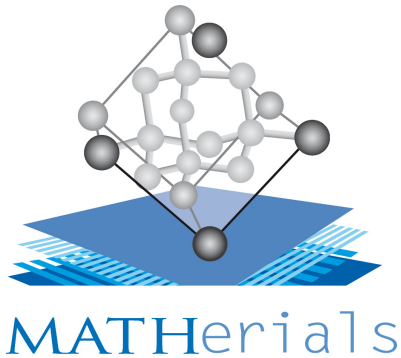
Marginal-constrained modified Wasserstein barycenters for Gaussian mixtures

Maxime Dalery[3], Geneviève Dusson[3], Virginie Ehrlacher[1,2]

[1] CERMICS, Ecole des Ponts ParisTech

[2] MATHERIALS team-project, INRIA Paris

[3] CNRS, Université de Bourgogne Franche-Comté

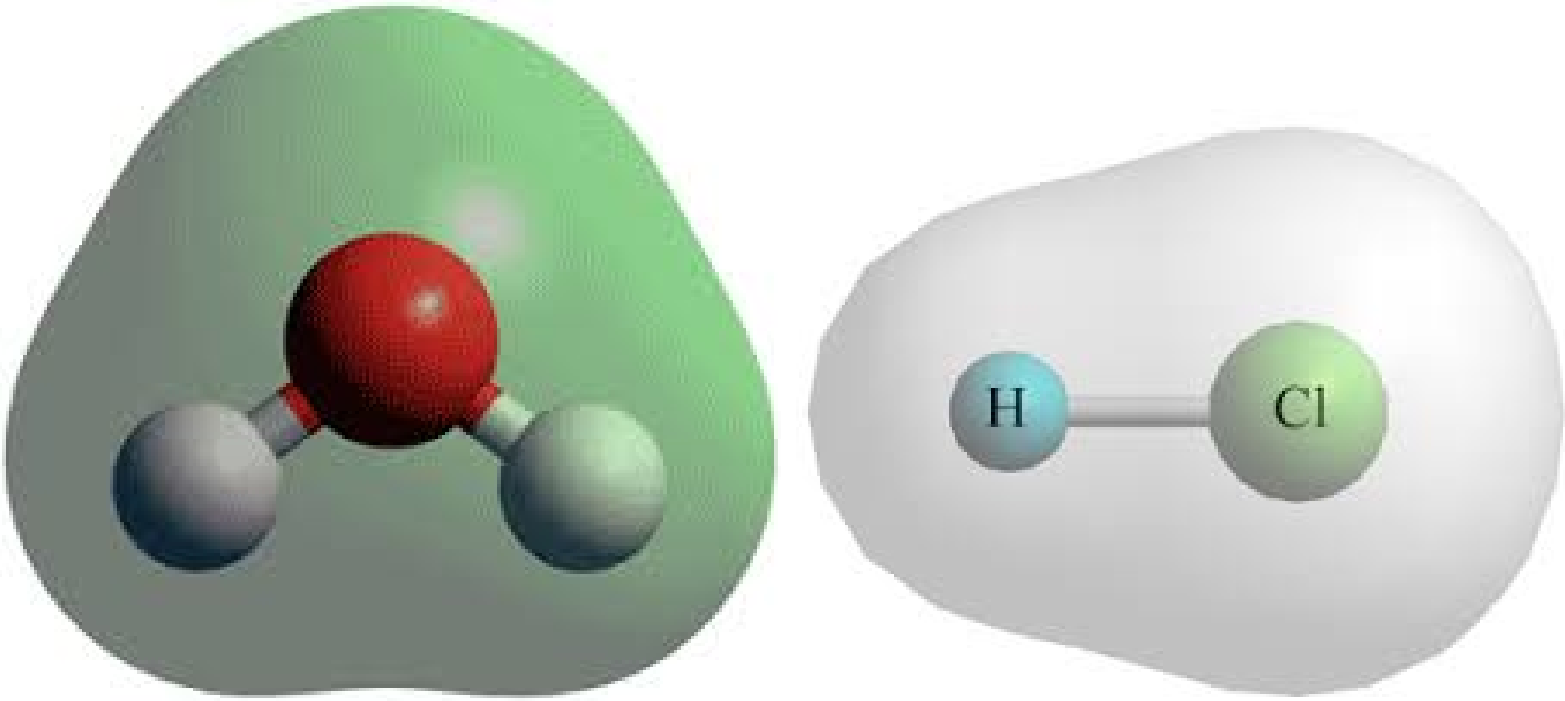


Journée SMAI-SIGMA, 4th December 2025

Outline of the talk

Outline of the talk

Electronic structure calculations for molecules



Computation of the **evolution in time of the state of the set of electrons** in a molecule: electrical, magnetical, optical properties...

Many-body Schrödinger model

For the sake of simplicity, atomic units will be used and the influence of spin will be neglected.

Many-body Schrödinger model

For the sake of simplicity, atomic units will be used and the influence of spin will be neglected.

Born-Oppenheimer approximation:

Many-body Schrödinger model

For the sake of simplicity, atomic units will be used and the influence of spin will be neglected.

Born-Oppenheimer approximation:

Let us consider a physical system composed of

Many-body Schrödinger model

For the sake of simplicity, atomic units will be used and the influence of spin will be neglected.

Born-Oppenheimer approximation:

Let us consider a physical system composed of

- P nuclei, that are assumed to be (fixed) classical point charges, whose positions and electric charges are denoted by $R_1, \ldots, R_P \in \mathbb{R}^3$ and $Z_1, \ldots, Z_P \in \mathbb{N}^*$ respectively;

Many-body Schrödinger model

For the sake of simplicity, atomic units will be used and the influence of spin will be neglected.

Born-Oppenheimer approximation:

Let us consider a physical system composed of

- P nuclei, that are assumed to be (fixed) classical point charges, whose positions and electric charges are denoted by $R_1, \ldots, R_P \in \mathbb{R}^3$ and $Z_1, \ldots, Z_P \in \mathbb{N}^*$ respectively;
- N electrons, considered as quantum particles: the ground (lowest energy) state of the electrons is represented by a complex-valued function $\psi : \mathbb{R}^{3N} \to \mathbb{C}$ which is antisymmetric with respect to permutations of the order of the electrons. The function ψ is called the ground state **wavefunction** of the system of electrons.

Many-body Schrödinger model

For the sake of simplicity, atomic units will be used and the influence of spin will be neglected.

Born-Oppenheimer approximation:

Let us consider a physical system composed of

- P nuclei, that are assumed to be (fixed) classical point charges, whose positions and electric charges are denoted by $R_1, \ldots, R_P \in \mathbb{R}^3$ and $Z_1, \ldots, Z_P \in \mathbb{N}^*$ respectively;
- N electrons, considered as quantum particles: the ground (lowest energy) state of the electrons is represented by a complex-valued function $\psi : \mathbb{R}^{3N} \to \mathbb{C}$ which is antisymmetric with respect to permutations of the order of the electrons. The function ψ is called the ground state **wavefunction** of the system of electrons.

Physical interpretation of the wavefunction:

For $x_1, \ldots, x_N \in \mathbb{R}^3$, the quantity $|\psi(x_1, \ldots, x_N)|^2$ represents the probability density at time t of the positions $x_1, \ldots, x_N$ of the N electrons.

For $B \subset \mathbb{R}^{3N}$,

$$\int_B |\psi|^2 \text{: probability that the electrons are located in the set } B.$$

Consequence: $\boxed{\int_{\mathbb{R}^{3N}} |\psi|^2 = \|\psi\|^2_{L^2(\mathbb{R}^{3N})} = 1}$

Electronic Schrödinger model: Born-Oppenheimer approximation

Let $\boldsymbol{R} = (R_1, \cdots, R_P) \in \mathcal{R} \subset \mathbb{R}^{3P}$ the set of positions of the nuclei.

$$V_{\boldsymbol{R}}(x) = -\sum_{p=1}^{P} \frac{Z_p}{|x - R_p|}, \quad x \in \mathbb{R}^3.$$

For a given value of $\boldsymbol{R}$, the corresponding ground state electronic wavefunction $\psi_{\boldsymbol{R}}$ is solution to the eigenvalue problem (electronic Schrödinger problem)

$$H_{\boldsymbol{R}}\psi_{\boldsymbol{R}} = E_{\boldsymbol{R}}\psi_{\boldsymbol{R}},$$

with

$$H_{\boldsymbol{R}} = -\Delta_{x_1,\ldots,x_N} + \sum_{i=1}^{N} V_{\boldsymbol{R}}(x_i) + \sum_{1 \leqslant i < j \leqslant N} \frac{1}{|x_i - x_j|}$$

Problem: Linear eigenvalue problem for functions defined on the high-dimensional space $\mathbb{R}^{3N}$.

