The multimarginal optimal transport with Coulomb cost

Séminaire du SAMM *March '25*

Rodrigue Lelotte — *CERMICS @ Ecole des ponts et chaussées*

The optimal transport in statistical physics

Motivation from statistical physics (1/2)

- N **identical (classical) particles** with positions $x_1, \dots, x_N$ in $\mathbb{R}^d$

- Particles $x_1, \dots, x_N$ are distributed along $\mathbb{P}(x_1, \dots, x_N) \in \mathscr{P}_{\text{sym}}((\mathbb{R}^d)^N)$

- Two-body **interaction potential** $w(|x-y|)$

Interaction energy of a configuration $(x_1, \dots, x_N)$ *:*

$$c(x_1, \dots, x_N) := \sum_{1 \le i < j \le N} w(|x_i - x_j|)$$

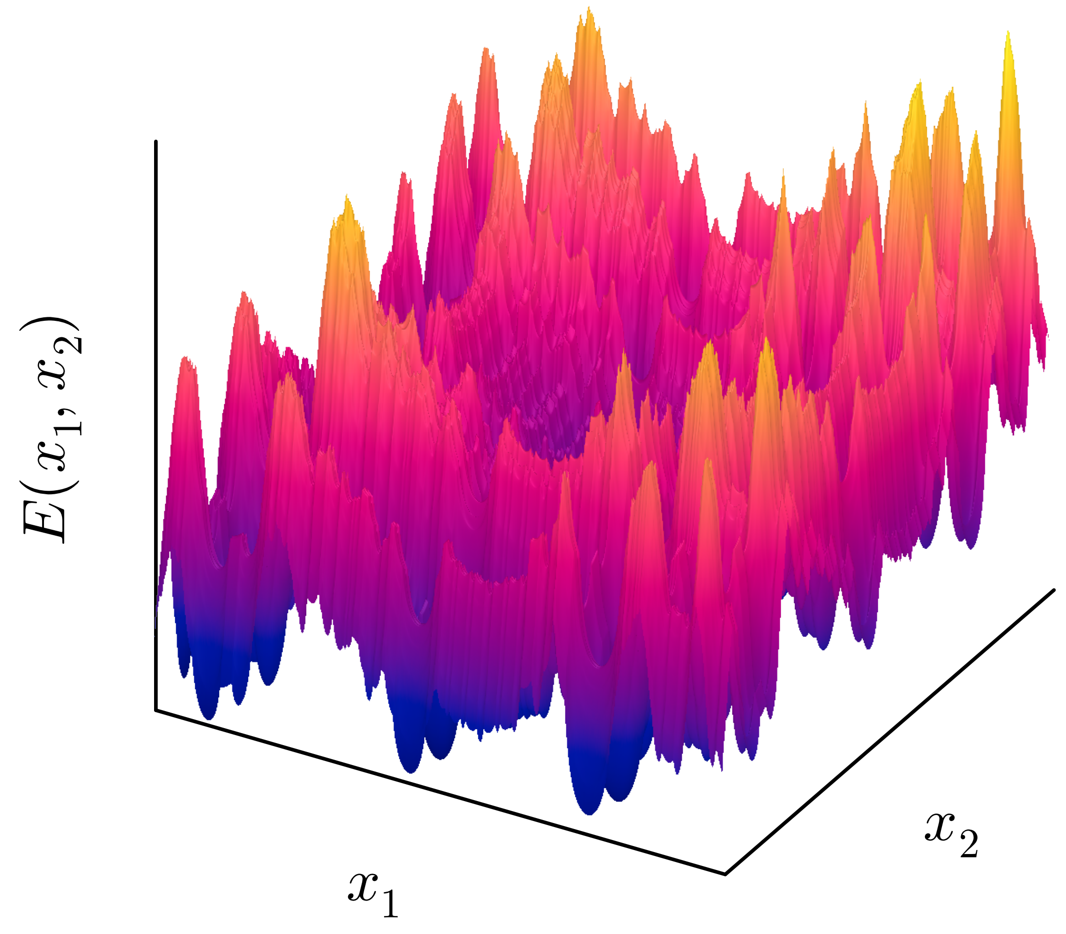
- One-body **external potential** $V_{\text{ex}} : \mathbb{R}^d \to \mathbb{R}$ (*e.g.* confining potential)

Ground-state/free energy:
$(T=0)$ $(T>0)$

$$F_N^{(T)}(V_{\text{ex}}) := \inf_{\mathbb{P} \in \mathscr{P}_{\text{sym}}((\mathbb{R}^d)^N)} \left\{ \underbrace{\int_{\mathbb{R}^{dN}} \left(c(x_1, \dots, x_N) + \sum_{i=1}^N V_{\text{ex}}(x_i) \right) d\mathbb{P}}_{\textit{Full interact° = interact° + external pot°}} + \overbrace{T}^{\textit{Temperature}} \cdot \mathit{Ent}(\mathbb{P}) \right\}$$

Motivation from statistical physics (2/2)

Computing this quantity is **hard** (*e.g.* many local minima $N \gg 1$) !



Two-step minimisation strategy [= **Density Functional Theory**]

Split *infimum* into two *infima*

$$F_N^{(T)}(V_{\mathrm{ex}}) = \inf_{\mathbb{P}\in\mathscr{P}_{\mathrm{sym}}((\mathbb{R}^d)^N)} \{\ldots\} = \inf_{\rho\in\mathscr{P}(\mathbb{R}^d)} \inf_{\rho_{\mathbb{P}}=\rho} \{\ldots\}$$

where $\rho_{\mathbb{P}}$ is marginal of $\mathbb{P}$:

$$\rho_{\mathbb{P}}(x) := \int_{\mathbb{R}^{d(N-1)}} \mathbb{P}(x, dx_2, \ldots, dx_N)$$

Multimarginal (entropic) OT

Ground-state/free energy = ***Legendre transform*** of OT

$$F_N^{(T)}(V_{\mathrm{ex}}) = \underbrace{\inf_{\rho\in\mathscr{P}(\mathbb{R}^d)}} \left\{ OT_N^{(T)}(\rho) + \underbrace{\int_{\mathbb{R}^{dN}} V_{\mathrm{ex}}\rho} \right\} \quad \text{where} \quad OT_N^{(T)}(\rho) := \inf_{\rho_{\mathbb{P}}=\rho} \left\{ \int_{\mathbb{R}^{dN}} c(x_1, \ldots, x_N) d\mathbb{P} + T \cdot Ent(\mathbb{P}) \right\}$$

😊 *« Smaller » search-space* 😊 *Linear dependence in ext° pot°*

But, solving OT is complicated 🤨 ? $\Longrightarrow$ Use **approximations** of $OT_N^{(T)}(\rho)$!

