Stochastic Control and Schrödinger Bridges Tutorial

Daniel D. Lee

CORNELL TECH

KIAS KOREA INSTITUTE FOR ADVANCED STUDY

Topics

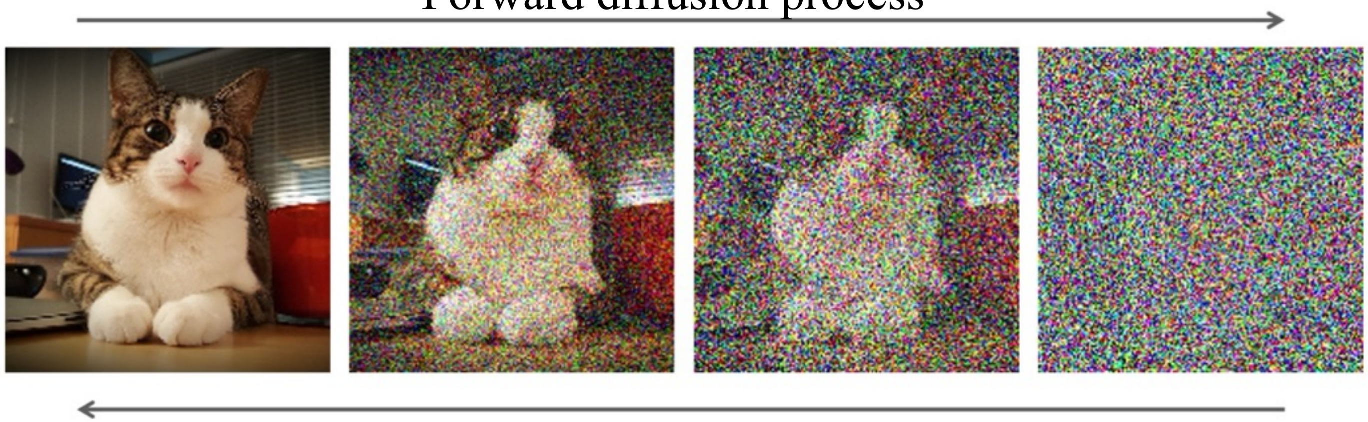
- Introduction to distributions and flows
- Optimal control and dynamic programming
- Stochastic systems
- Connection to optimal transport
- Schrödinger bridges

Caveats

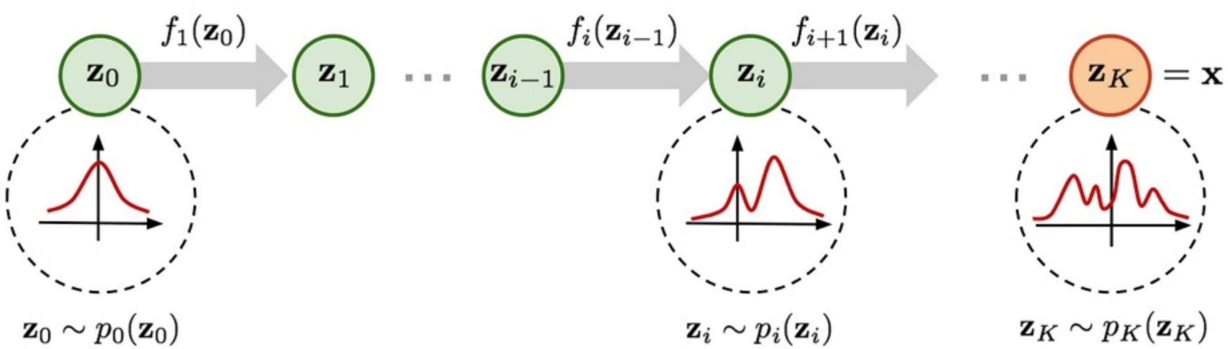
- Very active ML research area
- Can only cover a few basic topics
- My ignorance about other topics
- Physicist POV, sacrificing mathematical rigor for intuition
- Additional resources at KIAS,
  e.g. Dohyun Kwon, Yungkyun Noh, …

