Sample and Map from a Single Convex Potential: Generation using Conjugate Moment Measures

Nina Vesseron
CREST - ENSAE

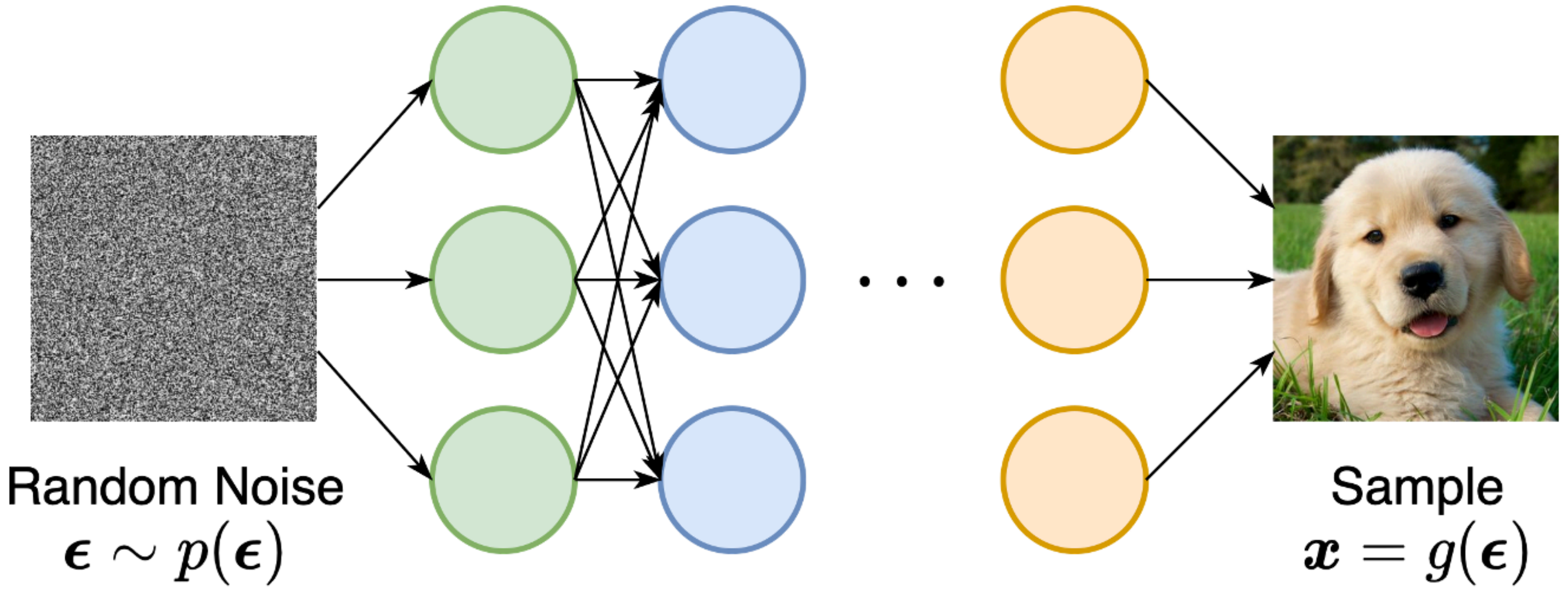
Louis Béthune
Apple

Marco Cuturi
CREST - ENSAE & Apple

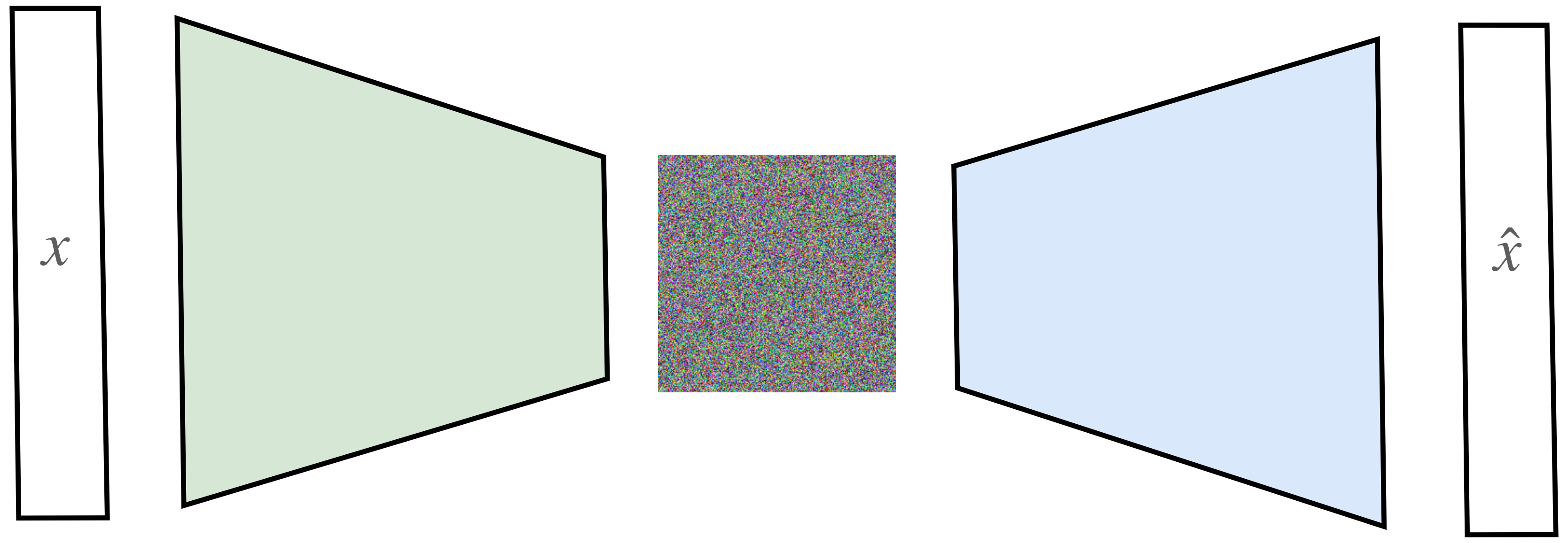
Generative Modeling

- Objective: Given samples $x_1, \ldots, x_n \sim \rho$, draw new samples from ρ.

GAN [1]



VAE



Diffusion models [1]

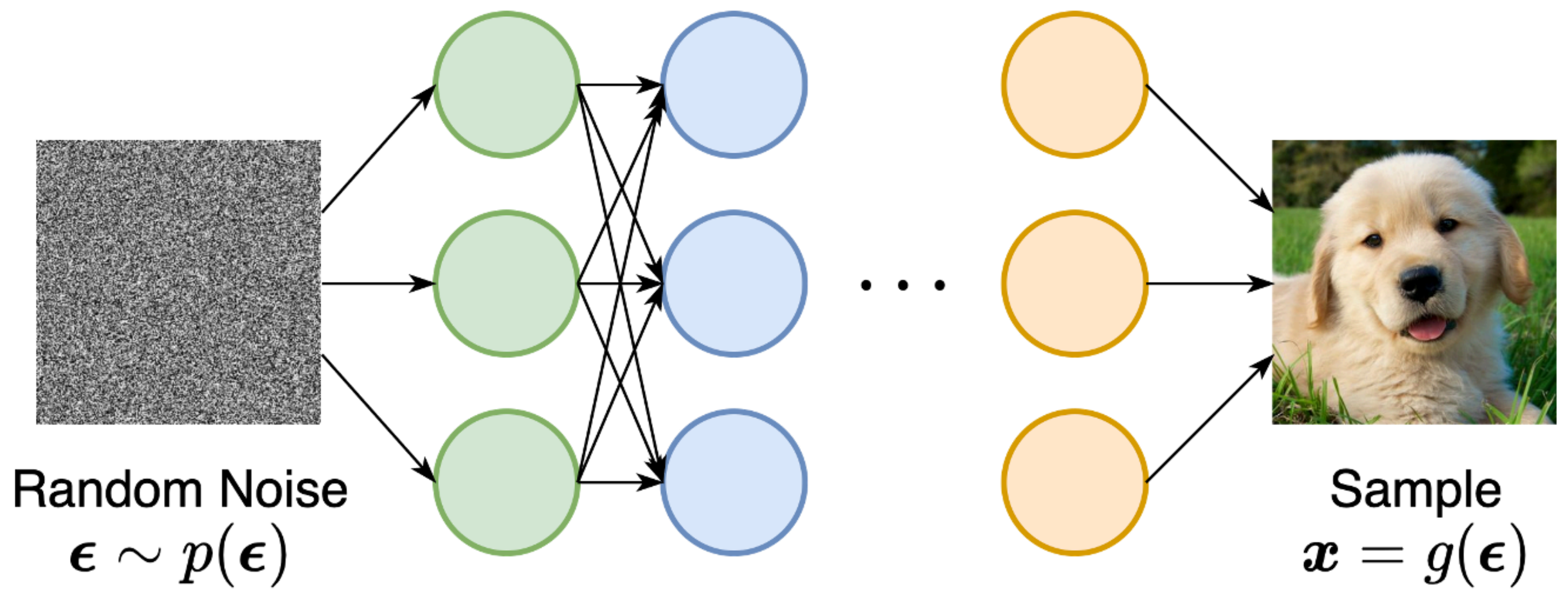
- Current solutions: **Sample a noise** z (e.g. Gaussian) **Learn a transport map** ϕ such that $\phi(z) \sim \rho$