What is my problem ?

What is my problem ?

$$OT_N^{(T)}(\rho) := \inf_{\rho_{\mathbb{P}}=\rho} \left\{ \int_{\mathbb{R}^{dN}} c(x_1,\dots,x_N)\,d\mathbb{P} + T\cdot Ent(\mathbb{P}) \right\}$$

- I want to solve **numerically** $OT_N^{(T)}(\rho)$ for $N \gg 1$
- More precisely, for **small** $0 < T \ll 1$ in order to approximate the **unregularized** OT, that is $OT_N^{(0)}(\rho)$
- I am going to use the Kantorovich duality [= **Bi-Legendre transform**]

$$OT_N^{(T)}(\rho) = \sup_{V:\mathbb{R}^d\to\mathbb{R}} \underbrace{\left\{ F_N^{(T)}(V) + N\int_{\mathbb{R}^d} V\rho \right\}}_{:= D_N^{(T)}(V)}$$

😊 **Strictly and smooth concave** max° problem !

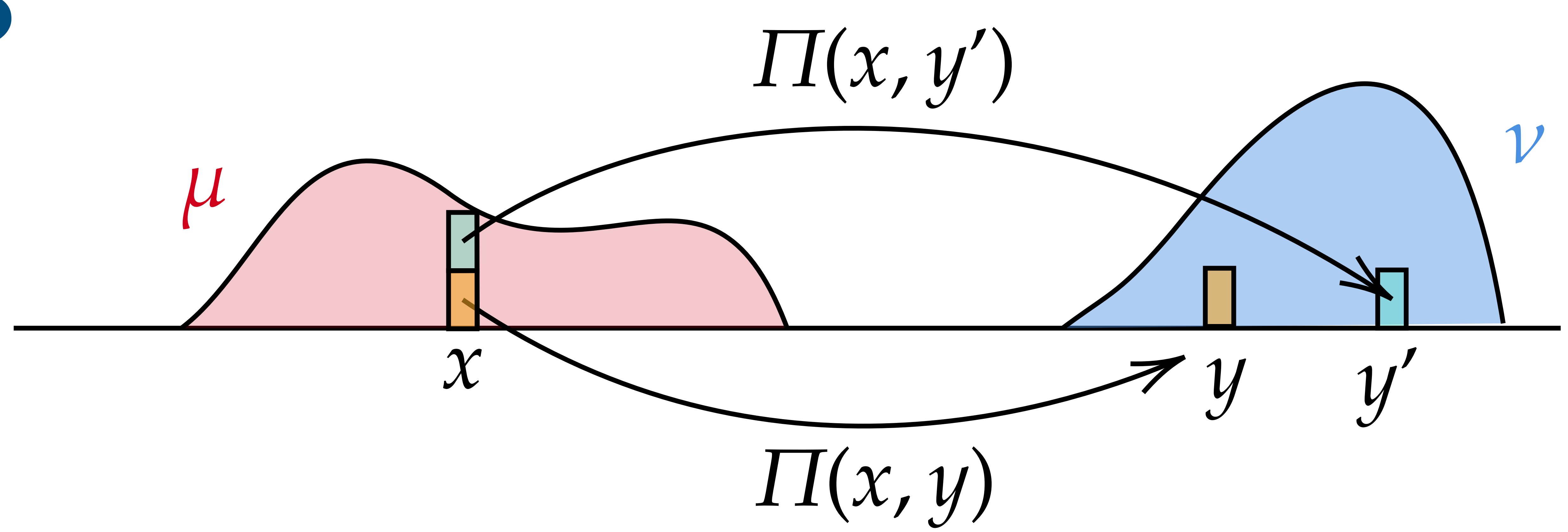
Contributions

- I am going to do a **Gradient Ascent** [≈ *Sinkhorn algorithm*] : $V_{t+1} \leftarrow V_t + \lambda \nabla D_N^{(T)}(V_t)$
- I am going to introduce a (natural & physically relevant) way to **discretize** this C^0 problem.
- …which is also amenable to get **quantitative error estimates**.

Multimarginal optimal transport

What is *Optimal Transport* ?

Given two probability measures $\mu, \nu \in \mathscr{P}(\mathbb{R}^d)$ how to *transport μ onto ν* while *minimizing a cost of transportation* $c(x, y)$?



A *transport plan* $\mathbb{P} \in \Pi(\mu, \nu)$ from μ to ν is a measure on $\mathbb{R}^d \times \mathbb{R}^d$ such that the first marginal (*resp.* second marginal) of $\mathbb{P}$ is μ (*resp.* ν)

« $\mathbb{P}(x, y)$ *is the amount of mass of* $\boldsymbol{\mu}$ *at position* $\boldsymbol{x}$ *sent to position* $\boldsymbol{y}$»

Moving an infinitesimal amount of mass from x to y costs $c(x, y)$, the *optimal* transport reads

$$\inf_{\mathbb{P} \in \Pi(\mu,\nu)} \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y) d\mathbb{P}(x, y) \right\}$$

😊 Under weak assumptions, there exists a minimiser $\mathbb{P}^*$

Is the transport induced by a Monge transport ? 🤔

$$\exists T : \mathbb{R}^d \to \mathbb{R}^d \qquad \mathbb{P}^*(x, y) = \mu(x) \otimes \delta(y - T(x))$$

« *The mass is not split : all the mass of μ at x is sent to a single location (i.e. y)* »

What is *Kantorovich duality* (and why caring) ?

