

A Global Optimization Approach for Multimarginal Optimal Transport Problems with Coulomb Cost

Yukuan Hu

State Key Laboratory of Scientific and Engineering Computing
Institute of Computational Mathematics and Scientific/Engineering Computing
Academy of Mathematics and Systems Science
Chinese Academy of Sciences, China

Joint works with **Huajie Chen** (BNU) and **Xin Liu** (AMSS)

Minisymposium on Eigenvalue Problems in Electronic Structure Calculations
ICIAM 2023, Waseda University, Tokyo, Japan
August 24, 2023



- Electronic minimization problem in wavefunction: **the curse of dimensionality.**
- Density functional theory (DFT).

[Hohenberg-Kohn 1964; Levy 1982; Lieb 1983; Zhou 2012/2019; ...]

$$E_0 = \inf_{\rho \in \mathcal{D}} \left\{ F_{HK}[\rho] + \int_{\mathbb{R}^d} v_{\text{ext}}(\mathbf{r}) \rho(\mathbf{r}) \, d\mathbf{r} \right\}.$$

- E_0 : ground state energy.
- $d \in \{1, 2, 3\}$: dimension of system.
- $N \in \mathbb{N}$: number of electrons.
- $\rho : \mathbb{R}^d \rightarrow \mathbb{R}$: single-particle density, with $\int \rho = N$.
- $\mathcal{D} \subseteq L^1(\mathbb{R}^d)$: feasible region.
- $v_{\text{ext}} : \mathbb{R}^d \rightarrow \mathbb{R}$: external potential.
- $F_{HK}[\rho]$: Hohenberg-Kohn functional (**kinetic energy + interaction**).

No tractable expression for $F_{HK}[\rho]!$



Approximation models for F_{HK}

- The Kohn-Sham (KS) model. [Kohn-Sham 1965]
the non-interaction limit of F_{HK} ; inexact if interaction dominates. [Cohen et al. 2008]

Numerical methods for the KS model:

[Saad et al. 2010; Lin et al. 2019]

[Roothaan 1951; McWeeny 1960; Pulay 1980; Schneider et al. 2009; Zhang et al. 2014; Dai et al. 2017;
Dai et al. 2019/2020; Gao et al. 2019; Bai et al. 2020/2022; Gao et al. 2022; ...]

Examples of strongly correlated systems:

[Cohen et al. 2008]

- transitional metal oxides; [Dagotto 2005]
- dissociating chemical bonds; [Grüning et al. 2003]
- low-density nanodevices. [Ghosal et al. 2007]
- ...

- The strictly correlated electrons (SCE) model¹. [Seidl et al. 1999; Seidl et al. 2000]
the strong-interaction limit of F_{HK} ; inexact if kinetic energy dominates.
- The KS-SCE model. [Seidl et al. 2007]
the combination of the KS and SCE models.

¹Other popular approaches for the strongly correlated systems include coupled cluster method [Bartlett-Musiał 2007], dynamical mean-field theory [Georges et al. 1996], density matrix renormalization group [White 1992], variational Monte Carlo [Pfau et al. 2020], etc.

The SCE Model as an Optimization Problem



Multimarginal optimal transport problem with Coulomb cost (MMOT)

[Buttazzo et al. 2012; Cotar et al. 2013]

$$\begin{aligned} V_{\text{ee}}^{\text{SCE}}[\rho] := \min_{\Pi \in \mathcal{P}(\Omega^N)} & \int_{\Omega^N} c_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) \, d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N), \\ \text{s. t. } & \int_{\Omega^{i-1} \times \Omega^{N-i}} d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_i, \dots, \mathbf{r}_N) = \frac{1}{N} \rho(\mathbf{r}_i), \\ & \forall \mathbf{r}_i \in \Omega, i = 1, \dots, N. \end{aligned}$$

- $\Omega \subseteq \mathbb{R}^d$: a bounded domain of interest.
- $c_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) := \sum_{i=1}^N \sum_{j>i} 1/|\mathbf{r}_i - \mathbf{r}_j|$: Coulomb cost function.
- $\mathcal{P}(\Omega^N)$: probability measure space over Ω^N .
- Π : probability measure over Ω^N .

$$\int_{\mathcal{A}_1 \times \dots \times \mathcal{A}_N} d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \text{Prob}\{\text{the } i\text{th electron lies in } \mathcal{A}_i, i = 1, \dots, N\}.$$

The SCE Model as an Optimization Problem (Cont.)



Multimarginal optimal transport problem with Coulomb cost (MMOT)

[Buttazzo et al. 2012; Cotar et al. 2013]

$$\begin{aligned} V_{\text{ee}}^{\text{SCE}}[\rho] := \min_{\Pi \in \mathcal{P}(\Omega^N)} & \int_{\Omega^N} c_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) \, d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N), \\ \text{s. t. } & \int_{\Omega^{i-1} \times \Omega^{N-i}} d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_i, \dots, \mathbf{r}_N) = \frac{1}{N} \rho(\mathbf{r}_i), \\ & \forall \mathbf{r}_i \in \Omega, \quad i = 1, \dots, N. \end{aligned}$$

Targets: to approximate

- the **SCE energy functional** $V_{\text{ee}}^{\text{SCE}}$;
- the **SCE potential** $u_{\text{SCE}} \sim \delta V_{\text{ee}}^{\text{SCE}} / \delta \rho$ (related to the optimal dual potentials of the MMOT [Chen et al. 2014; Di Marino-Gerolin 2020]).

Again, the curse of dimensionality!

Reformulations of the MMOT



Reformulations (tools)	SCE Pot. ²	Problem sizes	Weaknesses	References
<i>Kantorovich dual</i> (Gaussian mixture model + derivative-free optimization)	✓	$N \leq 6, K \leq 10$	high-dimensional inner problem & sensitivity to #Gaussians	[Mendl-Lin 2013]
<i>Monge-like ★</i> (linear programming solver)	✓	$N = 2, K \leq 4000$	trivial N	[Chen et al. 2014]
<i>Entropy regularization</i> (Sinkhorn algorithm)	✗	$N \leq 3, K \leq 1000$	numerical instability & high computational complexity	[Benamou et al. 2016]
<i>N-representability</i> (convex relaxation + semidefinite programming)	✗	$N \leq 13, K \leq 3500$	untight relaxation	[Khoo-Ying 2019] [Khoo et al. 2020]
<i>Extremal representation</i> (linear programming solver + genetic search-based column generation)	✓	$N \leq 30, K \leq 120$	validation only in the 1D context & heuristics	[Friescke et al. 2022]

K : number of grid points for discretization

²“SCE Pot.” stands for “SCE potential”.