Success of Generative AI

Forward diffusion process

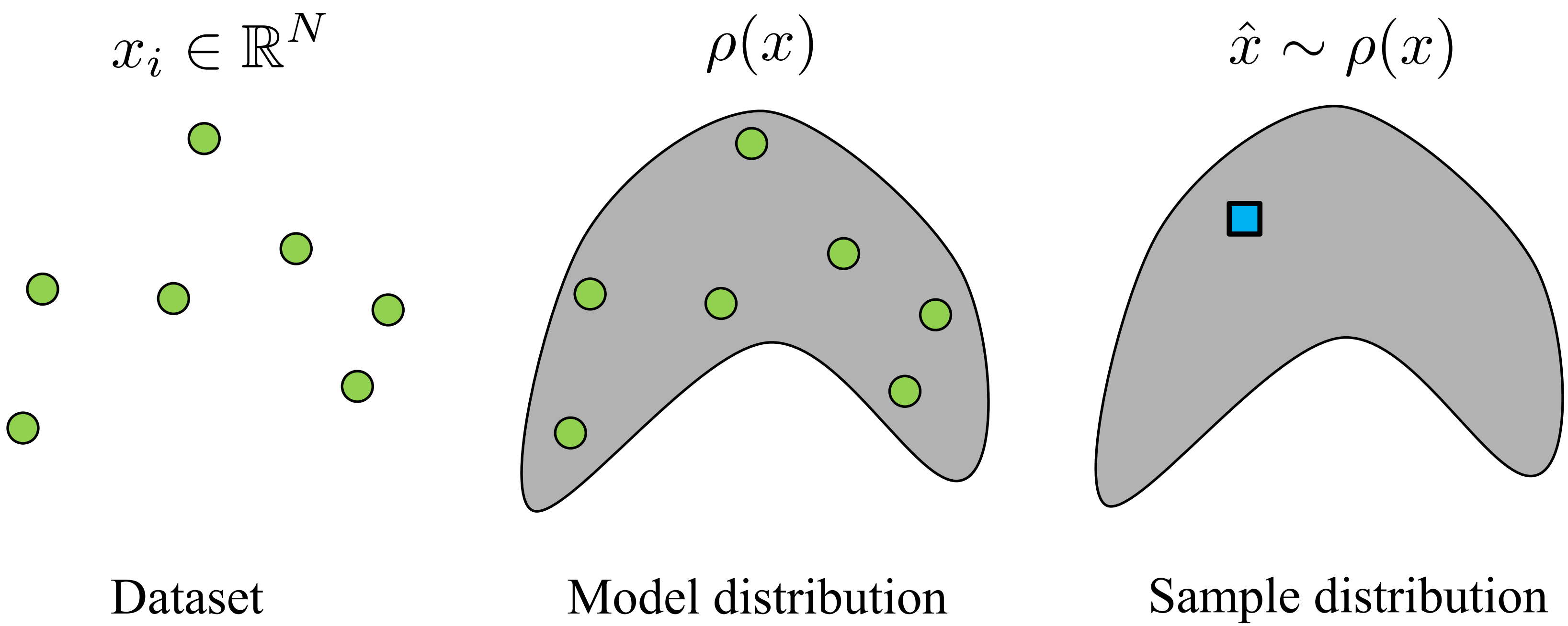


Reverse diffusion process



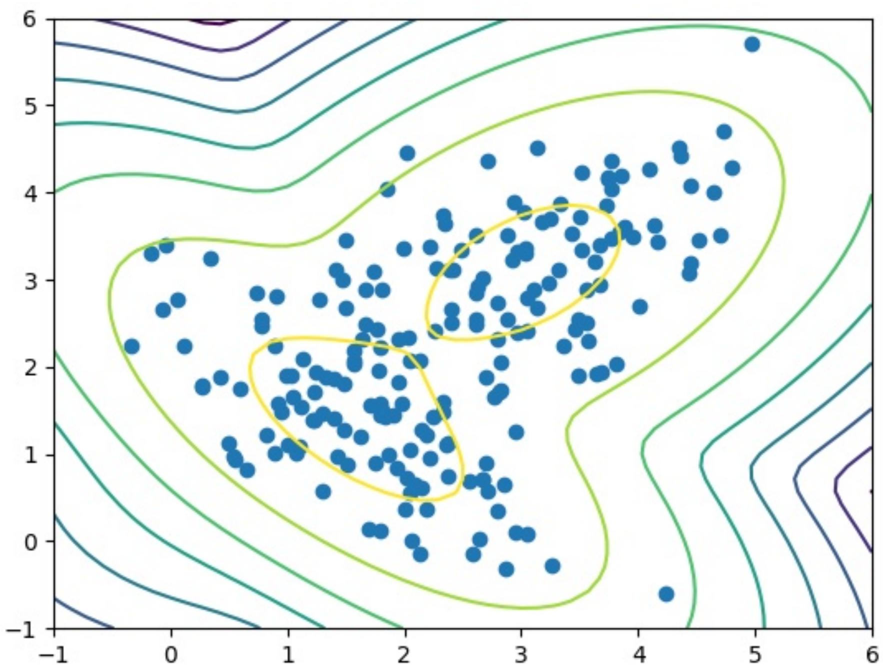
Normalizing flow models

Gen AI paradigm



- AI approaches to model data distribution and generate new samples as opposed to discriminative models

Data density distributions



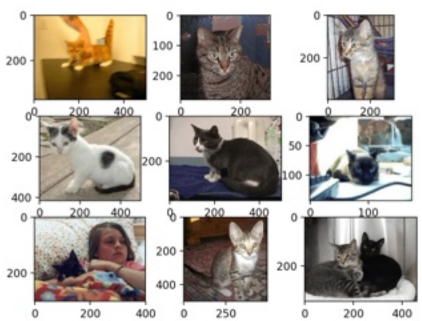
$$x \in \mathbb{R}^N$$

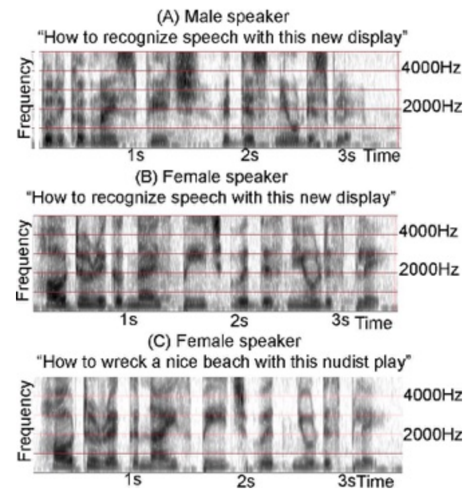
Probability density function

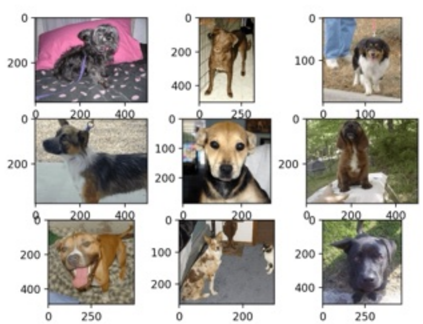
$$\rho(x) \geq 0$$

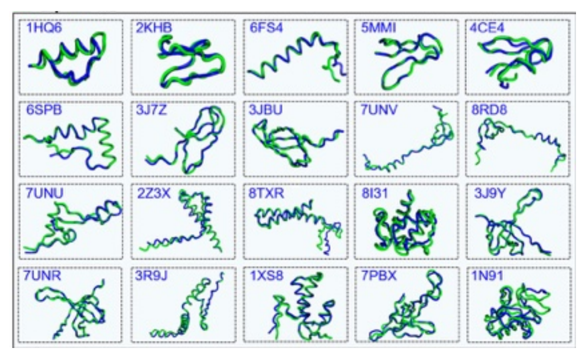
- Analyzing data distributions in high dimensional feature spaces is critical to AI and ML

Multiple data distributions



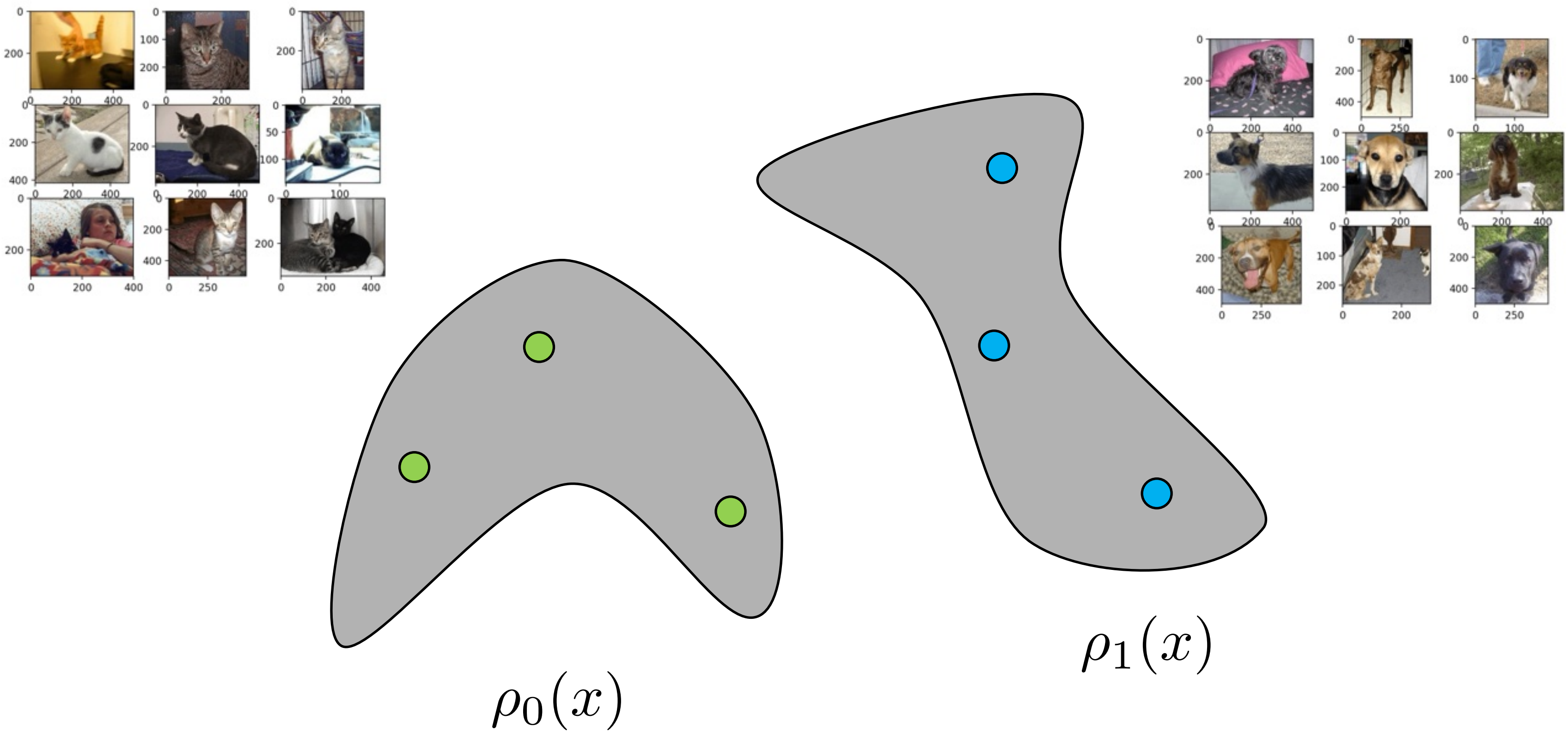






- How to compare and align different distributions?