Kantorovich duality

$$\inf_{\mathbb{P}\in\Pi(\mu,\nu)} \left\{ \int_{\mathbb{R}^d\times\mathbb{R}^d} c(x,y)d\mathbb{P}(x,y) \right\} = \sup_{\phi,\psi} \left\{ \int_{\mathbb{R}^d} \phi d\mu + \int_{\mathbb{R}^d} \psi d\nu \right\}$$

where $\phi, \psi : \mathbb{R}^d \to \mathbb{R}$ *are such that*

$$\phi(x) + \psi(y) \leq c(x,y)$$

Optimality conditions

ϕ, ψ are optimal (*Kantorovich potentials*) if and only if $\phi(x) + \psi(y) = c(x,y) \quad \mathbb{P}^* - a.e.(x,y)$

« *Points* (x,y) *on the support of* $\mathbb{P}^*$ *achieve the minimum of* $c(x,y) - \phi(x) - \psi(y)$ »

Existence and uniqueness of Monge map

This entails that on the support of $\mathbb{P}^*$: $\nabla_x \phi(x) = \nabla_x c(x,y)$ and $\nabla_y \psi(y) = \nabla_y c(x,y)$

Theorem — *If* $y \mapsto \nabla_x c(x,y)$ *is invertible, then* $\mathbb{P}^*$ *is the graph of* $T(x) = \nabla_x c(x,\cdot)^{-1}(\nabla\phi(x))$

Example — $c(x,y) = \frac{1}{2}|x-y|^2 \qquad \nabla\phi(x) = \nabla_x c(x,y) \implies \nabla\phi(x) = x - y \implies y = \underbrace{x - \nabla\phi(x)}_{:=T(x)}$

What is *multimarginal* optimal transport ?

N marginals $\mu_1, \ldots, \mu_N \in \mathscr{P}(\mathbb{R}^d)$

Cost of transportation $c(x_1, \ldots, x_N)$

Multimarginal OT

$$\inf_{\mathbb{P}\in\Pi(\mu_1,\ldots,\mu_N)} \int_{(\mathbb{R}^d)^N} c(x_1, \ldots, x_N) d\mathbb{P}$$

Open question *Under which assumptions there exists* $T_2, \ldots, T_N : \mathbb{R}^d \to \mathbb{R}^d$ *such that*

$$\mathbb{P}^*(x_1, \ldots, x_N) = \mu_1(x_1) \bigotimes_{i=2}^{N} \delta(x_i - T_i(x_1))$$

Almost completely open !

Optimality conditions

$V_1, \ldots, V_N$ are optimal if and only if

$$V_1(x_1) + \ldots + V_N(x_N) = c(x_1, \ldots, x_N) \qquad \mathbb{P}^* - a.e.$$

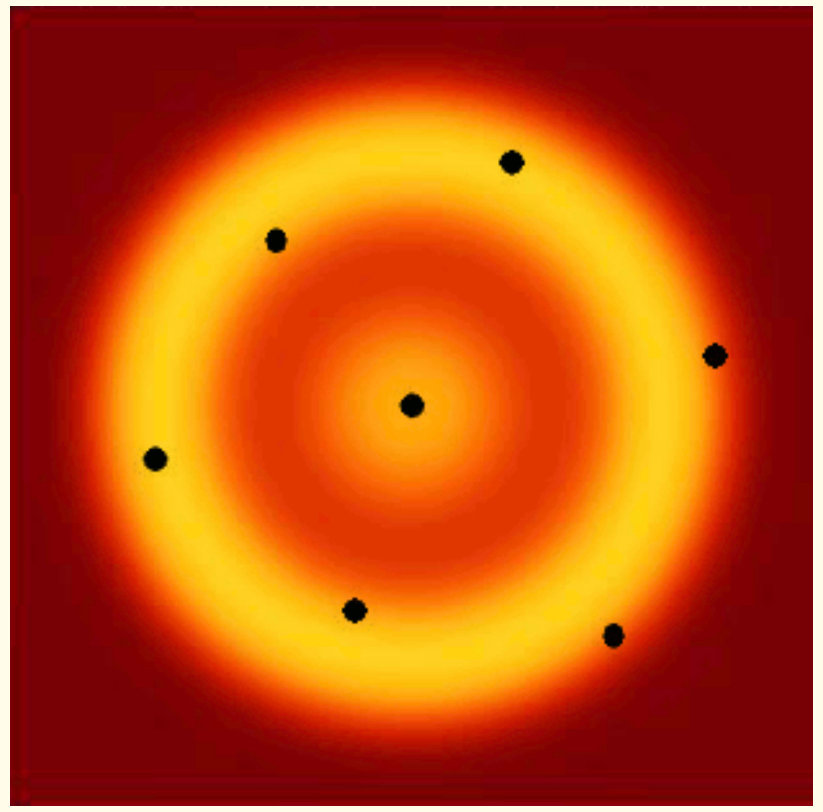
$$\implies \nabla_{x_i} V_i(x_i) = \nabla_{x_i} c(x_1, \ldots, x_i, \ldots, x_N)$$

Duality

$$\sup_{V_1,\ldots,V_N} \left\{ \sum_{i=1}^{N} \int_{\mathbb{R}^d} V_i \mu_i \right\}$$

where $V_1, \ldots, V_N : \mathbb{R}^d \to \mathbb{R}$ *are such that*

$$V_1(x_1) + \ldots + V_N(x_N) \leq c(x_1, \ldots, x_N)$$



For repulsive Coulomb-like costs $w(|x-y|) = |x-y|^{-s}$ it is believed to be generically true. Conjecture holds for more than two decades…