Electronic (one or two-body) density:

$$\rho(x) = \int_{\mathbb{R}^{3(N-1)}} |\psi(x, x_2, \ldots, x_N)|^2, \quad \tau(x, y) = \int_{\mathbb{R}^{3(N-2)}} |\psi(x, y, x_3, \ldots, x_N)|^2,$$

$$\boxed{\rho(x) = \int_{\mathbb{R}^3} \tau(x, y)\, dy} \qquad \int_{\mathbb{R}^3} \rho(x)\, dx = \int_{\mathbb{R}^6} \tau(x, y)\, dx\, dy = 1$$

Electronic (one or two-body) density:

$$\rho(x) = \int_{\mathbb{R}^{3(N-1)}} |\psi(x, x_2, \ldots, x_N)|^2, \quad \tau(x, y) = \int_{\mathbb{R}^{3(N-2)}} |\psi(x, y, x_3, \ldots, x_N)|^2,$$

$$\boxed{\rho(x) = \int_{\mathbb{R}^3} \tau(x, y)\, dy} \qquad \int_{\mathbb{R}^3} \rho(x)\, dx = \int_{\mathbb{R}^6} \tau(x, y)\, dx\, dy = 1$$

Instead of solving the full many-body Schrödinger electronic problem, quantum chemists prefer to solve **approximate problems**, which enable to obtain approximations of the electronic one-body density $\rho_{\boldsymbol{R}}(x)$ (or electronic two-body densities $\tau_{\boldsymbol{R}}(x, y)$).

Density (Matrix) Functional Theory

Electronic (one or two-body) density:

$$\rho(x) = \int_{\mathbb{R}^{3(N-1)}} |\psi(x, x_2, \ldots, x_N)|^2, \quad \tau(x, y) = \int_{\mathbb{R}^{3(N-2)}} |\psi(x, y, x_3, \ldots, x_N)|^2,$$

$$\boxed{\rho(x) = \int_{\mathbb{R}^3} \tau(x, y)\, dy} \qquad \int_{\mathbb{R}^3} \rho(x)\, dx = \int_{\mathbb{R}^6} \tau(x, y)\, dx\, dy = 1$$

Instead of solving the full many-body Schrödinger electronic problem, quantum chemists prefer to solve **approximate problems**, which enable to obtain approximations of the electronic one-body density $\rho_{\boldsymbol{R}}(x)$ (or electronic two-body densities $\tau_{\boldsymbol{R}}(x, y)$).

These approximate models boil down to solving **nonlinear eigenvalue problems** for functions defined on $\mathbb{R}^3$ (or $\mathbb{R}^6$).

Density (Matrix) Functional Theory

Electronic (one or two-body) density:

$$\rho(x) = \int_{\mathbb{R}^{3(N-1)}} |\psi(x, x_2, \ldots, x_N)|^2, \quad \tau(x,y) = \int_{\mathbb{R}^{3(N-2)}} |\psi(x, y, x_3, \ldots, x_N)|^2,$$

$$\boxed{\rho(x) = \int_{\mathbb{R}^3} \tau(x,y)\, dy} \qquad \int_{\mathbb{R}^3} \rho(x)\, dx = \int_{\mathbb{R}^6} \tau(x,y)\, dx\, dy = 1$$

Instead of solving the full many-body Schrödinger electronic problem, quantum chemists prefer to solve **approximate problems**, which enable to obtain approximations of the electronic one-body density $\rho_{\boldsymbol{R}}(x)$ (or electronic two-body densities $\tau_{\boldsymbol{R}}(x,y)$).

These approximate models boil down to solving **nonlinear eigenvalue problems** for functions defined on $\mathbb{R}^3$ (or $\mathbb{R}^6$).

Example of electronic one-body density approximate model: Find $\Phi_{\boldsymbol{R}}(x) = (\phi_{1,\boldsymbol{R}}(x), \cdots, \phi_{N,\boldsymbol{R}}(x))$ solution to

$$\begin{cases} (-\Delta_x + V_{\boldsymbol{R}}(x) + W[\rho_{\boldsymbol{R}}](x))\, \phi_{i,\boldsymbol{R}}(x) = \epsilon_{i,\boldsymbol{R}} \phi_{i,\boldsymbol{R}}(x) \\ \\ \rho_{\boldsymbol{R}}(x) = \frac{1}{N} \sum_{i=1}^N |\phi_{i,\boldsymbol{R}}(x)|^2 \\ W[\rho](x) = \int_{y \in \mathbb{R}^3} \frac{1}{|x-y|} \rho(y)\, dy \end{cases}$$

Model-order reduction for two-body electronic density models

Naturally, **two-body density** models are **more accurate** but **more expensive** than **one-body density models**...

Model-order reduction for two-body electronic density models

Naturally, **two-body density** models are **more accurate** but **more expensive** than **one-body density models**...

Objective: Build **new reduced-order models** to efficiently compute approximations $\widetilde{\tau}_{\boldsymbol{R}}$ of $\tau_{\boldsymbol{R}}$ for a large number of values of $\boldsymbol{R} \in \mathcal{R}$, out of the knowledge of $\rho_{\boldsymbol{R}}$ (computed from a one-body density model).

Model-order reduction for two-body electronic density models

Naturally, **two-body density** models are **more accurate** but **more expensive** than **one-body density models**...

Objective: Build **new reduced-order models** to efficiently compute approximations $\widetilde{\tau}_{\boldsymbol{R}}$ of $\tau_{\boldsymbol{R}}$ for a large number of values of $\boldsymbol{R} \in \mathcal{R}$, out of the knowledge of $\rho_{\boldsymbol{R}}$ (computed from a one-body density model).

$$\text{Compute } \rho_{\boldsymbol{R}} \text{ out of a one-body model} \Rightarrow \text{Compute } \widetilde{\tau}_{\boldsymbol{R}} \approx \tau_{\boldsymbol{R}}$$

Model-order reduction for two-body electronic density models

Naturally, **two-body density** models are **more accurate** but **more expensive** than **one-body density models**...

Objective: Build **new reduced-order models** to efficiently compute approximations $\widetilde{\tau}_{\boldsymbol{R}}$ of $\tau_{\boldsymbol{R}}$ for a large number of values of $\boldsymbol{R} \in \mathcal{R}$, out of the knowledge of $\rho_{\boldsymbol{R}}$ (computed from a one-body density model).

$$\text{Compute } \rho_{\boldsymbol{R}} \text{ out of a one-body model} \Rightarrow \text{Compute } \widetilde{\tau}_{\boldsymbol{R}} \approx \tau_{\boldsymbol{R}}$$

Important property 1: Link between one-body and two-body densities

$$\int_{\mathbb{R}^3} \tau_{\boldsymbol{R}}(x,y)\,dy = \rho_{\boldsymbol{R}}(x), \quad \int_{\mathbb{R}^3} \widetilde{\tau}_{\boldsymbol{R}}(x,y)\,dy = \rho_{\boldsymbol{R}}(x)$$

$$\Rightarrow \text{Marginal constraint}$$

Model-order reduction for two-body electronic density models

Naturally, **two-body density** models are **more accurate** but **more expensive** than **one-body density models**...

Objective: Build **new reduced-order models** to efficiently compute approximations $\widetilde{\tau}_{\boldsymbol{R}}$ of $\tau_{\boldsymbol{R}}$ for a large number of values of $\boldsymbol{R} \in \mathcal{R}$, out of the knowledge of $\rho_{\boldsymbol{R}}$ (computed from a one-body density model).

$$\text{Compute } \rho_{\boldsymbol{R}} \text{ out of a one-body model} \Rightarrow \text{Compute } \widetilde{\tau}_{\boldsymbol{R}} \approx \tau_{\boldsymbol{R}}$$

Important property 1: Link between one-body and two-body densities

$$\int_{\mathbb{R}^3} \tau_{\boldsymbol{R}}(x,y)\,dy = \rho_{\boldsymbol{R}}(x), \quad \int_{\mathbb{R}^3} \widetilde{\tau}_{\boldsymbol{R}}(x,y)\,dy = \rho_{\boldsymbol{R}}(x)$$

$$\Rightarrow \text{Marginal constraint}$$

Important property 2: Invariance by translation If $\boldsymbol{c} \in \mathbb{R}^3$ is a translation vector, it must hold that

$$\rho_{\boldsymbol{R}+\boldsymbol{c}} = \rho_{\boldsymbol{R}}(\cdot + \boldsymbol{c}), \quad \tau_{\boldsymbol{R}+\boldsymbol{c}} = \tau_{\boldsymbol{R}}(\cdot + \boldsymbol{c}, \cdot + \boldsymbol{c}), \quad \widetilde{\tau}_{\boldsymbol{R}+\boldsymbol{c}} = \widetilde{\tau}_{\boldsymbol{R}}(\cdot + \boldsymbol{c}, \cdot + \boldsymbol{c})$$

$$\Rightarrow \text{Optimal transport}$$

Optimal transport for model order reduction

Non-exhaustive list...