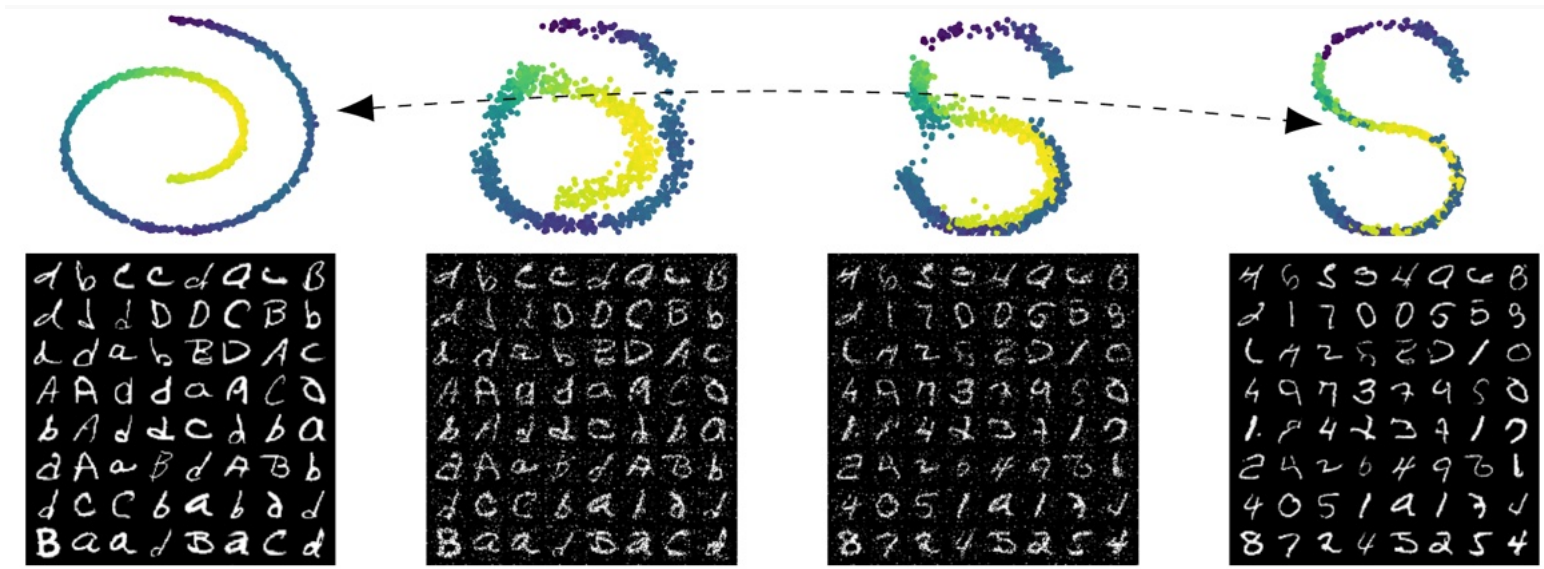
Two data distributions



- Given samples from two datasets with no other information about relations, how to compare them?

Unpaired data translation

$\rho_0(x)$ $\rho_1(x)$



De Bortoli et. al. (2021)

- How to map correspondences between two data distributions?

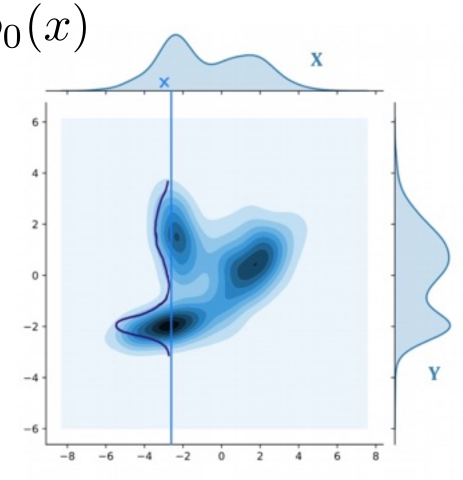
Optimal transport

Monge-Kantorovich: $$\min_{\rho(x,y)} \int dx\, dy\, \rho(x,y) c(x,y)$$

Constrained marginals: $$\int \rho(x,y) dy = \rho_0(x) \qquad \int \rho(x,y) dx = \rho_1(y)$$

$\rho_0(x)$

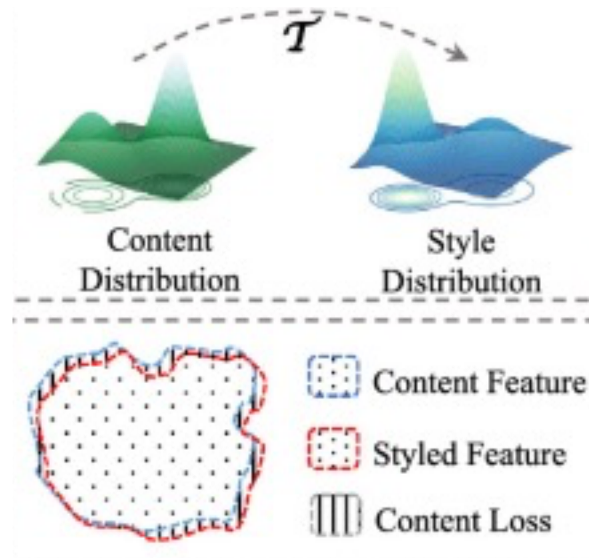
$\rho(x,y)$



$\rho_1(y)$

- Optimize joint distribution to find correspondences between marginal distributions

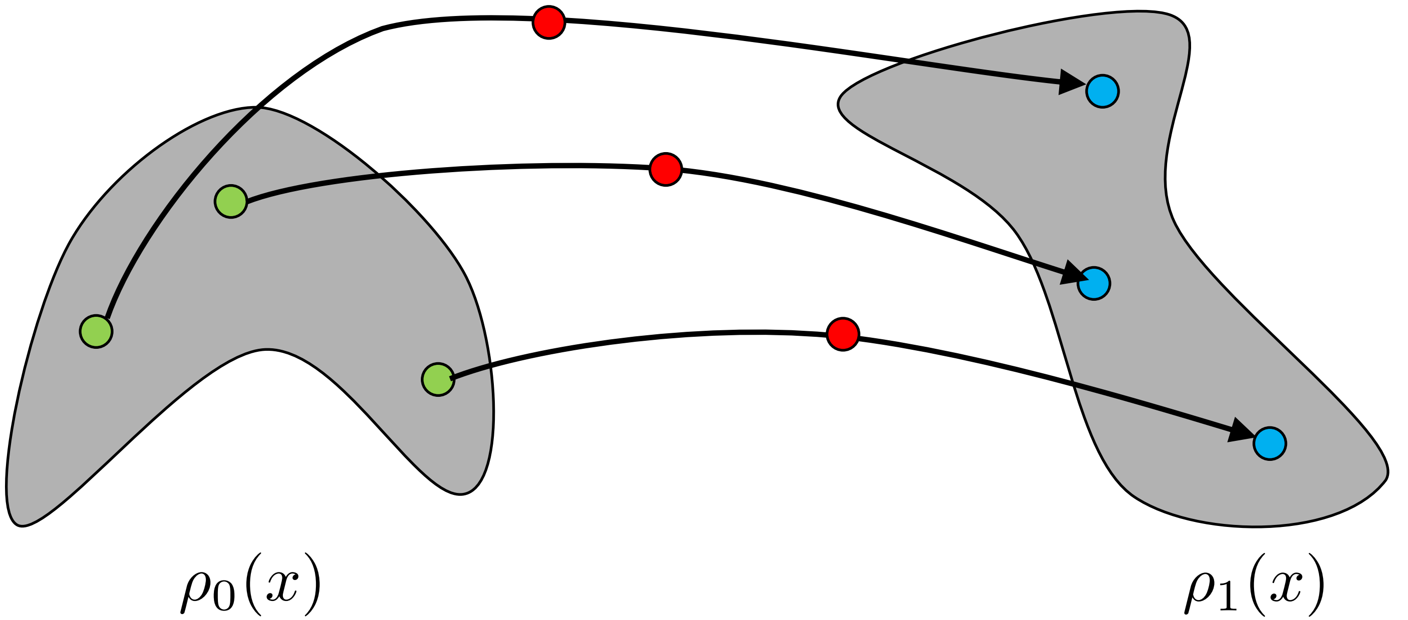
Image style transfer



Kolkin et. al. (2019)

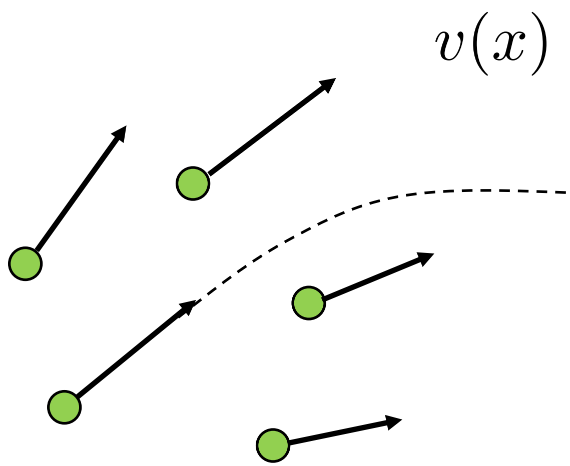
- Map distribution of image content features with distribution of image style features

Dynamic transport



- We also want to understand how samples move (interpolate) between distributions.

