldots, \Psi_{\mathbf{R}_{K-1}}$

Online part : compute solutions for **many** parameters $\mathbf{R} \in \mathcal{R}$ in the reduced space spanned by selected snapshots, i.e. in the basis spanned by the solutions $\Psi_{\mathbf{R}_1}, \ldots, \Psi_{\mathbf{R}_K}$.

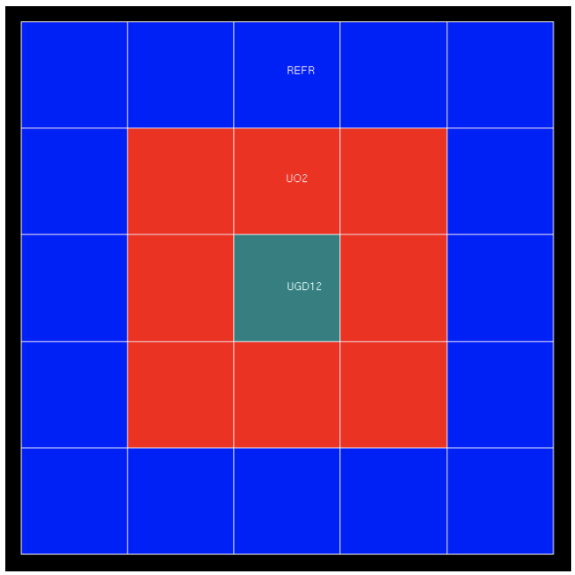
Barrault, Maday, Nguyen, Patera : An empirical interpolation method : application to efficient reduced-basis discretization of partial differential equations. C. R.(2004)

Hesthaven, Rozza, Stamm : Certified Reduced Basis Methods for Parametrized Partial Differential Equations. Springer (2016)

Many successful examples

Used for many industrial problems now

- Continuum mechanics
- Thermal equations
- ...



- Neutronics
- Non self-adjoint eigenvalue problem
- Parametrization in each cell of the nuclear core

Key point : The solutions should be well approximated by **linear combinations** of a few solutions.

Hesthaven, Rozza, Stamm : Certified Reduced Basis Methods for Parametrized Partial Differential Equations. Springer (2016)

Taumhas, D., Ehrlacher, Lelièvre, Madiot : Reduced basis method for non-symmetric eigenvalue problems : application to the multigroup neutron diffusion equations. arXiv :2307.05978.

An interesting notion : the Kolmogorov n-width

Definition for a Hilbert space $\mathbb{H}$: $\mathcal{M} := \{\Psi_{\mathbf{R}}, \quad \mathbf{R} \in \mathcal{R}\}$

$$\varepsilon_n(\mathcal{M}, \mathbb{H}) := \inf_{\substack{V_n \subset \mathbb{H} \\ \dim V_n = n}} \sup_{\mathbf{R} \in \mathcal{R}} \|\Psi_{\mathbf{R}} - \mathrm{P}_{V_n} \Psi_{\mathbf{R}}\|.$$

- Characterizes if the reduced basis method has a chance to work
- The faster the decay, the better !

Typical example where it works : elliptic equation

$$A_{\mathbf{R}} \Psi_{\mathbf{R}} = f,$$

with affine representation of $A_{\mathbf{R}}$:

$$A_{\mathbf{R}} = \sum_{q=1}^{Q} \theta_q(\mathbf{R}) A_q, \quad \text{for some } \theta_q \in \mathbb{R},\ A_q \text{ continuous operators}$$

Exponential decay of the Kolmogorov n-width :

$$\varepsilon_n(\mathcal{M}, \mathbb{H}) \leq C\ exp(-cn^{1/Q}).$$

Ohlberger, Rave : Reduced Basis Methods : Success, Limitations and Future Challenges (2015)

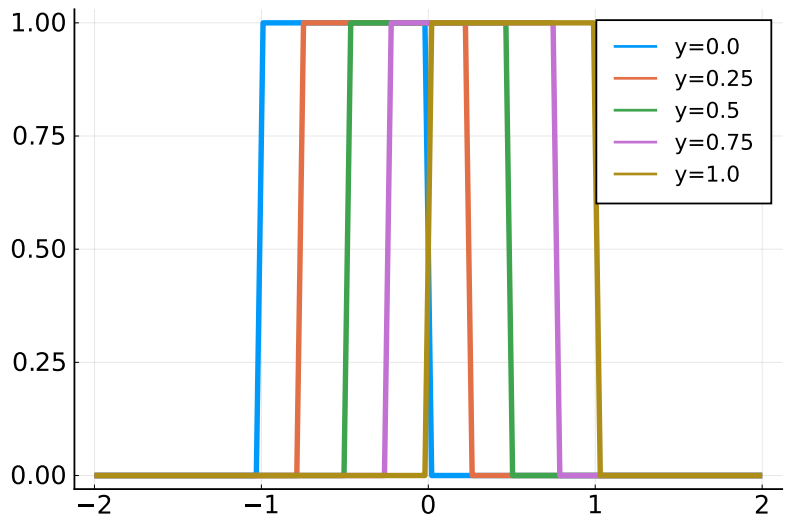
Cohen, DeVore : Approximation of high-dimensional parametric PDEs Acta Numerica (2015)

Less successful examples : Transport problems

Simple example : one-dimensional transport equation, $y \in [0,1]$,

$$\begin{cases} \partial_t \Psi_y(t,x) + y\partial_x \Psi_y(t,x) = 0, \quad x \in \mathbb{R}, t \in \mathbb{R}_+ \\ \Psi_y(0,x) = \mathbf{1}_{[-1,0]} \end{cases}$$

At $t = 1$, the solutions are $\Psi_y(t=1,x) = \mathbf{1}_{[y-1,y]}$



Kolmogorov n-width for $\mathcal{M} := \{\mathbf{1}_{[y-1,y]}, \quad y \in [0,1]\}$:

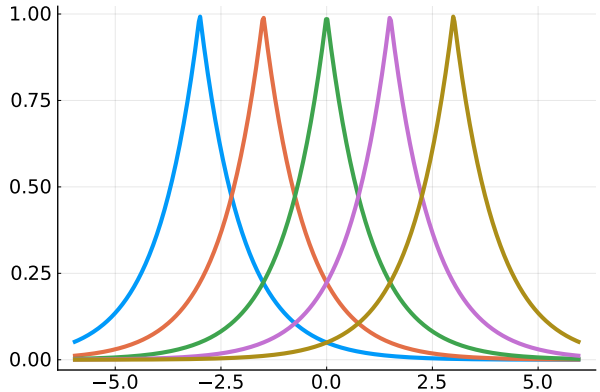
$$\varepsilon_n(\mathcal{M}, L^2(\Omega)) \geq cn^{-1/2}$$

Ohlberger, Rave : Reduced Basis Methods : Success, Limitations and Future Challenges (2015)

Cohen, DeVore : Approximation of high-dimensional parametric PDEs Acta Numerica (2015)

What about electronic structure ?

Similar behavior as the transport problem



Kolmogorov n-width for the 1D toy problem

For the problem with one nucleus, $\mathcal{M} = \{\Psi_R, R \in [-\bar{R}, \bar{R}]\}$, there exist positive constant $c_{\bar{R}}, C_{\bar{R}}$ such that

$$c_{\bar{R}} n^{-\frac{3}{2}} \leqslant \varepsilon_n(\mathcal{M}, L^2(\mathbb{R})) \leqslant C_{\bar{R}} n^{-\frac{3}{2}}.$$

For the problem with two nuclei, $\mathcal{M} = \{\Psi_{(R_1,R_2)}, R_1, R_2 \in [-\bar{R}, \bar{R}]^2\}$, there exists a positive constant $c_{\bar{R}}$ such that

$$c_{\bar{R}} n^{-\frac{3}{2}} \leqslant \varepsilon_n(\mathcal{M}, L^2(\mathbb{R})).$$

Dalery, Dusson, Ehrlacher, Lozinski : Nonlinear reduced basis using mixture Wasserstein barycenters : application to an eigenvalue problem inspired from quantum chemistry, arxiv :2307.15423

Alternative : Finding a good nonlinear transformation

Nonlinear space defined with n parameters

- Neural networks
  - Lee and Carlberg. Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders
- Nonlinear encoding-decoding map
  - Cohen, Farhat, Maday, Somacal : Nonlinear compressive reduced basis approximation for PDE's. Comptes Rendus Mécanique. 351, 1–18 (2023).
  - Lee and Carlberg. Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders
- Optimal transport-based transformations
  - Ehrlacher, Lombardi, Mula, Vialard : Nonlinear model reduction on metric spaces. Application to one-dimensional conservative PDEs in Wasserstein spaces. Esaim Math. Model (2020).
  - Iollo, Taddei : Mapping of coherent structures in parameterized flows by learning optimal transportation with Gaussian models. J. Comput. Phys. 471, 111671 (2022).

Key points :

- How many solutions to approximate accurately the solution for a new parameter ?
- Computational cost of a new solution

Replacing the Hilbert space by a metric space

Recall the definition for a Hilbert space $\mathbb{H}$: $\mathcal{M} := \{\Psi_{\mathbf{R}}, \quad \mathbf{R} \in \mathcal{R}\}$

$$\varepsilon_n(\mathcal{M}, \mathbb{H}) := \inf_{\substack{V_n \subset \mathbb{H} \\ \dim V_n = n}} \sup_{\mathbf{R} \in \mathcal{R}} \|\Psi_{\mathbf{R}} - \mathrm{P}_{V_n} \Psi_{\mathbf{R}}\|.$$

We need to

- replace the norm by a distance
- find an alternative for the projection
- linear combination replaced by **barycenter**

Given a metric space $\mathbb{M}$, with distance d, for convex parameters $\boldsymbol{t} = (t_1, \ldots, t_n)$, i.e. positive with $\sum_{i=1}^n t_i = 1$, and elements $\Psi_{\mathbf{R}_1}, \ldots, \Psi_{\mathbf{R}_n} \in \mathbb{M}$

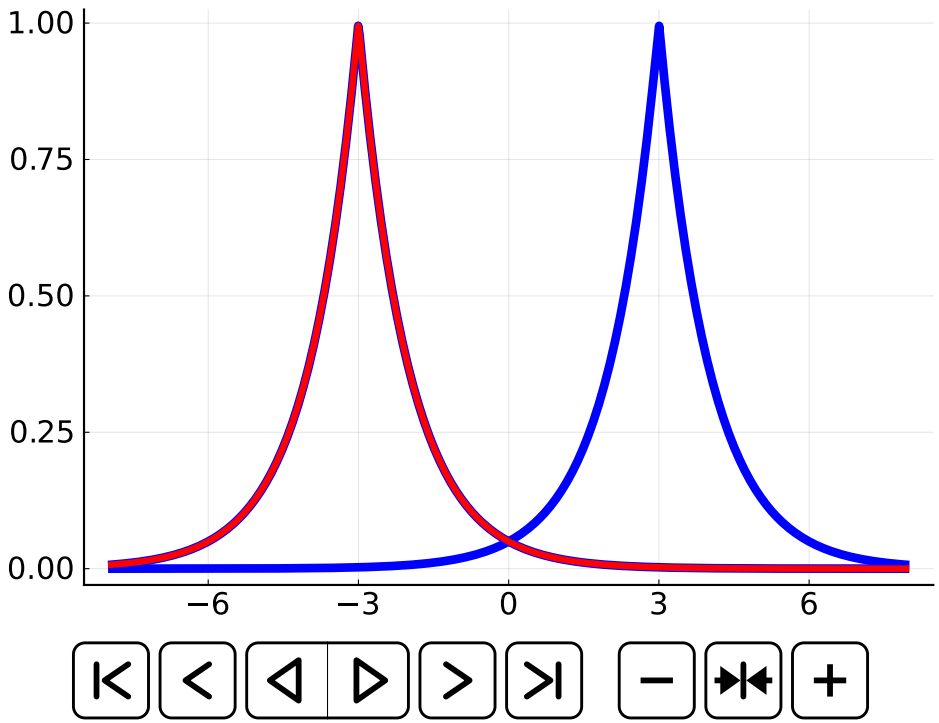
$$\mathrm{bar}(\boldsymbol{t}; \Psi_{\mathbf{R}_1}, \ldots, \Psi_{\mathbf{R}_n})) := \mathrm{argmin}_{u \in \mathbb{M}} \quad \sum_{i=1}^n t_i d(u, \Psi_{\mathbf{R}_i})^2.$$