- [Iollo, Lombardi, 2014]: transport maps computed as linear combinations of optimal transport maps
- [VE, Lombardi, Mula, Vialard, 2020]: use of Wasserstein barycenters (1d) with greedy algorithms
- [Iollo, Taddei, 2022]: use of Gaussian models
- [Do, Feydy, Mula,2023]: extension of the Wasserstein barycenter methodology to higher-dimensional settings
- [Rim, Peherstorfer, Mandli, 2023]: towards a Galerkin approach combined with optimal transport

Wasserstein metric

- $n \in \mathbb{N}^\star$
- $\mathcal{P}_2(\mathbb{R}^n)$: set of probability measures on $\mathbb{R}^n$ with finite second-order moments

Definition (2-Wasserstein distance)

For $\rho_0, \rho_1 \in \mathcal{P}_2(\mathbb{R}^n)$,

$$W_2(\rho_0, \rho_1) := \inf_{\pi \in \Pi(\rho_0, \rho_1)} \left(\int_{\mathbb{R}^n \times \mathbb{R}^n} \|x - y\|^2 \, d\pi(x, y) \right)^{1/2}, \tag{1}$$

where

- $\| \cdot \|$ denotes the euclidean norm of $\mathbb{R}^n$
- $\Pi(\rho_0, \rho_1)$: set of probability measures π over $\mathbb{R}^n \times \mathbb{R}^n$ with marginals ρ_0 and ρ_1

$$d\rho_0(x) = \int_{y \in \mathbb{R}^n} d\pi(x, y) \qquad d\rho_1(y) = \int_{x \in \mathbb{R}^n} d\pi(x, y)$$

(also called the set of transport plans between ρ_0 and ρ_1).

In the following, we will write

$$\rho_0 = \mathrm{marg}_x(\pi) \quad \text{and} \quad \rho_1 = \mathrm{marg}_y(\pi)$$

Wasserstein barycenter

For $M \in \mathbb{N}^\star$,

$$\Lambda_M = \left\{ (\lambda_1, \dots, \lambda_M) \in (\mathbb{R}_+)^M, \quad \sum_{m=1}^{M} \lambda_m = 1 \right\}$$

Definition (Wasserstein barycenter (Agueh, Carlier, 2010))

The Wasserstein barycenter of
a collection of M probability measures $\boldsymbol{\rho} := (\rho^1, \dots, \rho^M) \in \mathcal{P}_2(\mathbb{R}^n)^M$
associated to a set of barycentric weights $\boldsymbol{\lambda} := (\lambda_1, \dots, \lambda_M) \in \Lambda_M$ is defined as the unique solution to the problem

$$\inf_{\rho \in \mathcal{P}_2(\mathbb{R}^n)} \sum_{m=1}^{M} \lambda_m \, W_2(\rho, \rho^m)^2 \tag{2}$$

The unique minimizer of (2) is denoted by $\mathrm{Bar}^{\boldsymbol{\lambda}}_{W_2}(\boldsymbol{\rho})$.

Wasserstein barycenter

For $M \in \mathbb{N}^\star$,

$$\Lambda_M = \left\{ (\lambda_1, \dots, \lambda_M) \in (\mathbb{R}_+)^M, \quad \sum_{m=1}^M \lambda_m = 1 \right\}$$

Definition (Wasserstein barycenter (Agueh, Carlier, 2010))

The Wasserstein barycenter of
a collection of M probability measures $\boldsymbol{\rho} := (\rho^1, \dots, \rho^M) \in \mathcal{P}_2(\mathbb{R}^n)^M$
associated to a set of barycentric weights $\boldsymbol{\lambda} := (\lambda_1, \dots, \lambda_M) \in \Lambda_M$ is defined as the unique solution to the problem

$$\inf_{\rho \in \mathcal{P}_2(\mathbb{R}^n)} \sum_{m=1}^M \lambda_m \, W_2(\rho, \rho^m)^2 \tag{2}$$

The unique minimizer of (2) is denoted by $\mathrm{Bar}^{\boldsymbol{\lambda}}_{W_2}(\boldsymbol{\rho})$.

L^2 **barycenter:** Assume that $(\rho^1, \dots, \rho^M) \in L^2(\mathbb{R}^n)^M$, the L^2 barycenter is defined as the unique solution to

$$\inf_{\rho \in L^2(\mathbb{R}^n)} \sum_{m=1}^M \lambda_m \, \|\rho - \rho^m\|_{L^2}^2$$

which is given by

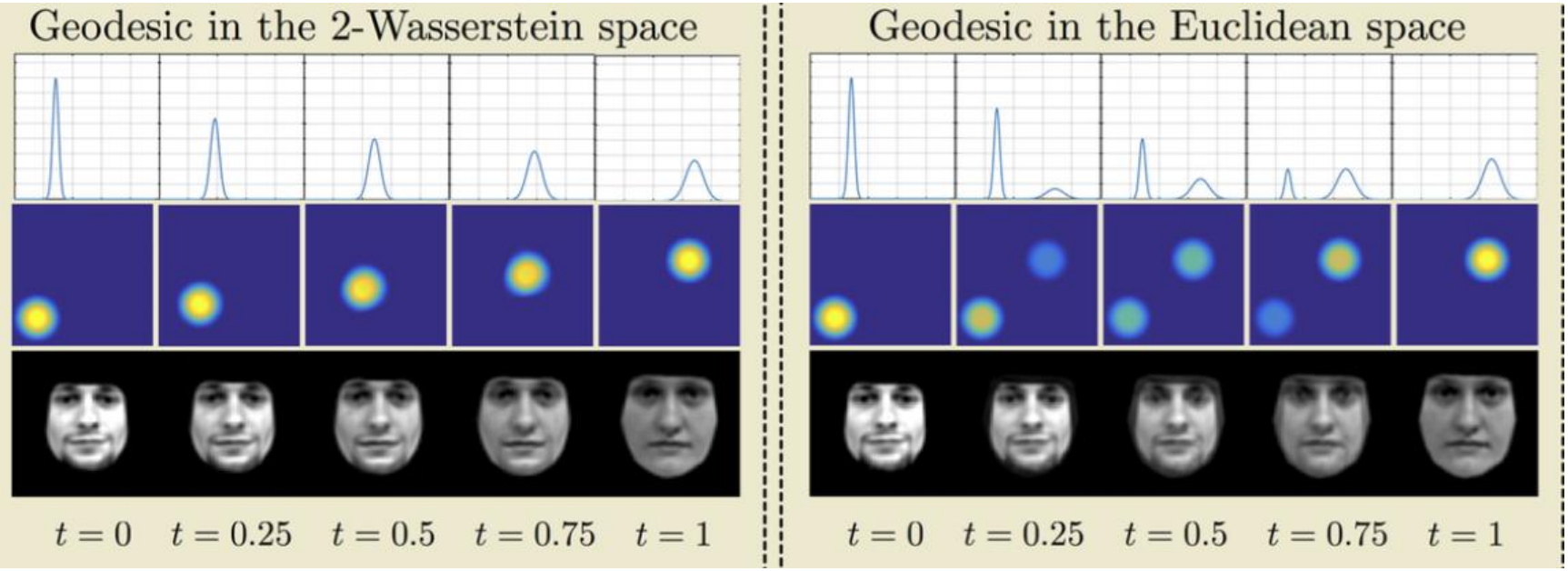
$$\mathrm{Bar}^{\boldsymbol{\lambda}}_{L^2}(\boldsymbol{\rho}) = \sum_{m=1}^M \lambda_m \rho_m$$

W_2 vs L^2 barycenter

[Kolouri et al. 2016]

$$M = 2, \qquad \boldsymbol{\rho} = (\rho_1, \rho_2), \qquad \boldsymbol{\lambda} = (t, 1-t) \text{ with } t \in [0,1]$$

$\mathrm{Bar}^{\boldsymbol{\lambda}}_{W_2}(\boldsymbol{\rho})$ $\mathrm{Bar}^{\boldsymbol{\lambda}}_{L^2}(\boldsymbol{\rho})$



Interesting property of the Wasserstein metric for transport-dominated problems:
If $\rho_2 = \rho_1(\cdot + \boldsymbol{c})$ for some $\boldsymbol{c} \in \mathbb{R}^n$, then it holds that

$$\mathrm{Bar}^{\boldsymbol{\lambda}}_{W_2}(\boldsymbol{\rho}) = \rho_1(\cdot + t\boldsymbol{c}), \qquad \mathrm{Bar}^{\boldsymbol{\lambda}}_{L^2}(\boldsymbol{\rho}) = (1-t)\rho_1 + t\rho_1(\cdot + \boldsymbol{c})$$

Model-order reduction for two-body electronic density models

Naturally, **two-body density** models are **more accurate** but **more expensive** than **one-body density models**...

Model-order reduction for two-body electronic density models

Naturally, **two-body density** models are **more accurate** but **more expensive** than **one-body density models**...

Objective: Build **new reduced-order models** to efficiently compute approximations $\widetilde{\tau}_{\boldsymbol{R}}$ of $\tau_{\boldsymbol{R}}$ for a large number of values of $\boldsymbol{R} \in \mathcal{R}$, out of the knowledge of $\rho_{\boldsymbol{R}}$ (computed from a one-body density model).

Model-order reduction for two-body electronic density models

Naturally, **two-body density** models are **more accurate** but **more expensive** than **one-body density models**...