← *Wikipedia [Strictly-correlated electrons]*

Physical interpretation of Kantorovich duality $(T = 0)$

MOT = $$\sup_{V_1,\ldots,V_N} \left\{ \sum_{i=1}^{N} \int_{\mathbb{R}^d} V_i \mu_i \right\}$$

where $V_1, \ldots, V_N : \mathbb{R}^d \to \mathbb{R}$ *are such that*

$$V_1(x_1) + \ldots + V_N(x_N) \le c(x_1, \ldots, x_N)$$

- This can be rewritten as an *unconstrained problem*: Given any $V : \mathbb{R}^d \to \mathbb{R}$, up to the addition of the constant $V \leftarrow V + E_N(V)/N$, the function is admissible, where

$$E_N(V) = \inf_{x_1,\ldots,x_N} \left\{ c(x_1, \ldots, x_N) - \sum_{i=1}^{N} V(x_i) \right\}$$

- Remark that $E_N(V) = F_N^{(0)}(V)$ = *ground-state energy*

- In our case, all marginals are the same $\rho = \mu_1 = \ldots = \mu_N$ and one can choose $V = V_1 = \ldots = V_N$

$$OT_N^{(0)}(\rho) = \sup_{V:\mathbb{R}^d \to \mathbb{R}} \left\{ E_N(V) + N \int_{\mathbb{R}^d} V\rho \right\}$$

V is optimal if and only if $\quad c(x_1, \ldots, x_N) - \sum_{i=1}^{N} V(x_i) = E_N(V) \qquad \mathbb{P}^* - a.e.\,(x_1, \ldots, x_N)$

« The optimal V is interpreted physically as (minus) the external potential which forces the particles to live at equilibrium inside the support of $\mathbb{P}^*$ »

Physical interpretation of Kantorovich duality $(T > 0)$

- Duality at positive temperature $T > 0$ is the same but one replaces the ground-state energy by the *free energy*

$$OT_N^{(T)}(\rho) = \sup_{V:\mathbb{R}^d \to \mathbb{R}} \left\{ F_N^{(T)}(V) + N \int_{\mathbb{R}^d} V\rho \right\}$$

Gibbs variational principle :

$$F_N^{(T)}(V) = -T \log \int_{(\mathbb{R}^d)^N} e^{-\frac{1}{T}(c(x_1,\ldots,x_N) - \sum_{i=1}^N V(x_i))} dx_1 \ldots dx_N$$

- *Smooth* (strictly concave) functional and $\frac{\delta F_N^{(T)}}{\delta V}(V_0)$ is given by (*minus*) the *marginal of Gibbs measure*

$$\mathbb{P}_{T,N}(V_0) \propto \exp\Big[-\frac{1}{T}(c(x_1, \ldots, x_N) - \sum_{i=1}^N V_0(x_i))\Big]$$

*« **Canonical ensemble** » — distribution of interacting particles with ext° pot° $-V_0$ in thermal equilibrium at temperature $T > 0$*

V_T^* is optimal if and only if $\frac{\delta F_N^{(T)}}{\delta V}(V_T^*) = \rho$

« The optimal V_T is interpreted physically as (minus) the external potential which forces the** canonical ensemble **to have marginal ρ»

A strategy to solve the optimal transport

Let's recall my problem

$$OT_N^{(T)}(\rho) := \inf_{\rho_{\mathbb{P}}=\rho} \left\{ \int_{\mathbb{R}^{dN}} c(x_1, \ldots, x_N) d\mathbb{P} + T \cdot Ent(\mathbb{P}) \right\}$$

- I want to solve **numerically** $OT_N^{(T)}(\rho)$ for $N \gg 1$
- More precisely, for **small** $0 < T \ll 1$ in order to approximate the **unregularized** OT, that is $OT_N^{(0)}(\rho)$
- I am going to use the **Kantorovich duality** [= **Bi-Legendre transform**]

$$OT_N^{(T)}(\rho) = \sup_{V:\mathbb{R}^d \to \mathbb{R}} \underbrace{\left\{ F_N^{(T)}(V) + N \int_{\mathbb{R}^d} V\rho \right\}}_{:= D_N^{(T)}(V)}$$

😊 **Strictly and smooth concave** max° problem !

- I am going to do a **Gradient Ascent** [≈ *Sinkhorn algorithm*] : $V_{t+1} \leftarrow V_t + \lambda \nabla D_N^{(T)}(V_t)$
- According to what precedes, we have $\nabla D_N^{(T)}(V_t) = N\rho - N\rho^T(V_t)$ where $\rho^{(T)}(V_t)$ is the marginal of the canonical ensemble with potential V_t
- … it remains to find a way to **discretize** this C^0 problem !

A simple idea

Idea to solve $OT_N^{(T)}(\rho)$

1. Reduced-order search-space $V \in \mathrm{Span}(\{\phi_i\}_{i=1,\ldots,M})$ for some finite basis $\{\phi_i\}_{i=1,\ldots,M}$

2. Solve with a **gradient ascent** the problem

$$\sup_{V \in \mathrm{Span}(\{\phi_i\}_{i=1,\ldots,N})} \left\{ F_N^{(T)}(V) - N \int_{\mathbb{R}^d} V\rho \right\} \quad (\star)$$

$\Longleftrightarrow$ Optimise over the weights $\omega_1, \ldots, \omega_M$ s.t. $V = \sum_{i=1}^{M} \omega_i \phi_i$

Derivative objective wrt to ω_i : $\quad N \underbrace{\int \phi_i \rho^{(T)}(V)}_{\text{MCMC methods}} - N \underbrace{\int \phi_i \rho}_{\text{Computed offline}}$

$(\star)$ = dual formulation of **Moment-Constrained Optimal Transport** (MCOT)