The Monge-Like Ansatz

[Chen et al. 2014]

$$\begin{cases} d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \rho(\mathbf{r}_1)\gamma_2(\mathbf{r}_1, \mathbf{r}_2) \cdots \gamma_N(\mathbf{r}_1, \mathbf{r}_N)/N \, d\mathbf{r}_1 \cdots d\mathbf{r}_N, \\ \int_{\Omega} \gamma_i(\mathbf{r}_1, \mathbf{r}_i) \, d\mathbf{r}_i = 1, \quad \int_{\Omega} \rho(\mathbf{r}_1)\gamma_i(\mathbf{r}_1, \mathbf{r}_i) \, d\mathbf{r}_1 = \rho(\mathbf{r}_i), \quad \gamma_i(\mathbf{r}_1, \mathbf{r}_i) \geq 0, \quad \forall i. \end{cases}$$

- $\gamma_i : \Omega \times \Omega \rightarrow \mathbb{R}$: coupling between the positions of the first and i th electrons.

$$\int_{\mathcal{A}_1 \times \mathcal{A}_i} \rho(\mathbf{r}_1)\gamma_i(\mathbf{r}_1, \mathbf{r}_i) \, d\mathbf{r}_1 \, d\mathbf{r}_N = \text{Prob}\{\text{the first electron in } \mathcal{A}_1, \text{ the } i\text{th electron in } \mathcal{A}_i\}.$$

- Degeneration to the Monge ansatz³ when the couplings become mappings $\{\mathcal{T}_i\}_{i=2}^N$ from \mathbb{R}^d to \mathbb{R}^d . [Seidl et al. 1999; Seidl et al. 1999; Seidl et al. 2000]

$$\mathcal{T}_i \xleftrightarrow{\gamma_i(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r}' - \mathcal{T}_i(\mathbf{r}))} \gamma_i.$$

- Physically meaningful.
- Spectacular dimension reduction.
- $\{\gamma_i\}_{i=2}^N \rightarrow \{\mathcal{T}_i\}_{i=2}^N \xrightarrow{\text{PDE}} \textcolor{red}{usce}$. [Seidl et al. 2007]

³The Monge ansatz is valid when $d = 1$ or for special systems with spherically symmetric densities.

[Seidl et al. 2007; Colombo et al. 2015; Seidl et al. 2017; Bindini et al. 2020; ...]

Discretized MMOT under the Monge-Like Ansatz



$$\begin{aligned} \min_X \quad & f(X_2, \dots, X_N) := \sum_{i=2}^N \langle X_i, \Lambda ECE \rangle + \sum_{j>i} \langle X_i, E \Lambda X_j ECE \rangle, \\ \text{s. t.} \quad & X_i \mathbf{e} = \mathbf{1}, \quad X_i^\top E \boldsymbol{\varrho} = \boldsymbol{\varrho}, \quad \text{Tr}(X_i) = 0, \quad X_i \geq 0, \quad i = 2, \dots, N, \\ & \langle X_i, X_j \rangle = 0, \quad \forall i, j \in \{2, \dots, N\} \text{ with } i \neq j. \end{aligned} \tag{P}$$

- $\mathcal{T} := \{e_k\}_{k=1}^K$: non-overlapping finite elements mesh, $\Omega = \cup_{k=1}^K e_k$.
- $\mathbf{e} := [|e_1|, \dots, |e_N|]^\top \in \mathbb{R}_+^K$: volumes of elements; $\mathbf{1} \in \mathbb{R}_+^K$: all-ones vector.
- $\boldsymbol{\varrho} := [\varrho_1, \dots, \varrho_K]^\top \in \mathbb{R}_+^K$ with $\varrho_k := \int_{e_k} \rho(\mathbf{r}) \, d\mathbf{r} / |e_k|$ for any k .
- $E := \text{Diag}(\mathbf{e})$, $\Lambda := \text{Diag}(\boldsymbol{\varrho}) \in \mathbb{R}^{K \times K}$.
- $C := (c_{jk})_{jk}$, $X_i := (x_{i,jk})_{jk} \in \mathbb{R}^{K \times K}$ ($i = 2, \dots, N$), where

$$c_{jk} := \begin{cases} \frac{1}{|e_j| \cdot |e_k|} \int_{e_k} \int_{e_j} \frac{1}{|\mathbf{r} - \mathbf{r}'|} \, d\mathbf{r} \, d\mathbf{r}', & \text{if } j \neq k; \\ 0, & \text{otherwise,} \end{cases}$$

$$x_{i,jk} := \frac{1}{|e_j| \cdot |e_k|} \int_{e_k} \int_{e_j} \gamma_i(\mathbf{r}, \mathbf{r}') \, d\mathbf{r} \, d\mathbf{r}'.$$

- $X := (X_i)_{i=2}^N \in (\mathbb{R}^{K \times K})^{N-1}$.

Discretized MMOT under the Monge-Like Ansatz (Cont.)



$$\begin{aligned} \min_X \quad & f(X_2, \dots, X_N) := \sum_{i=2}^N \langle X_i, \Lambda ECE \rangle + \sum_{j>i} \langle X_i, E \Lambda X_j ECE \rangle, \\ \text{s. t.} \quad & X_i \mathbf{e} = \mathbf{1}, \quad X_i^\top E \boldsymbol{\varrho} = \boldsymbol{\varrho}, \quad \text{Tr}(X_i) = 0, \quad \textcolor{blue}{X_i \geq 0}, \quad i = 2, \dots, N, \\ & \langle X_i, X_j \rangle = 0, \quad \forall i, j \in \{2, \dots, N\} \text{ with } i \neq j. \end{aligned} \tag{P}$$

ℓ_1 penalized optimization problem

$$\begin{aligned} \min_X \quad & f_\beta(X_2, \dots, X_N) := f(X_2, \dots, X_N) + \beta \sum_{j>i} \langle X_i, X_j \rangle, \\ \text{s. t.} \quad & X_i \in \mathcal{S} := \{W \in \mathbb{R}_+^{K \times K} : \mathcal{B}(W) = \mathbf{b}\}, \quad i = 2, \dots, N. \end{aligned} \tag{P}_\beta$$