Nonlinear Kolmogorov n-width

$$\varepsilon_n(\mathcal{M}, \mathbb{M}) := \inf_{\Psi_{\mathbf{R}_1}, \ldots, \Psi_{\mathbf{R}_n} \in \mathbb{M}} \sup_{\mathbf{R} \in \mathcal{R}} \inf_{\boldsymbol{t}} d\left(\Psi_{\mathbf{R}}, \mathrm{bar}(\boldsymbol{t}; \Psi_{\mathbf{R}_1}, \ldots, \Psi_{\mathbf{R}_n})\right))$$

Motivation behind using optimal transport

Use of **optimal transport** : at minima deals with the translations
Barycenter between two Slater functions : a translated Slater function

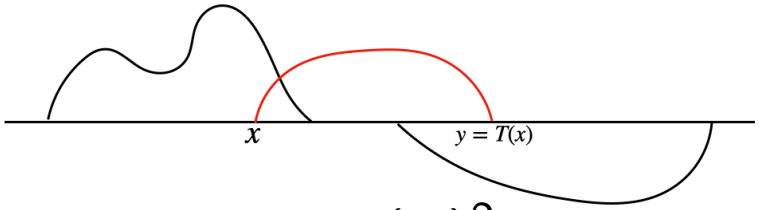


Simple one-nucleus problem : $\forall n > 1, \varepsilon_n(\mathcal{M}, (\mathcal{P}_2(\mathbb{R}), W_2)) = 0.$

Outline

Optimal transport in a nutshell

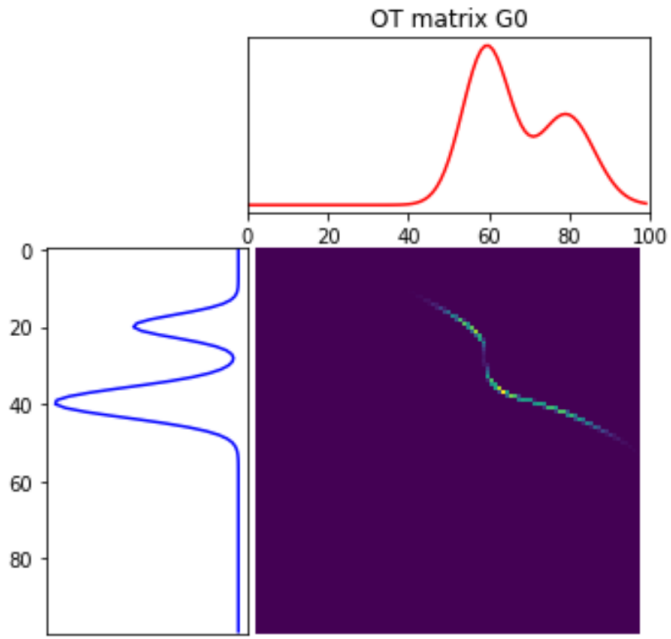
Originally introduced by Monge : moving a pile of sand efficiently to cover a sinkhole



Wasserstein distance : for $u, v \in \mathcal{P}_2(\Omega)^2$ as

$$W_2(u,v)^2 := \inf_{\pi \in \Pi(u,v)} \int_{\Omega^2} (x-y)^2 \, d\pi(x,y),$$

$\Pi(u,v)$: set of probability measures over Ω^2 with marginals u and v.



Delon, Desolneux : A Wasserstein-Type Distance in the Space of Gaussian Mixture Models. SIAM J. Imaging Sci. (2020).

Wasserstein barycenters

- n probability measures $u_1, \dots, u_n$
- n positive weights $t_1, \dots, t_n$ summing to 1

Barycenter is a solution to the problem

$$\inf_{u \in \mathcal{P}_2(\Omega)} \sum_{i=1}^{n} t_i W_2(u, u_i)^2.$$

Alternative formulation : multimarginal optimal transport problem

$$\inf_{\gamma \in \Pi(u_1, \dots, u_n)} \int_{\Omega^n} \frac{1}{2} \sum_{i,j=1}^{n} t_i t_j (x_i - x_j)^2 \, d\gamma(x_1, \dots, x_n),$$

where $\Pi(u_1, \dots, u_n)$ denotes the set of probability measures on Ω^n having $u_1, \dots, u_n$ as marginals.

$$\text{Then } \ \text{bar}(\boldsymbol{t}; u_1, \dots, u_n) = P_{\boldsymbol{t}} \# \gamma, \quad \text{ with } P_{\boldsymbol{t}}(x_1, \dots, x_n) = \sum_{i=1}^{n} t_i x_i.$$

Agueh, Carlier : Barycenters in the Wasserstein Space. SIAM J. Math. Anal. (2011).

Gangbo, Swiech : Optimal maps for the multidimensional Monge–Kantorovich problem. Commun. Pure Appl. Math. (1998)

A few examples : One-dimensional case

Cumulative distribution function (cdf) of an element $u \in \mathcal{P}_2(\mathbb{R})$ is

$$\mathrm{cdf}_u : x \in \mathbb{R} \longmapsto \int_{-\infty}^{x} \mathrm{d}[u],$$

Inverse cumulative distribution function (icdf) : generalized inverse of the cdf

$$\mathrm{icdf}_u : s \in [0,1] \longmapsto \mathrm{cdf}_u^{-1} := \inf\{x \in \mathbb{R}, \mathrm{cdf}_u(x) > s\}.$$

Then, for any $(u, v) \in \mathcal{P}_2(\mathbb{R})^2$, there holds

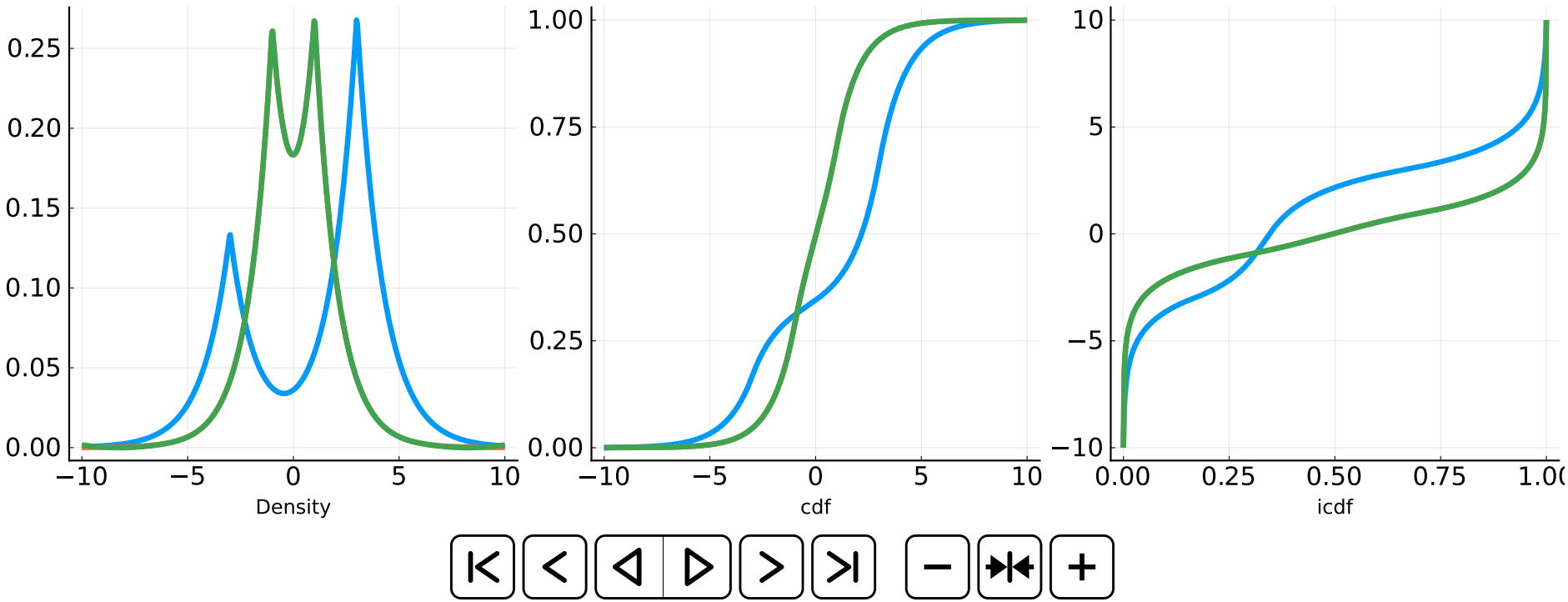
$$W_2(u, v) = \|\mathrm{icdf}_u - \mathrm{icdf}_v\|_{L^2([0,1])},$$

and for any set of barycentric weights $\boldsymbol{t} := (t_1, \ldots, t_n)$ and $\boldsymbol{u} := (u_1, \ldots, u_n)$,

the icdf of the barycenter $\mathrm{Bar}_{W_2}^{\boldsymbol{t}}(\boldsymbol{u})$ satisfies

$$\mathrm{icdf}_{\mathrm{Bar}_{W_2}^{\boldsymbol{t}}(\boldsymbol{u})} = \sum_{i=1}^{n} t_i \mathrm{icdf}_{u_i}.$$

Illustration



$$\mathrm{icdf}_{\mathrm{Bar}^{\boldsymbol{t}}_{W_2}(\boldsymbol{u})} = \lambda\, \mathrm{icdf}_{u_1} + (1-\lambda)\, \mathrm{icdf}_{u_2}.$$

A few examples : Location-scatter transforms

$$\mathcal{A} := \left\{ T\#a, \quad T : x \in \mathbb{R}^d \mapsto Ax + b,\ A \in \mathcal{S}_d,\ b \in \mathbb{R}^d \right\}.$$

All measures generated with translation and dilations of a single measure

Wasserstein distance : Let $a_0, a_1 \in \mathcal{A}$ havings means m_0, m_1 and covariance matrices Σ_0, Σ_1, such that the transport map

$$Tx = Ax + (m_0 - m_1), \quad \text{with } A = \Sigma_0^{-1/2} \left(\Sigma_0^{1/2} \Sigma_1 \Sigma_0^{1/2} \right)^{1/2} \Sigma_0^{-1/2},$$

is such that $T\#a_0 = a_1$. Then

$$W_2^2(a_0, a_1) = \|m_0 - m_1\|^2 + \mathrm{Tr}\left(\Sigma_0 + \Sigma_1 - 2(\Sigma_0^{1/2} \Sigma_1 \Sigma_0^{1/2})^{1/2} \right),$$

Wasserstein barycenter : Let $a_1, \ldots, a_n \in \mathcal{A}$ having mean m_j and covariance matrices Σ_j. For weights $\boldsymbol{t} := (t_1, \ldots, t_n)$, the barycenter is the atom $\mathrm{bar}a_{\boldsymbol{t}} = T_{\boldsymbol{t}}\#a$ with $T_{\boldsymbol{t}} = Sx + m$, where S is the only positive definite matrix satisfying