[Alfonsi, Coyaud, Ehrlacher & Lombardi '21]

$$\rho_{\mathbb{P}} = \rho \qquad v.s. \qquad \int \sum_{j=1}^{N} \phi_i(x_j) d\mathbb{P} = N \int \phi_i \rho \qquad \forall i = 1,\ldots,M.$$

« Marginal constraint is relaxed into moment constraints »

Some simple *quantitative bounds*

😊 Working with the dual of MCOT allows to derive simple quantitative bounds

Assume $supp(\mu_i) \subset S$ for all $i = 1,\ldots,N$. Denote $\mathcal{S} = (S_i)_{i=1}^M$ a partition of S

Theorem [L'] — *If the cost of transportation $c(x_1,\ldots,x_N)$ is $\alpha-$Hölder, and if the moment functions are taken to be piecewise constant functions on the partition $\mathcal{S}$, then the error between the MCOT and the true optimal transport is* ***bounded by*** $NC\epsilon^\alpha$ *where N is the number of marginal, where $C = \max_{i=1} \|V_i^*\|_{L^\infty}$ (where V_i^* and the optimal Kantorovich potential) and $\epsilon = \max_{i=1,\ldots M} \mathrm{diam}(S_i)$*

This gives a rough bound of $\mathcal{O}(M^{-\alpha/d})$ where M is the number of moment functions

If one can proves regularity on the Kantorovich potentials, then faster convergence rates :

Theorem [L'] — *If the optimal Kantorovich potential $V_1^* \ldots V_N^*$ are C^{k+1}, then considering the moment functions are piecewise polynomials of order up to k on the partition $\mathcal{S}$, the error between the MCOT and the true optimal transport is* ***bounded by*** $NC\epsilon^k$

The case of Coulomb(-like) costs

How to choose the ϕ_i's in the case of the *Coulomb cost* (1/2)

From now on, we consider **Coulomb interaction** $w(|x-y|) = \dfrac{1}{|x-y|}$ in dimension $d = 3$

Mean-field limit of OT at $T = 0$ (formally from [Cotar, Friesecke & Pass '14])

If $\rho_N = \rho$ is fixed for some $\rho \in \mathscr{P}(\mathbb{R}^d)$

Kantorovich potential at $T = 0$ $\left\{ \dfrac{V_N^{(0)}}{N} \longrightarrow \underbrace{-\rho \star |x|^{-1} = -\int \rho(y)|x-y|^{-1}\,dy}_{\textit{Electrostatic potential generated by } -\rho} \right.$

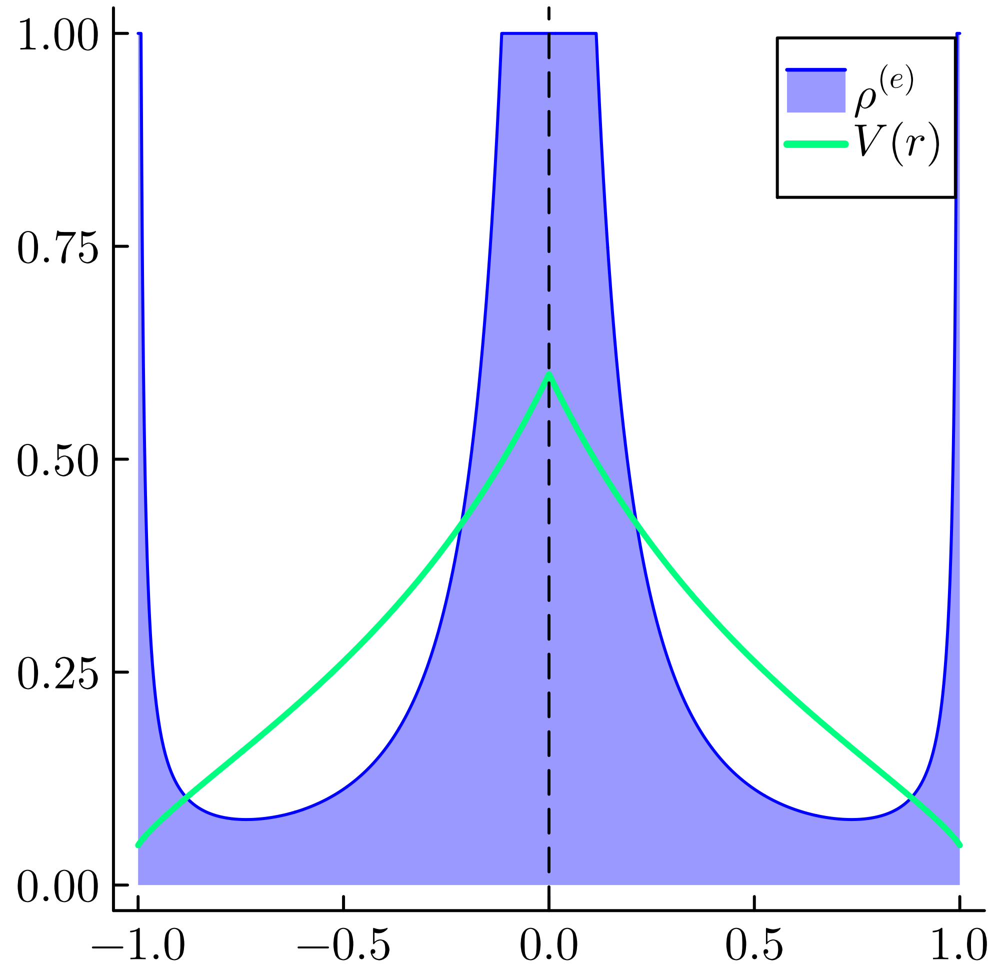
Takeaway message $V_N^{(0)} = -N\rho \star |x|^{-1} + \text{correction terms}$ as $N \to \infty$

How to choose the $\phi_i's$ in the case of the *Coulomb cost* (2/2)