[1] Drawings from Yang Song blog: https://yang-song.net/blog/2021/score/

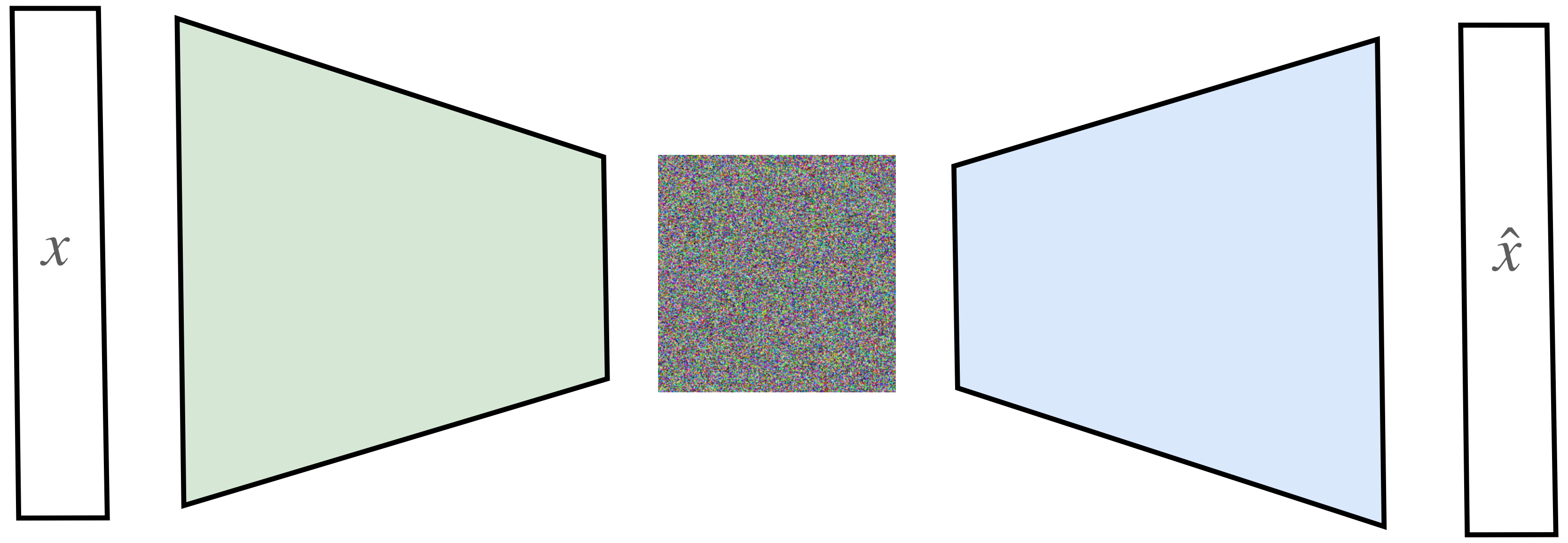
Generative Modeling

- Objective: Given samples $x_1, \ldots, x_n \sim \rho$, draw new samples from ρ.

GAN [1]



VAE



Diffusion models [1]

- Current solutions: **Sample a noise** z (e.g. Gaussian) **Learn a transport map** ϕ such that $\phi(z) \sim \rho$

Question: Can we tie Sampling and Transport ?

[1] Drawings from Yang Song blog: https://yang-song.net/blog/2021/score/

Moment measures

Moment measures factorization Theorem [1]

Given a probability measure ρ that admits a first moment, whose barycenter is 0, and which is not supported on a hyperplane, there exists a convex function $u : \mathbb{R}^d \to \mathbb{R}$ such that:

$$\boxed{\rho = \nabla u_\sharp e^{-u} \qquad \bigstar}$$

Groundbreaking works [1] [2]

Cordero-Erausquin and Klartag (2015)

$$\min_{u \in \mathscr{C}(\mathbb{R}^d)} \int u^* \mathrm{d}\rho - \ln\left(\int e^{-u}\right)$$

→ Solutions: u that verify $\bigstar$

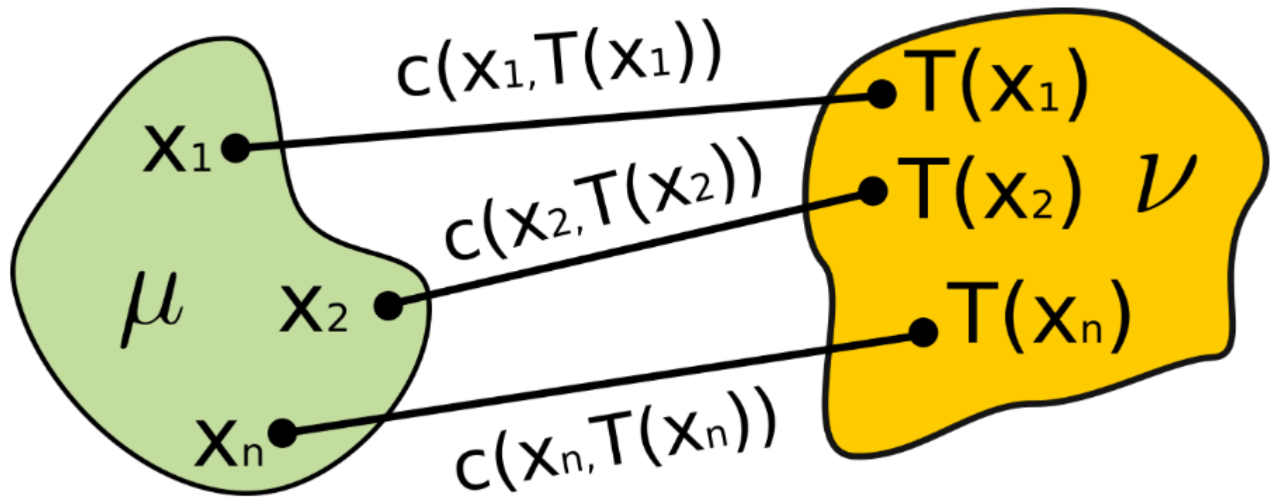
Santambrogio (2016)

$$\min_{\mu \in \mathscr{P}(\mathbb{R}^d)} \{\underbrace{\mathscr{E}(\mu)}_{\int \mu \ln(\mu)\,\mathrm{d}x} + \underbrace{\mathscr{T}(\mu,\rho)}_{\inf_{f \in \mathscr{C}(\mathbb{R}^d)} \int f\,\mathrm{d}\mu + \int f^*\,\mathrm{d}\rho} : \mu \in \mathscr{P}_1(\mathbb{R}^d)\}$$

→ Solutions: e^{-u} that verify $\bigstar$

[1] Dario Cordero-Erausquin and Bo'az Klartag. Moment measures. Journal of Functional Analysis, 268(12):3834–3866, 2015

[2]. Filippo Santambrogio Dealing with moment measures via entropy and optimal transport. Journal of Functional Analysis, 271(2):418–436, 2016



Links with Optimal Transport

Moment measures

$\rho = \nabla u_\sharp e^{-u}$

$u : \mathbb{R}^d \to \mathbb{R}$ convex

Monge Problem (1781)

The Monge problem[1] seeks a map $T : \mathbb{R}^d \to \mathbb{R}^d$ that transports μ on ν while minimizing the transport cost.

$$\mathscr{W}_2^2(\mu, \nu) := \inf_{\substack{T:\mathbb{R}^d \to \mathbb{R}^d \\ T_\# \mu = \nu}} \int_{\mathbb{R}^d} \frac{1}{2} \|x - T(x)\|^2 \mathrm{d}\mu(x)$$

Brenier's Theorem[2]

Let $u : \mathbb{R}^d \to \mathbb{R}$ be a convex function. For any μ, ∇u is the optimal $T^\star$ of Monge problem between μ and $\nabla u_\# \mu$.

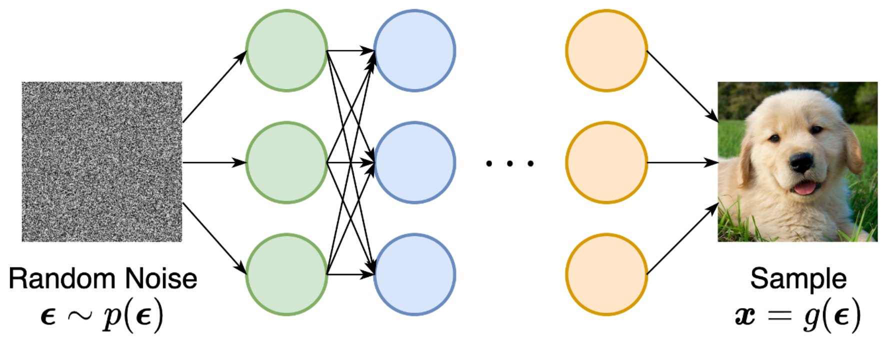
Link with Moment Measures: ∇u is the optimal $T^\star$ of Monge problem from $\mu = e^{-u}$ to $\nu = \rho$
as $\nabla u_\sharp e^{-u} = \rho$ and u is convex.

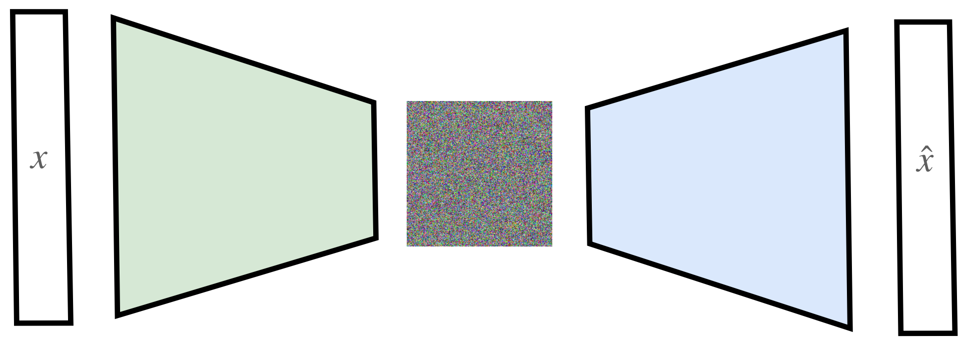
[1] Gaspard Monge. Mémoire sur la théorie des déblais et des remblais. *Histoire de l'Académie Royale des Sciences*, pages 666–704, 1781.

[2] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on pure and applied mathematic*, 44(4):375–417, 1991.

Contributions

Objective: Tie Sampling and Transport in the generative modeling process.