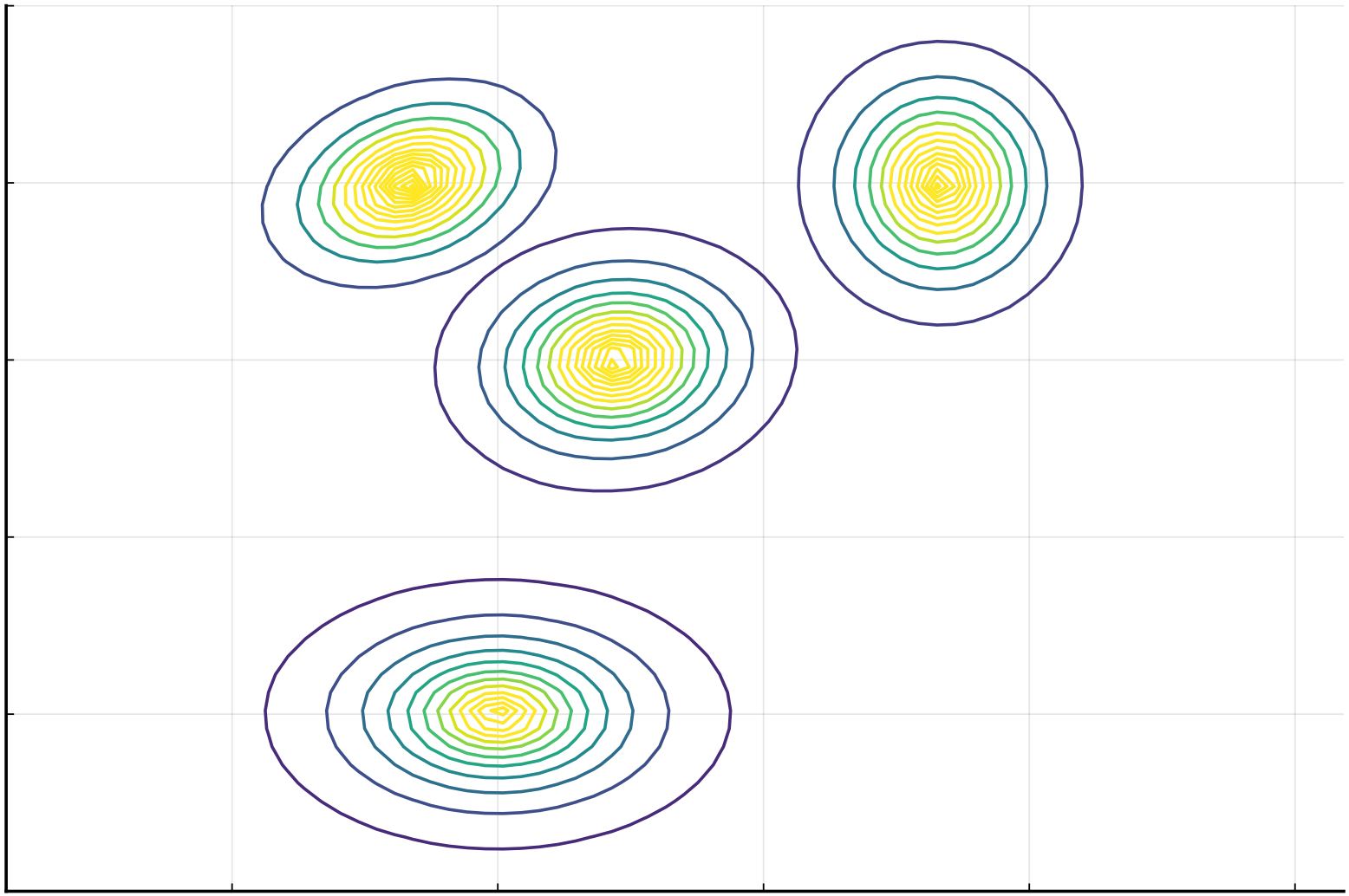
$$S = \sum_{j=1}^{n} t_j (S^{1/2} \Sigma_j S^{1/2})^{1/2}, \quad \text{and} \quad m = \sum_{j=1}^{n} t_j m_j.$$

Center is the mean of centers, small equation to solve for the covariance.

Alvarez-Esteban, del Barrio, Cuesta-Albertos, Matran : A fixed-point approach to barycenters in Wasserstein space. (2016).

Illustration

Barycenter between three Slater distributions



Kolmogorov n-width for the Wasserstein distance

Solution manifold : $\mathcal{M} = \{\Psi_{\mathbf{R}}, \mathbf{R} \in [-\bar{R}, \bar{R}]^M\}$

Since $W_2(u, v) = \|\mathrm{icdf}_u - \mathrm{icdf}_v\|_{L^2(0,1)}$, we consider

$$\varepsilon_n(\mathcal{M}, W_2) := \varepsilon_n(\mathrm{icdf}(\mathcal{M}), L^2(0,1))$$

Simple one-nucleus problem : $\forall n > 1, \varepsilon_n(\mathcal{M}, W_2) = 0$

Case of two nuclei in $[-\bar{R}, \bar{R}]^2$ **:** There exists a constant $C_{\bar{R}} > 0$ such that for all $n \geq 1$,

$$\varepsilon_n(\mathcal{M}, W_2) \leq C_{\bar{R}} n^{-5/2}.$$

To compare with the linear Kolmogorov width :

$$c_{\bar{R}} n^{-3/2} \leqslant \varepsilon_n(\mathcal{M}, L^2(\mathbb{R}))$$

Dalery, Dusson, Ehrlacher, Lozinski : Nonlinear reduced basis using mixture Wasserstein barycenters : application to an eigenvalue problem inspired from quantum chemistry, arxiv :2307.15423

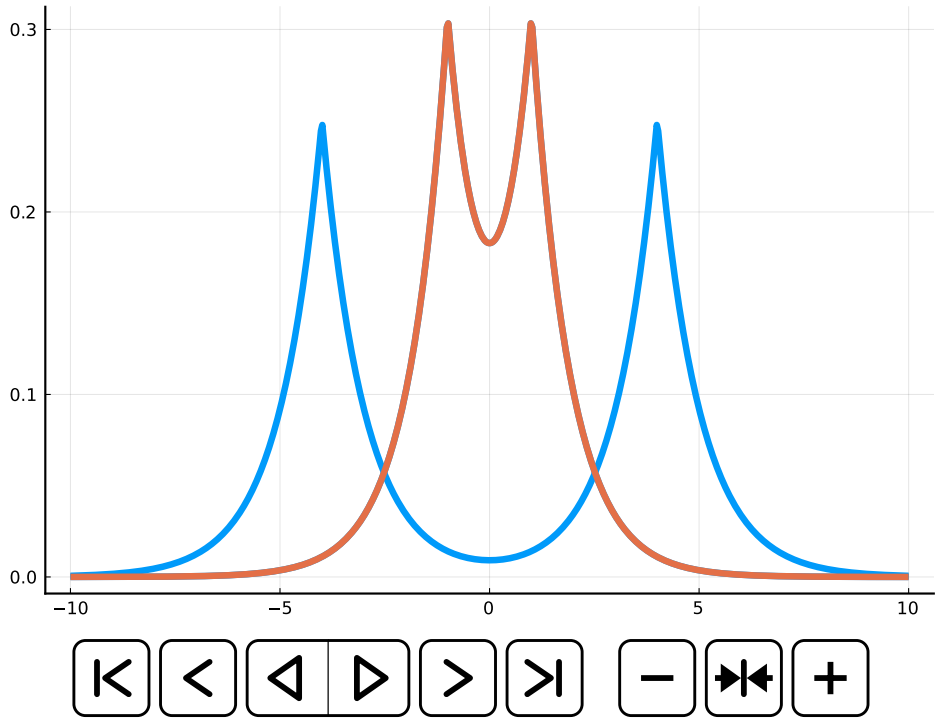
Wasserstein barycenter between two solutions

Constrained convex optimization problem

$$\min_{w:=(w_{jk})_{\substack{1\le j\le Ngrid,\\ 1\le k\le Ngrid}}\in\Pi(\Lambda_0,\Lambda_1)} \sum_{j=1}^{Ngrid}\sum_{k=1}^{Ngrid} w_{jk}\|x_j - x_k\|^2,$$

$$\text{with}\quad \Pi(\Lambda_0,\Lambda_1) := \left\{ w := (w_{jk})_{\substack{1\le j\le Ngrid,\\ 1\le k\le Ngrid}} \in \mathbb{R}_+^{Ngrid\times Ngrid},\quad \forall 1\le j\le Ngrid,\ \sum_{k=1}^{K} w_{jk} = \lambda_0^j,\right.$$

$$\left.\forall 1\le k\le Ngrid,\ \sum_{j=1}^{J} w_{jk} = \lambda_1^k \right\}.$$



Limitations :

- High computational cost
- Smoothing of the barycenter
- Bad scaling with the dimension
- Multi-marginal problem

A modified distance

$\mathcal{A} \subset \mathcal{P}(\mathbb{R})$: dictionary of atoms (Slater functions, gaussians, etc.)

Definition : for all mixtures $\mu_0 = \sum_{j=1}^{J} \lambda_0^j a_0^j \in \mathcal{M}(\mathcal{A})$ and $\mu_1 = \sum_{k=1}^{K} \lambda_1^k a_1^k \in \mathcal{M}(\mathcal{A})$, we define

$$mW_2(\mu_0, \mu_1)^2 := \min_{w:=(w_{jk})_{\substack{1\leq j\leq J,\\ 1\leq k\leq K}} \in \Pi(\Lambda_0,\Lambda_1)} \sum_{j=1}^{J}\sum_{k=1}^{K} w_{jk} W_2^2(a_0^j, a_1^k),$$

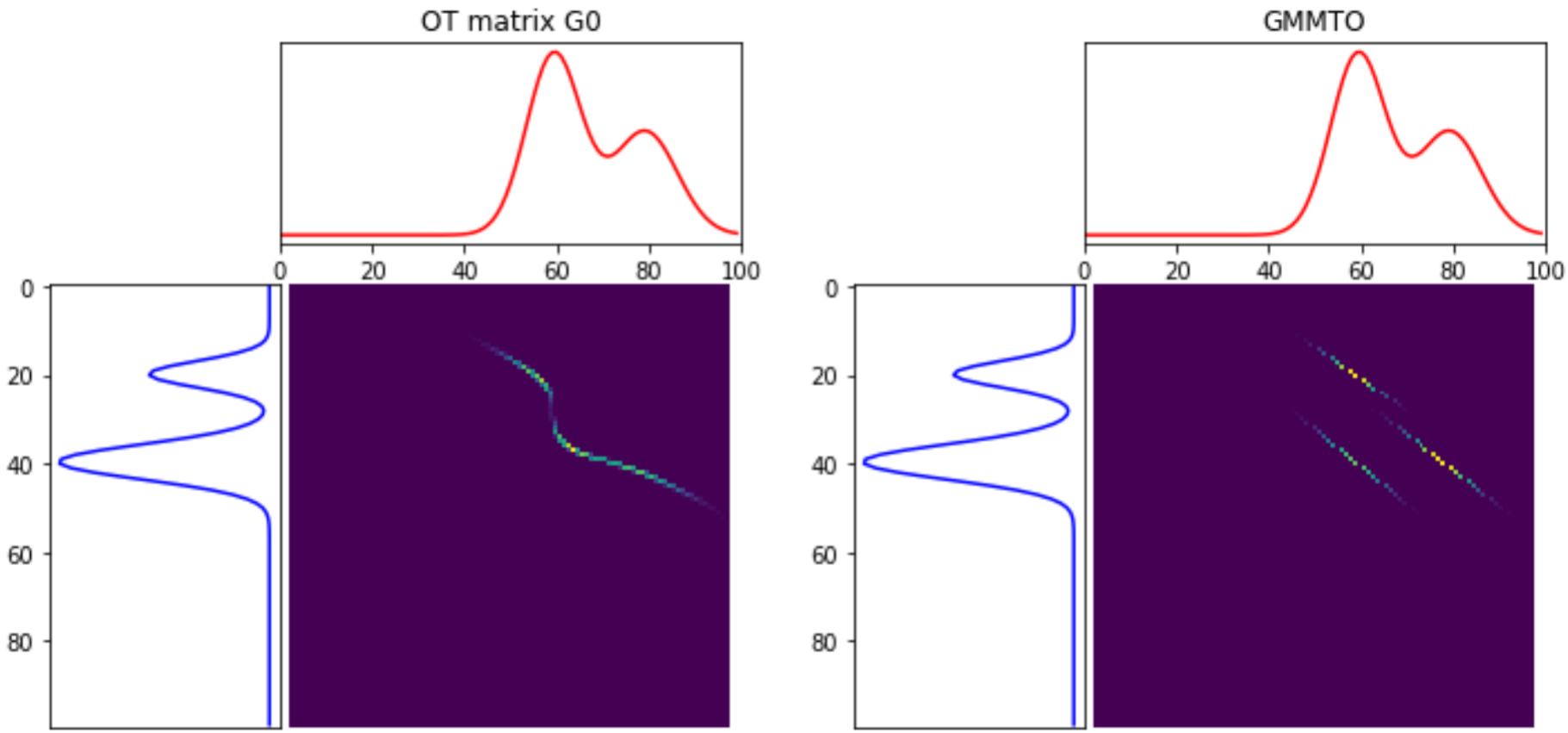
$$\text{with} \quad \Pi(\Lambda_0, \Lambda_1) := \Bigg\{ w := (w_{jk})_{1\leq j\leq J, 1\leq k\leq K} \in \mathbb{R}_+^{J\times K},$$

$$\forall 1 \leq j \leq J, \sum_{k=1}^{K} w_{jk} = \lambda_0^j, \quad \forall 1 \leq k \leq K, \sum_{j=1}^{J} w_{jk} = \lambda_1^k \Bigg\}.$$