Dynamic flows



Velocity flow field

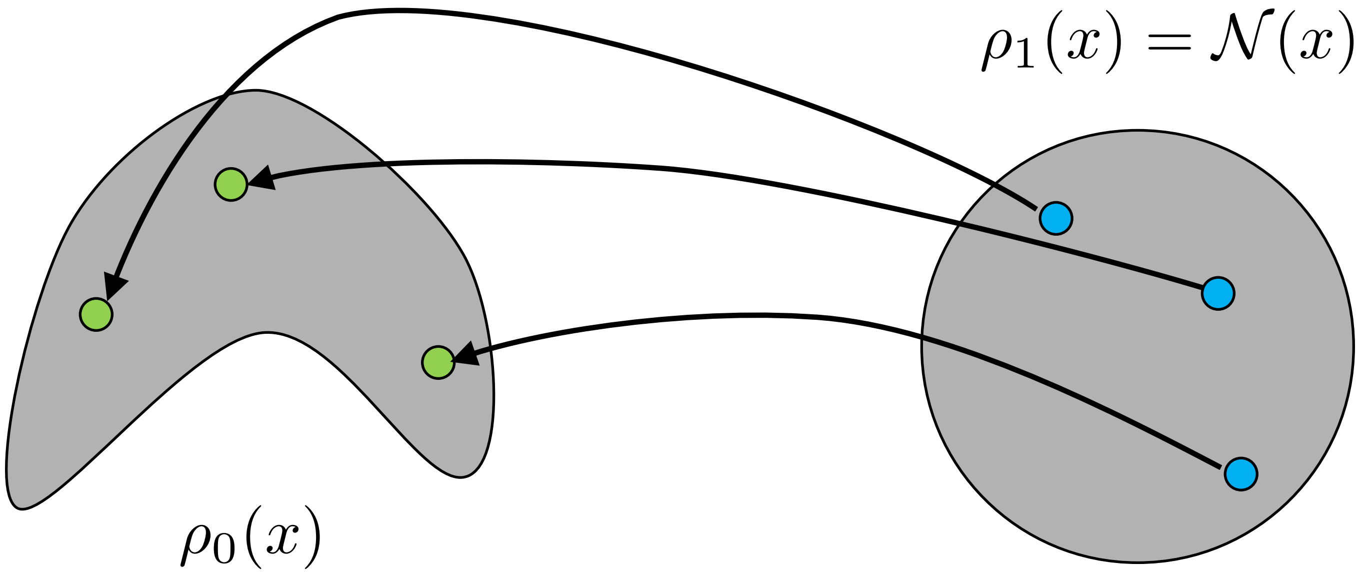
$$v : \mathbb{R}^N \to \mathbb{R}^N$$

Particle motion described by differential equation:

$$dx_t = v(x_t)dt$$

- Dynamical motion of particles described by velocity flows and differential equations

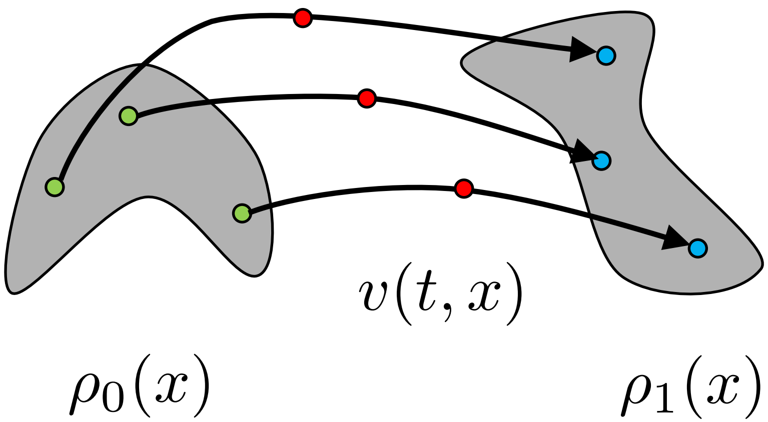
Reverse diffusion models

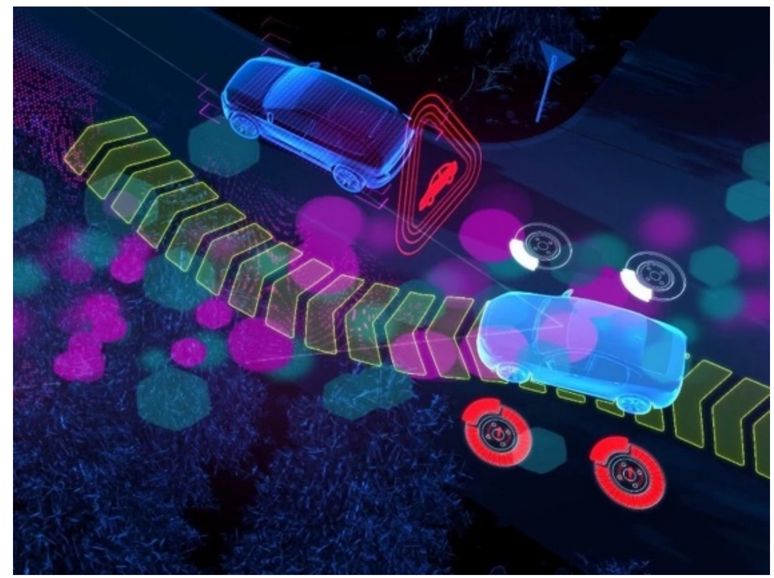


Score function: $v(t, x) = \nabla \log \rho_t(x)$

- Score function is a time-dependent velocity flow mapping from standard Gaussian distribution to source distribution

Control



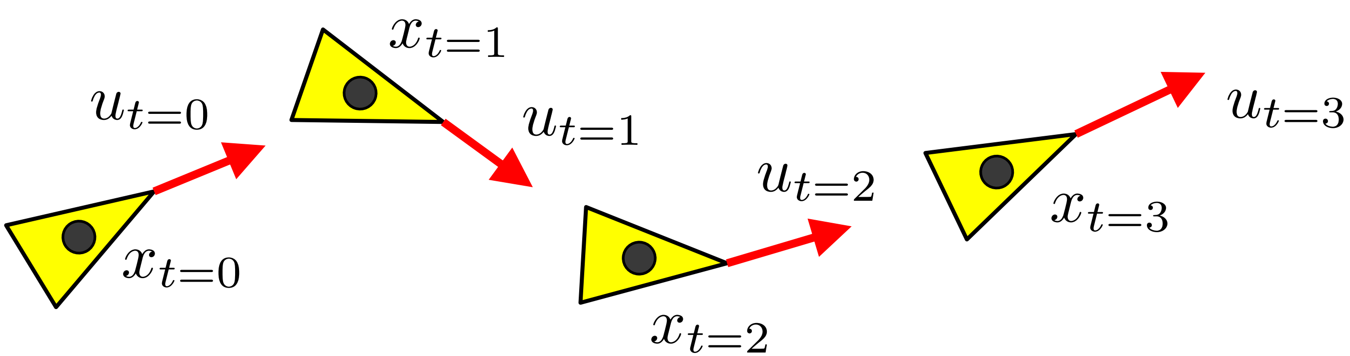


- Controlling flow from one distribution to another is analogous to steering individual particles to goal positions