Moment measures factorization

Given a probability measure ρ, there exists a convex function $u : \mathbb{R}^d \to \mathbb{R}$ such that: $\boxed{\rho = \nabla u_\sharp e^{-u} \quad \bigstar}$

- We show why moment measures are unsuited for generative modeling.
- We demonstrate a new factorization: $\boxed{\text{Conjugate moment measure } \ \rho = \nabla w^*{}_\sharp e^{-w} \quad \blacklozenge}$
- We propose a method to estimate w of $\blacklozenge$ from samples $x_1, \ldots, x_n \sim \rho$.
- We tested our approach on generative tasks.

Moment measures

Moment measures

$\rho = \nabla u_\sharp e^{-u}$

$u : \mathbb{R}^d \to \mathbb{R}$ convex

Cordero-Erausquin and Klartag (2015) [1]

$$\min_{u\in\mathscr{C}(\mathbb{R}^d)} \int u^* \mathrm{d}\rho - \ln\left(\int e^{-u}\right)$$

$$\mathsf{KL}(\rho\|\mathfrak{P}_{u^*}) = \mathscr{E}(\rho) + \int u^*\mathrm{d}\rho + \ln\left(\int e^{-u^*}\right)$$

$$\mathfrak{P}_{u^*} = \frac{e^{-u^*}}{\int e^{-u^*}}$$

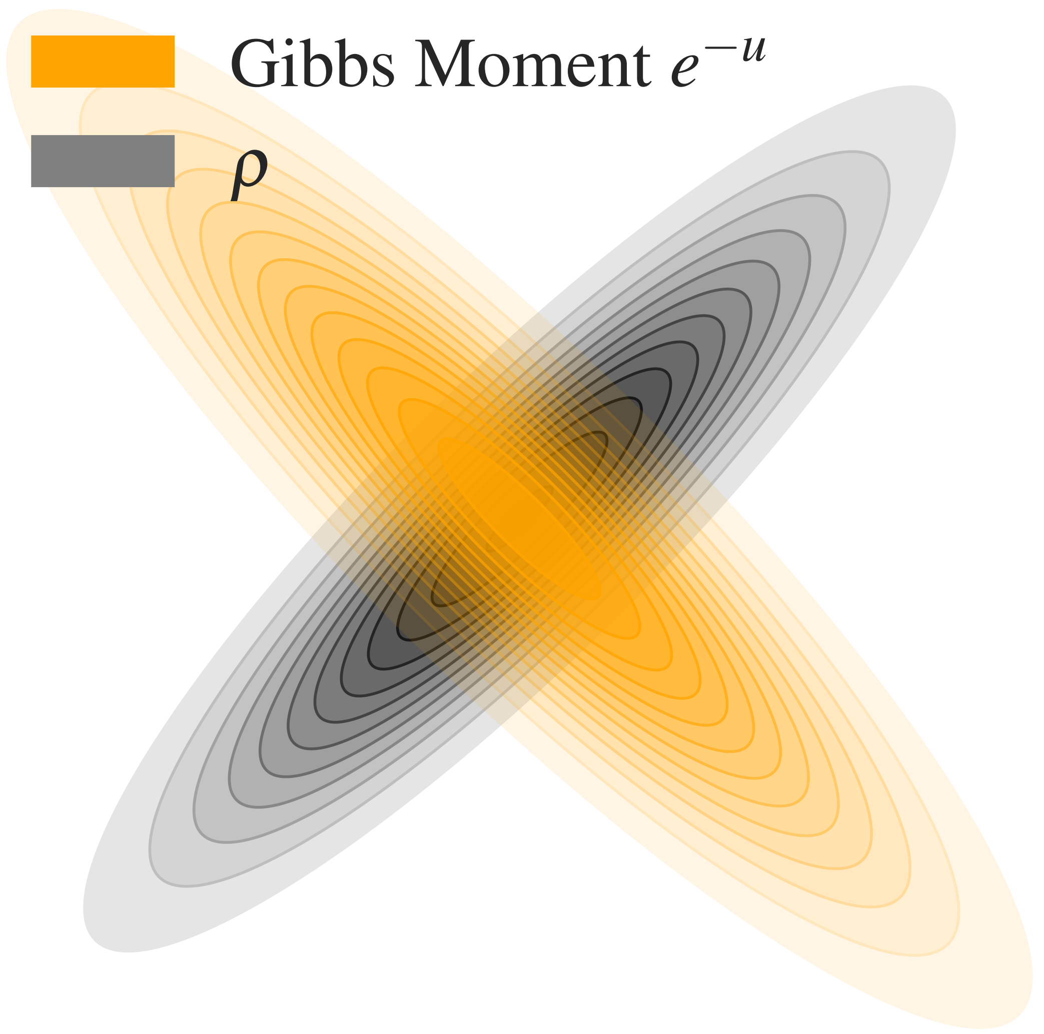
$$\min_{u\in\mathscr{C}(\mathbb{R}^d)} \mathsf{KL}(\rho\|\mathfrak{P}_{u^*}) - \ln\left(\int e^{-u}\int e^{-u^*}\right)$$

Legendre transform:

$$u^*(y) := \sup_{x\in\mathbb{R}^d} \langle x, y\rangle - u(x)$$

Proposition Gaussian case

Let $\rho = \mathcal{N}(0_{\mathbb{R}^d}, \Sigma)$. If Σ is non degenerate, the solutions of the moment measures factorisation are $u_m(x) = \frac{1}{2}(x-m)^T\Sigma(x-m)$ with $m \in \mathbb{R}^d$. The associated Gibbs distribution is $e^{-u_m} = \mathcal{N}(m, \Sigma^{-1})$.



$\rho = \mathcal{N}\left(0, \begin{bmatrix} 2 & 1.8 \\ 1.8 & 2 \end{bmatrix}\right)$

[1] Dario Cordero-Erausquin and Bo'az Klartag. Moment measures. Journal of Functional Analysis, 268(12):3834–3866, 2015

Moment measures

$\rho = \nabla u_\sharp e^{-u}$

$u : \mathbb{R}^d \to \mathbb{R}$ convex

Conjugate moment measures

Conjugate Moment Measures

$\rho = \nabla w^*{}_\sharp e^{-w}$

$w : \mathbb{R}^d \to \mathbb{R}$ convex

Conjugate Moment Measure Theorem

Let ρ be an **absolutely continuous** probability measure **supported on a compact, convex** set.

There exists a convex function w such that $\rho = \nabla w^*{}_\sharp e^{-w}$ $\bigstar$

Sketch of Proof

Using $\nabla w \circ \nabla w^* = i_d$, $\bigstar$ is equivalent to $\nabla w_\sharp \rho = e^{-w}$.

→ ∇w is the OT map from ρ to e^{-w}

→ w is the Brenier potential $\mathfrak{B}(\rho, e^{-w})$ from ρ to e^{-w}

→ w is a fixed point of $v \to \mathfrak{B}(\rho, e^{-v})$

We prove that $v \to \mathfrak{B}(\rho, e^{-v})$ admits a fixed point using Schauder's fixed point Theorem.

Moment measures

$\rho = \nabla u_\sharp e^{-u}$

$u : \mathbb{R}^d \to \mathbb{R}$ convex

Gaussian case

Conjugate Moment Measures

$\rho = \nabla {w^*}_\sharp e^{-w}$

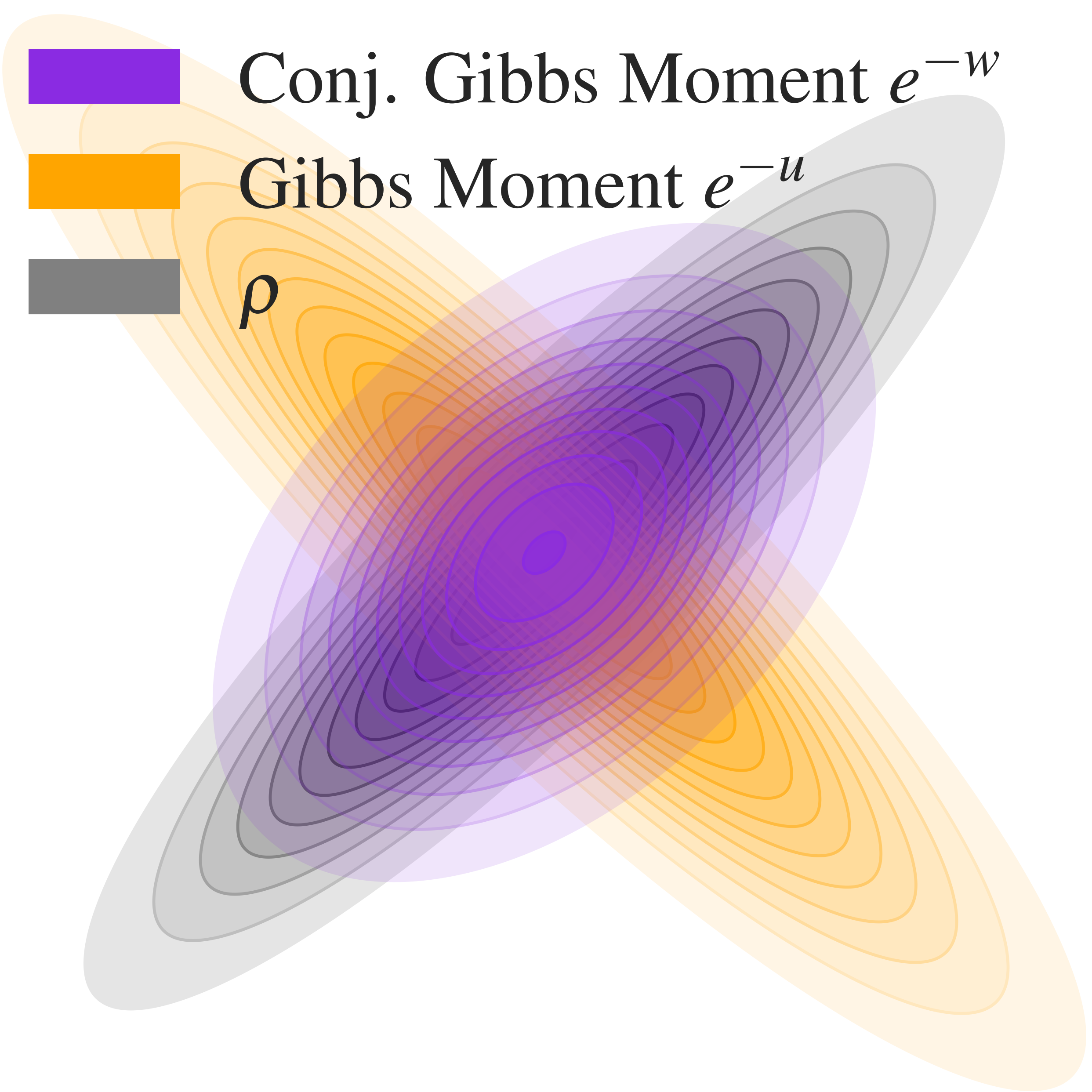
$w : \mathbb{R}^d \to \mathbb{R}$ convex