Theorem (L '24) — For all $T \geq 0$, there exists a positive measure $\rho_T^{(e)}$ such that the (entropic) Kantorovich potential V_T rewrites as the electrostatic potential generated by the charge density $\rho_T^{(e)}$:

$$V_T(x) = -\rho_T^{(e)} \star |x|^{-1} = -\int \rho^{(e)}(y)|x-y|^{-1}dy$$

We call $\rho_T^{(e)}$ the external dual charge. Moreover, $\rho_T^{(e)}(\mathbb{R}^d) = N - 1.$



Interpretation — *« The charge density $\rho_T^{(e)}$ attracts the electrons into the optimal transport plan $\mathbb{P}^*$ »*

Idea to discretize the (entropic) Kantorovich potential

Reduced-order search-space $\rho_T^{(e)} \in \mathrm{Span}(\{\mu_i\}_{i=1,\ldots,N})$ for finite basis of measures $\{\mu_i\}_{i=1,\ldots,M}$

Otherwise stated, potential search-space is $V \in \mathrm{Span}(\{\phi_i\}_{i=1,\ldots,M})$ where $\phi_i := \mu_i \star |x|^{-1}$

Vague claim (L' 24 for $d = 1$) — *If $\rho_N = \rho$ is fixed for some $\rho \in \mathscr{P}(\mathbb{R}^d)$, then* $\dfrac{\rho_0^{(e)}}{N} \xrightarrow[narrow]{N\to\infty} \rho$

Numerics for *uniform droplets*

- **Uniform droplets**

$$N \in \mathbb{N}, \quad \rho_N = N^{-1}\chi_{B_N} \quad \text{with } B_N = B(0, r_N) \subset \mathbb{R}^3 \quad \text{s.t.} \quad |B_N| = N$$

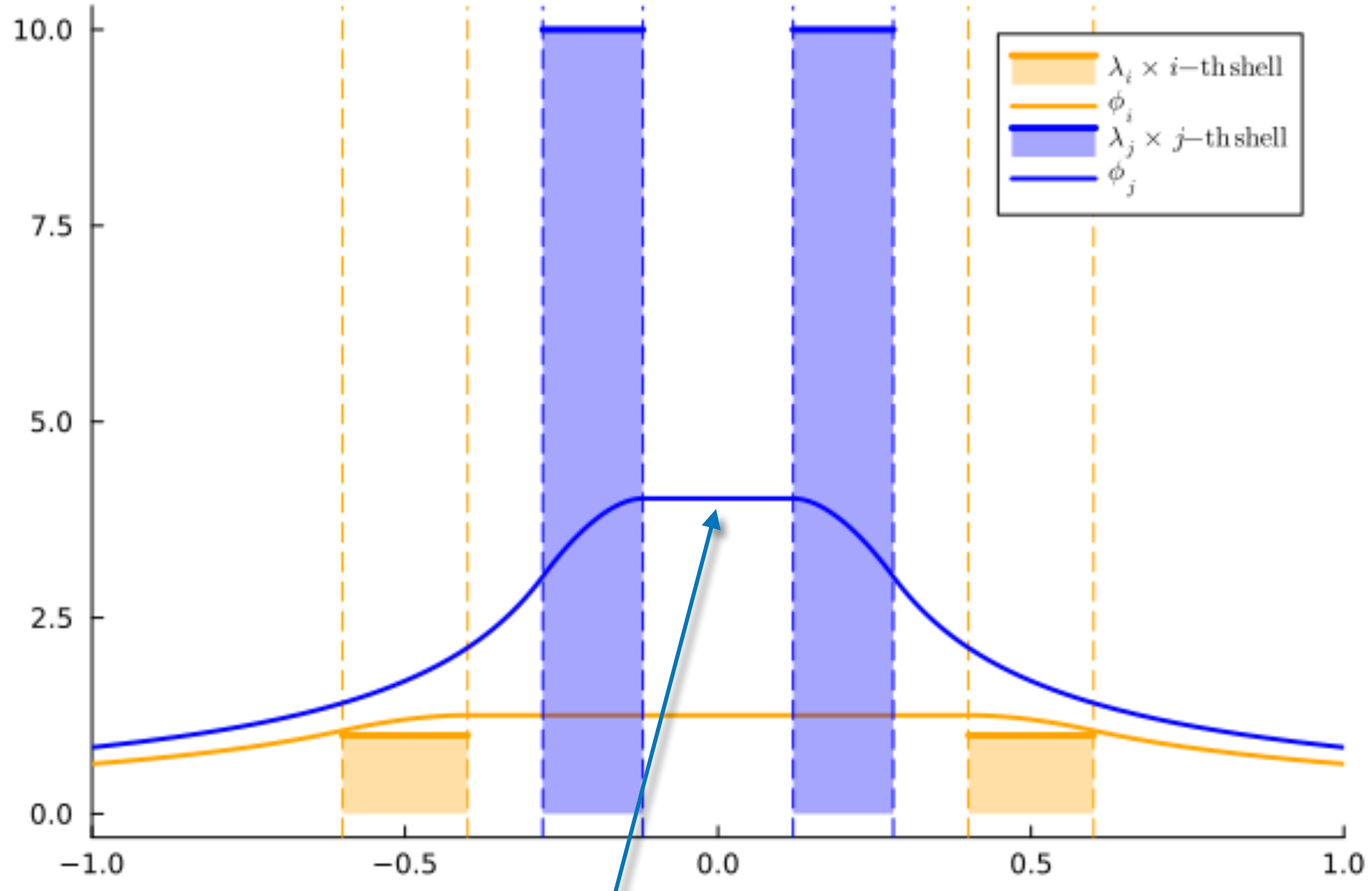
- Ball B_N discretized into M **concentric shells** S_i

$$B_N = \cup_{i=1}^{M} S_i \quad \text{where } S_i = B(r_i)\backslash B(r_{i-1}), \quad 0 = r_0 < r_1 < \cdots < r_M = r_N$$

- μ_i is **indicator of** S_i**'s**

$$\widehat{V}(\omega_1, \ldots, \omega_N) = \sum_{i=1}^{M} \omega_i \mu_i \star |x|^{-1} \quad \text{where } \mu_i = \chi_{S_i}$$

- Initialized on **mean-field limit**, *i.e.* $\omega_i^0 = 1$ for $i = 1, \ldots, M$.