Discrete time system

Discrete time deterministic dynamical system:

$$x_{t+1} = x_t + f(t, x_t, u_t) \qquad t = 0, 1, \ldots, T-1$$



State vectors: $x_t \in \mathbb{R}^N$ Control actions: $u_t \in \mathbb{R}^M$

- Dynamics describe evolution of state vectors in time

Optimal control

$$x_{t+1} = x_t + f(t, x_t, u_t) \qquad t = 0, 1, \ldots, T-1$$

State vectors: $x_t \in \mathbb{R}^N$ Control actions: $u_t \in \mathbb{R}^M$

Given initial state $x_{t=0}$ and controls $u_{0:T-1} = u_0, u_1, \ldots, u_{T-1}$

we can compute states $x_{1:T}$

Finite time horizon cost:

$$J(x_0, u_{0:T-1}) = \phi(x_T) + \sum_{t=0}^{T-1} R(t, x_t, u_t)$$

- Problem of optimal control is to find sequence $u_{0:T-1}$ that minimizes $J(x_0, u_{0:T-1})$

Naïve algorithm

$$J(x_0, u_{0:T-1}) = \phi(x_T) + \sum_{t=0}^{T-1} R(t, x_t, u_t)$$

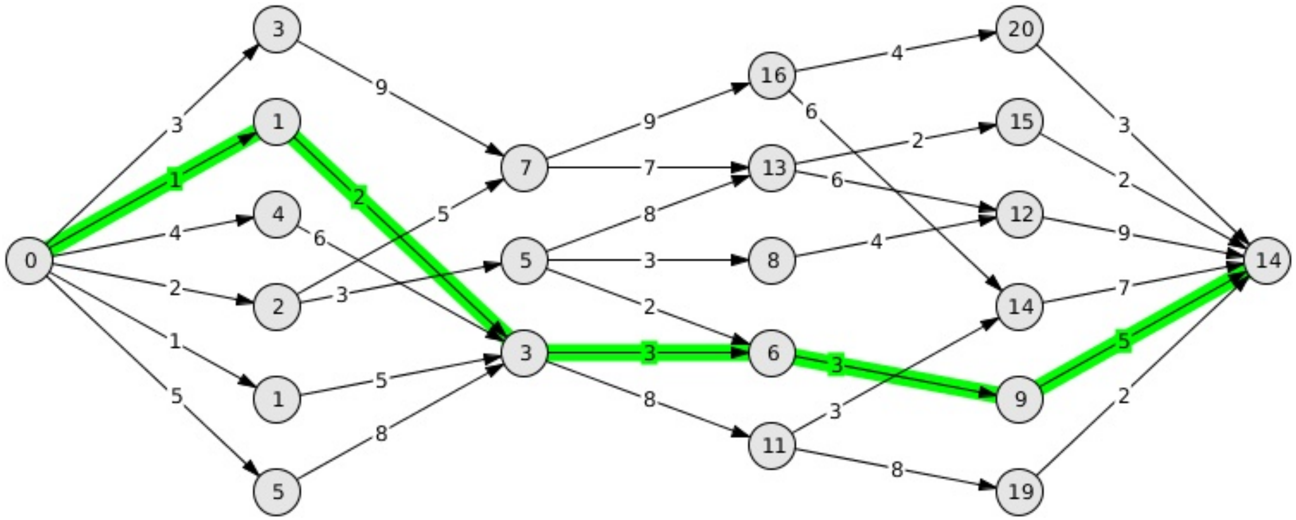
$$\min_{u_{0:T-1}} J(x_0, u_{0:T-1})$$

Evaluate cost function over all possible control vectors: $u_{0:T-1} = u_0, u_1, \ldots, u_{T-1}$

Time complexity: $O(N_u^T)$

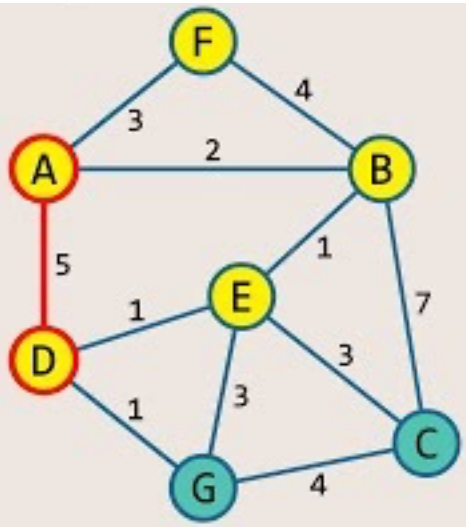
- Exponential time complexity is intractable for most problems

Principle of optimality



- Principle of Optimality: An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision. (Bellman, 1957, Chap. III.3.)

Dynamic programming



Node	Cost	Previous
A	0	
B	2	A
C	6	E
D	4	E
E	3	B
F	3	A
G	6	E

- Dijkstra's algorithm recursively computes cost-to-go function to determine shortest paths

Optimal cost-to-go

$$J(x_0, u_{0:T-1}) = \phi(x_T) + \sum_{t=0}^{T-1} R(t, x_t, u_t)$$

Optimal cost-to-go (value) function:

$$V(t, x_t) = \min_{u_{t:T-1}} \left[\phi(x_T) + \sum_{s=t}^{T-1} R(s, x_s, u_s) \right]$$

Solves optimal control problem from an intermediate time and states until terminal time T

$$V(T, x) = \phi(x)$$

$$V(0, x) = \min_{u_{0:T-1}} C(x, u_t)$$

Bellman equation

Recursion relation:

$$
\begin{aligned}
V(t, x_t) &= \min_{u_{t:T-1}} \left[\phi(x_T) + \sum_{s=t}^{T-1} R(s, x_s, u_s) \right] \\
&= \min_{u_t} \left[R(t, x_t, u_t) + \min_{u_{t+1:T-1}} \left(\phi(x_T) + \sum_{s=t+1}^{T-1} R(s, x_s, u_s) \right) \right] \\
&= \min_{u_t} \left[R(t, x_t, u_t) + V(t+1, x_{t+1}) \right] \\
&= \min_{u_t} \left[R(t, x_t, u_t) + V(t+1, x_t + f(t, x_t, u_t)) \right]
\end{aligned}
$$

- This is called the Bellman equation which can compute the optimal cost-to-go for all intermediate times and states

Optimal control policy

Initialization:

$$V(T, x) = \phi(x)$$

Backwards: $t = T - 1, \ldots, 0$

$$u_t^\star(x) = \arg\min_u \left[R(t, x, u) + V(t+1, x + f(t, x, u))\right]$$

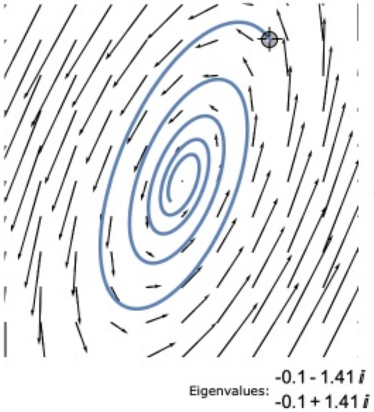
$$V(t, x) = R(t, x, u_t^\star(x)) + V(t+1, x + f(x, u_t^\star(x)))$$

Forwards: $t = 0, \ldots, T - 1$

$$x_{t+1}^\star = x_t^\star + f(t, x_t^\star, u_t^\star(x_t^\star))$$

Linear system example

$$x_{t+1} = Ax_t + Bu_t$$



Quadratic cost function:

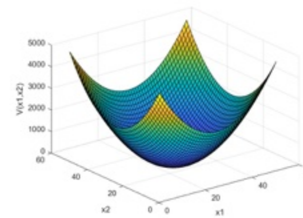
$$C = x_T^\top Q_T x_T + \sum_{s=0}^{T-1} x_s^\top Q x_s + u_s^\top R u_s$$