- $\beta > 0$: penalty parameter.
- $\mathcal{B}(W) := [\mathbf{e}^\top W^\top, \boldsymbol{\varrho}^\top E W, \text{Tr}(W)]^\top \in \mathbb{R}^{2K+1}$ for any $W \in \mathbb{R}^{K \times K}$.
- $\mathbf{b} := [\mathbf{1}^\top, \boldsymbol{\varrho}^\top, 0]^\top \in \mathbb{R}^{2K+1}$.
- Equivalence to using *effective cost functions*. [Giuliani-Vignale 2005]
- **Exactness** of the ℓ_1 penalty function. [H.-Liu 2023]
- Nonconvex quadratic program, **NP-hard**. [Pardalos-Vavasis 1991]

Our Route and Contributions



Route: $F_{\text{HK}}[\rho] \xrightarrow{\text{SCE}} \text{MMOT} \xrightarrow[\tau]{\text{ansatz}} \text{MPGCC } (\mathbf{P}) \xrightarrow{\ell_1 \text{ penalty}} \text{problem } (\mathbf{P}_\beta).$

Contributions

- (i) A **global** optimization approach for the problem (\mathbf{P}_β) with an efficient local solver.
- (ii) Numerical simulations with the **first visualization** of $\{\mathcal{T}_i\}_{i=2}^N$ in 2D contexts.

A Global Optimization Approach—GGR

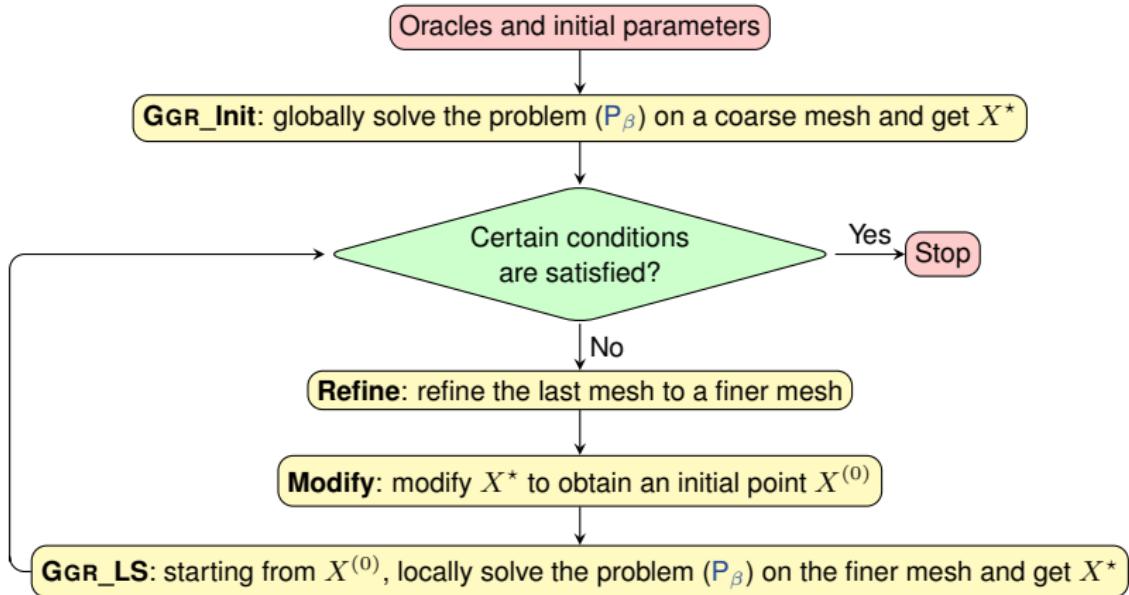


Figure 1: The GGR flowchart.

“G”: global optimization; “GR”: grid refinements.



Framework 1: GGR

- Input:** Oracle returning C , e , ϱ in proper dimensions; global solver; initialization subroutine; local solver; initial mesh $\mathcal{T}^{(0)}$ with $K^{(0)}$ elements; initial penalty parameter $\beta^{(0)} > 0$.
- 1 Set $\ell := 0$.
 - 2 **GGR_Init**: the **global solver** solves the problem (P_β) with size $K^{(0)}$ to get $X^{(0,\star)}$.
 - 3 **while** certain criteria not satisfied **do**
 - 4 Refine the last mesh $\mathcal{T}^{(\ell)}$ to $\mathcal{T}^{(\ell+1)}$ with $K^{(\ell+1)}$ elements.
 - 5 Modify $X^{(\ell,\star)}$ using the **initialization subroutine** to obtain $X^{(\ell+1,0)}$.
 - 6 Update $\beta^{(\ell)}$ to $\beta^{(\ell+1)} > 0$ if necessary.
 - 7 **GGR_LS** $(\ell + 1)$: the **local solver** starts from $X^{(\ell+1,0)}$ and solves the problem (P_β) with size $K^{(\ell+1)}$ to get $X^{(\ell+1,\star)}$.
 - 8 Set $\ell := \ell + 1$.
 - 9 **end**

Output: $X^{(\ell)} \in \left(\mathbb{R}^{K^{(\ell)} \times K^{(\ell)}}\right)^{N-1}$.

Initialization Illustration (1D)

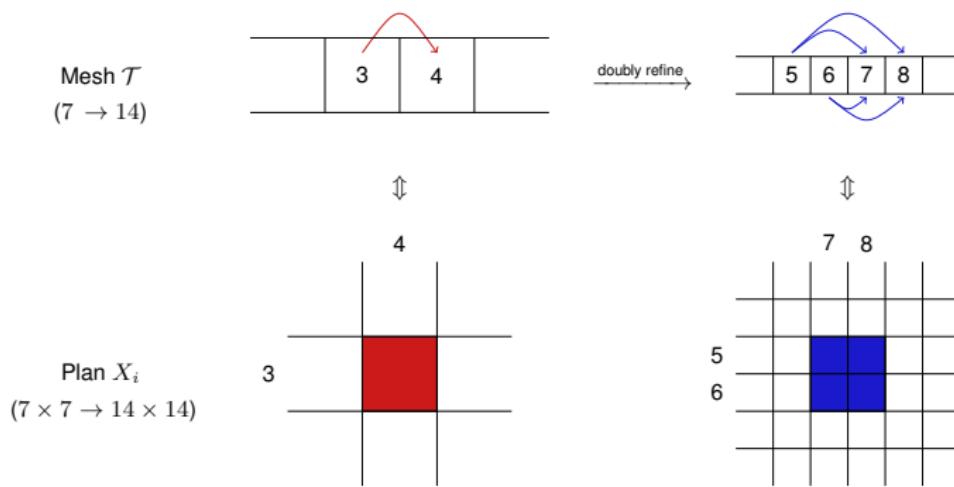


Figure 2: Initialization in 1D context (from $K^{(\ell)} = 7$ to $K^{(\ell+1)} = 14$).