Objective: Build **new reduced-order models** to efficiently compute approximations $\widetilde{\tau}_{\boldsymbol{R}}$ of $\tau_{\boldsymbol{R}}$ for a large number of values of $\boldsymbol{R} \in \mathcal{R}$, out of the knowledge of $\rho_{\boldsymbol{R}}$ (computed from a one-body density model).

$$\text{Compute } \rho_{\boldsymbol{R}} \text{ out of a one-body model} \Rightarrow \text{Compute } \widetilde{\tau}_{\boldsymbol{R}} \approx \tau_{\boldsymbol{R}}$$

Model-order reduction for two-body electronic density models

Naturally, **two-body density** models are **more accurate** but **more expensive** than **one-body density models**...

Objective: Build **new reduced-order models** to efficiently compute approximations $\widetilde{\tau}_{\boldsymbol{R}}$ of $\tau_{\boldsymbol{R}}$ for a large number of values of $\boldsymbol{R} \in \mathcal{R}$, out of the knowledge of $\rho_{\boldsymbol{R}}$ (computed from a one-body density model).

$$\text{Compute } \rho_{\boldsymbol{R}} \text{ out of a one-body model} \Rightarrow \text{Compute } \widetilde{\tau}_{\boldsymbol{R}} \approx \tau_{\boldsymbol{R}}$$

Important property 1: Link between one-body and two-body densities

$$\mathrm{marg}_x \tau_{\boldsymbol{R}} = \mathrm{marg}_y \tau_{\boldsymbol{R}} = \rho_{\boldsymbol{R}}, \quad \mathrm{marg}_x \widetilde{\tau}_{\boldsymbol{R}} = \mathrm{marg}_y \widetilde{\tau}_{\boldsymbol{R}} = \rho_{\boldsymbol{R}}$$

$$\Rightarrow \text{Marginal constraint}$$

Model-order reduction for two-body electronic density models

Naturally, **two-body density** models are **more accurate** but **more expensive** than **one-body density models**...

Objective: Build **new reduced-order models** to efficiently compute approximations $\widetilde{\tau}_{\boldsymbol{R}}$ of $\tau_{\boldsymbol{R}}$ for a large number of values of $\boldsymbol{R} \in \mathcal{R}$, out of the knowledge of $\rho_{\boldsymbol{R}}$ (computed from a one-body density model).

$$\text{Compute } \rho_{\boldsymbol{R}} \text{ out of a one-body model} \Rightarrow \text{Compute } \widetilde{\tau}_{\boldsymbol{R}} \approx \tau_{\boldsymbol{R}}$$

Important property 1: Link between one-body and two-body densities

$$\mathrm{marg}_x \tau_{\boldsymbol{R}} = \mathrm{marg}_y \tau_{\boldsymbol{R}} = \rho_{\boldsymbol{R}}, \quad \mathrm{marg}_x \widetilde{\tau}_{\boldsymbol{R}} = \mathrm{marg}_y \widetilde{\tau}_{\boldsymbol{R}} = \rho_{\boldsymbol{R}}$$

$$\Rightarrow \text{Marginal constraint}$$

Important property 2: Invariance by translation If $\boldsymbol{c} \in \mathbb{R}^3$ is a translation vector, it must hold that

$$\rho_{\boldsymbol{R}+\boldsymbol{c}} = \rho_{\boldsymbol{R}}(\cdot + \boldsymbol{c}), \quad \tau_{\boldsymbol{R}+\boldsymbol{c}} = \tau_{\boldsymbol{R}}(\cdot + \boldsymbol{c}, \cdot + \boldsymbol{c}), \quad \widetilde{\tau}_{\boldsymbol{R}+\boldsymbol{c}} = \widetilde{\tau}_{\boldsymbol{R}}(\cdot + \boldsymbol{c}, \cdot + \boldsymbol{c})$$

$$\Rightarrow \text{Optimal transport}$$

Principle for reduced-order model for electronic structure calculations

[Dalery, Dusson, VE, Lozinski, 2024]

- **Offline phase:**
  1. Choose values $\boldsymbol{R}^1, \ldots, \boldsymbol{R}^K \in \mathcal{R}$ and compute

$$\rho_{\boldsymbol{R}^k} \quad \text{and} \quad \tau_{\boldsymbol{R}^k}, \quad k = 1, \ldots, K \qquad \text{(snapshots)}$$

  2. Select $k_1, k_2, \cdots, k_M \in \{1, \ldots, K\}$ with $M \in \mathbb{N}^\star$ (small) out of a greedy algorithm so that, denoting by $\boldsymbol{\rho} = (\rho_{k_1}, \ldots, \rho_{k_M})$ and $\boldsymbol{\tau} = (\tau_{k_1}, \ldots, \tau_{k_M})$, we have

$$\forall k = 1, \ldots, K, \quad \tau_{\boldsymbol{R}^k} \approx \mathrm{Bar}^{\boldsymbol{\lambda}^k}_{W_2}(\boldsymbol{\tau}) \text{ for some } \boldsymbol{\lambda}^k \in \Lambda_M. \tag{3}$$

Principle for reduced-order model for electronic structure calculations

[Dalery, Dusson, VE, Lozinski, 2024]

- **Offline phase:**
  1. Choose values $\boldsymbol{R}^1, \ldots, \boldsymbol{R}^K \in \mathcal{R}$ and compute

$$\rho_{\boldsymbol{R}^k} \quad \text{and} \quad \tau_{\boldsymbol{R}^k}, \quad k = 1, \ldots, K \qquad \text{(snapshots)}$$

  2. Select $k_1, k_2, \cdots, k_M \in \{1, \ldots, K\}$ with $M \in \mathbb{N}^\star$ (small) out of a greedy algorithm so that, denoting by $\boldsymbol{\rho} = (\rho_{k_1}, \ldots, \rho_{k_M})$ and $\boldsymbol{\tau} = (\tau_{k_1}, \ldots, \tau_{k_M})$, we have

$$\forall k = 1, \ldots, K, \quad \tau_{\boldsymbol{R}^k} \approx \mathrm{Bar}^{\boldsymbol{\lambda}^k}_{W_2}(\boldsymbol{\tau}) \text{ for some } \boldsymbol{\lambda}^k \in \Lambda_M. \tag{3}$$

- **Online phase:** For a new value $\boldsymbol{R} \in \mathcal{R}$,
  1. Compute $\rho_{\boldsymbol{R}}$ (out of a one-body density model)
  2. Compute $\boldsymbol{\lambda}_{\boldsymbol{R}} \in \Lambda_M$ so that

$$\rho_{\boldsymbol{R}} \approx \mathrm{Bar}^{\boldsymbol{\lambda}_{\boldsymbol{R}}}_{W_2}(\boldsymbol{\rho}) \text{ for some } \boldsymbol{\lambda}_{\boldsymbol{R}} \in \Lambda_M.$$

  3. Compute $\widetilde{\tau}_{\boldsymbol{R}} \in \Lambda_M$ as

$$\widetilde{\tau}_{\boldsymbol{R}} = \mathrm{Bar}^{\boldsymbol{\lambda}_{\boldsymbol{R}}}_{W_2}(\boldsymbol{\tau}). \tag{4}$$

Principle for reduced-order model for electronic structure calculations

[Dalery, Dusson, VE, Lozinski, 2024]

- **Offline phase:**
  1. Choose values $\boldsymbol{R}^1, \ldots, \boldsymbol{R}^K \in \mathcal{R}$ and compute

$$\rho_{\boldsymbol{R}^k} \quad \text{and} \quad \tau_{\boldsymbol{R}^k}, \quad k = 1, \ldots, K \qquad \text{(snapshots)}$$

  2. Select $k_1, k_2, \cdots, k_M \in \{1, \ldots, K\}$ with $M \in \mathbb{N}^\star$ (small) out of a greedy algorithm so that, denoting by $\boldsymbol{\rho} = (\rho_{k_1}, \ldots, \rho_{k_M})$ and $\boldsymbol{\tau} = (\tau_{k_1}, \ldots, \tau_{k_M})$, we have

$$\forall k = 1, \ldots, K, \quad \tau_{\boldsymbol{R}^k} \approx \mathrm{Bar}^{\boldsymbol{\lambda}^k}_{W_2}(\boldsymbol{\tau}) \text{ for some } \boldsymbol{\lambda}^k \in \Lambda_M. \tag{3}$$

- **Online phase:** For a new value $\boldsymbol{R} \in \mathcal{R}$,
  1. Compute $\rho_{\boldsymbol{R}}$ (out of a one-body density model)
  2. Compute $\boldsymbol{\lambda}_{\boldsymbol{R}} \in \Lambda_M$ so that

$$\rho_{\boldsymbol{R}} \approx \mathrm{Bar}^{\boldsymbol{\lambda}_{\boldsymbol{R}}}_{W_2}(\boldsymbol{\rho}) \text{ for some } \boldsymbol{\lambda}_{\boldsymbol{R}} \in \Lambda_M.$$

  3. Compute $\widetilde{\tau}_{\boldsymbol{R}} \in \Lambda_M$ as

$$\widetilde{\tau}_{\boldsymbol{R}} = \mathrm{Bar}^{\boldsymbol{\lambda}_{\boldsymbol{R}}}_{W_2}(\boldsymbol{\tau}). \tag{4}$$

Two issues related to marginals in Steps (3) and (4)

Issues related to marginal constraints

(3): We would like that

$$\tau_{\boldsymbol{R}^k} \approx \mathrm{Bar}^{\boldsymbol{\lambda}^k}_{W_2}(\boldsymbol{\tau}) \Rightarrow \rho_{\boldsymbol{R}^k} \approx \mathrm{Bar}^{\boldsymbol{\lambda}^k}_{W_2}(\boldsymbol{\rho})$$

Not true in general!