- This is called the Linear-Quadratic Regulator (LQR) control problem

Riccati equation

Quadratic cost-to-go function:

$$V(t,x) = x^\top P_t x$$



Initialization: $V(T,x) = x^\top Q_T x \quad P_T = Q_T$

$$V(t,x) = \min_u \left[x^\top Q x + u^\top R u + V(t+1, Ax + Bu) \right]$$

$$P_t = Q + A^\top P_{t+1} A - A^\top P_{t+1} B (R + B^\top P_{t+1} B)^{-1} B^\top P_{t+1} A$$

- Recursive solution to Bellman equation for LQR is known as the Riccati equation

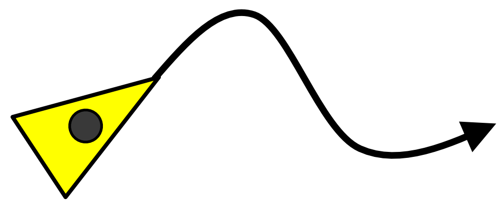
Continuous time

Take the continuous time limit: $\Delta t \to 0$

$$x_{t+\Delta t} = x_t + f(t, x_t, u_t)\Delta t$$

Deterministic ODE:

$$dx_t = f(t, x_t, u_t)dt$$



Integrated cost:

$$J(x_0, u_{0\to T}) = \phi(x_T) + \int_0^T d\tau\, R(\tau, x(\tau), u(\tau))$$

- Control problem is to find optimal controls $u^\star(t)$ that minimize cost

Hamilton-Jacobi

Value function:

$$V(t,x) = \min_u \left[R(t,x,u)dt + V(t+dt, x+f(t,x,u)dt)\right]$$

$$\approx \min_u \left[R(t,x,u)dt + V(t,x) + \partial_t V(t,x)dt + \partial_x V(t,x) f(t,x,u)dt\right]$$

$$\boxed{-\partial_t V(t,x) = \min_u \left[R(t,x,u) + \partial_x V(t,x) f(t,x,u)\right]}$$

Boundary condition: $V(T,x) = \phi(x)$

- Hamilton-Jacobi-Bellman (HJB) is a nonlinear PDE that solves *V(t,x)* backwards for all intermediate time and states

Optimal HJB policy

Initialization:

$$V(T, x) = \phi(x)$$

Backwards PDE: $t = T \to 0$

$$u_t(x) = \arg\min_u \left[R(t, x, u) + \partial_x V(t, x) f(t, x, u)\right]$$

$$-\partial_t V(t, x) = R(t, x, u) + \partial_x V(t, x) f(t, x, u_t^\star(x))$$

Forwards ODE: $t = 0 \to T$

$$x_{t=0}^\star = x_0$$

$$dx_t^\star = f(t, x_t^\star, u_t^\star(x_t^\star))dt$$

Continuous time LQR

Linear ODE: $\dot{x} = Ax_t + Bu_t$

Quadratic cost: $J = x_T^\top Q_T x_T + \int_0^T d\tau \, (x_\tau^\top Q x_\tau + u_\tau^\top R u_\tau)$

Value function: $V(t,x) = x^\top P_t x$

Initialization: $P_T = Q_T$

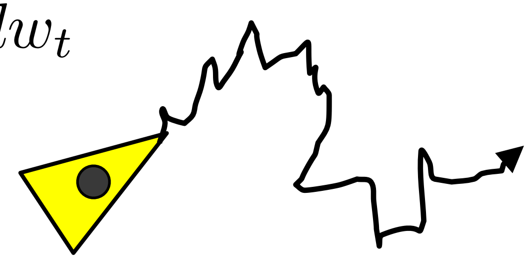
$$-\dot{P}_t = A^\top P_t + P_t A - P_t B R^{-1} B^\top P_t + Q$$

- Continuous-time Riccati equation is matrix ODE running backwards in time

Stochastic dynamics

Stochastic differential equation (SDE) dynamics:

$$dx_t = f(t, x_t, u_t)dt + \nu_t dw_t$$



Wiener process (Brownian noise):

$$dw_t \sim \mathcal{N}(0, I\, dt)$$

Cost becomes an expectation over random variables:

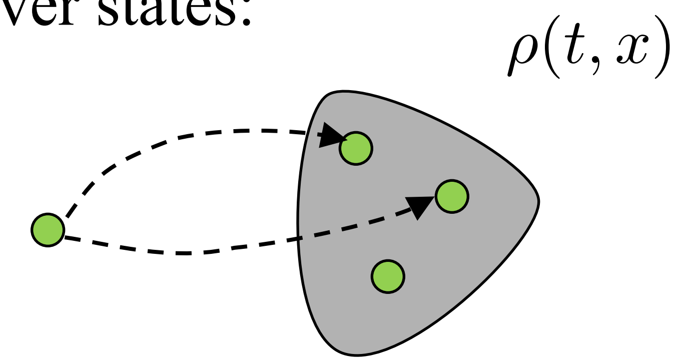
$$J(x_0, u) = \left\langle \phi(x_T) + \int_0^T d\tau\, R(t, x_t, u_t) \right\rangle$$

Belief state

SDE dynamics:

$$dx_t = f(t, x_t, u_t)dt + \nu_t dw_t$$

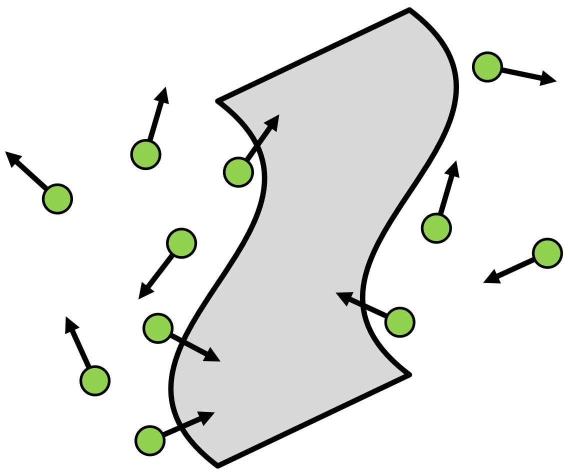
Distribution over states:



- Because of stochastic dynamics, need to track distribution of states (belief state) over time

Fokker-Planck equation

$$\partial_t \rho(t,x) = -\nabla \cdot [\rho(t,x) f(t,x,u)] + \frac{1}{2}\nu_t \nabla^2 \rho(t,x)$$



Continuity equation (conservation of mass):

$$\frac{\partial \rho_t}{\partial t} = -\nabla \cdot \vec{J}$$

Flux:

$$\vec{J}_v = \rho_t(x) f(t,x,u)$$

$$\vec{J}_D = -\frac{1}{2}\nu_t \nabla \rho_t(x)$$

- Fokker-Planck equation describes time evolution of state distribution

HJB equation

Value function recursion:

$$V(t, x_t) = \min_{u_t} R(t, x_t, u_t) + \langle V(t + dt, x_{t+dt}) \rangle$$

$$\langle V(t + dt, x_{t+dt}) \rangle = \int dx_{t+dt} \mathcal{N}(x_{t+dt} | x_t, \nu_t dt) V(t + dt, x_{t+dt})$$

$$= V(t, x_t) + \partial_t V(t, x_t) dt + \langle dx \rangle \partial_x V(t, x_t) + \frac{1}{2} \langle dx^2 \rangle \partial_x^2 V(t, x_t)$$

$$\langle dx \rangle = f(t, x, u) dt \quad \langle dx^2 \rangle = \nu_t dt$$

$$-\partial_t V(t, x) = \min_u \left[R(t, x, u) + f(t, x, u) \partial_x V(t, x) + \frac{1}{2} \nu_t \partial_x^2 V(x, t) \right]$$