Initialization Illustration (2D)

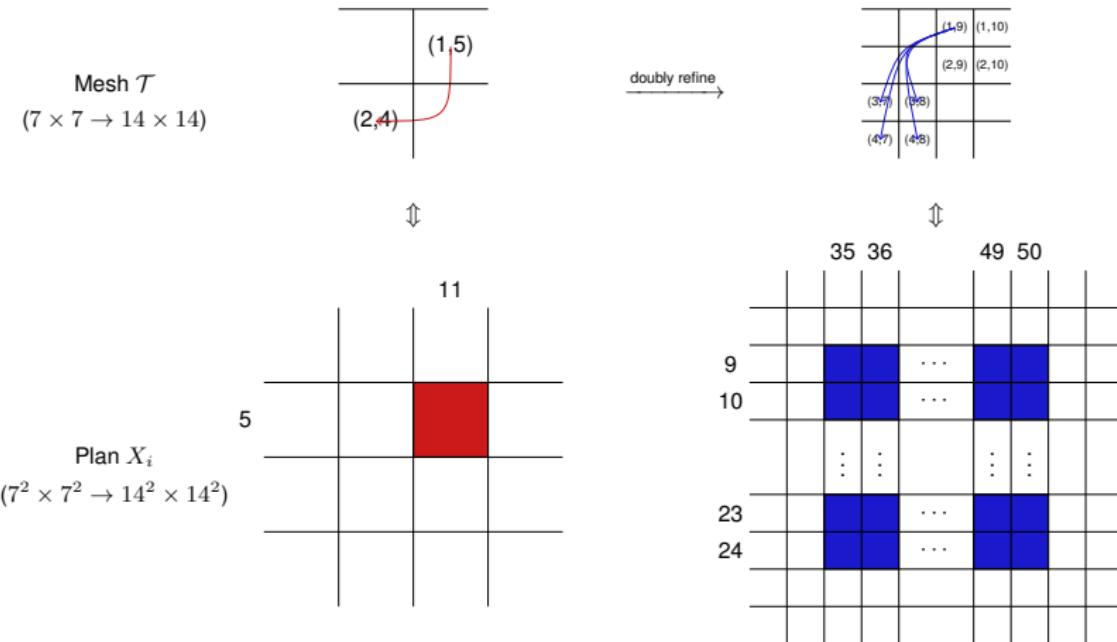


Figure 3: Initialization in 2D context (from $K^{(\ell)} = 7 \times 7$ to $K^{(\ell+1)} = 14 \times 14$).

Initialization Subroutine Based on GR



Algorithm 2: GR for initialization.

Input: Coarse mesh $\mathcal{T}^{(\ell)}$ with $K^{(\ell)}$ elements and refined mesh $\mathcal{T}^{(\ell+1)}$ with $K^{(\ell+1)}$ elements; solution of the previous step $X^{(\ell,*)}$; scaling factor $r > 0$.

```
1 for i = 2, ..., N do
2   for j = 1, ..., K do
3     for k = 1, ..., K do
4       if  $x_{i,jk}^{(\ell,*)} > 0$  then
5         Find  $e_{j_t}^{(\ell+1)}$ ,  $t = 1, \dots, s_j$  such that  $e_j^{(\ell)} = \cup_{t=1}^{s_j} e_{j_t}^{(\ell+1)}$ .
6         Find  $e_{k_u}^{(\ell+1)}$ ,  $u = 1, \dots, s_k$  such that  $e_k^{(\ell)} = \cup_{u=1}^{s_k} e_{k_u}^{(\ell+1)}$ .
7         Set  $x_{i,j_t k_u}^{(\ell+1,0)} = r \cdot x_{i,jk}^{(\ell,*)}$  for  $t \in \{1, \dots, s_j\}$ ,  $u \in \{1, \dots, s_k\}$ .
8       end
9     end
10   end
11 end
```

Output: $X^{(\ell+1,0)} := (X_i^{(\ell+1,0)})_{i=2}^N \in (\mathbb{R}^{K^{(\ell+1)} \times K^{(\ell+1)}})^{N-1}$.



The **block structure** of the problem (\mathbf{P}_β).

(omitting the outer loop superscript)

$$\min_X f_\beta(X_2, \dots, X_N), \text{ s. t. } X_i \in \mathcal{S}, i = 2, \dots, N.$$

⇒ Proximal alternating linearized minimization method (**PALM**). [Bolte et al. 2014]

Algorithm 3: PALM for solving the problem (\mathbf{P}_β).

Input: $X^{(0)} := (X_i^{(0)})_{i=2}^N \in (\mathbb{R}^{K \times K})^{N-1}, \{\sigma_i^{(0)} > 0\}_{i=2}^N, \beta > 0$.

```

1 Set  $k := 0$ .
2 while certain conditions not satisfied do
3   for  $i = 2, \dots, N$  do
4     Invoke a subsolver to solve the  $i$ th proximal linearized subproblem
        
$$\min_{X_i \in \mathcal{S}} \left\langle \nabla_{X_i} f_\beta(X_{<i}^{(k+1)}, X_{\geq i}^{(k)}), X_i - X_i^{(k)} \right\rangle + \frac{\sigma_i^{(k)}}{2} \|X_i - X_i^{(k)}\|_F^2 \quad (1)$$

        to obtain  $X_i^{(k+1)} \in \mathbb{R}^{K \times K}$  fulfilling certain inexact criterion.
        Update the  $i$ th proximal parameter  $\sigma_i^{(k)}$  to  $\sigma_i^{(k+1)} > 0$  if necessary.
5   end
6   Set  $k := k + 1$ .
7 end
```

Output: $X^{(k)} = (X_i^{(k)})_{i=2}^N \in (\mathbb{R}^{K \times K})^{N-1}$.