Proposition Gaussian case

Let $\rho = \mathcal{N}(0_{\mathbb{R}^d}, \Sigma)$. If Σ is non degenerate, the solutions of the moment measures factorisation are $u_m(x) = \frac{1}{2}(x-m)^T\Sigma(x-m)$ with $m \in \mathbb{R}^d$. The associated Gibbs distribution is $e^{-u_m} = \mathcal{N}(m, \Sigma^{-1})$.

Proposition Gaussian case

Let $\rho = \mathcal{N}(m, \Sigma)$. If Σ is non degenerate, a solution of the conjugate moment measures factorisation is $w(x) = \frac{1}{2}(x-r)^T\Sigma^{-1/3}(x-r)$ with $r = (I_d + \Sigma^{1/3})^{-1}m$. The associated Gibbs distribution is $e^{-w} = \mathcal{N}(r, \Sigma^{1/3})$.



Estimate w

Goal

Estimate w of $\bigstar$ from samples $x_1, \ldots, x_n \sim \rho$.

Fixed-point algorithm

Ideas:

- ∇w is the OT map from ρ to e^{-w}
- w is the Brenier potential $\mathfrak{B}(\rho, e^{-w})$ from ρ to e^{-w}
- w is a fixed point of $v \to \mathfrak{B}(\rho, e^{-v})$

Algorithm:

$$w_0 := \underbrace{\tfrac{1}{2}\|\cdot\|^2}_{e^{-w_0} = \mathcal{N}(0, I_d)}; \quad \forall t \geq 1, \quad w_{t+1} := \mathfrak{B}(\rho, e^{-w_t})$$

Estimate w

Conjugate Moment Measures

$\rho = \nabla w^*{}_\sharp e^{-w} \quad \bigstar$

$w : \mathbb{R}^d \to \mathbb{R}$ convex

Goal

Estimate w of $\bigstar$ from samples $x_1, \ldots, x_n \sim \rho$.

Fixed-point algorithm

Ideas:

- ∇w is the OT map from ρ to e^{-w}
- w is the Brenier potential $\mathfrak{B}(\rho, e^{-w})$ from ρ to e^{-w}
- w is a fixed point of $v \to \mathfrak{B}(\rho, e^{-v})$

Algorithm:

$$w_0 := \frac{1}{2}\|\cdot\|^2; \quad \forall t \geq 1, \quad w_{t+1} := \mathfrak{B}(\rho, e^{-w_t})$$

$$e^{-w_0} = \mathcal{N}(0, I_d)$$

1D case

Theorem: The OT map between density functions μ and ν is $\nabla \mathfrak{B}(\mu, \nu) = C_\nu^{-1} \circ C_\mu$.

with $C_\mu : \mathbb{R} \to [0,1]$ the cumulative distribution function $C_\mu(x) := \int_{-\infty}^x \mathrm{d}\mu$ and $C_\mu^{-1} : [0,1] \to \mathbb{R} \cup \{-\infty\}$ the quantile function $C_\mu^{-1}(r) := \min\{x \in \mathbb{R} \cup \{-\infty\} : C_\mu(x) \geq r\}$.

Moment measures

$\rho = \nabla u_{\sharp} e^{-u}$

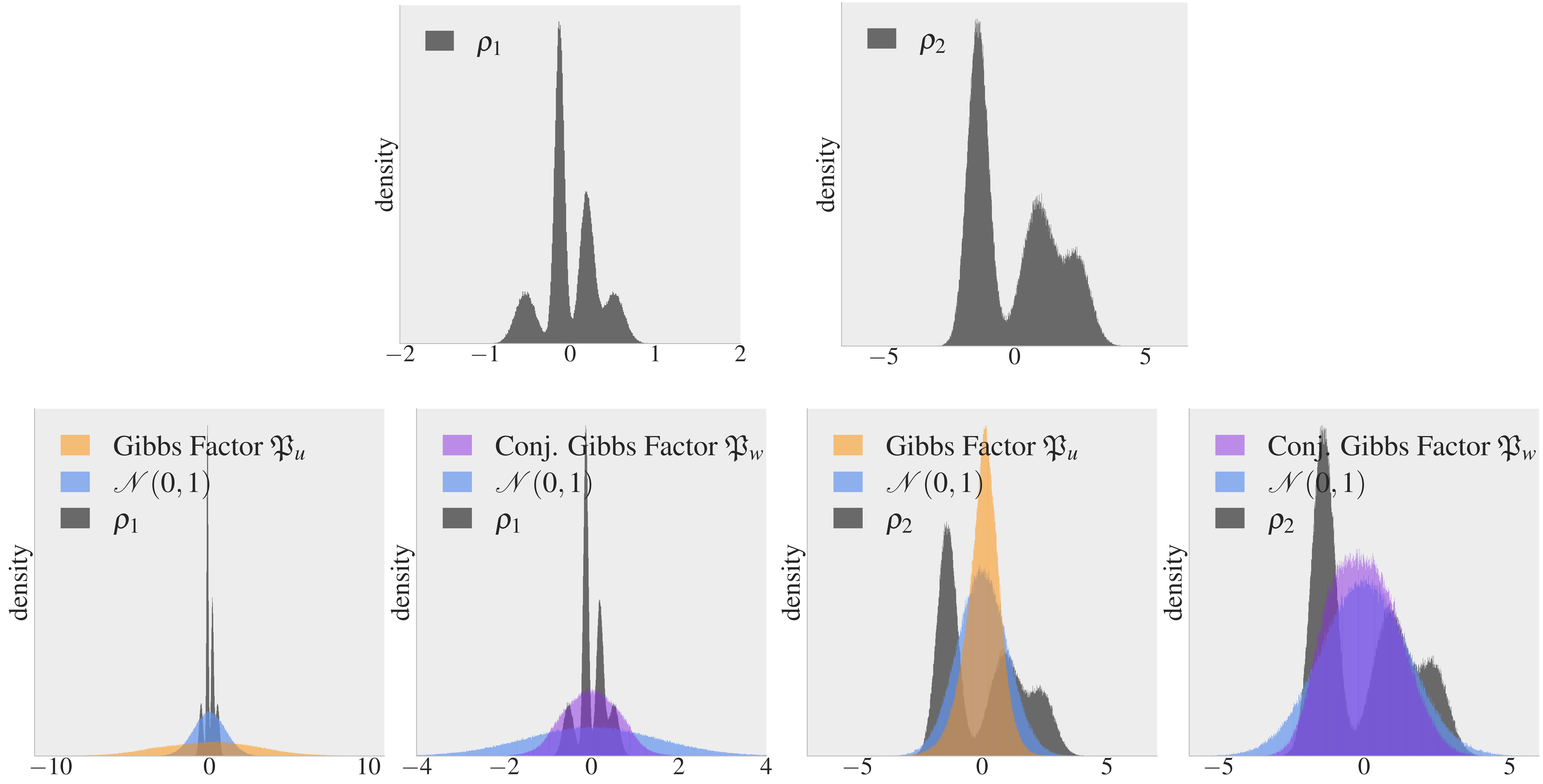
$u : \mathbb{R}^d \to \mathbb{R}$ convex

Estimate w

Conjugate Moment Measures

$\rho = \nabla w^*{}_{\sharp} e^{-w}$

$w : \mathbb{R}^d \to \mathbb{R}$ convex



Estimate w when $d > 1$

Conjugate Moment Measures

$\rho = \nabla w^*{}_\sharp e^{-w}$

$w : \mathbb{R}^d \to \mathbb{R}$ convex

Algorithm: $w_0 := \frac{1}{2}\|\cdot\|^2; \quad \forall t \geq 1, \quad w_{t+1} := \mathfrak{B}(\rho, e^{-w_t})$

Methodology:

- Parameterize w with an input convex neural network (ICNN)[1] w_θ.
- Use an optimal transport solver[2] to estimate $\mathfrak{B}(\rho, e^{-w_\theta})$.

[1] Brandon Amos, Lei Xu, and J Zico Kolter. Input convex neural networks. In International Conference on Machine Learning, pages 146–155. PMLR, 2017.

[2] Brandon Amos. On amortizing convex conjugates for optimal transport. In The Eleventh International Conference on Learning Representations, 2023.

Estimate w when $d > 1$

Conjugate Moment Measures

$\rho = \nabla w^*{}_\sharp e^{-w}$

$w : \mathbb{R}^d \to \mathbb{R}$ convex

Algorithm: $w_0 := \frac{1}{2}\|\cdot\|^2; \quad \forall t \geq 1, \quad w_{t+1} := \mathfrak{B}(\rho, e^{-w_t})$

Methodology:

- Parameterize w with an input convex neural network (ICNN)[1] w_θ. $\approx$ MLP with non negative weights + convex activation functions
- Use an optimal transport solver[2] to estimate $\mathfrak{B}(\rho, e^{-w_\theta})$.