Boundary condition: $V(T, x) = \phi(x)$

Effect of noise

Fokker-Planck (forwards in time):

$$\partial_t \rho(t,x) = -\nabla \cdot [\rho(t,x) f(t,x,u)] + \frac{1}{2}\nu_t \nabla^2 \rho(t,x)$$

$$\rho(0,x) = \rho_0(x)$$

Hamilton-Jacobi-Bellman (backwards in time):

$$-\partial_t V(t,x) = \min_u \left[R(t,x,u) + f(t,x,u) \cdot \nabla V(t,x) + \frac{1}{2}\nu_t \nabla^2 V(x,t) \right]$$

$$V(T,x) = \phi(x)$$

- PDE evolution of belief state and value function contain Laplacian terms describe influence of noise

LQG example

System State and Measurement

$$\dot{\mathbf{x}}(t) = \mathbf{F}(t)\mathbf{x}(t) + \mathbf{G}(t)\mathbf{u}(t) + \mathbf{L}(t)\mathbf{w}(t)$$
$$\mathbf{z}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{n}(t)$$

Control Law

$$\mathbf{u}(t) = -\mathbf{C}(t)\hat{\mathbf{x}}(t) + \mathbf{C}_F(t)\mathbf{y}_C(t)$$

State Estimate

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{F}(t)\hat{\mathbf{x}}(t) + \mathbf{G}(t)\mathbf{u}(t) + \mathbf{K}(t)\left[\mathbf{z}(t) - \mathbf{H}\hat{\mathbf{x}}(t)\right]$$
$$= \left[\mathbf{F}(t) - \mathbf{G}(t)\mathbf{C}(t) - \mathbf{K}(t)\mathbf{H}\right]\hat{\mathbf{x}}(t) + \mathbf{G}(t)\mathbf{C}_F(t)\mathbf{y}_C(t) + \mathbf{K}(t)\mathbf{z}(t)$$

Estimator Gain and State Covariance Estimate

$$\mathbf{K}(t) = \mathbf{P}(t)\mathbf{H}^T\mathbf{N}^{-1}(t)$$
$$\dot{\mathbf{P}}(t) = \mathbf{F}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}^T(t) + \mathbf{L}(t)\mathbf{W}(t)\mathbf{L}^T(t) - \mathbf{P}(t)\mathbf{H}^T\mathbf{N}^{-1}(t)\mathbf{H}\mathbf{P}(t)$$

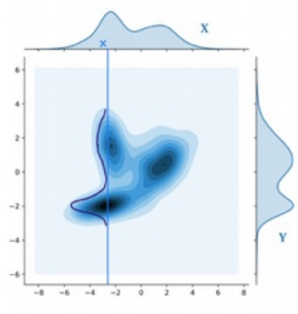
Control Gain and Adjoint Covariance Estimate

$$\mathbf{C}(t) = \mathbf{R}^{-1}(t)\mathbf{G}^T(t)\mathbf{S}(t)$$
$$\dot{\mathbf{S}}(t) = -\mathbf{Q}(t) - \mathbf{F}(t)^T\mathbf{S}(t) - \mathbf{S}(t)\mathbf{F}(t) + \mathbf{S}(t)\mathbf{G}(t)\mathbf{R}^{-1}(t)\mathbf{G}^T(t)\mathbf{S}(t)$$

- Linear-Quadratic-Gaussian: Linear dynamics, Quadratic cost function, Gaussian (Brownian) noise

Static optimal transport

$$\min_{\rho(x,y)} \int dx\, dy\, \rho(x,y) c(x,y)$$



$\rho(x,y)$

Euclidean distance cost: $c(x,y) = \|y-x\|^2$

Constrained marginals:

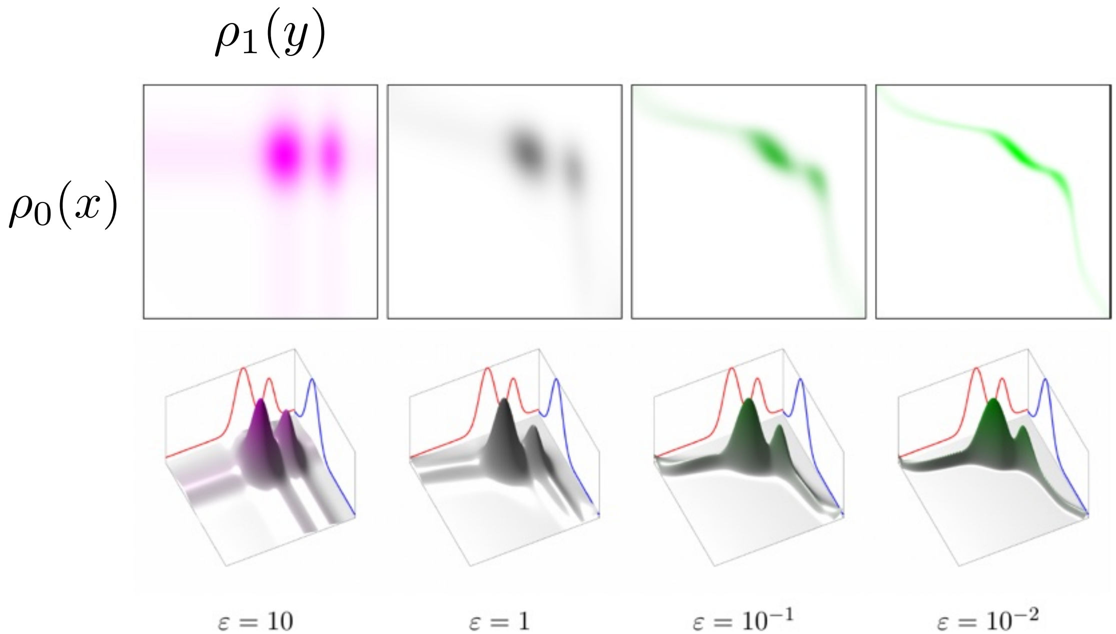
$$\int \rho(x,y)dy = \rho_0(x)$$

$$\int \rho(x,y)dx = \rho_1(y)$$

- Minimum gives squared Wasserstein-2 distance between two distributions

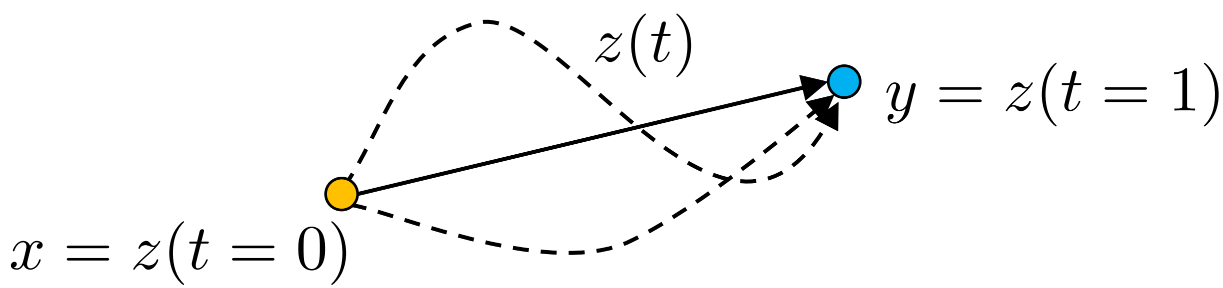
Entropy regularization

$$\min_{\rho(x,y)} \int dx\, dy\, \rho(x,y) \|y - x\|^2 + \varepsilon D_{KL}[\rho(x,y) \| \rho_0(x)\rho_1(y)]$$