Subsolver

Semismooth Newton-CG method (**SSNCG**) for the dual subproblem. [Li et al. 2020]

Inexact criterion

Given a nonnegative sequence $\{\varepsilon^{(k)}\}$, stop SSNCG if

$$\max \left\{ - \left\langle \boldsymbol{\lambda}_i^{(k+1)}, \mathcal{B}(X_i^{(k+1)}) - \mathbf{b} \right\rangle, 0 \right\} + \|\mathcal{B}(X_i^{(k+1)}) - \mathbf{b}\|_{\infty} \leq \varepsilon^{(k)},$$

where $\boldsymbol{\lambda}_i^{(k+1)} := [\boldsymbol{\lambda}_{i,\mathbf{1}}^{(k+1)\top}, \boldsymbol{\lambda}_{i,\varrho}^{(k+1)\top}, \mu^{(k+1)}]^\top \in \mathbb{R}^{2K+1}$ is an approximate dual solution given by SSNCG, $\boldsymbol{\lambda}_{i,\mathbf{1}}^{(k+1)}, \boldsymbol{\lambda}_{i,\varrho}^{(k+1)} \in \mathbb{R}^K, \mu^{(k+1)} \in \mathbb{R}$.



Let $L = \frac{N(N-1)}{2} (\|\boldsymbol{\varrho}\|_\infty \|\mathbf{e}\|_\infty^3 \|C\|_2 + \beta)$.

Theorem 1 (Convergence under iterate infeasibility [H.-Liu 2023])

Let $\{X^{(k)}\}$ be the sequence generated by PALM with a nonnegative $\{\varepsilon^{(k)}\}$ and $\{\sigma_i^{(k)}\}_{i=2}^N \subseteq (L, \gamma L]$, where $\gamma > 1$.

- (i) If $\{\varepsilon^{(k)}\}$ is square summable, then $\{X^{(k)}\}$ has at least a limit point and any limit point of $\{X^{(k)}\}$ is a KKT point of the problem (P_β) .
- (ii) If $\{\varepsilon^{(k)}\}$ and $\{k(\varepsilon^{(k)})^{2p}\}$ are summable for some $p \in (0, 1)$, then $\{X^{(k)}\}$ converges to a KKT point of the problem (P_β) .
- (iii) If $X^{(0)}$ is feasible and sufficiently close to some optimal solution X^* of the problem (P_β) , both $\sum_k \varepsilon^{(k)}$ and $\sum_k (\varepsilon^{(k)})^{2p}$ are sufficiently small for some $p \in (0, 1)$, then $\{X^{(k)}\}$ converges to an optimal solution of the problem (P_β) .



Global solvers

- 1D: random multi-start. [Hickernell-Yuan 1997]
- 2D: optimization software BARON. [Fourer et al. 1990; Sahinidis 2021]

Parameters

- GR: $r = 1$.
- PALM: $\sigma_i^{(k)} \equiv \sigma = 10^{-4}$.
- Penalty parameter: $\beta \sim \mathcal{O}(1/K)$.

Table 1: The value of β for different K .

K	(0, 10)	[10, 36)	[36, 80)	[80, 160)	[160, 320)
β	2^2	2^1	2^0	2^{-2}	2^{-3}
K	[320, 640)	[640, 1280)	[1280, 2560)	[2560, 5120)	[5120, ∞)
β	2^{-4}	2^{-5}	2^{-6}	2^{-7}	2^{-8}



Initialization

- SSNCG: all-zero initialization \rightsquigarrow **warm start**.
- PALM: initial points provided by GR.

Stopping criteria

- SSNCG: $\varepsilon^{(k)} \equiv 10^{-9}$.
- PALM: $\sqrt{\sigma} \|X^{(\ell,k+1)} - X^{(\ell,k)}\|_F \leq \varepsilon_{\text{outer}}$.

$$\varepsilon_{\text{outer}} = \begin{cases} 10^{-8}, & K^{(\ell+1)} \in (0, 200]; \\ 10^{-6}, & K^{(\ell+1)} \in (200, 2000]; \\ 10^{-5}, & K^{(\ell+1)} \in (2000, 10000]; \\ 10^{-4}, & K^{(\ell+1)} \in (10000, \infty). \end{cases}$$

Running environment

Intel Xeon Gold 6242R CPU @ 3.10GHz with 510GB RAM under Ubuntu 20.04.



- Energy ≥ 0 : **repulsive energy** f .
- $\{\hat{\mathcal{T}}_i\}_{i=2}^N : \mathbb{R}^d \rightarrow \mathbb{R}^d$: **approximated mappings** between electrons.

$$\hat{\mathcal{T}}_i(\mathbf{a}_j) := \sum_{1 \leq k \leq K} \mathbf{a}_k x_{i,jk} |e_k|, \quad j = 1, \dots, K, \quad i = 2, \dots, N,$$

where $\{\mathbf{a}_k\}_{k=1}^K$ are the barycenters of the elements $\{e_k\}_{k=1}^K$.

- $\hat{\lambda}_{\varrho} \in \mathbb{R}^K$: **approximated SCE potential**.

$$\text{Subsolver} \rightarrow \{\lambda_{i,\varrho}\}_{i=2}^N \rightarrow \lambda_{\varrho} := \frac{1}{N-1} \sum_{i=2}^N \lambda_{i,\varrho} \rightarrow \hat{\lambda}_{\varrho} := \lambda_{\varrho} - \min_{k=1}^K \{\lambda_{\varrho,k}\} \cdot \mathbf{1}.$$

- $\text{err} \geq 0$: **average error** from the given optimal mappings $\{\mathcal{T}_i\}_{i=2}^N$.

$$\text{err}(K, \Omega) := \frac{1}{K |\Omega|} \sum_{k=1}^K \sum_{i=2}^N |\mathcal{T}_i(\mathbf{a}_k) - \hat{\mathcal{T}}_i(\mathbf{a}_k)|.$$



Problem settings

$$\rho_1(x) = c_1(\cos(\pi x) + 1), \quad \Omega = [-1, 1];$$

$$\rho_2(x) = c_2(2e^{-6(x+0.5)^2} + 1.5e^{-4(x-0.5)^2}), \quad \Omega = [-1.5, 1.5];$$

$$\rho_3(x) = c_3 e^{-|x|}, \quad \Omega = [-5, 5].$$

- $c_i, i = 1, 2, 3$: normalizing factors such that $\int_{\Omega} \rho_i(x) dx = 3$.
- Initial equimass discretization ($K^{(0)} = 12$) + uniform mesh refinements.

Numerical Results on 1D Systems ($N = 3$) (Cont.)



Table 2: Output energies and calculated average errors of GGR
($d = 1, N = 3$).