Issues related to marginal constraints

(3): We would like that

$$\tau_{\boldsymbol{R}^k} \approx \mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}^k}(\boldsymbol{\tau}) \Rightarrow \rho_{\boldsymbol{R}^k} \approx \mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}^k}(\boldsymbol{\rho})$$

Not true in general!

(4): We would like that

$$\mathrm{marg}_x(\widetilde{\tau}_{\boldsymbol{R}}) = \mathrm{marg}_x(\mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}_{\boldsymbol{R}}}(\boldsymbol{\tau})) = \rho_{\boldsymbol{R}}$$

Also not true in general!

Issues related to marginal constraints

(3): We would like that

$$\tau_{\boldsymbol{R}^k} \approx \mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}^k}(\boldsymbol{\tau}) \Rightarrow \rho_{\boldsymbol{R}^k} \approx \mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}^k}(\boldsymbol{\rho})$$

Not true in general!

(4): We would like that

$$\mathrm{marg}_x(\widetilde{\tau}_{\boldsymbol{R}}) = \mathrm{marg}_x(\mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}_{\boldsymbol{R}}}(\boldsymbol{\tau})) = \rho_{\boldsymbol{R}}$$

Also not true in general!

Idea: Replace the exact W_2 barycenters $\mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}^k}(\boldsymbol{\tau})$ and $\mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}_{\boldsymbol{R}}}(\boldsymbol{\tau})$ by **marginal-preserving and marginal-constrained modified Wasserstein barycenters**

Issues related to marginal constraints

(3): We would like that

$$\tau_{\boldsymbol{R}^k} \approx \mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}^k}(\boldsymbol{\tau}) \Rightarrow \rho_{\boldsymbol{R}^k} \approx \mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}^k}(\boldsymbol{\rho})$$

Not true in general!

(4): We would like that

$$\mathrm{marg}_x(\widetilde{\tau}_{\boldsymbol{R}}) = \mathrm{marg}_x(\mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}_{\boldsymbol{R}}}(\boldsymbol{\tau})) = \rho_{\boldsymbol{R}}$$

Also not true in general!

Idea: Replace the exact W_2 barycenters $\mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}^k}(\boldsymbol{\tau})$ and $\mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}_{\boldsymbol{R}}}(\boldsymbol{\tau})$ by **marginal-preserving and marginal-constrained modified Wasserstein barycenters**

So far (and talk of today): marginal-preserving and marginal-constrained modified Wasserstein barycenters for **gaussian distributions and gaussian mixtures**

Issues related to marginal constraints

(3): We would like that

$$\tau_{\boldsymbol{R}^k} \approx \mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}^k}(\boldsymbol{\tau}) \Rightarrow \rho_{\boldsymbol{R}^k} \approx \mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}^k}(\boldsymbol{\rho})$$

Not true in general!

(4): We would like that

$$\mathrm{marg}_x(\widetilde{\tau}_{\boldsymbol{R}}) = \mathrm{marg}_x(\mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}_{\boldsymbol{R}}}(\boldsymbol{\tau})) = \rho_{\boldsymbol{R}}$$

Also not true in general!

Idea: Replace the exact W_2 barycenters $\mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}^k}(\boldsymbol{\tau})$ and $\mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}_{\boldsymbol{R}}}(\boldsymbol{\tau})$ by **marginal-preserving and marginal-constrained modified Wasserstein barycenters**

So far (and talk of today): marginal-preserving and marginal-constrained modified Wasserstein barycenters for **gaussian distributions and gaussian mixtures**

Extension to general distributions is work in progress!

Issues related to marginal constraints

(3): We would like that

$$\tau_{\boldsymbol{R}^k} \approx \mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}^k}(\boldsymbol{\tau}) \Rightarrow \rho_{\boldsymbol{R}^k} \approx \mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}^k}(\boldsymbol{\rho})$$

Not true in general!

(4): We would like that

$$\mathrm{marg}_x(\widetilde{\tau}_{\boldsymbol{R}}) = \mathrm{marg}_x(\mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}_{\boldsymbol{R}}}(\boldsymbol{\tau})) = \rho_{\boldsymbol{R}}$$

Also not true in general!

Idea: Replace the exact W_2 barycenters $\mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}^k}(\boldsymbol{\tau})$ and $\mathrm{Bar}_{W_2}^{\boldsymbol{\lambda}_{\boldsymbol{R}}}(\boldsymbol{\tau})$ by **marginal-preserving and marginal-constrained modified Wasserstein barycenters**

So far (and talk of today): marginal-preserving and marginal-constrained modified Wasserstein barycenters for **gaussian distributions and gaussian mixtures**

Extension to general distributions is work in progress!

Previous attempt in this direction: [Abraham, Abraham, Bergounioux, Carlier, 2017] marginal-constraint enforced using penalization

Outline of the talk

Optimal transport between Gaussian measures

- $n \in \mathbb{N}^\star$
- $\mathcal{S}^n_{+,\star}$: set of positive definite symmetric matrices of $\mathbb{R}^{n\times n}$
- For $\mu \in \mathbb{R}^n$ and $S \in \mathcal{S}^n_{+,\star}$, we denote by $\mathcal{N}(\mu, S)$ the gaussian probability measure with mean μ and covariance matrix S

Optimal transport between Gaussian measures

- $n \in \mathbb{N}^\star$
- $\mathcal{S}^n_{+,\star}$: set of positive definite symmetric matrices of $\mathbb{R}^{n\times n}$
- For $\mu \in \mathbb{R}^n$ and $S \in \mathcal{S}^n_{+,\star}$, we denote by $\mathcal{N}(\mu, S)$ the gaussian probability measure with mean μ and covariance matrix S

If $\rho_0 = \mathcal{N}(\mu_0, S_0)$ and $\rho_1 = \mathcal{N}(\mu_1, S_1)$, it holds that

$$W_2^2(\rho_0, \rho_1) = \|\mu_0 - \mu_1\|^2 + \mathcal{W}_2(S_0, S_1)^2$$

where $\mathcal{W}_2(S_0, S_1)$ is the **Bures-Wasserstein metric** between S_0 and S_1, defined as

$$\mathcal{W}_2(S_0, S_1)^2 = \mathrm{Tr}\left(S_0 + S_1 - 2\left(\sqrt{S_0}S_1\sqrt{S_0}\right)^{1/2}\right)$$

Wasserstein barycenters between Gaussian measures

For any $M \in \mathbb{N}^\star$, $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_M) \in \Lambda_M$ and $\boldsymbol{\rho} = (\rho_1, \dots, \rho_M) \in \mathcal{P}_2(\mathbb{R}^n)^M$ so that for all $i \in \{1, \dots, M\}$, $\rho_i = \mathcal{N}(\mu_i, S_i)$ for some $\mu_i \in \mathbb{R}^n$ and $S_i \in \mathcal{S}^n_{+,\star}$,

$$\mathrm{Bar}^{\boldsymbol{\lambda}}_{W_2}(\boldsymbol{\rho}) = \mathcal{N}(\mu_\star, S_\star)$$

where

$$\mu_\star = \sum_{m=1}^{M} \lambda_m \mu_m$$

and $S_\star \in \mathcal{S}^n_{+,\star}$ is the unique symmetric positive definite matrix solution to the following equation

$$\sum_{m=1}^{M} \lambda_m \left(\sqrt{S_\star} S_m \sqrt{S_\star}\right)^{1/2} = S_\star.$$

In the sequel, we will denote $S_\star$ by $\mathrm{Bar}^{\boldsymbol{\lambda}}_{\mathcal{W}_2}(\boldsymbol{S})$ where $\boldsymbol{S} := (S_1, \dots, S_M)$.

Marginals of Gaussian measures

Let us assume now that $n = n_x + n_y$ for some $n_x, n_y \in \mathbb{N}^\star$.

Marginals of Gaussian measures

Let us assume now that $n = n_x + n_y$ for some $n_x, n_y \in \mathbb{N}^\star$.

If $\rho = \mathcal{N}(\mu, S)$ with $\mu = (\mu_x, \mu_y)$ for some $\mu_x \in \mathbb{R}^{n_x}$, $\mu_y \in \mathbb{R}^{n_y}$, and $S \in \mathcal{S}_{+,\star}^{n_x+n_y}$, then we can write S as a block matrix

$$S = \begin{pmatrix} S_x & S_{xy} \\ S_{xy}^T & S_y \end{pmatrix}$$

with $S_x \in \mathbb{R}^{n_x \times n_x}$, $S_y \in \mathbb{R}^{n_y \times n_y}$ and $S_{xy} \in \mathbb{R}^{n_x \times n_y}$,

Marginals of Gaussian measures

Let us assume now that $n = n_x + n_y$ for some $n_x, n_y \in \mathbb{N}^\star$.