[1] Brandon Amos, Lei Xu, and J Zico Kolter. Input convex neural networks. In International Conference on Machine Learning, pages 146–155. PMLR, 2017.

[2] Brandon Amos. On amortizing convex conjugates for optimal transport. In The Eleventh International Conference on Learning Representations, 2023.

Estimate w

Algorithm: $w_0 := \frac{1}{2}\|\cdot\|^2; \quad \forall t \geq 1, \quad w_{t+1} := \mathfrak{B}(\rho, e^{-w_t})$

Methodology:
- Parameterize w with an input convex neural network (ICNN) w_θ.
- Use an optimal transport solver[2] to estimate $\mathfrak{B}(\rho, e^{-w_\theta})$.

From theory ...

$\mathfrak{B}(\mu, \nu)$ is the solution of the dual objective:

$$\mathfrak{B}(\mu, \nu) \in \underset{f \in CVX(\mathbb{R}^d)}{\arg\inf} \int_{\mathbb{R}^d} f \, d\mu + \int_{\mathbb{R}^d} f^* \, d\nu$$

where the f^* is the convex conjugate of f:

$$f^*(y) := \sup_{x \in \mathbb{R}^d} \langle x, y \rangle - f(x) = \langle x^\star, y \rangle - f(x^\star) \, .$$

with $x^\star = \nabla f^*(y)$ **Danskin. Theorem**[1]

[1] John M Danskin. The theory of max-min, with applications. SIAM Journal on Applied Mathematics, 14(4):641–664, 1966.

[2] Brandon Amos. On amortizing convex conjugates for optimal transport. In The Eleventh International Conference on Learning Representations, 2023.

Estimate w

Conjugate Moment Measures

$\rho = \nabla w^*{}_\sharp e^{-w}$

$w : \mathbb{R}^d \to \mathbb{R}$ convex

Algorithm: $w_0 := \frac{1}{2}\|\cdot\|^2; \quad \forall t \geq 1, \quad w_{t+1} := \mathfrak{B}(\rho, e^{-w_t})$

Methodology:
- Parameterize w with an input convex neural network (ICNN) w_θ.
- Use an optimal transport solver[2] to estimate $\mathfrak{B}(\rho, e^{-w_\theta})$.

From theory ...

$\mathfrak{B}(\mu,\nu)$ is the solution of the dual objective:

$$\mathfrak{B}(\mu,\nu) \in \underset{f \in CVX(\mathbb{R}^d)}{\arg\inf} \int_{\mathbb{R}^d} f \,\mathrm{d}\mu + \int_{\mathbb{R}^d} f^* \,\mathrm{d}\nu$$

where the f^* is the convex conjugate of f:

$$f^*(y) := \sup_{x \in \mathbb{R}^d} \langle x, y\rangle - f(x) = \langle x^\star, y\rangle - f(x^\star)\,.$$

with $x^\star = \nabla f^*(y)$ **Danskin. Theorem**[1]

... To practice

A surrogate MLP V_ϕ is used to approximate ∇w_θ^*

w_θ and V_ϕ are optimized using:

$$\mathscr{L}_{\text{Monge}}(\theta) = \frac{1}{n}\sum_{i=1}^{n} w_\theta(x_i) + \frac{1}{n}\sum_{j=1}^{n} \langle V_\phi(y_j), y_j\rangle - w_\theta(V_\phi(y_j))$$

$$\mathscr{L}_{\text{convex-dual}}(\phi) = \frac{1}{n}\sum_{j=1}^{n} \|V_\phi(y_j) - \boxed{\nabla w_\theta^*}(y_j)\|^2$$

Computed with a conjugate solver

[1] John M Danskin. The theory of max-min, with applications. SIAM Journal on Applied Mathematics, 14(4):641–664, 1966.

[2] Brandon Amos. On amortizing convex conjugates for optimal transport. In The Eleventh International Conference on Learning Representations, 2023.

Estimate w / Sample from ρ

Conjugate Moment Measures

$\rho = \nabla w^*{}_\sharp e^{-w}$

$w : \mathbb{R}^d \to \mathbb{R}$ convex

Algorithm: $w_0 := \frac{1}{2}\|\cdot\|^2; \quad \forall t \geq 1, \quad w_{t+1} := \mathfrak{B}(\rho, e^{-w_t})$

Estimate w / Sample from ρ

Conjugate Moment Measures

$\rho = \nabla w^*{}_\sharp e^{-w}$

$w : \mathbb{R}^d \to \mathbb{R}$ convex

Algorithm: $w_0 := \frac{1}{2}\|\cdot\|^2; \quad \forall t \geq 1, \quad w_{t+1} := \mathfrak{B}(\rho, e^{-w_t})$

Estimate w with w_θ

Algorithm

Initialize w_θ such that $w_\theta \approx \frac{1}{2}\|\cdot\|^2$ → Initialization $w_0 := \frac{1}{2}\|\cdot\|^2$

while not converged **do**

Draw n i.i.d samples $x_i \sim \rho$ → Sample from ρ

Draw $y_1, \ldots, y_n \sim e^{-w_\theta}$ using LMC algorithm → Sample from e^{-w_t} with adaptative γ

$\mathcal{L}_\theta \leftarrow \frac{1}{n}\sum_{i=1}^n w_\theta(x_i) - \frac{1}{n}\sum_{i=1}^n w_\theta(\tilde{x}(y_i))$

Update w_θ with $\nabla \mathcal{L}_\theta$ → One gradient step to estimate $\mathfrak{B}(\rho, e^{-w_t})$

end while

Sample from e^{-w_θ} using Langevin Monte Carlo (LMC) $\quad x^{(k+1)} = x^{(k)} - \gamma \nabla w_\theta(x^{(k)}) + \sqrt{2\gamma} z^{(k)}, \quad z^{(k)} \sim \mathcal{N}(0, I_d)$

Estimate w / Sample from ρ

Conjugate Moment Measures

$\rho = \nabla w^*{}_\sharp e^{-w}$

$w : \mathbb{R}^d \to \mathbb{R}$ convex

Algorithm: $w_0 := \frac{1}{2}\|\cdot\|^2;\quad \forall t \geq 1,\quad w_{t+1} := \mathfrak{B}(\rho, e^{-w_t})$

Estimate w with w_θ

Algorithm

Initialize w_θ such that $w_\theta \approx \frac{1}{2}\|\cdot\|^2$ → Initialization $w_0 := \frac{1}{2}\|\cdot\|^2$

while not converged **do**

Draw n i.i.d samples $x_i \sim \rho$ → Sample from ρ

Draw $y_1, \ldots, y_n \sim e^{-w_\theta}$ using LMC algorithm → Sample from e^{-w_t}

$$\left.\begin{array}{l} \mathcal{L}_\theta \leftarrow \frac{1}{n}\sum_{i=1}^n w_\theta(x_i) - \frac{1}{n}\sum_{i=1}^n w_\theta(\tilde{x}(y_i)) \\ \text{Update } w_\theta \text{ with } \nabla \mathcal{L}_\theta \end{array}\right\}$$

→ One gradient step to estimate $\mathfrak{B}(\rho, e^{-w_t})$

end while

Sample from $\rho \approx \nabla w_\theta^*{}_\sharp e^{-w_\theta}$

Sample from e^{-w_θ} using Langevin Monte Carlo (LMC) $\quad x^{(k+1)} = x^{(k)} - \gamma \nabla w_\theta(x^{(k)}) + \sqrt{2\gamma} z^{(k)},\ \ z^{(k)} \sim \mathcal{N}(0, I_d)$

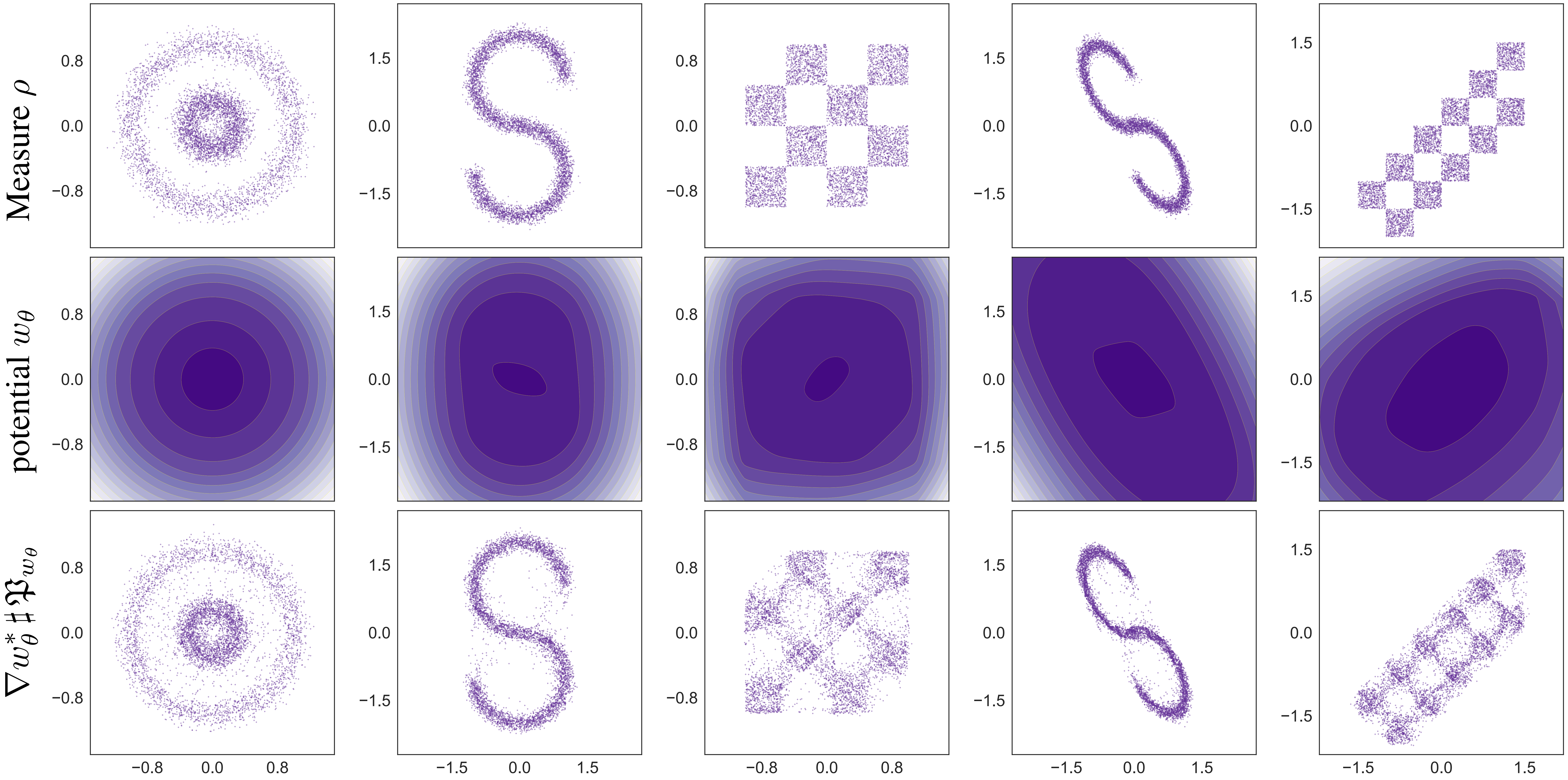
Transport using ∇w_θ^* with a gradient ascent algorithm $\quad \nabla w_\theta^*(y) = \arg\sup_x \langle x, y\rangle - w_\theta(x)$