Step	System 1				System 2				System 3			
	K	Energy	err_s	err_e	K	Energy	err_s	err_e	K	Energy	err_s	err_e
GGR_Init	12	18.114	-	0.031	12	10.695	-	0.034	12	5.935	-	0.040
GGR_LS(1)	24	18.911	0.049	0.013	24	11.301	0.053	0.016	24	6.275	0.053	0.018
GGR_LS(2)	48	19.004	0.022	0.009	48	11.362	0.026	0.011	48	6.346	0.027	0.013
GGR_LS(3)	96	19.019	0.014	0.004	96	11.370	0.016	0.007	96	6.356	0.019	0.012
GGR_LS(4)	192	19.021	0.007	0.003	192	11.372	0.011	0.004	192	6.360	0.013	0.001
GGR_LS(5)	384	19.022	0.007	0.002	384	11.373	0.006	0.002	384	6.361	0.003	0.000
GGR_LS(6)	768	19.022	0.004	0.001	768	11.373	0.003	0.000	768	6.361	0.001	0.000

- err_s: average error of the initial point.
- err_e: average error of the solution.
- $\text{err}_e \downarrow$ as $K \uparrow \Rightarrow$ GGR and *theoretical predictions.* ✓
- $\text{err}_s \downarrow$ as $K \uparrow \Rightarrow$ GR. ✓
- $\text{err}_e < \text{err}_s \Rightarrow$ PALM. ✓

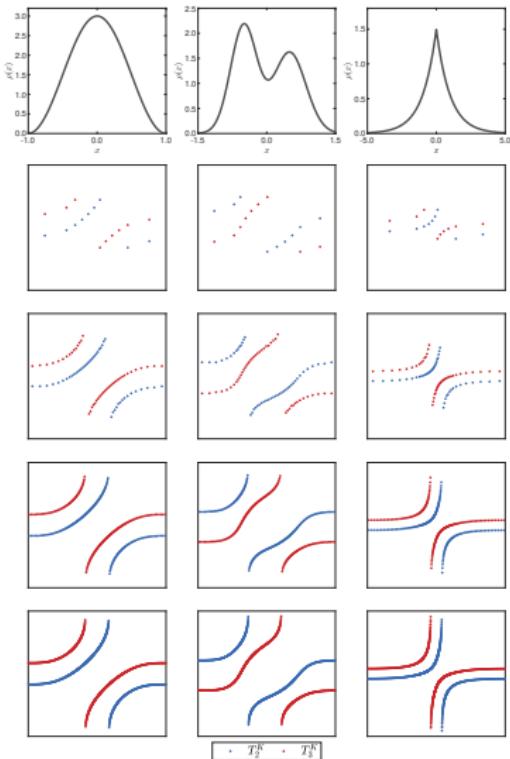


Figure 4: The marginals and approximated mappings in 1D $N = 3$ systems.
Left to right: ρ_1, ρ_2, ρ_3 ; up to down: $K = 12, 48, 192, 768$.



Comparison Between $\hat{\lambda}_\varrho$ and u_{SCE} ($d = 1, N = 3$)

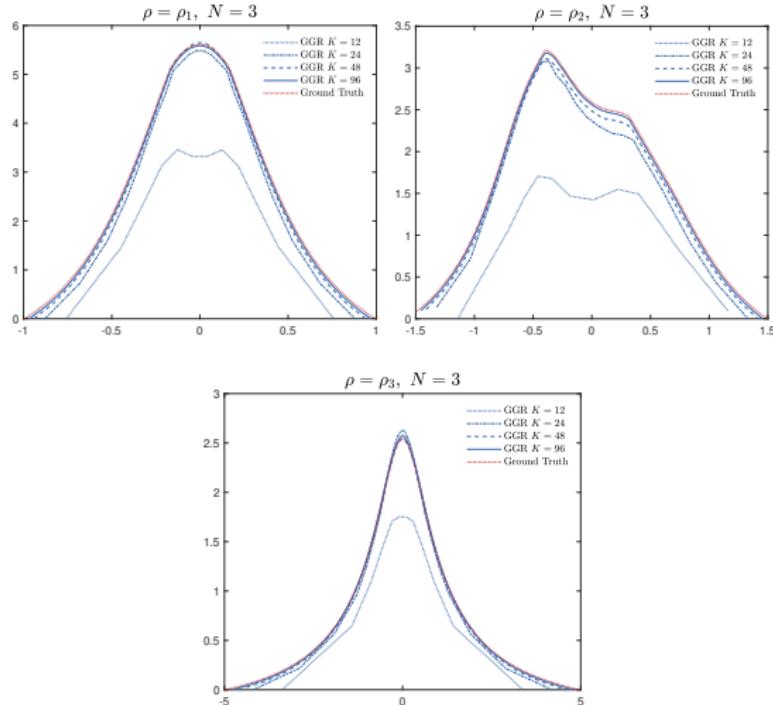


Figure 5: The evolution of $\hat{\lambda}_\varrho$ and the true SCE potential u_{SCE} when $d = 1, N = 3$.



Problem settings

$$\rho_4(x) = c_4 e^{-x^2/\sqrt{\pi}}, \quad \Omega = [-3, 3];$$

$$\rho_5(x) = c_5 \left(e^{-(x+2)^2} + 5e^{-2x^2} + e^{-(x-2)^2} \right), \quad \Omega = [-4, 4];$$

$$\begin{aligned} \rho_6(x) = c_6 & \left(e^{-4(x+2)^2} + e^{-4(x+1.5)^2} + e^{-4(x+1)^2} + e^{-4(x+0.5)^2} \right. \\ & \left. + e^{-4(x-2/3)^2} + e^{-4(x-4/3)^2} + e^{-4(x-2)^2} \right), \quad \Omega = [-3, 3]. \end{aligned}$$

- $c_i, i = 4, 5, 6$: normalizing factors such that $\int_{\Omega} \rho_i(x) dx = 7$.
- Initial equimass discretization ($K^{(0)} = 14$) + uniform mesh refinements.

Numerical Results on 1D Systems ($N = 7$) (Cont.)



Table 3: Output energies and calculated average errors of GGR
($d = 1, N = 7$).