If $\rho = \mathcal{N}(\mu, S)$ with $\mu = (\mu_x, \mu_y)$ for some $\mu_x \in \mathbb{R}^{n_x}$, $\mu_y \in \mathbb{R}^{n_y}$, and $S \in \mathcal{S}_{+,\star}^{n_x+n_y}$, then we can write S as a block matrix

$$S = \begin{pmatrix} S_x & S_{xy} \\ S_{xy}^T & S_y \end{pmatrix}$$

with $S_x \in \mathbb{R}^{n_x \times n_x}$, $S_y \in \mathbb{R}^{n_y \times n_y}$ and $S_{xy} \in \mathbb{R}^{n_x \times n_y}$,

and it then holds that

$$\mathrm{marg}_x(\rho) = \mathcal{N}(\mu_x, S_x) \quad \text{and} \quad \mathrm{marg}_y(\rho) = \mathcal{N}(\mu_y, S_y)$$

Marginals of Gaussian measures

Let us assume now that $n = n_x + n_y$ for some $n_x, n_y \in \mathbb{N}^\star$.

If $\rho = \mathcal{N}(\mu, S)$ with $\mu = (\mu_x, \mu_y)$ for some $\mu_x \in \mathbb{R}^{n_x}$, $\mu_y \in \mathbb{R}^{n_y}$, and $S \in \mathcal{S}_{+,\star}^{n_x+n_y}$, then we can write S as a block matrix

$$S = \begin{pmatrix} S_x & S_{xy} \\ S_{xy}^T & S_y \end{pmatrix}$$

with $S_x \in \mathbb{R}^{n_x \times n_x}$, $S_y \in \mathbb{R}^{n_y \times n_y}$ and $S_{xy} \in \mathbb{R}^{n_x \times n_y}$,

and it then holds that

$$\mathrm{marg}_x(\rho) = \mathcal{N}(\mu_x, S_x) \quad \text{and} \quad \mathrm{marg}_y(\rho) = \mathcal{N}(\mu_y, S_y)$$

Question related to marginal constraint: Given a fixed Gaussian measure $\rho_{\mathrm{ref}} = \mathcal{N}(\nu, T)$, how can we find another Gaussian measure $\rho = \mathcal{N}(\mu, S)$ with prescribed marginals

$$\mathrm{marg}_x(\rho) = \mathcal{N}(\mu_x, S_x) \quad \text{and} \quad \mathrm{marg}_y(\rho) = \mathcal{N}(\mu_y, S_y)$$

so that $W_2(\rho_{\mathrm{ref}}, \rho)^2$ is as small as possible?

Marginals of Gaussian measures

Let us assume now that $n = n_x + n_y$ for some $n_x, n_y \in \mathbb{N}^\star$.

If $\rho = \mathcal{N}(\mu, S)$ with $\mu = (\mu_x, \mu_y)$ for some $\mu_x \in \mathbb{R}^{n_x}$, $\mu_y \in \mathbb{R}^{n_y}$, and $S \in \mathcal{S}_{+,\star}^{n_x+n_y}$, then we can write S as a block matrix

$$S = \begin{pmatrix} S_x & S_{xy} \\ S_{xy}^T & S_y \end{pmatrix}$$

with $S_x \in \mathbb{R}^{n_x \times n_x}$, $S_y \in \mathbb{R}^{n_y \times n_y}$ and $S_{xy} \in \mathbb{R}^{n_x \times n_y}$,

and it then holds that

$$\mathrm{marg}_x(\rho) = \mathcal{N}(\mu_x, S_x) \quad \text{and} \quad \mathrm{marg}_y(\rho) = \mathcal{N}(\mu_y, S_y)$$

Question related to marginal constraint: Given a fixed Gaussian measure $\rho_{\mathrm{ref}} = \mathcal{N}(\nu, T)$, how can we find another Gaussian measure $\rho = \mathcal{N}(\mu, S)$ with prescribed marginals

$$\mathrm{marg}_x(\rho) = \mathcal{N}(\mu_x, S_x) \quad \text{and} \quad \mathrm{marg}_y(\rho) = \mathcal{N}(\mu_y, S_y)$$

so that $W_2(\rho_{\mathrm{ref}}, \rho)^2$ is as small as possible?

Partial answer: Necessarily, $\mu = (\mu_x, \mu_y)$ and $S = \begin{pmatrix} S_x & S_{xy} \\ S_{xy}^T & S_y \end{pmatrix}$ for some $S_{xy} \in \mathbb{R}^{n_x \times n_y}$ so that $\mathcal{W}_2(T, S)^2$ is as small as possible.

Definition (Geometric mean of covariance matrices)

For $S, T \in \mathcal{S}^n_{+,\star}$, the geometric mean of S and T is given by

$$S \# T := S^{1/2} \left(S^{1/2} T^{-1} S^{1/2} \right)^{-1/2} S^{1/2}$$

Lemma (Bhatia, 2009)

It holds that

(i) $S \# T$ *is the unique matrix* $C \in \mathcal{S}^n_{+,\star}$ *solution to the equation* $C S^{-1} C = T$;

(ii) $S \# T = T \# S$;

(iii) $(S \# T)^{-1} = S^{-1} \# T^{-1}$.

Definition (Geometric mean of covariance matrices)

For $S, T \in \mathcal{S}^n_{+,\star}$, the geometric mean of S and T is given by

$$S \# T := S^{1/2} \left(S^{1/2} T^{-1} S^{1/2} \right)^{-1/2} S^{1/2}$$

Lemma (Bhatia, 2009)

It holds that

(i) $S \# T$ *is the unique matrix* $C \in \mathcal{S}^n_{+,\star}$ *solution to the equation* $CS^{-1}C = T$;

(ii) $S \# T = T \# S$;

(iii) $(S \# T)^{-1} = S^{-1} \# T^{-1}$.

Important remark:

$$S = \begin{pmatrix} S_x & S_{xy} \\ S_{xy}^T & S_y \end{pmatrix} \in \mathcal{S}^{n_x+n_y}_{+,\star} \iff S_x \in \mathcal{S}^{n_x}_{+,\star},\ S_y \in \mathcal{S}^{n_y}_{+,\star},\ \|S_x^{-1/2} S_{xy} S_y^{-1/2}\|_2 < 1$$

with $\|\cdot\|_2$ denoting the operator norm

Theorem (Dalery, Dusson, VE, 2025)

Let $n_x, n_y \in \mathbb{N}^*$ and let $T, S \in \mathcal{S}^{n_x+n_y}_{+,\star}$ with block decomposition

$$T = \begin{pmatrix} T_x & T_{xy} \\ T_{xy}^\intercal & T_y \end{pmatrix} \qquad \text{and} \qquad S = \begin{pmatrix} S_x & S_{xy} \\ S_{xy}^\intercal & S_y \end{pmatrix} \tag{5}$$

with $S_x, T_x \in \mathcal{S}^{n_x}_{+,*}$, $S_y, T_y \in \mathcal{S}^{n_y}_{+,*}$ and $S_{xy}, T_{xy} \in \mathbb{R}^{n_x \times n_y}$. Let

$$\forall Z \in \mathcal{C}_{S_x,S_y} := \left\{ Z \in \mathbb{R}^{n_x \times n_y},\ \|S_x^{-1/2} Z S_y^{-1/2}\|_2 < 1 \right\}, \qquad S(Z) := \begin{pmatrix} S_x & Z \\ Z^\intercal & S_y \end{pmatrix} \tag{6}$$

The function F defined as

$$F : \mathcal{C}_{S_x,S_y} \ni Z \longmapsto \mathcal{W}_2(T, S(Z))^2 = \mathrm{Tr}\left(T + S(Z) - 2\left(\sqrt{T} S(Z) \sqrt{T}\right)^{1/2} \right) \tag{7}$$

is strictly convex. Moreover, the minimization problem

$$Z^*_{T,S} \in \underset{Z \in \mathcal{C}_{S_x,S_y}}{\mathrm{argmin}}\ \mathcal{W}_2^2(T, S(Z)) \tag{8}$$

has a unique minimizer which is given by

$$Z^*_{T,S} = ({T_x}^{-1} \# S_x)\, T_{xy}\, ({T_y}^{-1} \# S_y) \tag{9}$$

Marginal-constrained modified Wasserstein barycenter

The matrix

$$S(Z^*_{T,S}) = \begin{pmatrix} S_x & Z^*_{T,S} \\ (Z^*_{T,S})^\intercal & S_y \end{pmatrix}$$

is then the closest matrix to T in the Bures-Wasserstein sense with prescribed marginals S_x and S_y.

Marginal-constrained modified Wasserstein barycenter

The matrix

$$S(Z^*_{T,S}) = \begin{pmatrix} S_x & Z^*_{T,S} \\ (Z^*_{T,S})^\intercal & S_y \end{pmatrix}$$

is then the closest matrix to T in the Bures-Wasserstein sense with prescribed marginals S_x and S_y.