2D Experiments

Conjugate Moment Measures

$\rho = \nabla w^*{}_\sharp e^{-w}$

$w : \mathbb{R}^d \to \mathbb{R}$ convex



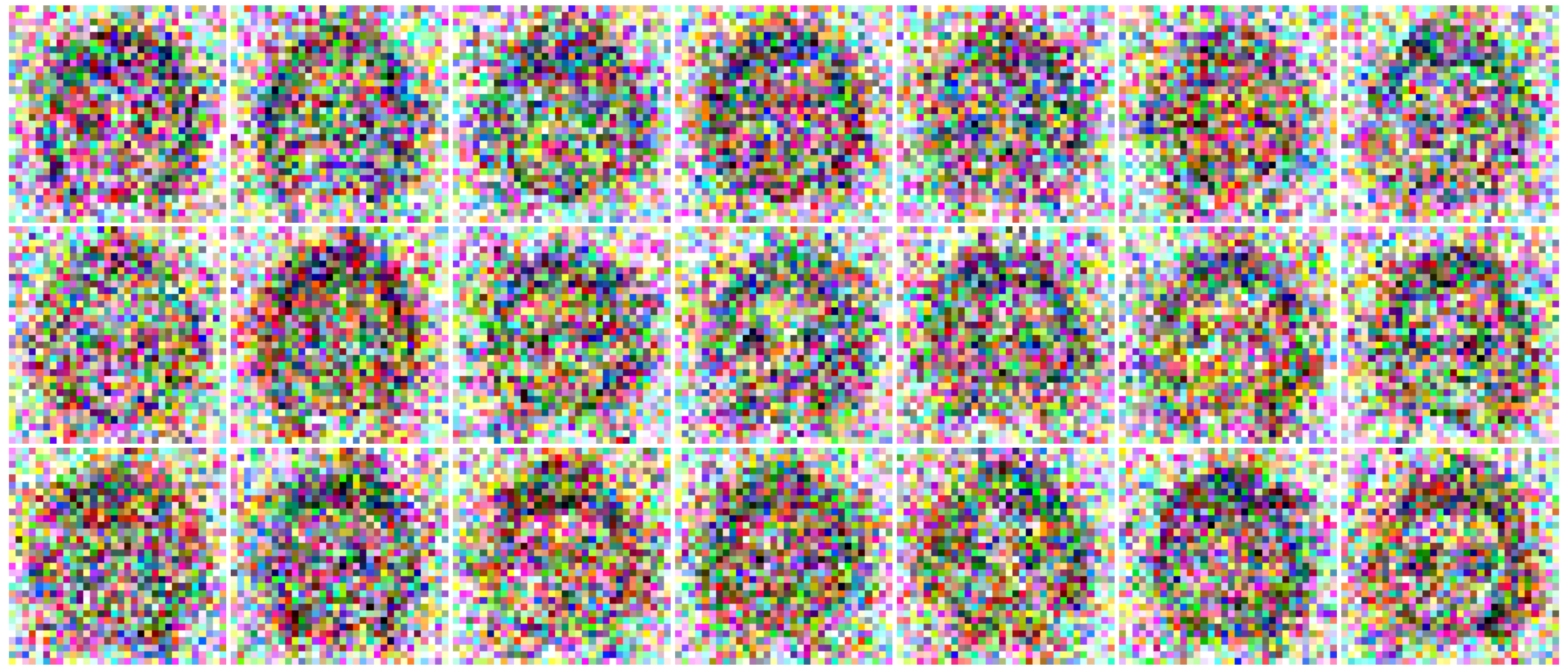
Cartoon Experiments $d = 3072$

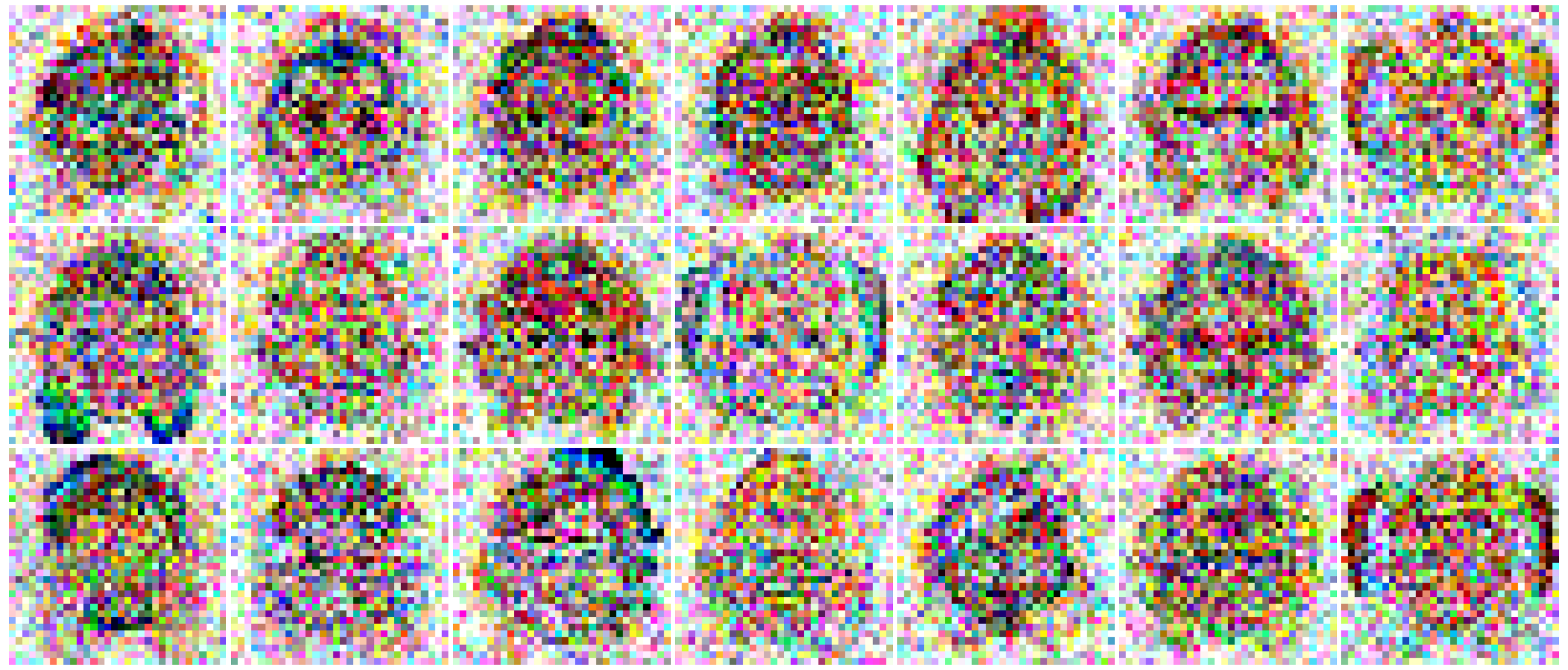
Conjugate Moment Measures

$\rho = \nabla w^*{}_\sharp e^{-w}$

$w : \mathbb{R}^d \to \mathbb{R}$ convex

Gibbs $\mathfrak{P}_{w_\theta}$ obtained from Langevin dynamics $\approx$ Gibbs $\mathfrak{P}_{w_\theta}$ obtained from $\nabla w_\theta \sharp \rho$





Conjugate $\nabla w_\theta^*(y) = \arg\sup_x \langle x, y\rangle - w(x)$

Transportation map $\nabla w_\theta = (\nabla w_\theta^*)^{-1}$

Generated images $\nabla w_\theta^* \sharp \mathfrak{P}_{w_\theta}$ $\approx$ Data distribution ρ

Monge-Ampère equation

Probability measures

$$\rho = \nabla w^*{}_\sharp e^{-w}$$

Densities

$$\rho(x) = e^{-\mathcal{E}_w(x)} \qquad \mathcal{E}_w(x) = w(\nabla w(x)) - \ln(\det H_w(x)) + \ln(C_w)$$