Step	System 4				System 5				System 6			
	K	Energy	err_s	err_e	K	Energy	err_s	err_e	K	Energy	err_s	err_e
GGR_Init	14	216.212	-	0.052	14	151.891	-	0.039	14	111.964	-	0.030
GGR_LS(1)	28	181.474	0.045	0.019	28	158.797	0.037	0.028	28	117.223	0.030	0.010
GGR_LS(2)	56	181.929	0.018	0.025	56	158.507	0.023	0.026	56	117.050	0.011	0.008
GGR_LS(3)	112	181.989	0.019	0.012	112	158.317	0.019	0.011	112	116.914	0.007	0.008
GGR_LS(4)	224	181.954	0.014	0.013	224	158.267	0.008	0.008	224	116.876	0.008	0.006
GGR_LS(5)	448	181.942	0.007	0.002	448	158.255	0.010	0.004	448	116.864	0.005	0.003
GGR_LS(6)	896	181.939	0.002	0.001	896	158.254	0.005	0.002	896	116.861	0.003	0.001

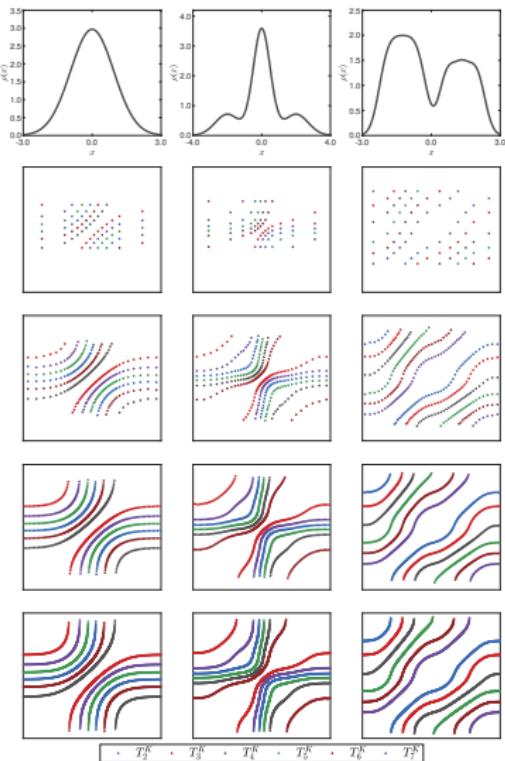


Figure 6: The marginals and approximated mappings in 1D $N = 7$ systems.
Left to right: ρ_4 , ρ_5 , ρ_6 ; up to down: $K = 14, 56, 224, 896$.



Comparison Between $\hat{\lambda}_\varrho$ and u_{SCE} ($d = 1, N = 7$)

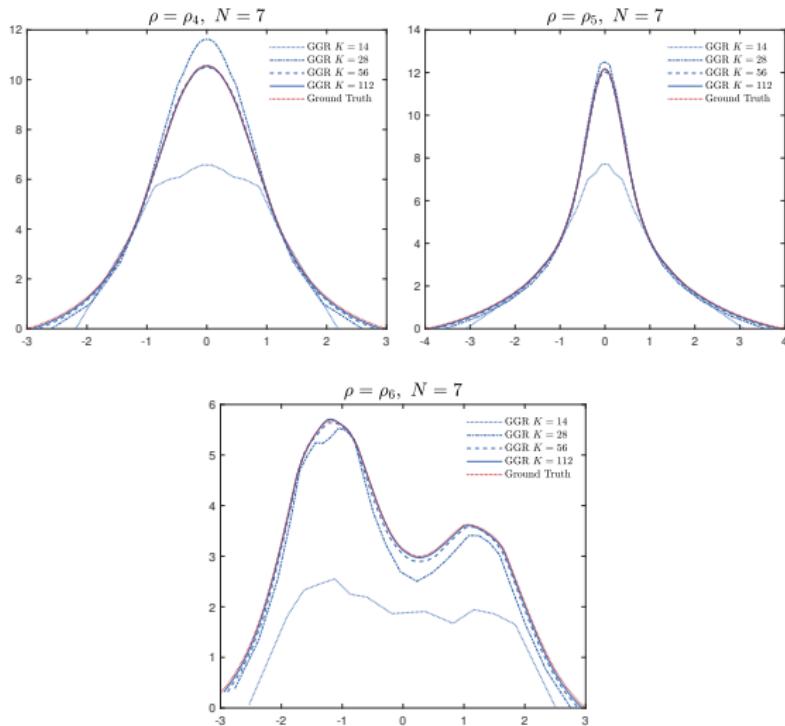


Figure 7: The evolution of $\hat{\lambda}_\varrho$ and the true SCE potential u_{SCE} when $d = 1, N = 7$.



Problem settings

$$\rho_7(x, y) = c_7 \left(e^{-2.5|(x, y) - (-1.5, 0)|^2} + 0.5e^{-2.5|(x, y) - (1.5, 0)|^2} \right), \quad \Omega = [-3, 3] \times [-2, 2];$$

$$\begin{aligned} \rho_8(x, y) = c_8 & \left(e^{-2.5|(x, y) - (-1.032, -0.84)|^2} + e^{-2.5|(x, y) - (0, 0.96)|^2} \right. \\ & \left. + e^{-2.5|(x, y) - (1.032, -0.84)|^2} \right), \quad \Omega = [-2.5, 2.5]^2. \end{aligned}$$

- c_i , $i = 7, 8$: the normalizing factors such that $\int_{\Omega} \rho_i(x, y) \, dx \, dy = 3$.
- Initial approximate equimass discretization ($K^{(0)} = 240, 170$).
the discretization is performed with FREEFEM. [Hecht 2012]
- Uniform mesh refinements.

Numerical Results on 2D Systems ($N = 3$) (Cont.)



Table 4: Output energies of GGR ($d = 2, N = 3$)⁴.

Step	System 7				System 8	
	K	Energy		K	Energy	
GGR_Init	240	4.073		170	4.068	
GGR_LS(1)	960	4.104		680	4.086	
GGR_LS(2)	3840	4.113		2720	4.090	
GGR_LS(3)	15360	4.116		10880	4.091	

⁴Since no optimal mappings are available, the average errors cannot be computed.

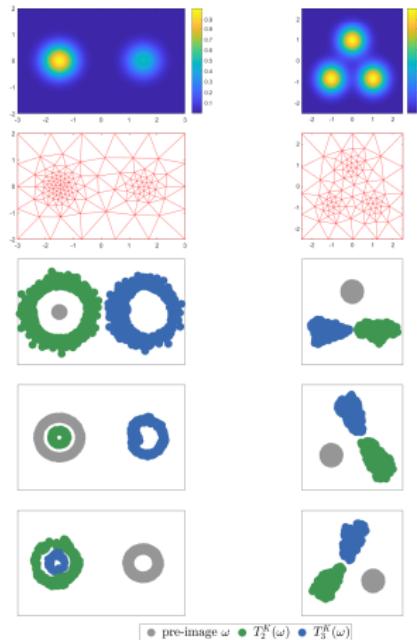


Figure 8: The marginals and approximated mappings in 2D $N = 3$ systems.
Left to right: ρ_7 ($K = 15460$), ρ_8 ($K = 10880$).