Aim : Exploit the compact form of the solution as a mixture of Slater functions

Delon, Desolneux : A Wasserstein-Type Distance in the Space of Gaussian Mixture Models. SIAM J. Imaging Sci. (2020).
Dusson, Ehrlacher, Nouaime : A Wasserstein-type metric for generic mixture models, including location-scatter and group invariant measures (2023)

Illustration of the modification of the transport plan



- Transport map replaced by affine lines
- Number of components cannot exceed $J + K - 1$

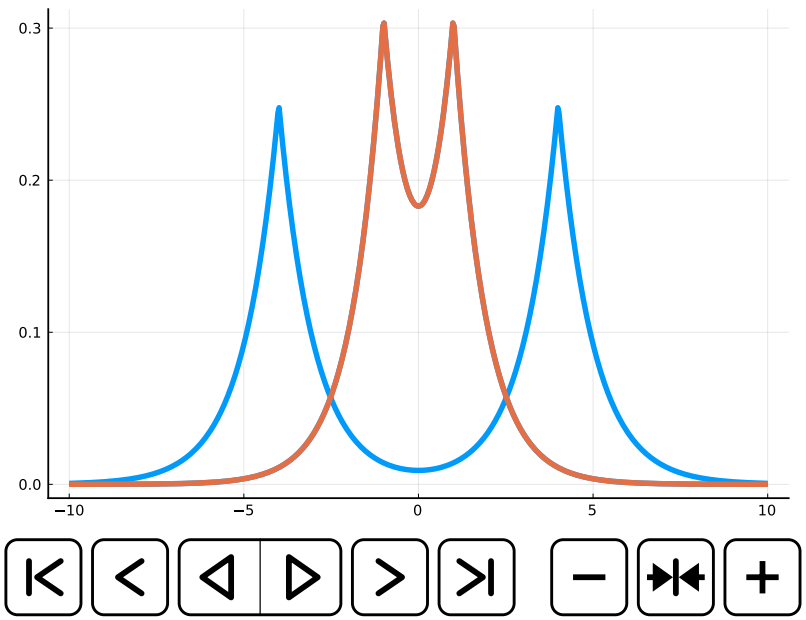
Delon, Desolneux : A Wasserstein-Type Distance in the Space of Gaussian Mixture Models. SIAM J. Imaging Sci. (2020).

Mixture barycenter between two solutions

- Probability measures $(u_1, \ldots, u_n)$
- Positive weights $\boldsymbol{t} = (t_1, \ldots, t_n)$.

Unique solution to the problem

$$\inf_{u \in \mathcal{M}(\mathcal{A})} \sum_{i=1}^{n} t_i \, mW_2^2(u, u_i)^2.$$



Interesting features :

- Way better approximation compared to W_2 barycenter
- Computational cost independent of the dimension

Formula for the barycenter :

$$\mathrm{Bar}^{\boldsymbol{t}}_{\mathrm{MW}_2}(u_1, \ldots, u_n) = \sum_{\mathbf{k} \in \mathbf{K}} w^*_{\mathbf{k}} \, \mathrm{Bar}^{\boldsymbol{t}}_{W_2}\left(u_1^{k^1}, \ldots, u_N^{k^n}\right),$$

A few properties

Valid for a large number of probability distributions

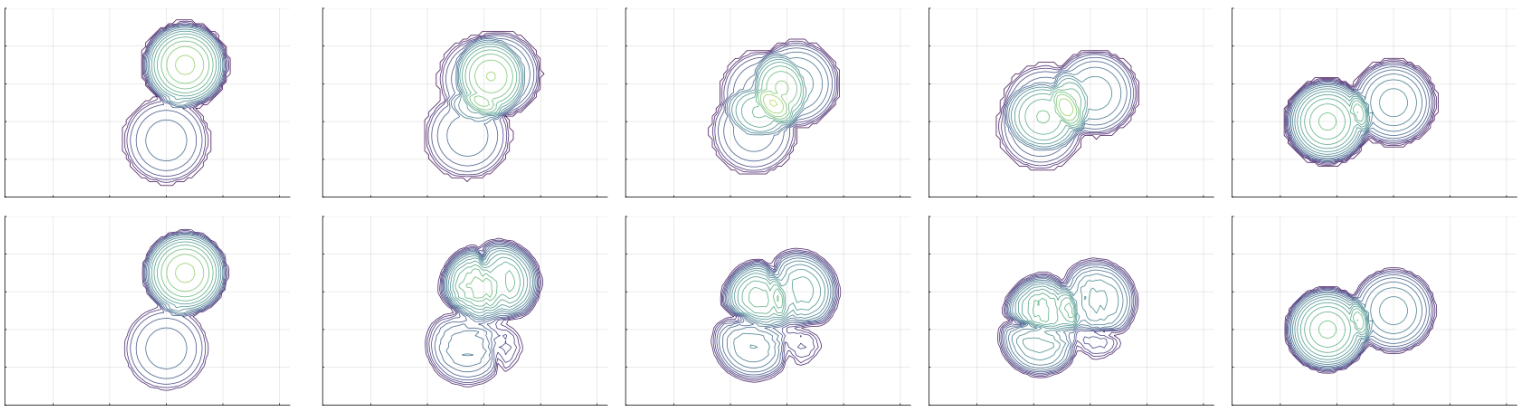
Mathematically :

- needs a geodesic space for the atoms (space with distance + geodesic)
- identifiability

Computationally : barycenters need to be easily computable (best if explicit !)

Examples :

- Elliptic distributions (Slater, gaussians, Wigner semicircle)
- Location-scatter (dilations+translations)
- Group-invariant distributions (invariance put into the distance between the atom distributions)



(a) $t = 0$ (b) $t = 0.25$ (c) $t = 0.5$ (d) $t = 0.75$ (e) $t = 1$

Dusson, Ehrlacher, Nouaime : A Wasserstein-type metric for generic mixture models, including location-scatter and group invariant measures (2023)

Back to Kolmogorov n-width

Definition of nonlinear Kolmogorov n-width :
The Kolmogorov n-width of the set $\mathcal{M} \subset \mathbb{M}$ is defined by

$$\varepsilon_n(\mathcal{M}, \mathbb{M}) = \inf_{\boldsymbol{m} \in \mathbb{M}^n} \sup_{\mathbf{R} \in \mathcal{R}} \inf_{\boldsymbol{t} \in \Omega(\boldsymbol{m})} mW_2(u_{\mathbf{R}}, \mathrm{Bar}^{\boldsymbol{t}}(\boldsymbol{m})).$$

Theorem : (Dalery, D., Ehrlacher, Lozinski) for a system with two nuclei with identical charges, for $n > 1$,

$$\varepsilon_n(\mathcal{M}, mW_2) = 0.$$

- Exact representation of all solutions in this case

Dalery, Dusson, Ehrlacher, Lozinski : Nonlinear reduced basis using mixture Wasserstein barycenters : application to an eigenvalue problem inspired from quantum chemistry (2023)

Outline

Linear reduced-order modelling

Optimal transport : a few properties

Practical strategy and numerical results for a 1D toy problem

An example in 2D : Approximating the pair-density

Practical strategy

Nonlinear version

Offline part : Select accurate solutions for a few **wisely chosen** parameters $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_K \in \mathcal{R}$

- Generate a training set of snapshots for parameters $\mathbf{R} \in \mathcal{R}_{\text{train}}$
- Select "good" snapshots with a greedy algorithm
  - Select two parameters $\mathbf{R}_1, \mathbf{R}_2 \in \mathcal{R}$ that are as far as possible
  - at each iteration $K \geq 3$, select the snapshot that is worse approximated **as a barycenter** in the **set** of the previously selected snapshots $\Psi_{\mathbf{R}_1}, \ldots, \Psi_{\mathbf{R}_{K-1}}$

Online part : compute solutions for **many** parameters $\mathbf{R} \in \mathcal{R}$ **as a barycenter of** selected snapshots, i.e. in the set of the solutions $\Psi_{\mathbf{R}_1}, \ldots, \Psi_{\mathbf{R}_K}$.

Details on the greedy algorithm

Input : Training set $\mathcal{M}_{\text{train}}$, number of elements to select Q
Select m^1 and m^2 solutions to $\text{argmax}_{(m^1,m^2)\in\mathcal{M}_{\text{train}}} MW_2(m^1, m^2)$.
$\mathcal{B} := \{m^1, m^2\}$
for $q = 3, \ldots, Q$ **do**

Select

$$m^k \in \text{argmax}_{m\in\mathcal{M}_{\text{train}}} \min_{\boldsymbol{t}\in\Omega_{q-1}} mW_2(m, \text{Bar}^{\boldsymbol{t}}_{MW2}(m^1, \ldots, m^{q-1}))^2, \quad (1)$$

where

$$\Omega_{q-1} = \left\{ \boldsymbol{t} \in \mathbb{R}^{q-1}, \ \sum_{i=1}^{q-1} \frac{t_i}{\zeta^i} > 0 \right\}.$$

$\mathcal{B} = \mathcal{B} \cup \{m^q\}$
end for
Output : Reduced basis $\mathcal{B} \subset \mathcal{M}_{tr}$

Details on the greedy algorithm - 2

Mixtures : $m = \sum_{k=1}^{K} \pi_k m_k$ and for $i \in \{1, \dots, n\}$, $m^i = \sum_{k^i=1}^{K^i} \pi^i_{k^i} m^i_{k^i}$

$$\textbf{Solve :} \quad \min_{\boldsymbol{t} \in \Omega_q} m W_2(m, \mathrm{Bar}^{\boldsymbol{t}}_{MW2}(m^1, \dots, m^q))^2,$$

$$\min_{\boldsymbol{t} \in \Omega_q} \min_{w \in \Pi(\boldsymbol{\pi}, w^*)} \sum_{\mathbf{k} \in \mathbf{K}} \sum_{k=1}^{K} w_{\mathbf{k},k} W_2(m_k, \mathrm{Bar}^{\boldsymbol{t}}_{W_2}(m^1_{k^1}, \dots, m^n_{k^n}))^2,$$

with

$$\begin{aligned} W_2(m_k, \mathrm{Bar}^{\boldsymbol{t}}_{W_2}(m^1_{k^1}, \dots, m^n_{k^n}))^2 &= \left(r_k - \sum_{i=1}^{n} \lambda_i r^i_{k^i} \right)^2 + 2 \left(\frac{1}{\zeta_k} - \sum_{i=1}^{n} \frac{t_i}{\zeta^i_{k^i}} \right)^2 \\ &= \boldsymbol{t}^\mathsf{T} A_{\mathbf{k}} \boldsymbol{t} + b^\mathsf{T}_{\mathbf{k},k} \boldsymbol{t} + c_k, \end{aligned}$$