Answer to the question related to marginal constraint: Given a fixed Gaussian measure $\rho_{\mathrm{ref}} = \mathcal{N}(\nu, T)$, how can we find another Gaussian measure $\rho = \mathcal{N}(\mu, S)$ with prescribed marginals

$$\mathrm{marg}_x(\rho) = \mathcal{N}(\mu_x, S_x) \quad \text{and} \quad \mathrm{marg}_y(\rho) = \mathcal{N}(\mu_y, S_y)$$

so that $W_2(\rho_{\mathrm{ref}}, \rho)^2$ is as small as possible?

Marginal-constrained modified Wasserstein barycenter

The matrix

$$S(Z^*_{T,S}) = \begin{pmatrix} S_x & Z^*_{T,S} \\ (Z^*_{T,S})^\intercal & S_y \end{pmatrix}$$

is then the closest matrix to T in the Bures-Wasserstein sense with prescribed marginals S_x and S_y.

Answer to the question related to marginal constraint: Given a fixed Gaussian measure $\rho_{\mathrm{ref}} = \mathcal{N}(\nu, T)$, how can we find another Gaussian measure $\rho = \mathcal{N}(\mu, S)$ with prescribed marginals

$$\mathrm{marg}_x(\rho) = \mathcal{N}(\mu_x, S_x) \quad \text{and} \quad \mathrm{marg}_y(\rho) = \mathcal{N}(\mu_y, S_y)$$

so that $W_2(\rho_{\mathrm{ref}}, \rho)^2$ is as small as possible?

Full answer: Necessarily, $\mu = (\mu_x, \mu_y)$ and $S = \begin{pmatrix} S_x & Z^*_{T,S} \\ (Z_{T,S})^*)^T & S_y \end{pmatrix}$

Marginal-constrained modified Wasserstein barycenter

The matrix

$$S(Z^*_{T,S}) = \begin{pmatrix} S_x & Z^*_{T,S} \\ (Z^*_{T,S})^\intercal & S_y \end{pmatrix}$$

is then the closest matrix to T in the Bures-Wasserstein sense with prescribed marginals S_x and S_y.

Answer to the question related to marginal constraint: Given a fixed Gaussian measure $\rho_{\mathrm{ref}} = \mathcal{N}(\nu, T)$, how can we find another Gaussian measure $\rho = \mathcal{N}(\mu, S)$ with prescribed marginals

$$\mathrm{marg}_x(\rho) = \mathcal{N}(\mu_x, S_x) \quad \text{and} \quad \mathrm{marg}_y(\rho) = \mathcal{N}(\mu_y, S_y)$$

so that $W_2(\rho_{\mathrm{ref}}, \rho)^2$ is as small as possible?

Full answer: Necessarily, $\mu = (\mu_x, \mu_y)$ and $S = \begin{pmatrix} S_x & Z^*_{T,S} \\ (Z_{T,S})^*)^T & S_y \end{pmatrix}$

Marginal-constrained modified Wasserstein barycenter: Apply the same formula with

$$\nu = \mu_\star \quad \text{and} \quad T = S_\star$$

Marginal-constrained modified Wasserstein barycenter

The matrix

$$S(Z^*_{T,S}) = \begin{pmatrix} S_x & Z^*_{T,S} \\ (Z^*_{T,S})^\intercal & S_y \end{pmatrix}$$

is then the closest matrix to T in the Bures-Wasserstein sense with prescribed marginals S_x and S_y.

Answer to the question related to marginal constraint: Given a fixed Gaussian measure $\rho_{\mathrm{ref}} = \mathcal{N}(\nu, T)$, how can we find another Gaussian measure $\rho = \mathcal{N}(\mu, S)$ with prescribed marginals

$$\mathrm{marg}_x(\rho) = \mathcal{N}(\mu_x, S_x) \quad \text{and} \quad \mathrm{marg}_y(\rho) = \mathcal{N}(\mu_y, S_y)$$

so that $W_2(\rho_{\mathrm{ref}}, \rho)^2$ is as small as possible?

Full answer: Necessarily, $\mu = (\mu_x, \mu_y)$ and $S = \begin{pmatrix} S_x & Z^*_{T,S} \\ (Z_{T,S})^*)^T & S_y \end{pmatrix}$

Marginal-constrained modified Wasserstein barycenter: Apply the same formula with

$$\nu = \mu_\star \quad \text{and} \quad T = S_\star$$

Marginal-preserving modified Wasserstein barycenter: Apply the same formula with

$$\nu = \mu_\star, \quad T = S_\star, \quad \mu_x = (\mu_x)_\star, \quad S_x = (S_x)_\star, \quad \mu_y = (\mu_y)_\star, \quad S_y = (S_y)_\star$$

Extension to Gaussian mixtures

The previous result enables to define marginal-constraint and marginal-preserving modified Wasserstein barycenters between an arbitrary number of Gaussian measures.

Extension to Gaussian mixtures

The previous result enables to define marginal-constraint and marginal-preserving modified Wasserstein barycenters between an arbitrary number of Gaussian measures.

This definition can be easily extended to define marginal-constraint and marginal-preserving modified Wasserstein barycenters between an arbitrary number of probability distributions between Gaussian mixtures, i.e. when

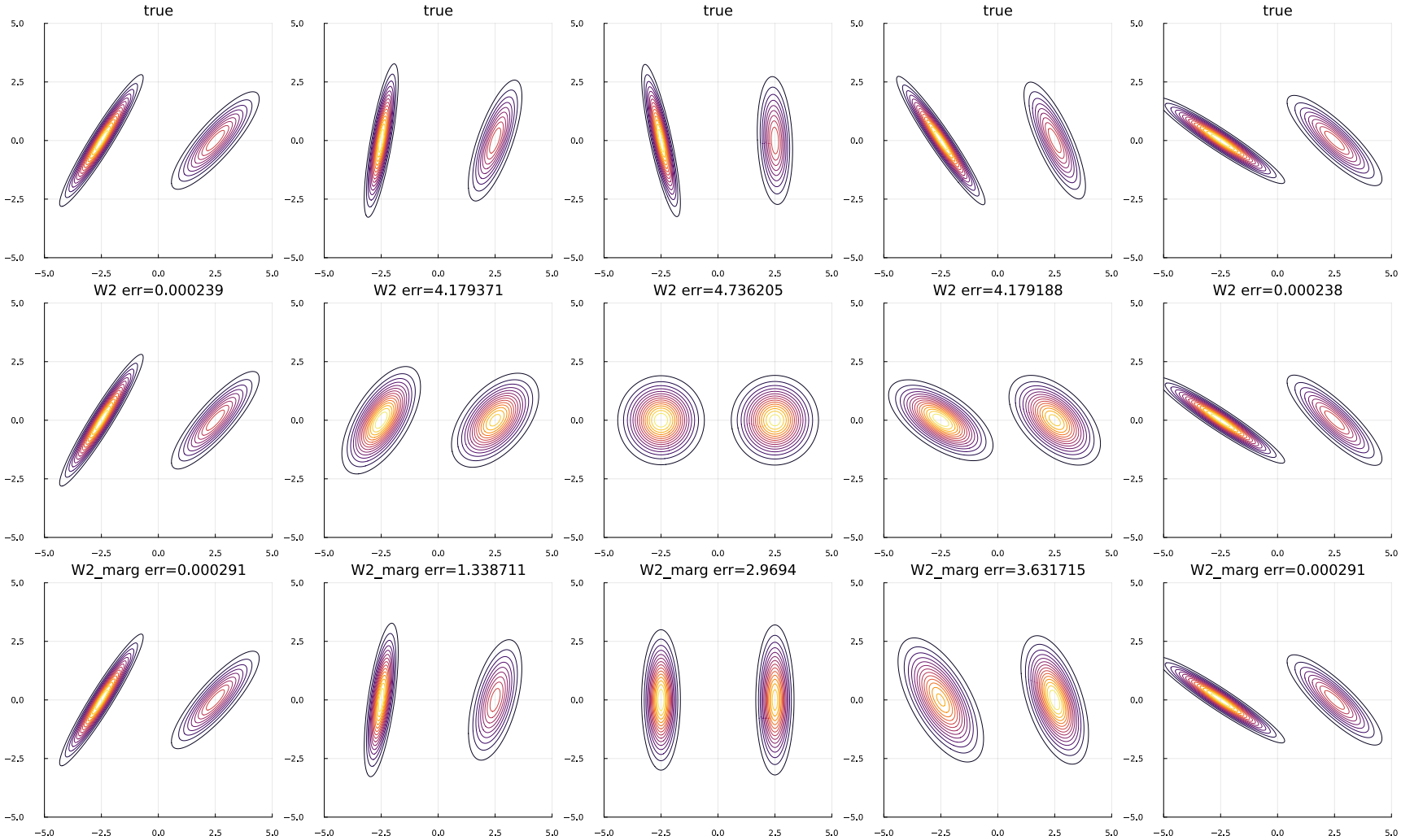
$$\forall 1 \leqslant m \leqslant M, \quad \rho_m = \sum_{k_m=1}^{K_m} \lambda_{k_m}^{(m)} \mathcal{N}(\mu_{k_m}^{(m)}, S_{k_m}^{(m)})$$

with $(\lambda_1^{(m)}, \ldots, \lambda_{K_m}^{(m)}) \in \Lambda_{K_m}$ for some $K_m \in \mathbb{N}^*$.