$\Rightarrow \left\{ \begin{array}{l} \text{GGR and } \textit{physical intuitions. ✓} \\ \text{The first visualization of 2D mappings. ✓} \end{array} \right.$

Evolution of $\hat{\lambda}_\varrho$ ($d = 2, N = 3$)

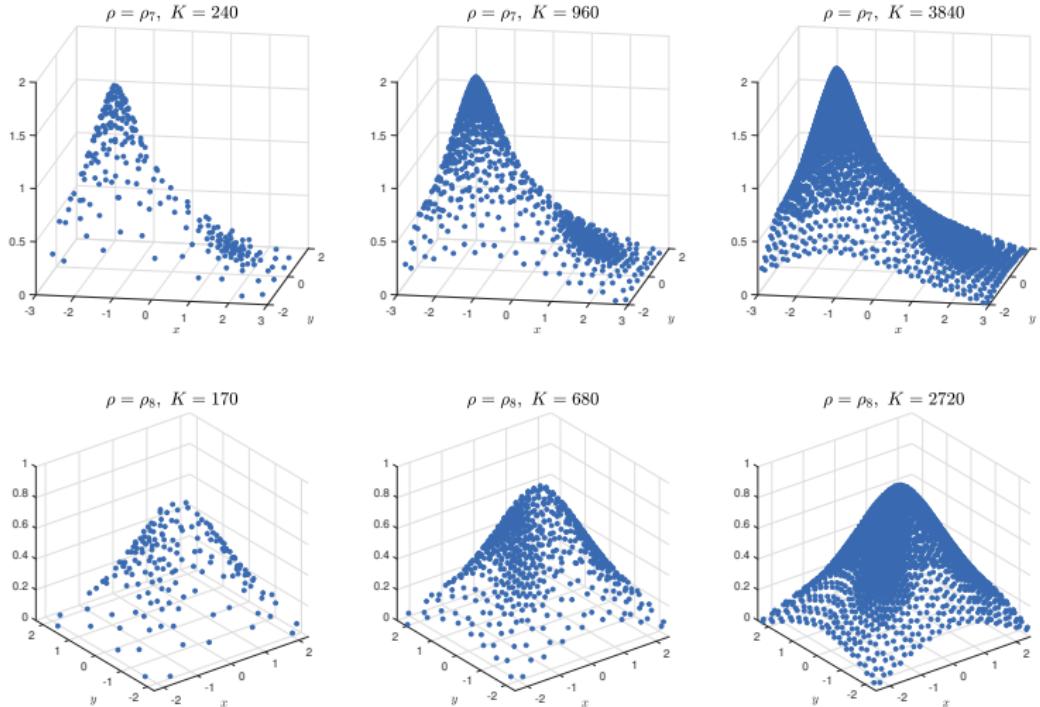


Figure 9: The evolution of $\hat{\lambda}_\varrho$ when $d = 2, N = 3$.



Conclusions

- Route: $F_{HK}[\rho] \xrightarrow{\text{SCE}} \text{MMOT} \xrightarrow[\tau]{\text{ansatz}} \text{MPGCC } (\mathbf{P}) \xrightarrow{\ell_1 \text{ penalty}} \text{problem } (\mathbf{P}_\beta)$.
- GGR: global optimization + local solver + hierarchical initialization.
- Numerical experiments:
 - larger systems, finer meshes;
 - results conforming to theoretical predictions and physical intuitions;
 - the first visualization of mappings between electrons in 2D contexts.

Future work

- Faster local solver⁵.
- MMOT under quasi-Monge ansatz. [Friesecke-Vögler 2018]
- Relation between $\hat{\lambda}_e$ and u_{SCE} .
- Combination with the KS model.

⁵Our recent work achieves acceleration via matrix entrywise sampling, enabling simulations in 3D contexts. Please refer to H., M. Li, X. Liu, and C. Meng. *Sampling-based approaches for multimarginal optimal transport problems with Coulomb cost*. arXiv preprint arXiv: 2306.16763, Jun. 2023.

Our Recent Work

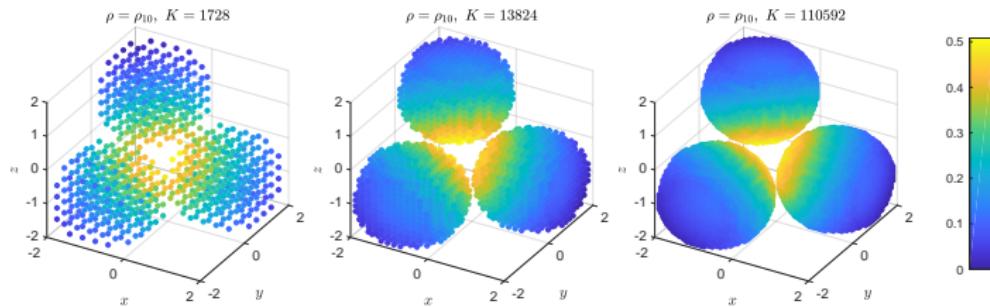
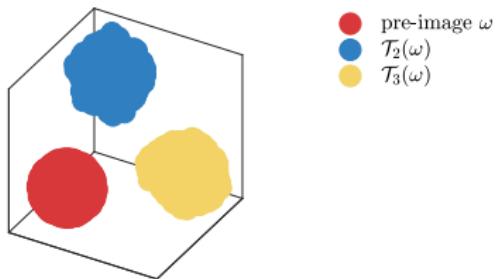


Figure 10: Approximated mappings (up) and the evolution of $\hat{\lambda}_\varrho$ (down) in 3D $N = 3$ systems.



- H., H. Chen, and X. Liu. A global optimization approach for multimarginal optimal transport problems with Coulomb cost. *SIAM Journal on Scientific Computing*, 2023, 45(3): A1214-A1238. (Main reference of this talk, GGR)
- H. and X. Liu. The convergence properties of infeasible inexact proximal alternating linearized minimization. *Science China Mathematics*, in press, DOI: 10.1007/s11425-022-2074-7. (Convergence properties of PALM)
- H. and X. Liu. The exactness of the ℓ_1 penalty function for a class of mathematical programs with generalized complementarity constraints. *Fundamental Research*, in press, DOI: 10.1016/j.fmre.2023.04.006. (Exactness of the ℓ_1 penalty function)

Thanks for your attention!

Email: ykhu@lsec.cc.ac.cn

Homepage: <https://huyukuan.github.io>