So

$$\min_{w \in \Pi(\boldsymbol{\pi}, w^*)} \min_{\boldsymbol{t} \in \Omega_k} \boldsymbol{t}^\mathsf{T} A \boldsymbol{t} + b^\mathsf{T}_w \boldsymbol{t} + c = \min_{w \in \Pi(\boldsymbol{\pi}, w^*)} -\frac{1}{4} b^\mathsf{T}_w A^{-1} b_w + c,$$

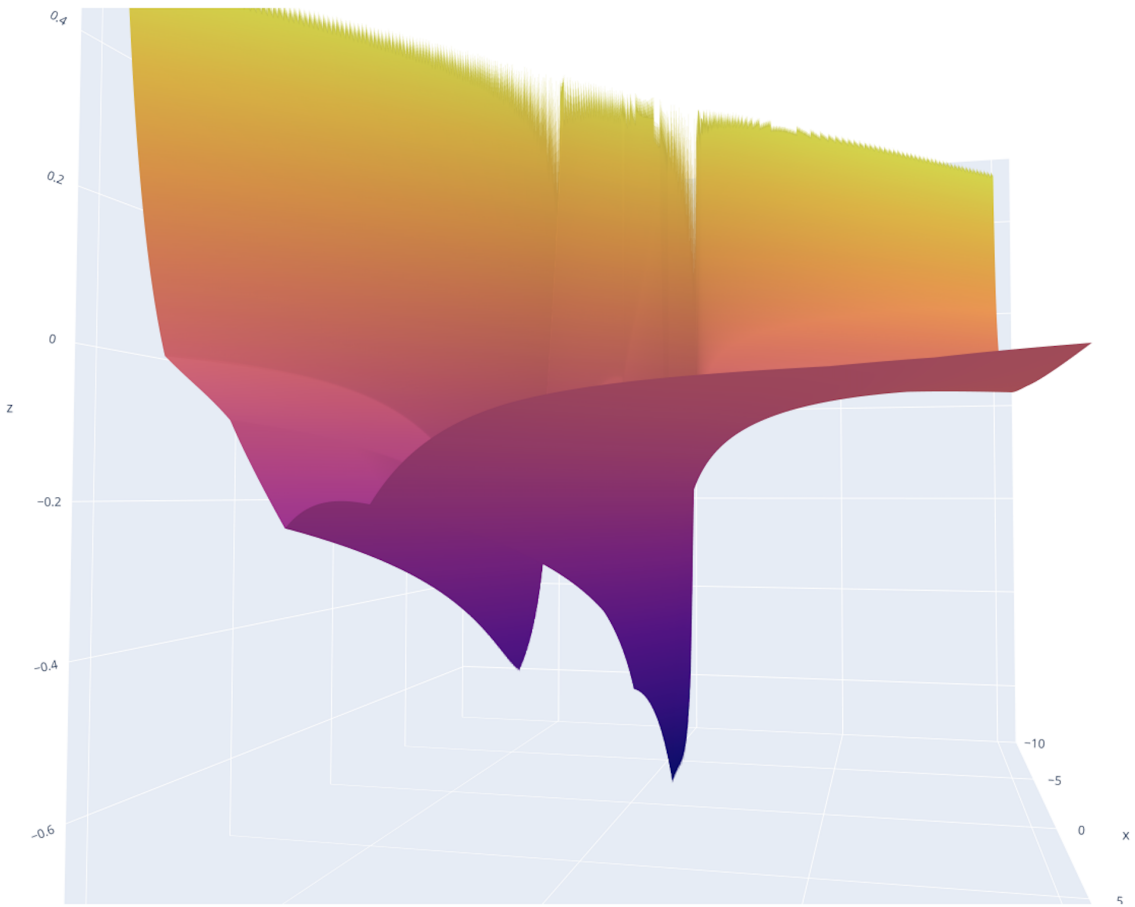
Non convex optimization problem to solve

Online algorithm : energy minimization

Given $\Psi_{\mathbf{R}_1}, \ldots, \Psi_{\mathbf{R}_n}$, solve

$$\inf_{\boldsymbol{t}} E(\mathrm{bar}(\boldsymbol{t}; \Psi_{\mathbf{R}_1}, \ldots, \Psi_{\mathbf{R}_n}))$$

- Nonlinear problem, but in low dimension
- Using quasi-Newton method starting from different initial guesses

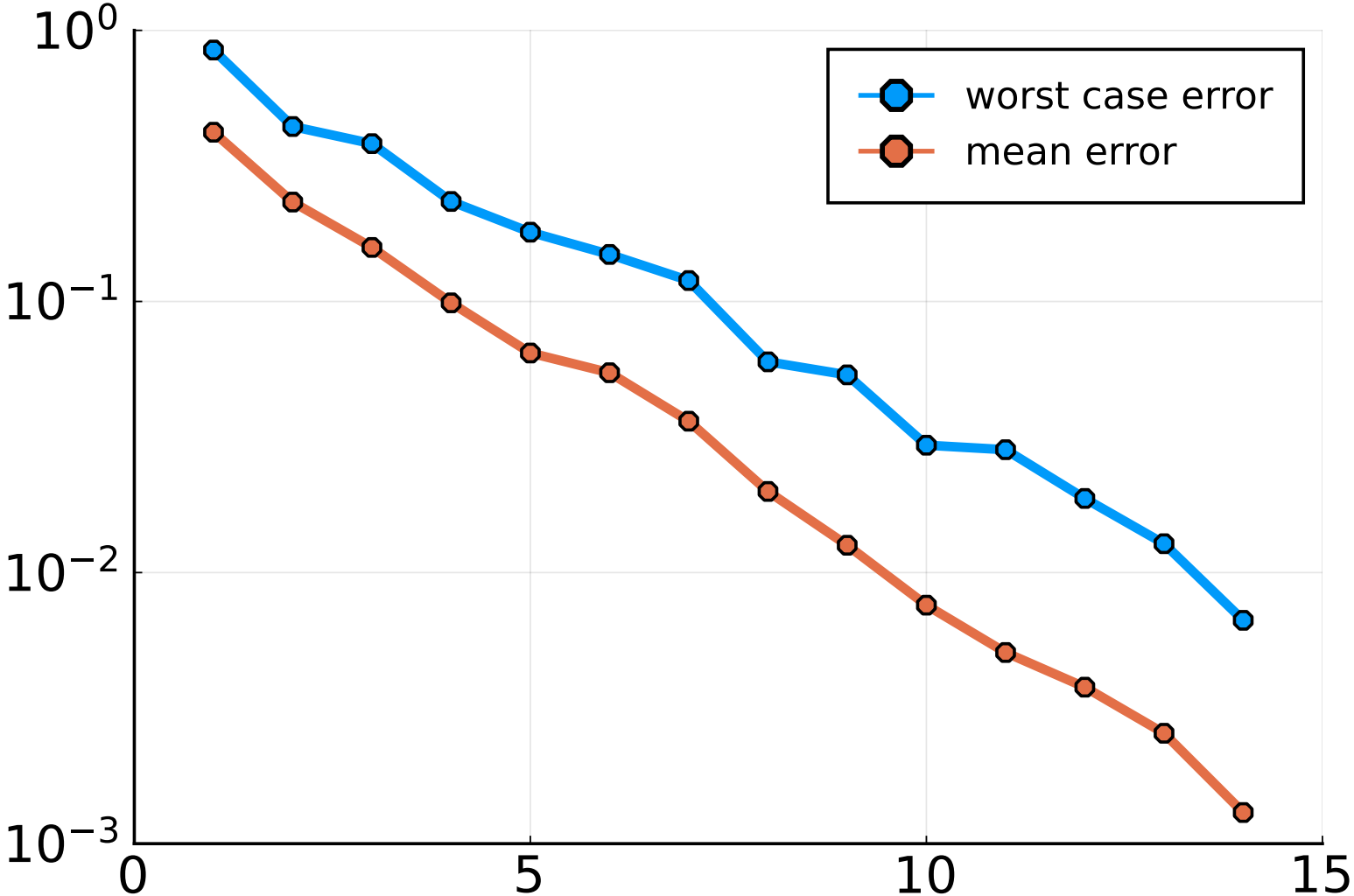


Numerical results : greedy algorithm

Charges : $(0.8, 1.1)$.

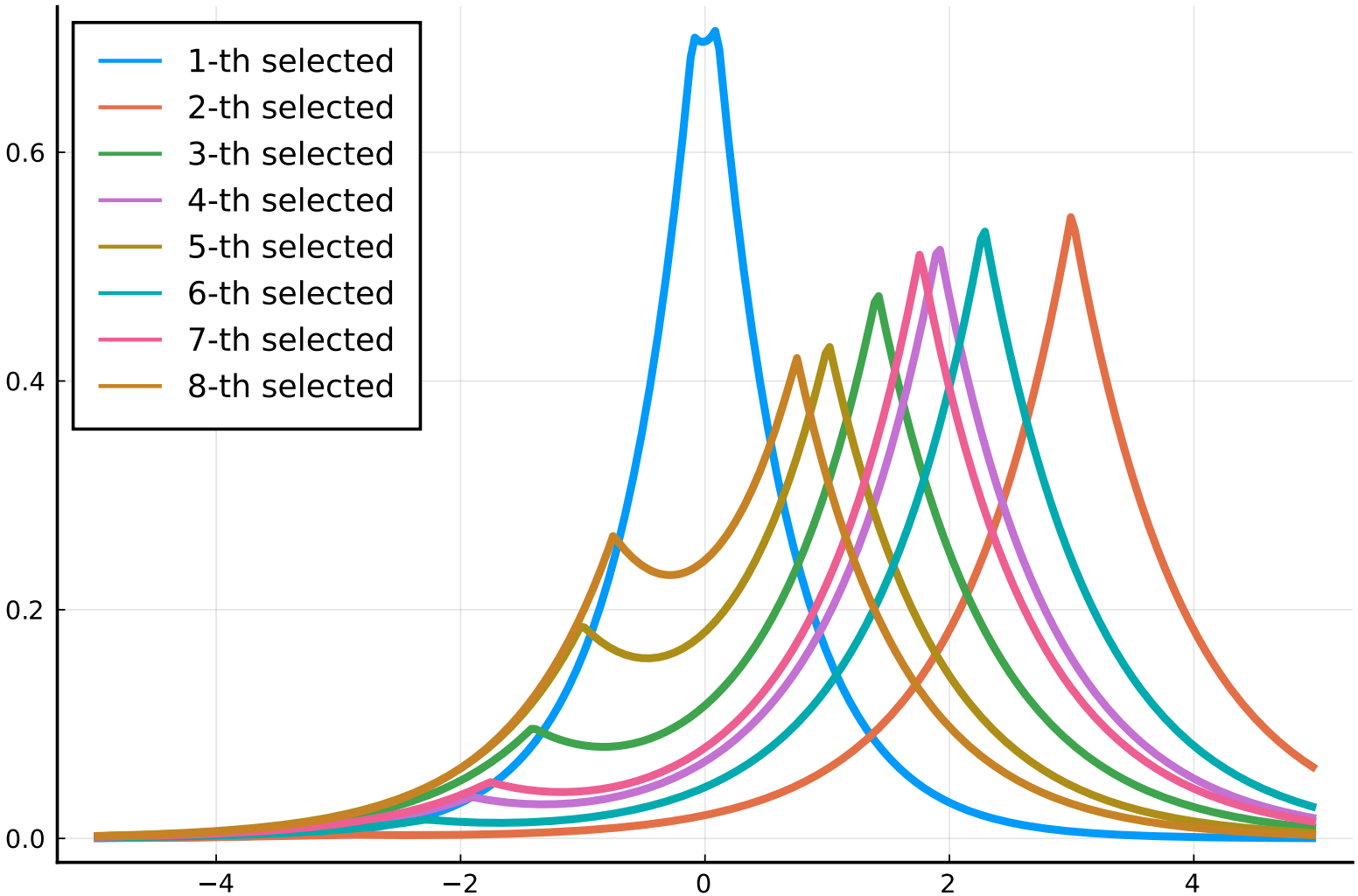
291 solutions in the training set.