Monge-Ampère equation

Conjugate Moment Measures

$\rho = \nabla w^*{}_\sharp e^{-w}$

$w : \mathbb{R}^d \to \mathbb{R}$ convex

Probability measures

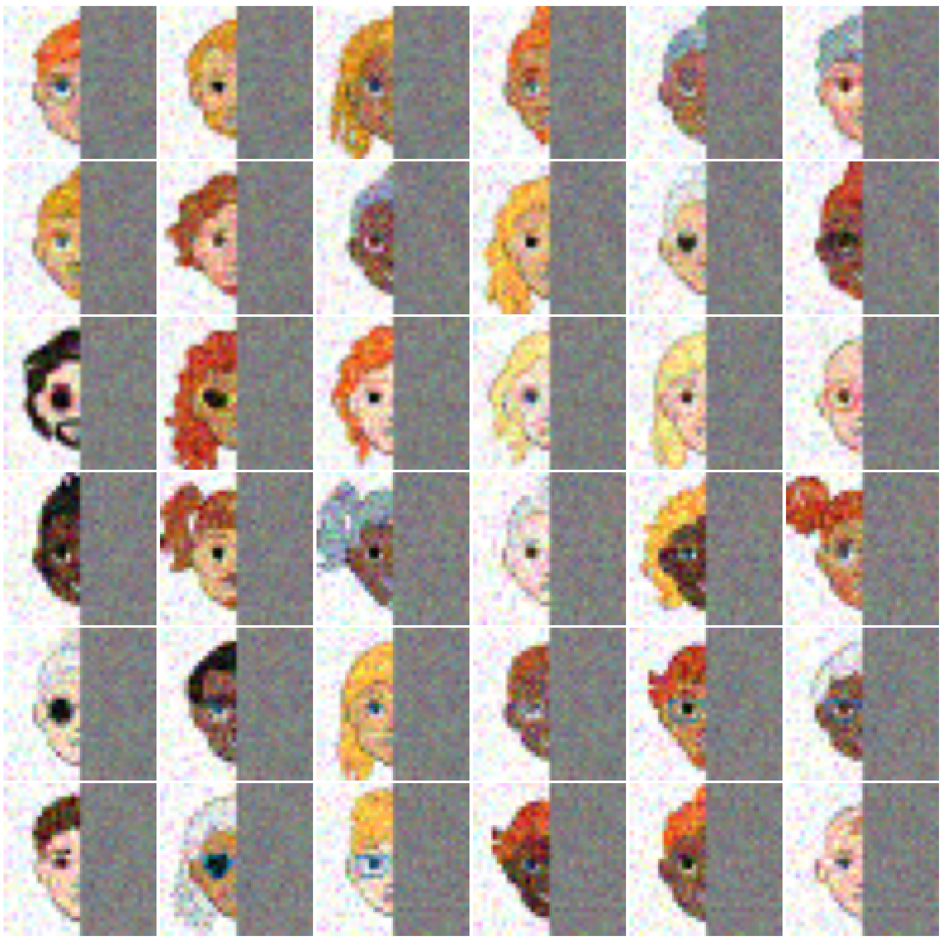
$$\rho = \nabla w^*{}_\sharp e^{-w}$$

Densities

$$\rho(x) = e^{-\mathscr{E}_w(x)} \qquad \mathscr{E}_w(x) = w(\nabla w(x)) - \ln(\det H_w(x)) + \ln(C_w)$$

Applications

Image reconstruction



Monge-Ampère equation

Conjugate Moment Measures

$\rho = \nabla w^*{}_\sharp e^{-w}$

$w : \mathbb{R}^d \to \mathbb{R}$ convex

Probability measures

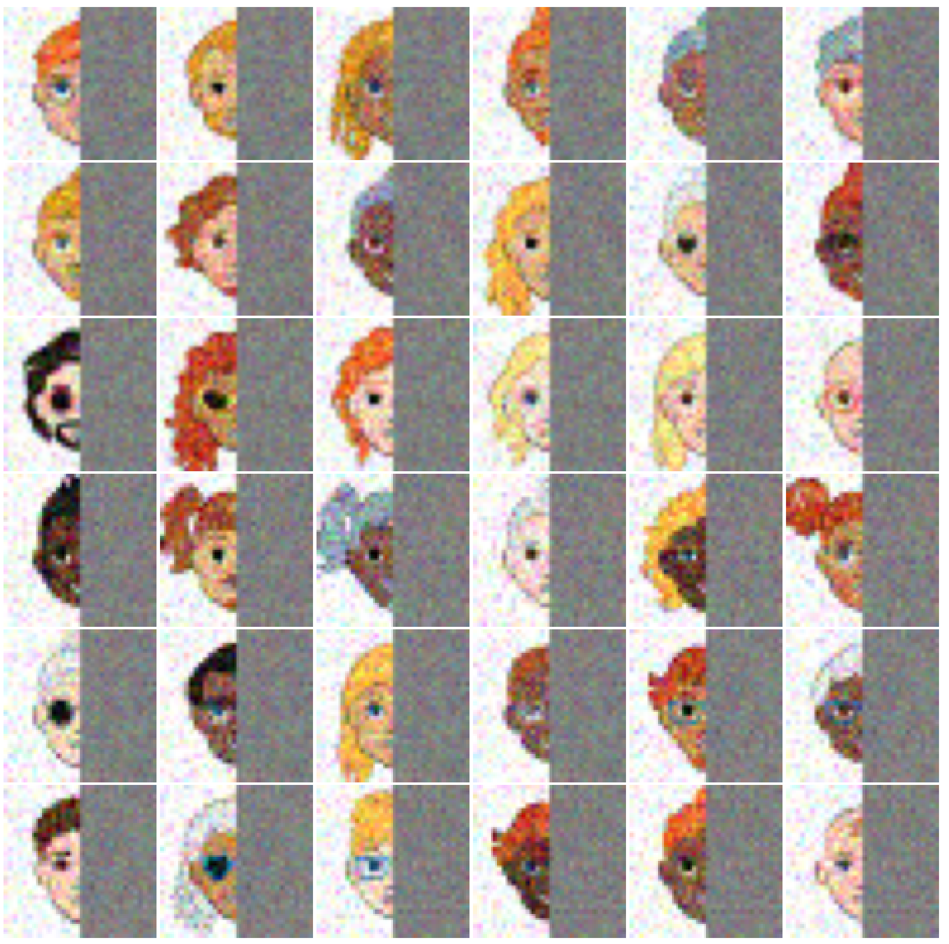
$$\rho = \nabla w^*{}_\sharp e^{-w}$$

Densities

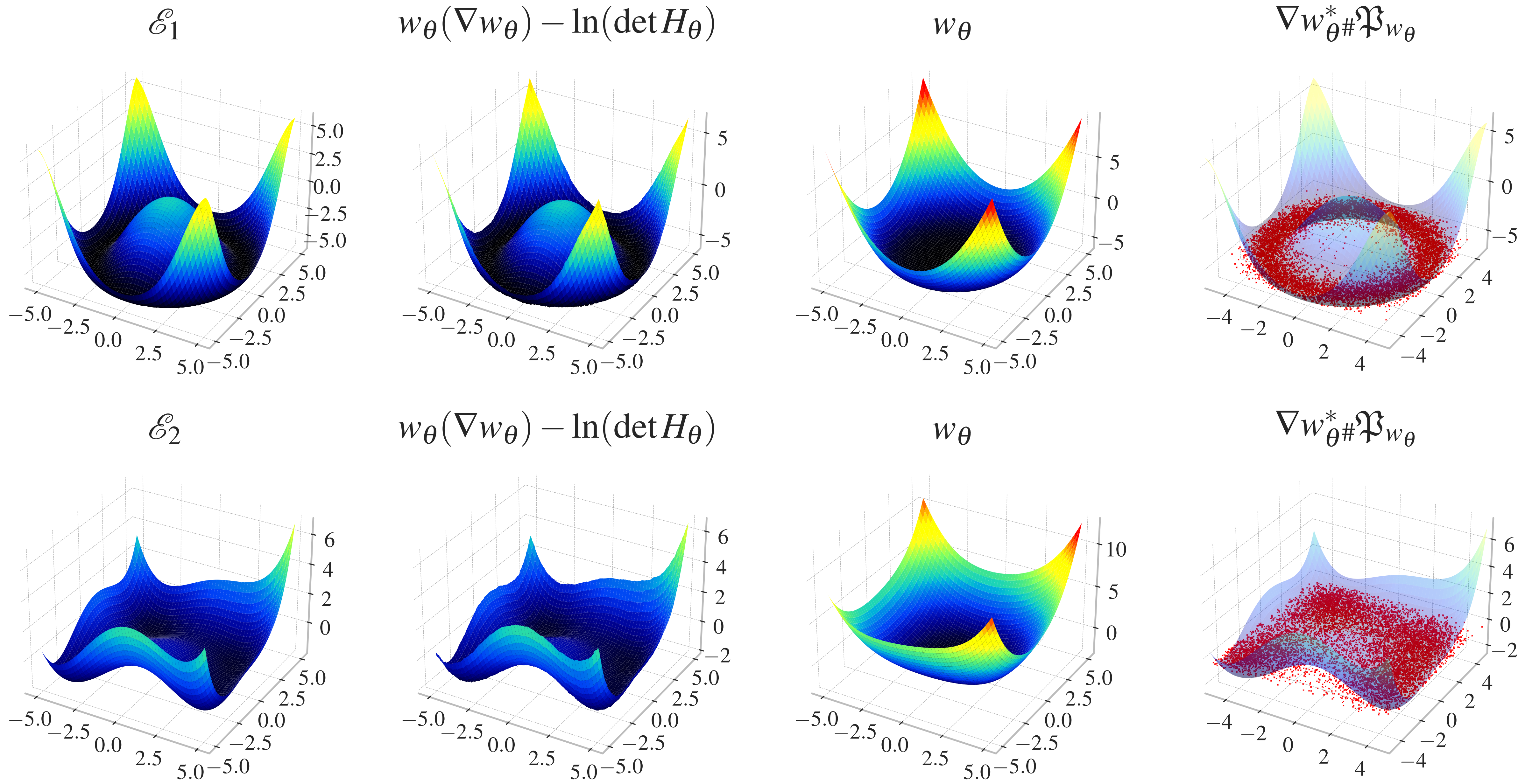
$$\rho(x) = e^{-\mathscr{E}_w(x)} \qquad \mathscr{E}_w(x) = w(\nabla w(x)) - \ln(\det H_w(x)) + \ln(C)$$