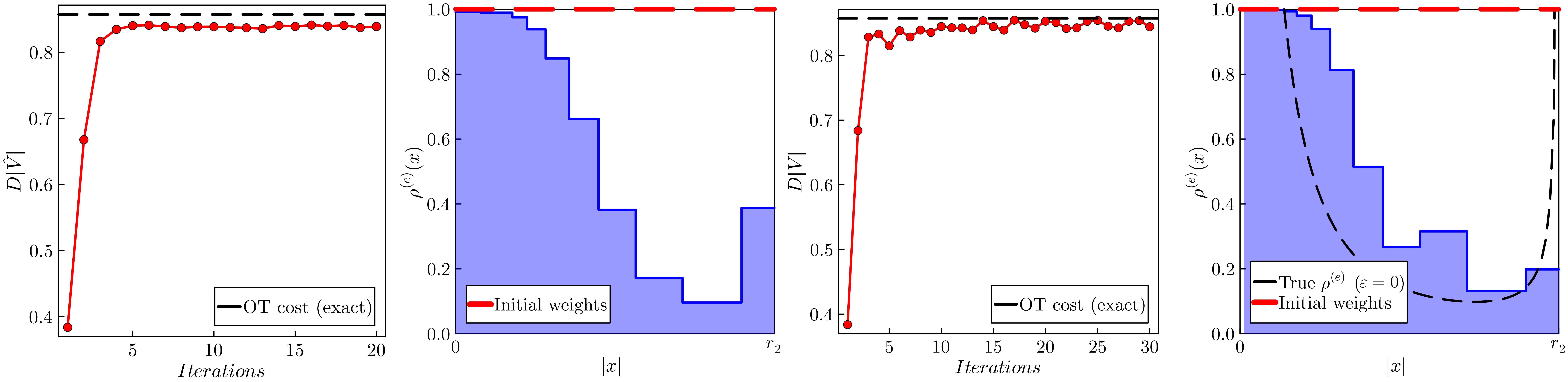
Example of $\phi_i = \mu_i \star |x|^{-1}$

$N = 2$ with $M \in \{10{,}20\}$ and $T = \{50^{-1}{,}500^{-1}\}$

= *Lower bound to OT cost*

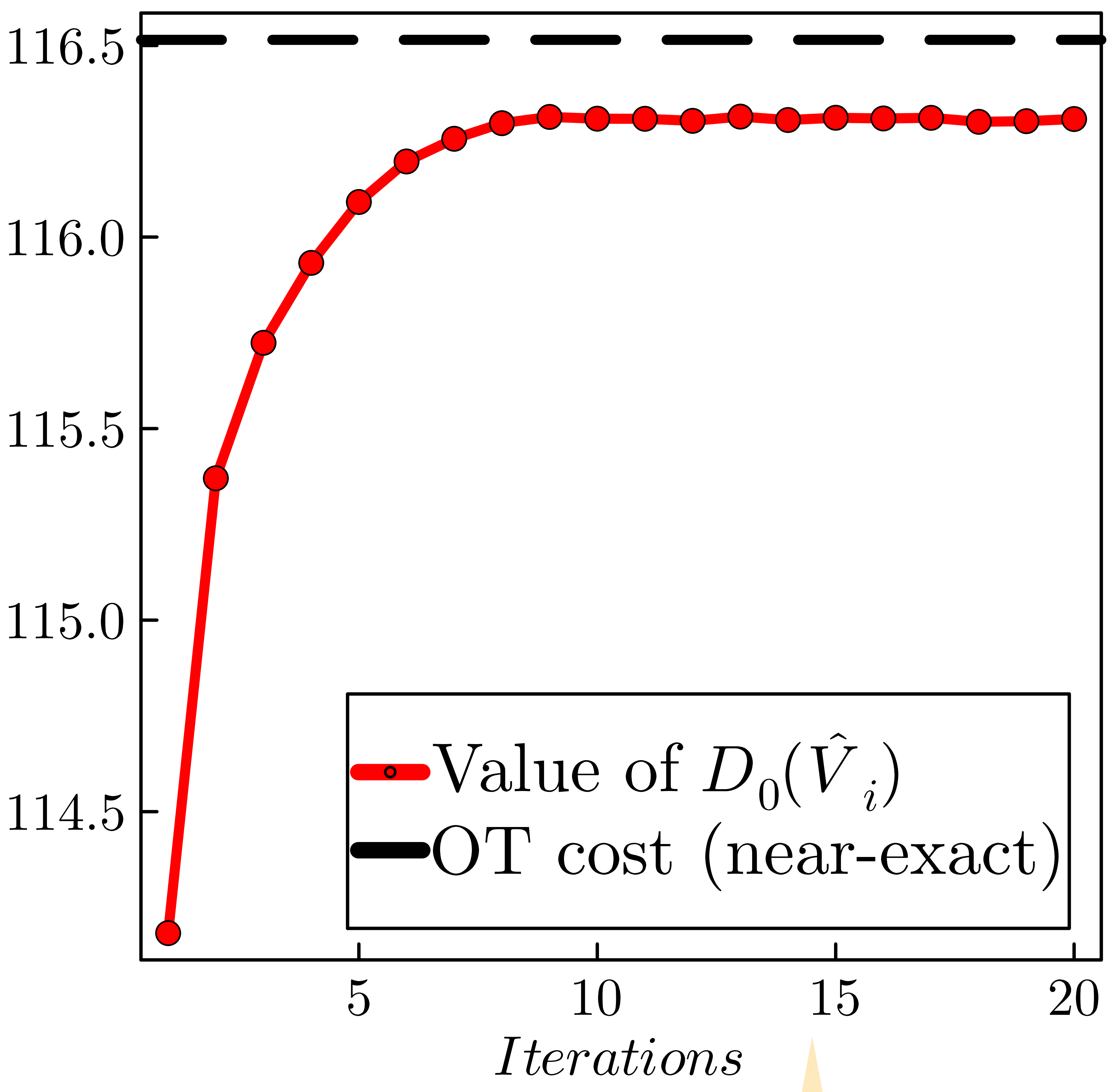
Optimized $V(\omega_1^*, \ldots, \omega_M^*)$ is then plugged as a **trial state** in the **unregularized OT** dual :