Error decrease with respect to number of selected snapshots



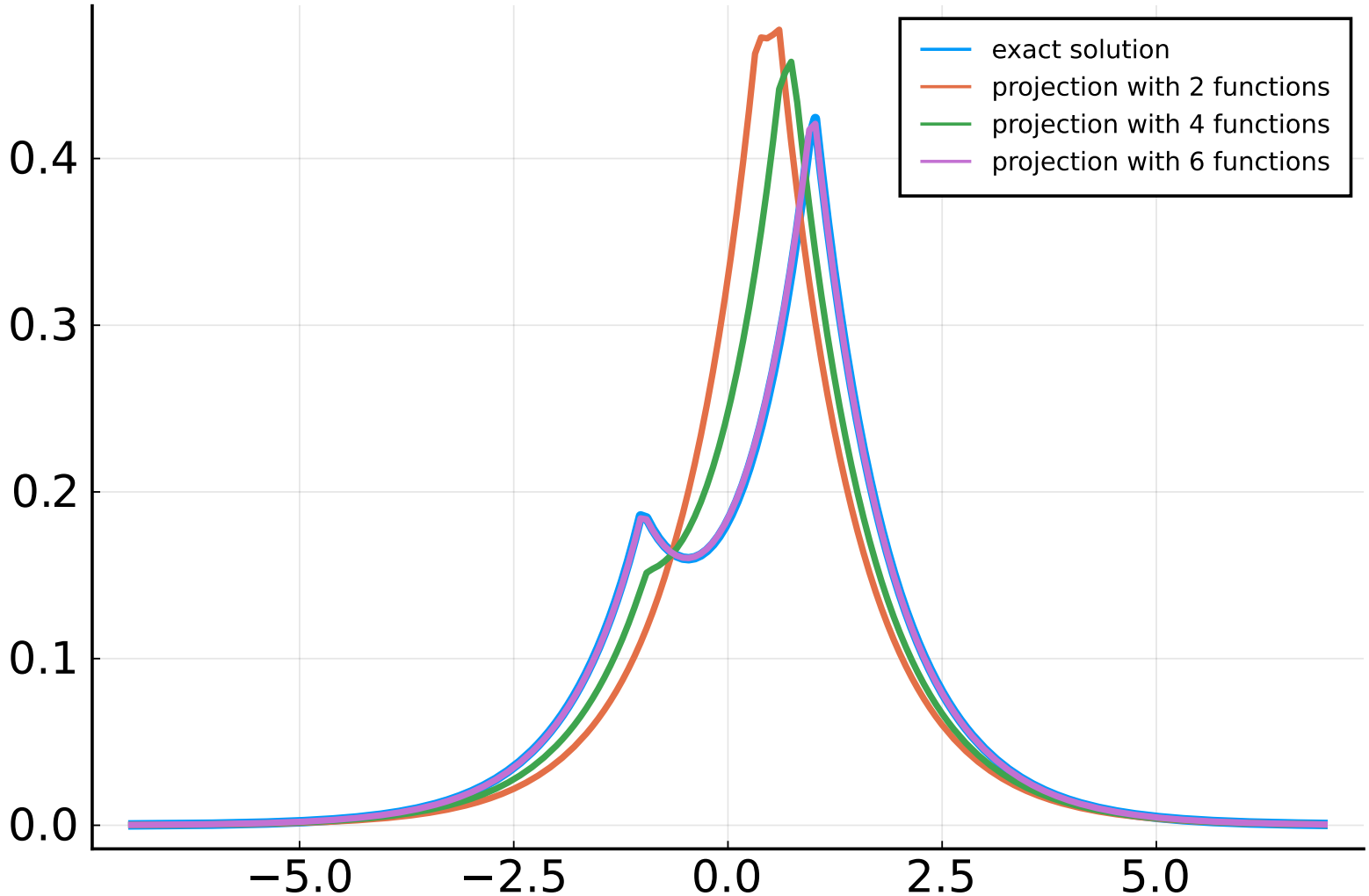
Numerical results : greedy algorithm

First eight selected solutions

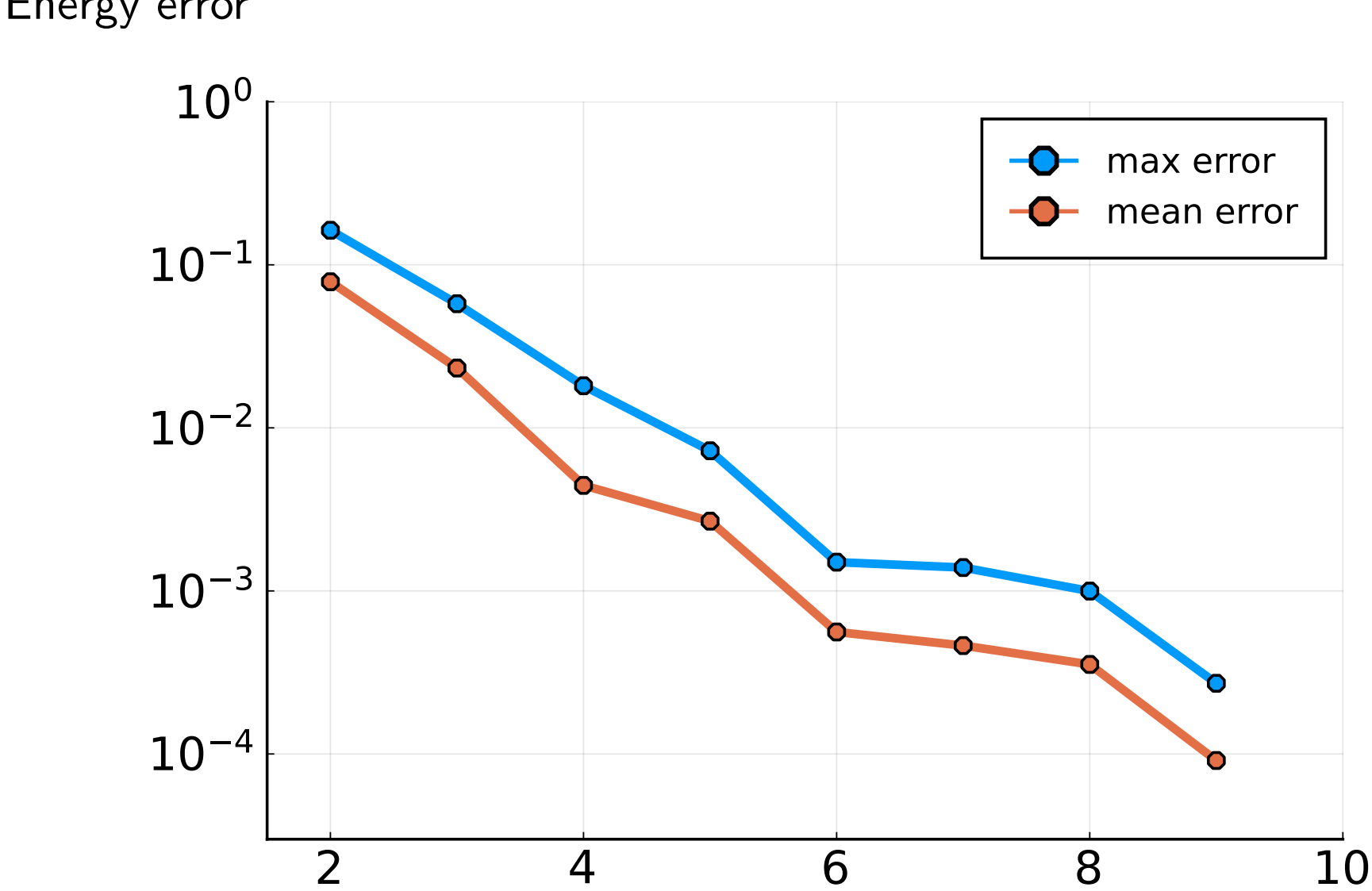


Numerical results : greedy algorithm

Projection example

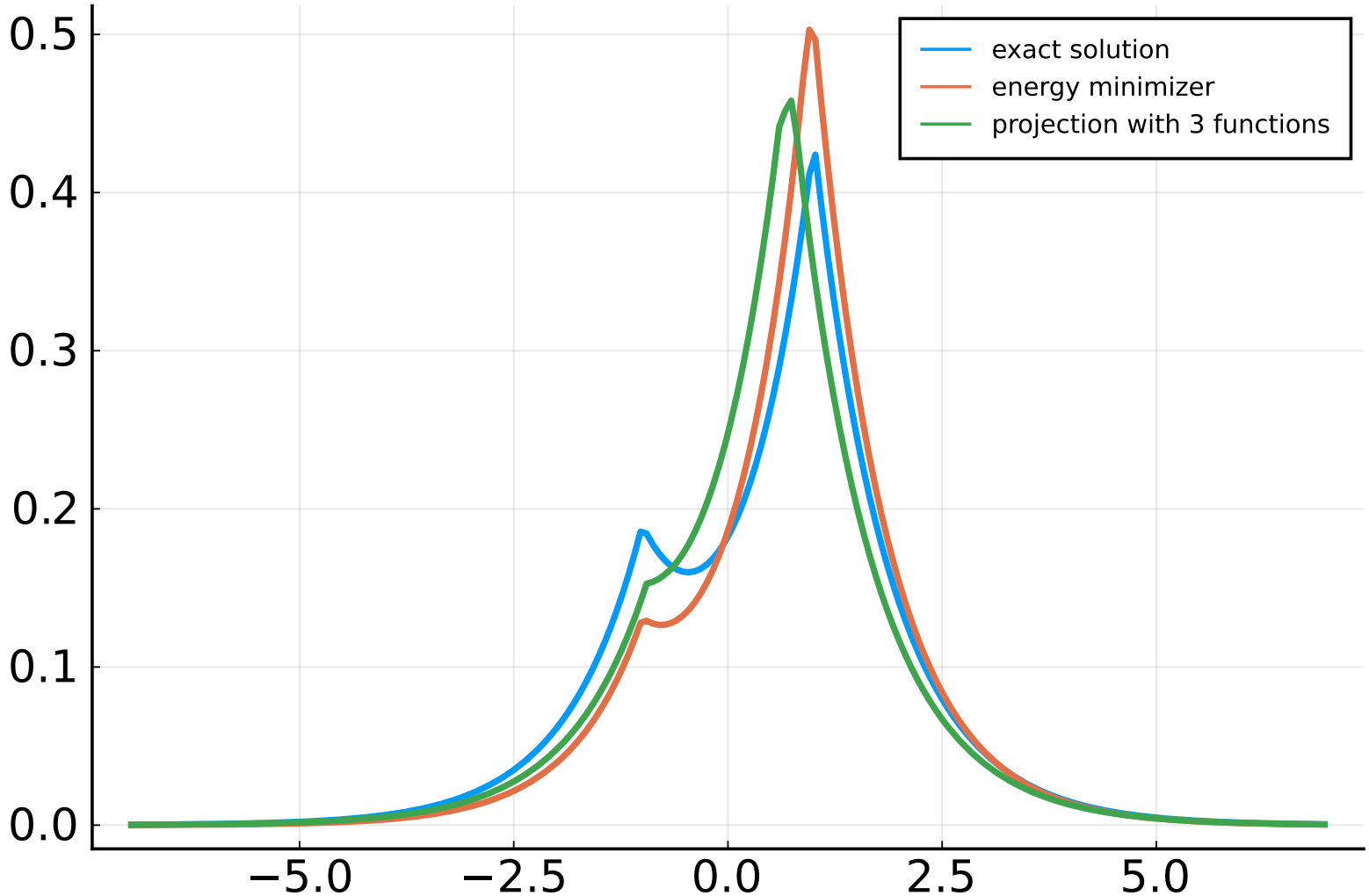


Numerical results : online energy minimization



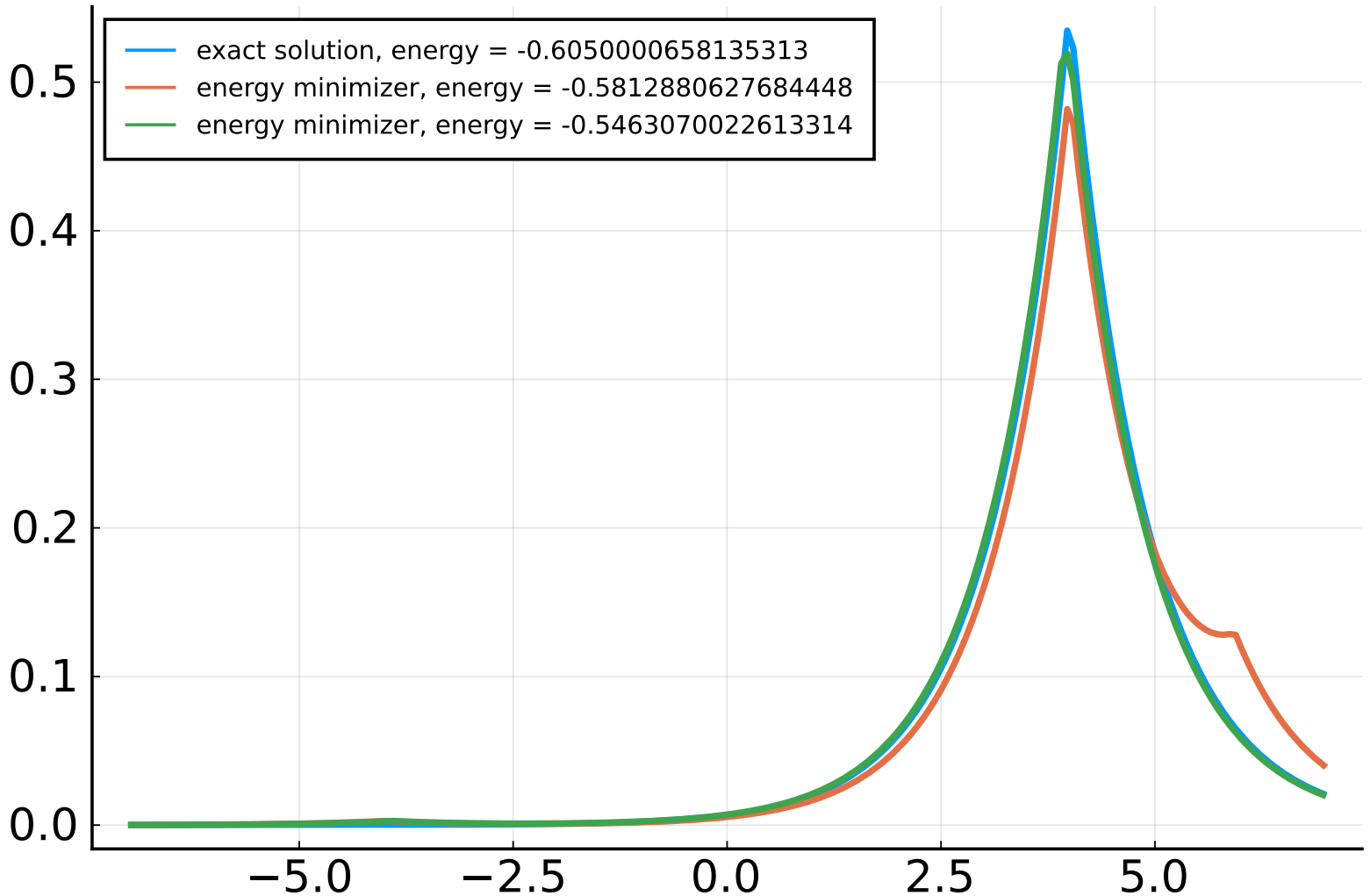
Numerical results

Comparison between projection and energy minimization



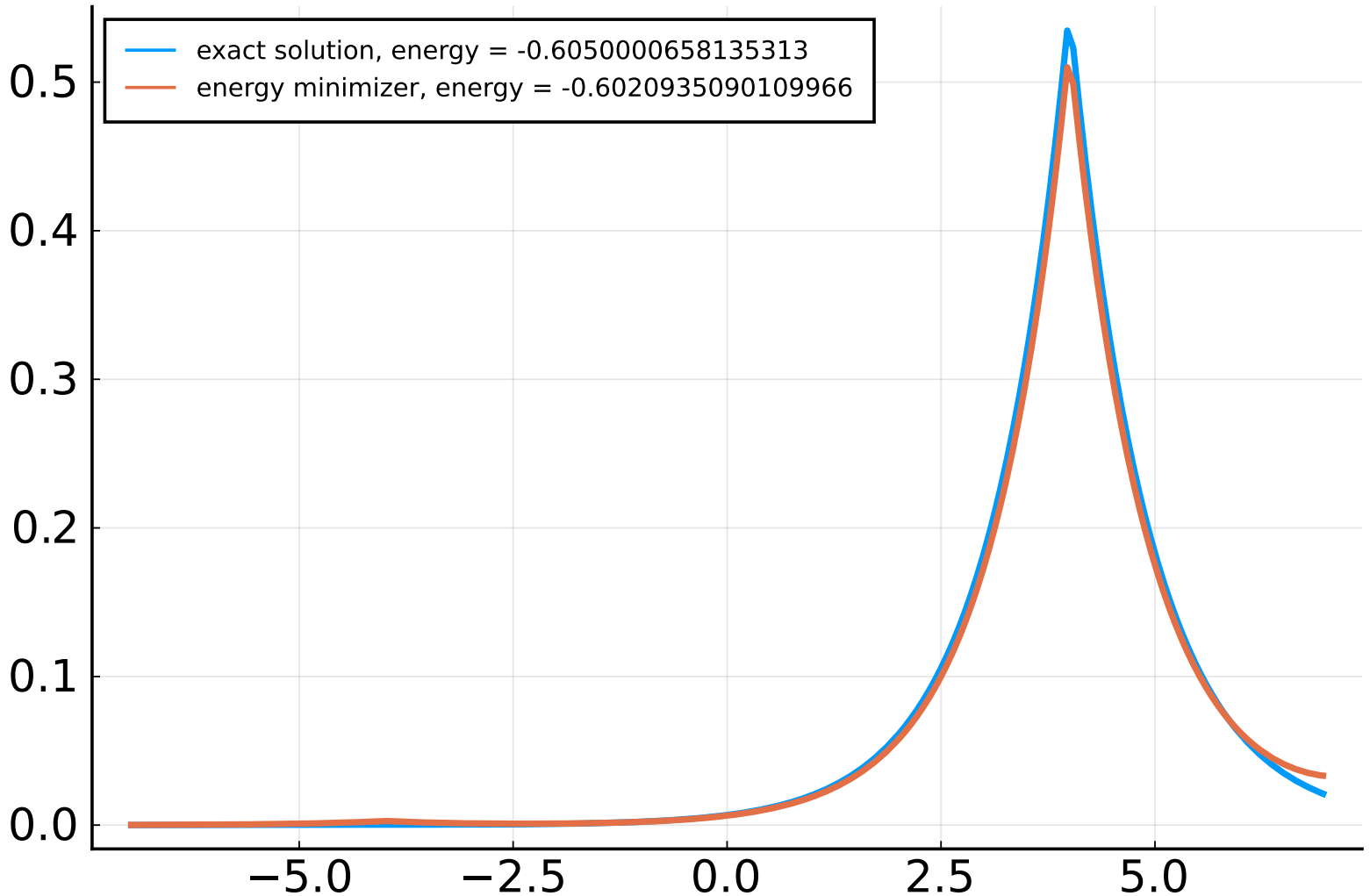
First extrapolation example

Using a basis with 3 functions