Applications

Image reconstruction



Learn w when ρ is known up to a normalizing constant



In this work

- We introduce a new factorization $\rho = \nabla w^*{}_\sharp e^{-w}$ ★
- We propose a method to estimate w when ρ is known either through samples or up to a normalizing constant
- We apply it to generative modeling

Next

- Extend theoretical results of ★ (unicity of w?, relax assumptions)
- Scale the method to more challenging datasets

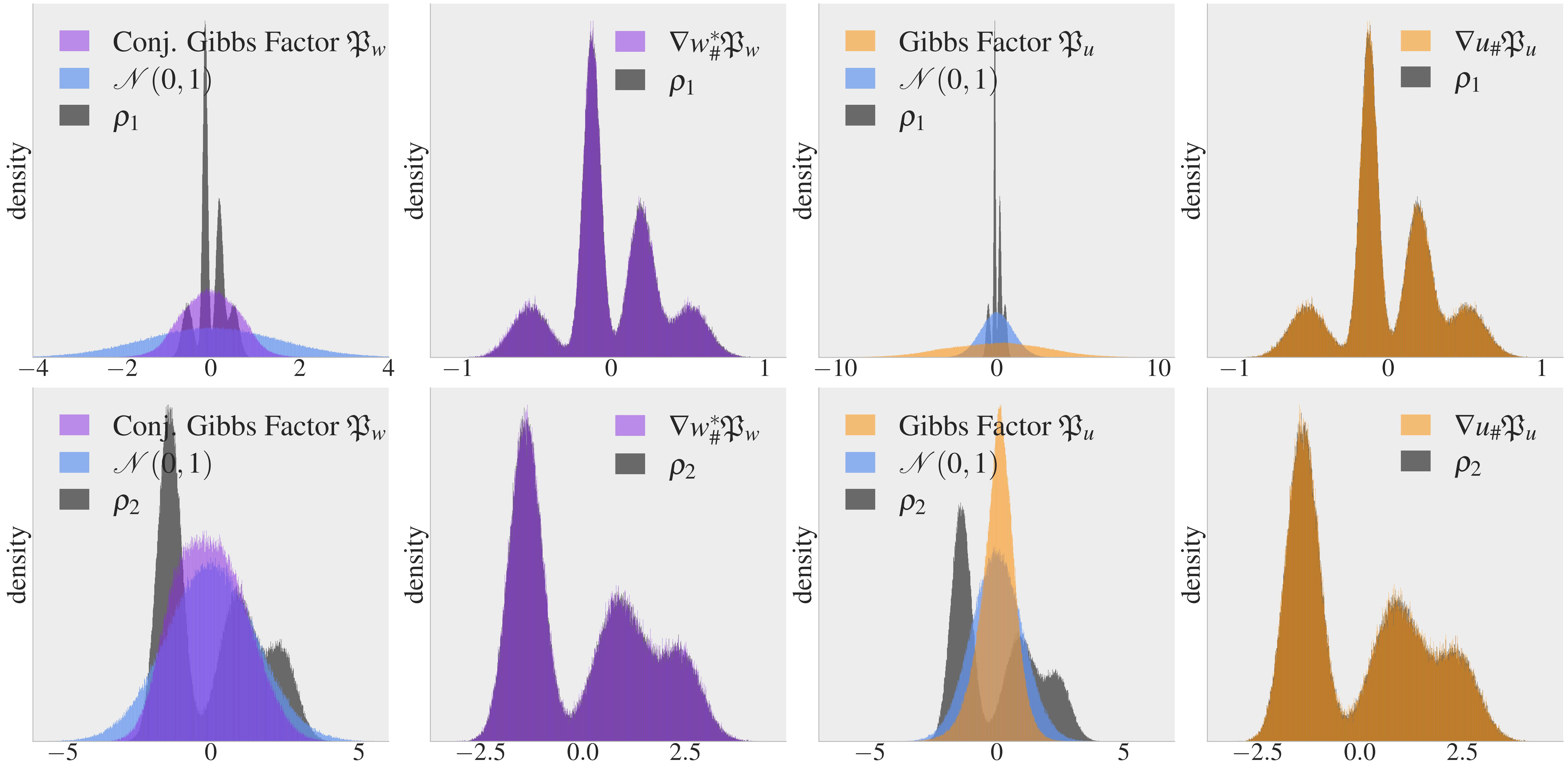
Thank You !

Estimate w

Conjugate Moment Measures

$\rho = \nabla w^*{}_\sharp e^{-w}$

$w : \mathbb{R}^d \to \mathbb{R}$ convex



Conjugate moment measures

Conjugate Moment Measure Theorem

Let ρ be an **absolutely continuous** probability measure **supported on a compact, convex** set.
There exists a convex function w such that $\rho = \nabla w^* \sharp\, e^{-w}$ ★

Sketch of Proof

Using $\nabla w \circ \nabla w^* = i_d$, ★ is equivalent to $\nabla w_\sharp \rho = e^{-w}$.

→ ∇w is the OT map from ρ to e^{-w}

→ w is the Brenier potential $\mathfrak{B}(\rho, e^{-w})$ from ρ to e^{-w}

→ w is a fixed point of $v \to \mathfrak{B}(\rho, e^{-v})$

Conjugate moment measures

Conjugate Moment Measure Theorem

Let ρ be an **absolutely continuous** probability measure **supported on a compact, convex** set.
There exists a convex function w such that $\rho = \nabla w^* \sharp\, e^{-w}$ ★

Sketch of Proof

Using $\nabla w \circ \nabla w^* = i_d$, ★ is equivalent to $\nabla w_\sharp \rho = e^{-w}$.

→ ∇w is the OT map from ρ to e^{-w}

→ w is the Brenier potential $\mathfrak{B}(\rho, e^{-w})$ from ρ to e^{-w}

→ w is a fixed point of $v \to \mathfrak{B}(\rho, e^{-v})$

We define $\mathfrak{P}_v^\Omega = \dfrac{1_\Omega e^{-v}}{\int_\Omega e^{-v}}$ where Ω is the compact, convex support of ρ and $G_\rho^\Omega(v) := \mathfrak{B}(\rho, \mathfrak{P}_v^\Omega)$.

We show that G_ρ^Ω admits a fixed point v and w defined as $w(x) := \begin{cases} v(x) & \text{if } x \in \Omega \\ +\infty & \text{else} \end{cases}$ solves ★ .

Conjugate moment measures

Conjugate Moment Measures

$\rho = \nabla w^*{}_\sharp e^{-w}$

$w : \mathbb{R}^d \to \mathbb{R}$ convex

Schauder's fixed point Theorem

Let $(X, \|.\|)$ be a Banach space and $\mathscr{M} \subset X$ is compact, convex, and nonempty. Any continuous operator $A : \mathscr{M} \to \mathscr{M}$ has at least one fixed point.

In this proof

- $(X, \|.\|) = (C(\Omega), \| \|_\infty)$
- $\mathscr{M} = \{f \in C(\Omega)$ such that $\forall x, y \in \Omega, |f(x) - f(y)| \leq R\|x - y\|_2$ and $f(0_{\mathbb{R}^d}) = 0\}$

  with R such that, $\Omega \subset \mathscr{B}(0,R) = \{x \in \mathbb{R}^d, \|x\|_2 \leq R\}$
- $A = G_\rho^\Omega : v \to \mathfrak{B}(\rho, \mathfrak{P}_v^\Omega)$ with $\mathfrak{P}_v^\Omega = \dfrac{1_\Omega e^{-v}}{\int_\Omega e^{-v}}$.

Conjugate moment measures

Conjugate Moment Measures

$\rho = \nabla w^*{}_\sharp e^{-w}$

$w : \mathbb{R}^d \to \mathbb{R}$ convex

Schauder's fixed point Theorem

Let $(X, \|.\|)$ be a Banach space and $\mathscr{M} \subset X$ is compact, convex, and nonempty. Any continuous operator $A : \mathscr{M} \to \mathscr{M}$ has at least one fixed point.

In this proof

- $(X, \|.\|) = (C(\Omega), \|\ \|_\infty)$
- $\mathscr{M} = \{f \in C(\Omega)$ such that $\forall x, y \in \Omega, |f(x) - f(y)| \leq R\|x - y\|_2$ and $f(0_{\mathbb{R}^d}) = 0\}$

  with R such that, $\Omega \subset \mathscr{B}(0,R) = \{x \in \mathbb{R}^d, \|x\|_2 \leq R\}$
- $A = G_\rho^\Omega : v \to \mathfrak{B}(\rho, \mathfrak{P}_v^\Omega)$ with $\mathfrak{P}_v^\Omega = \dfrac{1_\Omega e^{-v}}{\int_\Omega e^{-v}}$.

$\mathscr{M}$ is a non-empty, compact, convex set of X ✔ (Arzela-Ascoli theorem for compactness)

$A : \mathscr{M} \to \mathscr{M}$ The OT map $\nabla \mathfrak{B}(\rho, \mathfrak{P}_v^\Omega)$ transports ρ on $\mathfrak{P}_v^\Omega$ supported on $\Omega \subset \mathscr{B}(0,R) \implies \|\nabla \mathfrak{B}(\rho, \mathfrak{P}_v^\Omega)\| \leq R$

$\implies \mathfrak{B}(\rho, \mathfrak{P}_v^\Omega)$ is R Lipschitz

$A = G_\rho^\Omega$ is continuous Based on the Theorem: $\nu_n \xrightarrow{\mathbb{P}} \nu \implies \mathfrak{B}(\rho, \nu_n) \xrightarrow{\text{unif.}} \mathfrak{B}(\rho, \nu)$