[Dalery, Dusson, VE, 2025]

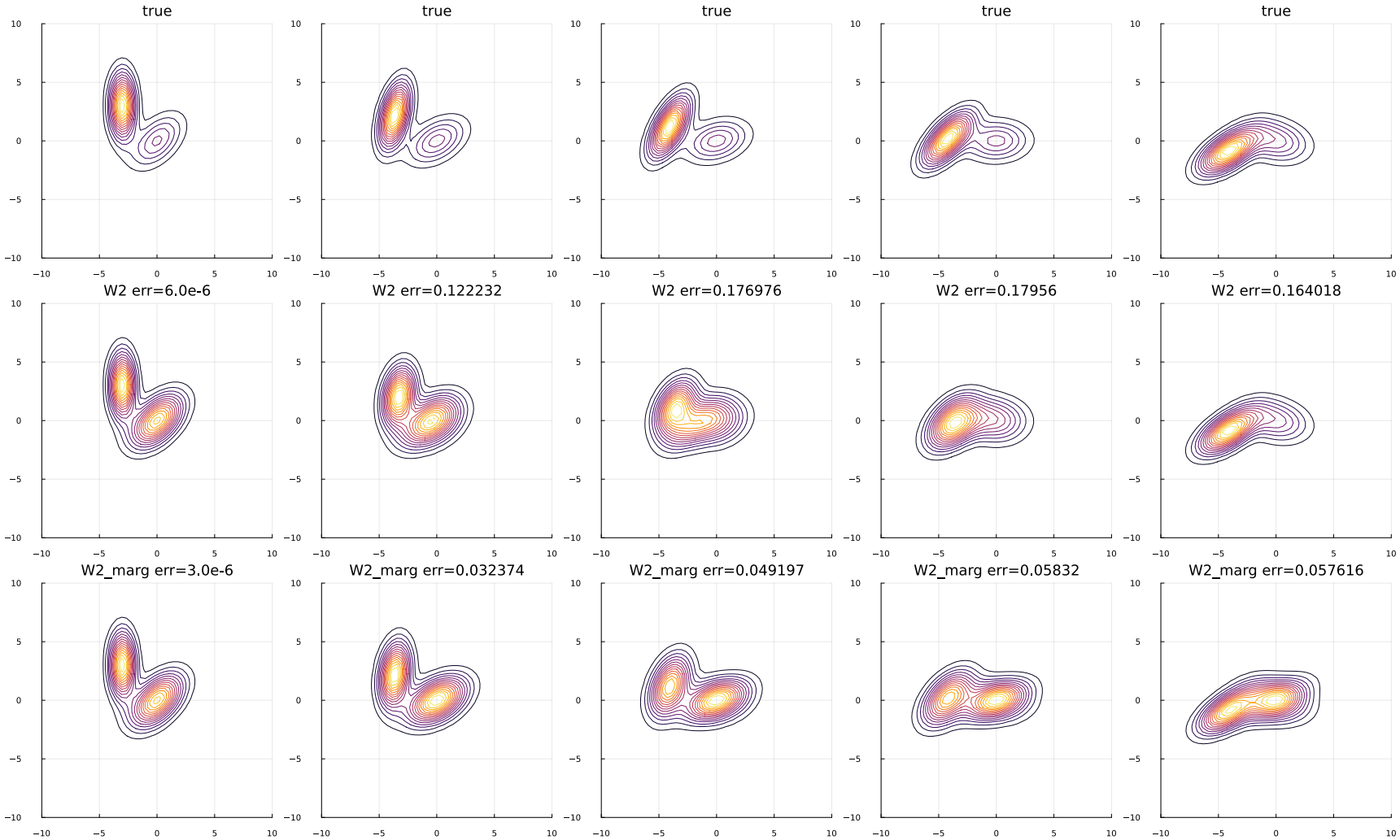
Outline of the talk

Toy gaussian model

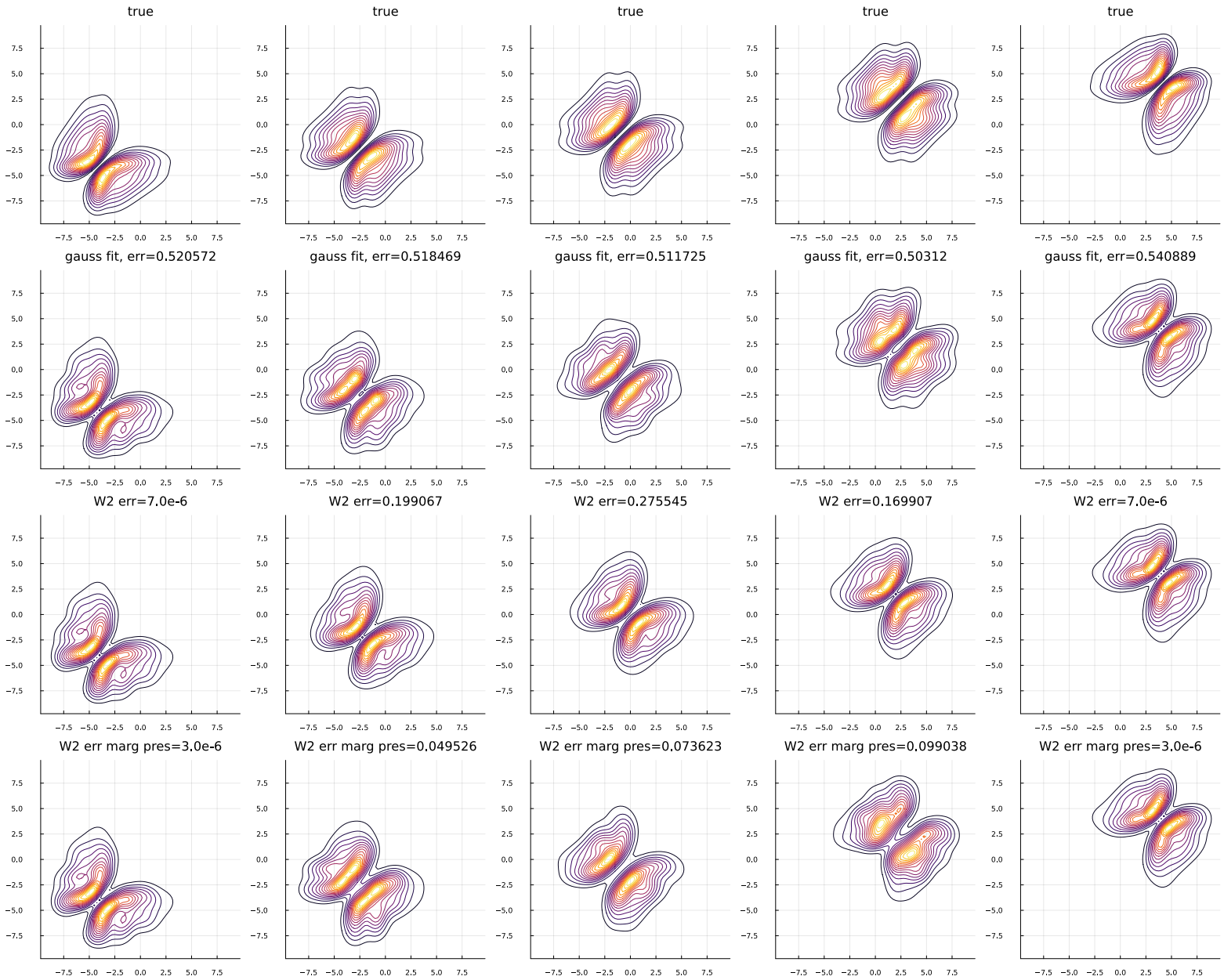


Fokker-Planck equation

$$\frac{\partial \rho}{\partial t} = -\nabla_{x,y} \cdot \left(A \begin{pmatrix} x \\ y \end{pmatrix} \rho \right) + D \Delta_{x,y} \psi \quad \text{avec } D > 0 \text{ et } A = \Omega \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$



Preliminary results for electronic structure calculations



Summary and perspectives

Summary:

- Definition of marginal-constrained (and marginal-preserving) modified Wasserstein barycenters between Gaussian measures and Gaussian mitures which can be analytically computed
- The definition heavily rely on new results concerning the geometric mean of covariance matrices
- Encouraging preliminary towards the design of new reduced-order models results for electronic structure calculations

Perspectives:

- Definition of marginal-constrained (and marginal-preserving) modified Wasserstein barycenters between Gaussian measures and Gaussian mitures which can be analytically computed
- Gaussian fit of electronic one or two-body densities lead to quite large errors!
  - Improve on gaussian fit algorithms
  - Extend the definition/computation of marginal-constrained modified Wasserstein barycenters to arbitrary measures

Thank you for your attention!