First extrapolation example

Using a basis with 5 functions



Outline

Approximating the pair density

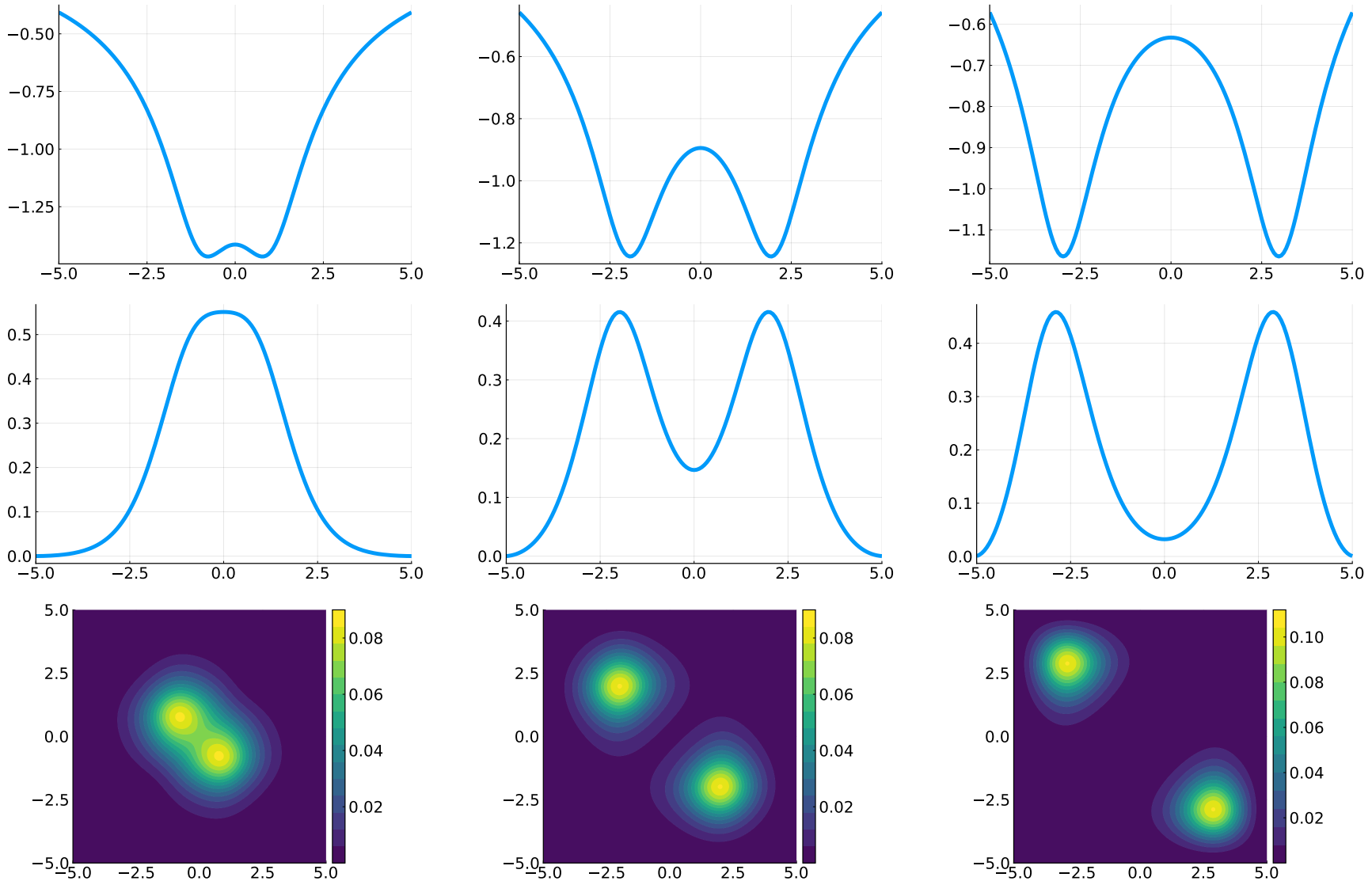
Defined as

$$\rho_2^{\Psi}(x,y) = \binom{N}{2} \int_{(\mathbb{R}^d)^{N-2}} |\Psi(x,y,x_3,\cdots,x_N)|^2 dx_3 \cdots dx_N.$$

- Propose practical approximations of the pair density in order to efficiently compute

$$V_{ee}[\Psi] = \int_{\mathbb{R}^{2d}} \frac{\rho_2^{\Psi}(x,y)}{|x-y|} dx\, dy.$$

Two particles systems



(a) $a = 0.5$ (b) $a = 1.5$ (c) $a = 2.5$

Finite element code for 1D particles developed with Xue Quan and Huajie Chen (in Julia).