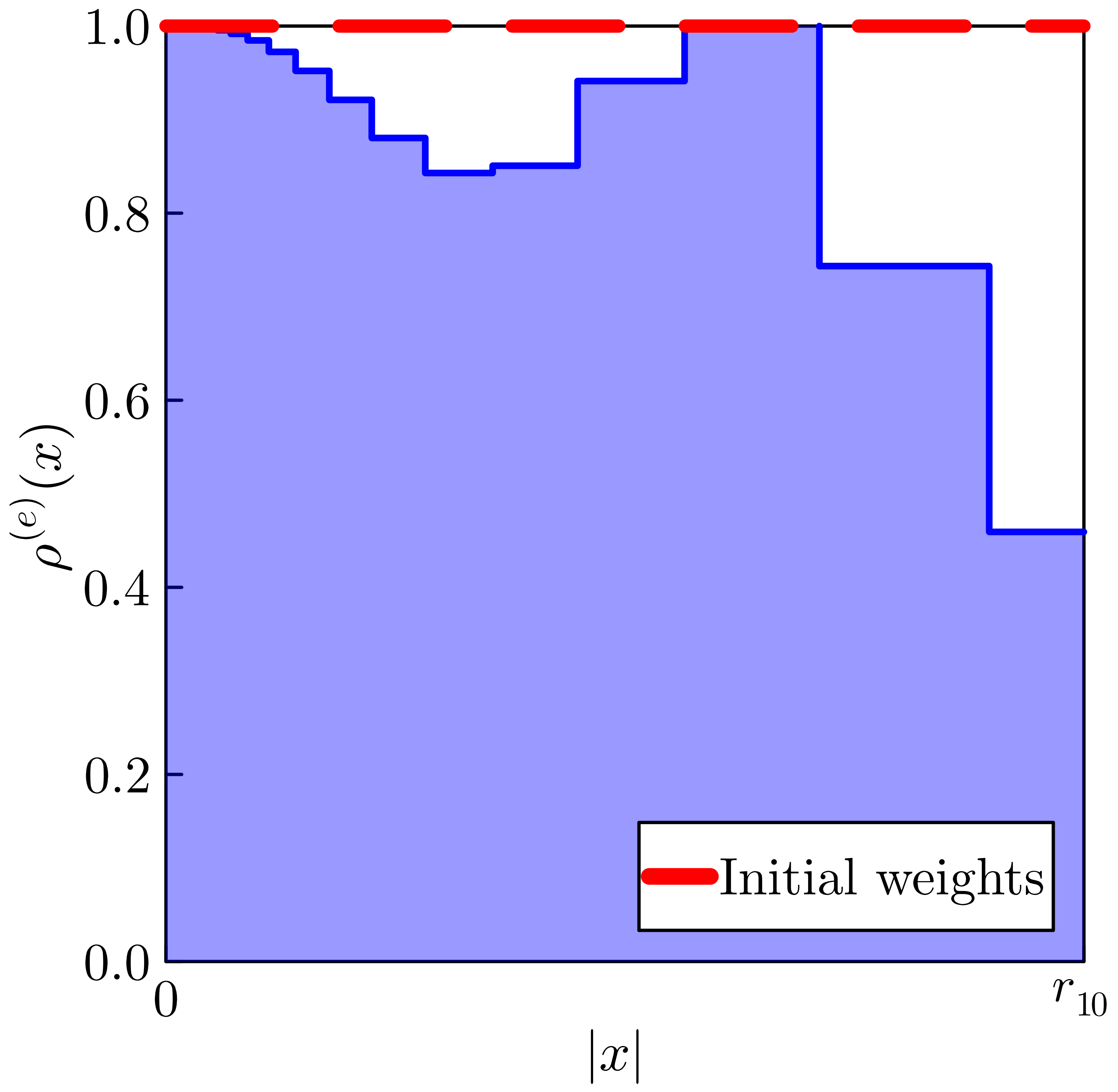
$$D_0(V) = \underbrace{\inf_{x_1,\ldots,x_N \in \mathbb{R}^3} \left\{ c(x_1, \ldots, x_N) - \sum_{i=1}^{N} V(x_i) \right\}}_{E_N(V)} + N \int_{\mathbb{R}^3} V\rho$$



$N = 20$ with $M = 20$ and $T = 150^{-1}$

Compared with **upper bounds on OT** of `[Räsänen, Gori–Giorgi & Seidl '16]`





N	$F_{\mathrm{SCE}}[\nu]$	In [34]
3	2.300	2.327
4	4.922	4.935
5	8.519	8.626
10	43.022	43.140
14	90.454	90.808
20	195.607	196.198
30	462.423	463.807

Of the order of 1 minute !

Error of order $\leq -N \times T \times \ln T$

Conclusion

I presented…

- How the (multimarginal) optimal transport arises in statistical physics
- A **general strategy** to solve numerically the MOT (= **dual of MCOT**) (with quantitative estimates)
- **Efficient discretization** of Kantorovich potentials for **Coulomb(-like) cost**

Outlook

- Lots of room for **optimisation/algorithmic improvement** — MCMC methods etc.
- *In which sense* $V_0 \simeq -N\rho \star |x|^{-1}$ *? Can we give* $1^{rst}/2^{nd}$ *order correction ?*
- Reference : `[L'24, An external dual charge approach for the OT with Coulomb cost, ESAIM COCV]`