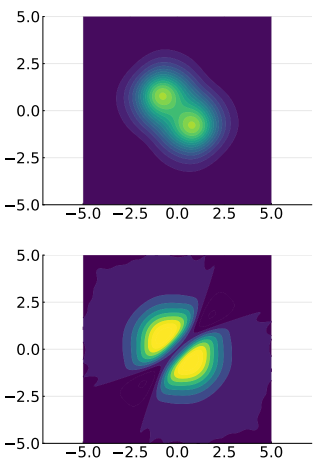
Approximation using Local Density Approximation

Relative error computed with respect to the Coulomb energy

$$V_{ee}[\Psi] = \int_{\mathbb{R}^2} \frac{\rho_2^{\Psi}(x,y)}{|x-y|} dx\, dy.$$

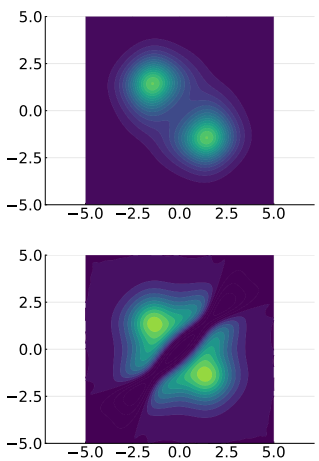
For LDA : ρ_2 such that

$$V_{ee}[\rho] = \frac{1}{2} \int \frac{\rho_2^{\Psi}(x,y)}{|x-y|} dx\, dy - c_x \int \rho(x)^{4/3} dx.$$



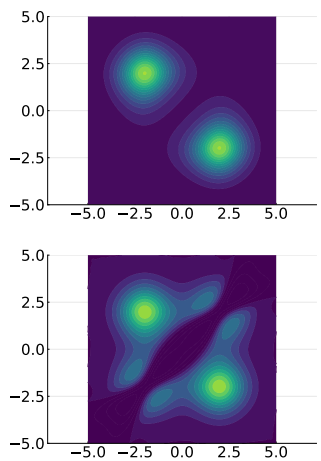
$Err = 0.29$

(d) $a = 1.$



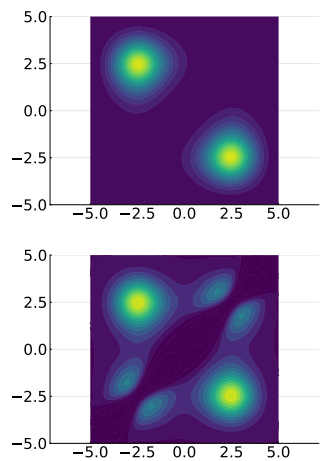
$Err = 0.22$

(e) $a = 1.5$



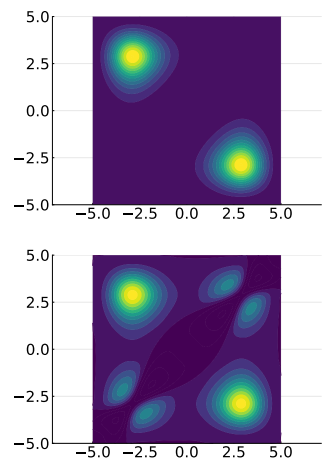
$Err = 0.14$

(f) $a = 2$



$Err = 0.09$

(g) $a = 2.5$



$Err = 0.07$

(h) $a = 3$

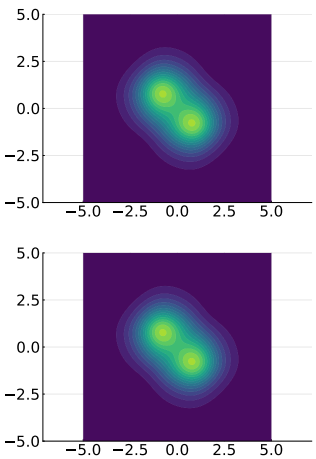
Chen, Friesecke : Pair Densities in Density Functional Theory. Multiscale Model. Simul. (2015).

Learning the pair density

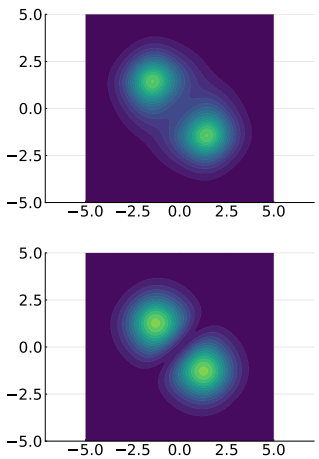
Main ideas :

- Construct a database of densities and pair densities
- Select most representative pair densities using a greedy algorithm
- For a new density,
  1. fit the density as a barycenter of selected densities
  2. approximate the pair density as a barycenter of corresponding pair densities

Simple example using a W_2 Wasserstein barycenter :

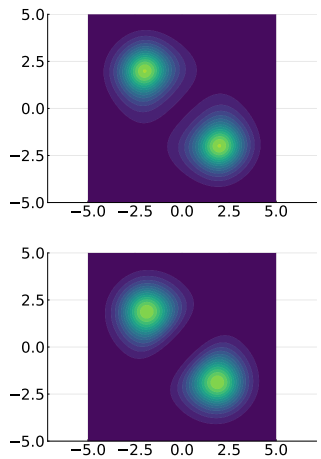


(i) $a = 1.$



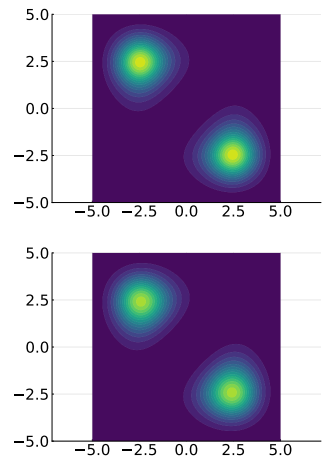
$Err = 1.2.10^{-3}$

(j) $a = 1.5$



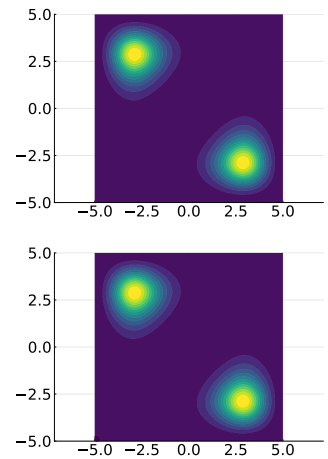
$Err = 8.1.10^{-3}$

(k) $a = 2$



$Err = 6.6.10^{-3}$

(l) $a = 2.5$

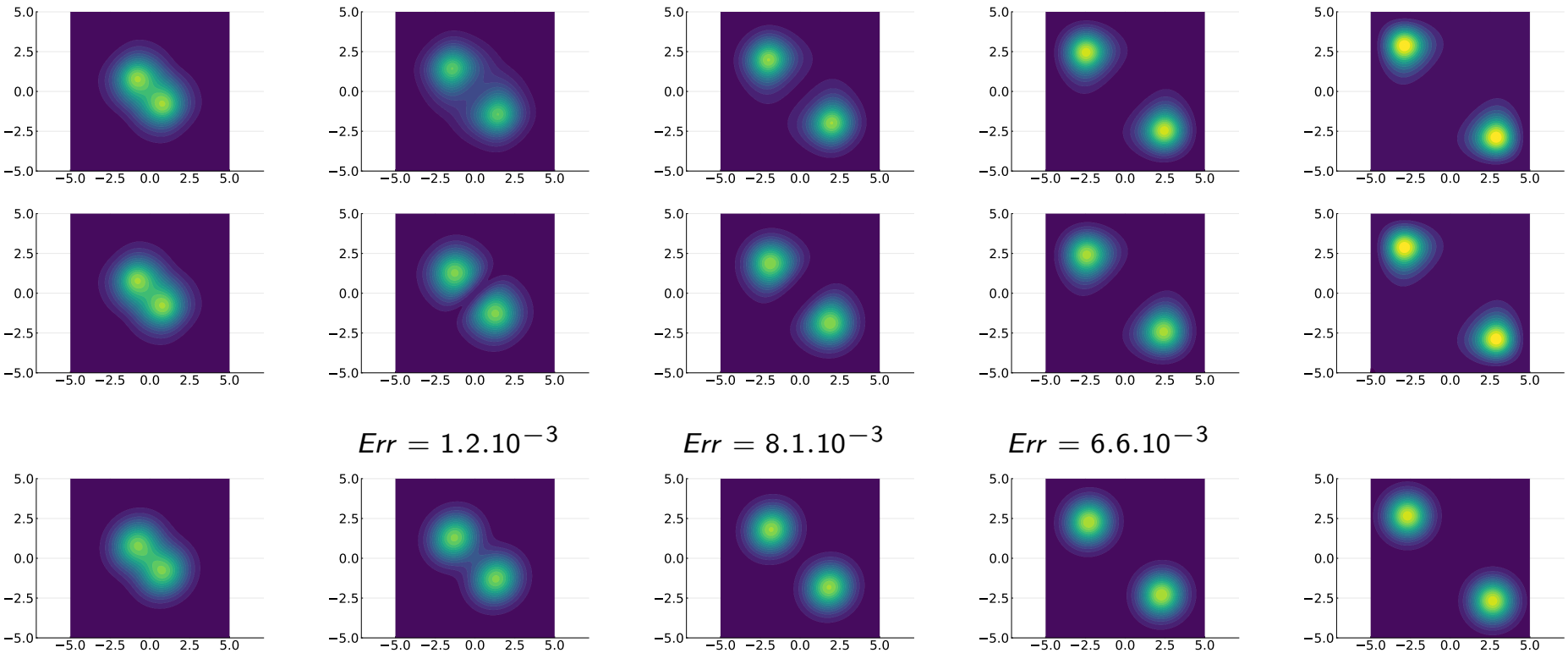


(m) $a = 3$

Using mixture distance : fitting the pair density

Computational cost limits the use of Wasserstein barycenters

- Pair densities can be fitted using a few gaussian mixtures - sparse representation for efficiently computing barycenters



$Err = 1.2.10^{-3}$ $Err = 8.1.10^{-3}$ $Err = 6.6.10^{-3}$

$Err = 4.8.10^{-2}$ $Err = 5.0.10^{-2}$ $Err = 4.4.10^{-2}$

(n) $a = 1.$ (o) $a = 1.5$ (p) $a = 2$ (q) $a = 2.5$ (r) $a = 3$

Conclusion and perspectives

Nonlinear reduced model based on mixture Wasserstein barycenters

Key points :

- Computation of barycenters independent of the underlying dimension
- Problem size depends on number of functions in the mixtures

Limitations :

- Needs to consider probability distributions
- Multi-marginal problem hard in general, but hope with new algorithms

  Friesecke, Penka : The GenCol algorithm for high-dimensional optimal transport : general formulation and application to barycenters and Wasserstein splines, http ://arxiv.org/abs/2209.09081, (2022).

Extensions :

- Consider orthogonal projectors problems using Quantum OT
- Accelerate online calculations via learning of the parameter map
- 3D simulations, error bounds

Thank you !