- Regularization for smooth solutions to optimization

Dynamic formulation



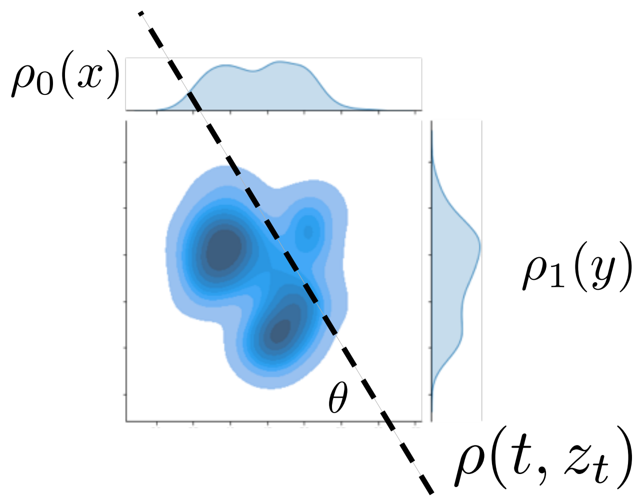
$$\|y-x\|^2 = \int_0^1 dt \int Dz_t \delta\left(z_t - [ty + (1-t)x]\right) \|\dot{z}_t\|^2$$
$$= \min_{z(t)} \int_0^1 dt \, \|\dot{z}_t\|^2$$

- Rewrite Euclidean cost as minimum over paths with constrained endpoints

Dynamic density distribution

$$\int dx\, dy\, \rho(x,y) \|y - x\|^2 = \int_0^1 dt \int Dz_t\, \rho(t, z_t) \|\dot{z}_t\|^2$$

$$\rho(t, z_t) = \int dx\, dy\, \rho(x,y) \delta\left(z_t - [ty + (1-t)x]\right)$$



- Can view time-dependent distribution as marginalization across diagonal line in joint distribution function

Benamou-Brenier formulation

$$\min_{(\rho,v)} \int_0^1 dt \int \frac{1}{2} \|v(t,x)\|^2 \rho(t,x) dx$$

Constraints:

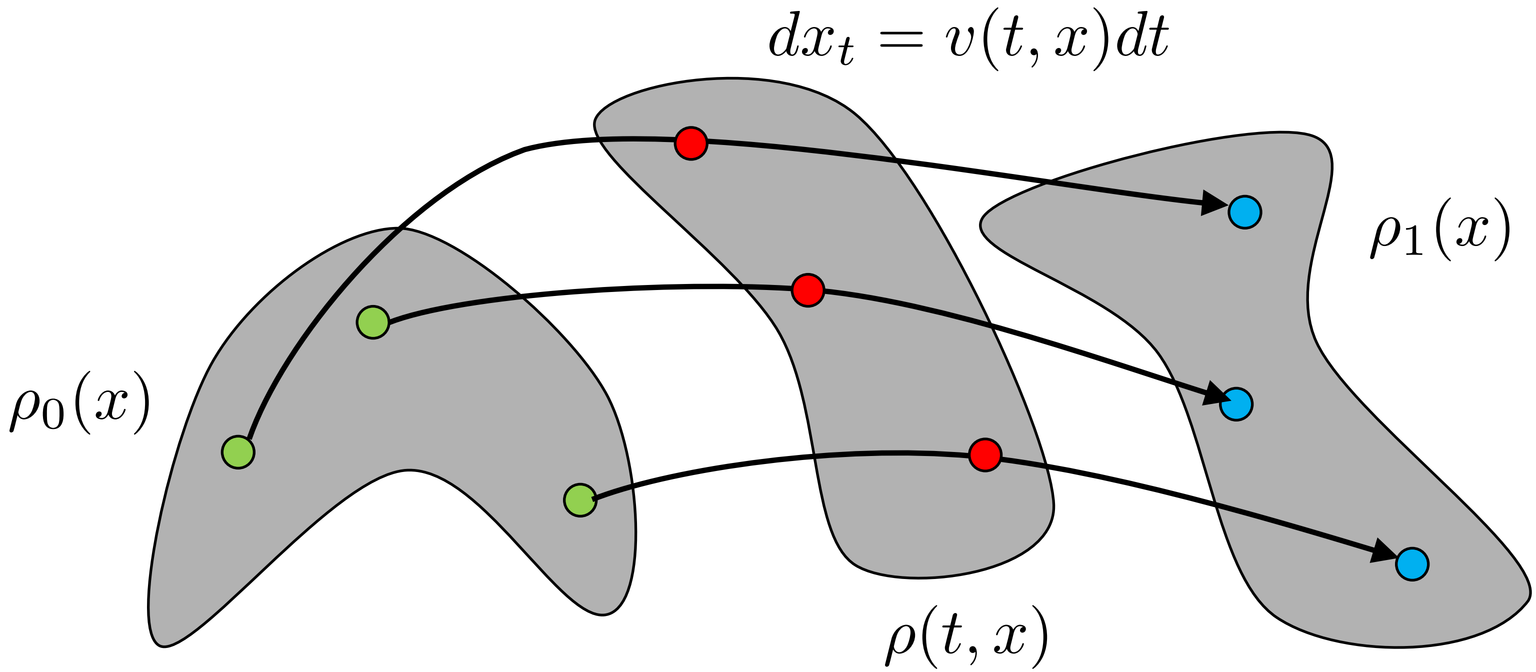
$$\frac{\partial \rho}{\partial t} + \nabla \cdot (v\rho) = 0$$

$$\rho(0,x) = \rho_0(x)$$

$$\rho(1,x) = \rho_1(x)$$

- Solution is equivalent to Monge-Kantorovich static optimal transport problem

Benamou-Brenier dynamics



$$\min_{(\rho,v)} \int_0^1 dt \int \frac{1}{2}\|v(t,x)\|^2 \rho(t,x) dx$$

Optimal control

Lagrangian with constraints:

$$\mathcal{L}(\rho, v) = \int_0^1 dt \int \left[\frac{1}{2}\|v(t,x)\|^2 \rho(t,x) + \lambda(t,x)\left(\frac{\partial \rho}{\partial t} + \nabla \cdot (v\rho)\right)\right] dx$$

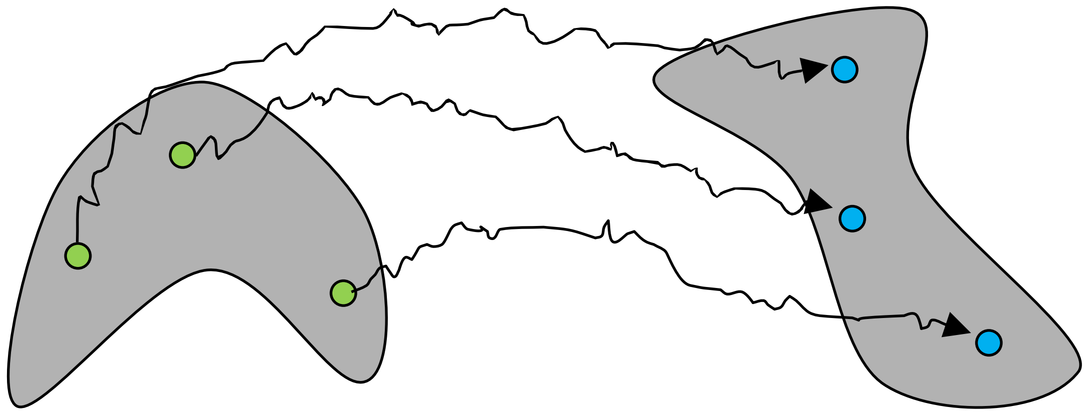
$$= \int_0^1 dt \int \left[\frac{1}{2}\|v(t,x)\|^2 - \left(\frac{\partial \lambda}{\partial t} + \nabla\lambda \cdot v\right)\right] \rho(t,x) dx$$

$$v^\star(t,x) = \nabla\lambda(t,x) \qquad \frac{\partial \lambda}{\partial t} + \frac{1}{2}\|\nabla\lambda\|^2 = 0$$

- Optimal flow field is gradient of value function satisfying HJB equation

Schrödinger problem (1931/32)

Erwin Schrödinger



- How to discover forces on hot gas particles consistent with observed particle distribution?

Schrödinger bridge

SDE dynamics: $$dx_t = v(t,x)dt + \epsilon_t dw_t$$

$$\min_{(\rho,v)} \int_0^1 dt \int \frac{1}{2}\|v(t,x)\|^2 \rho(t,x)dx$$

Constraints: $$\frac{\partial \rho}{\partial t} + \nabla \cdot (v\rho) - \frac{1}{2}\epsilon_t \nabla^2 \rho = 0$$

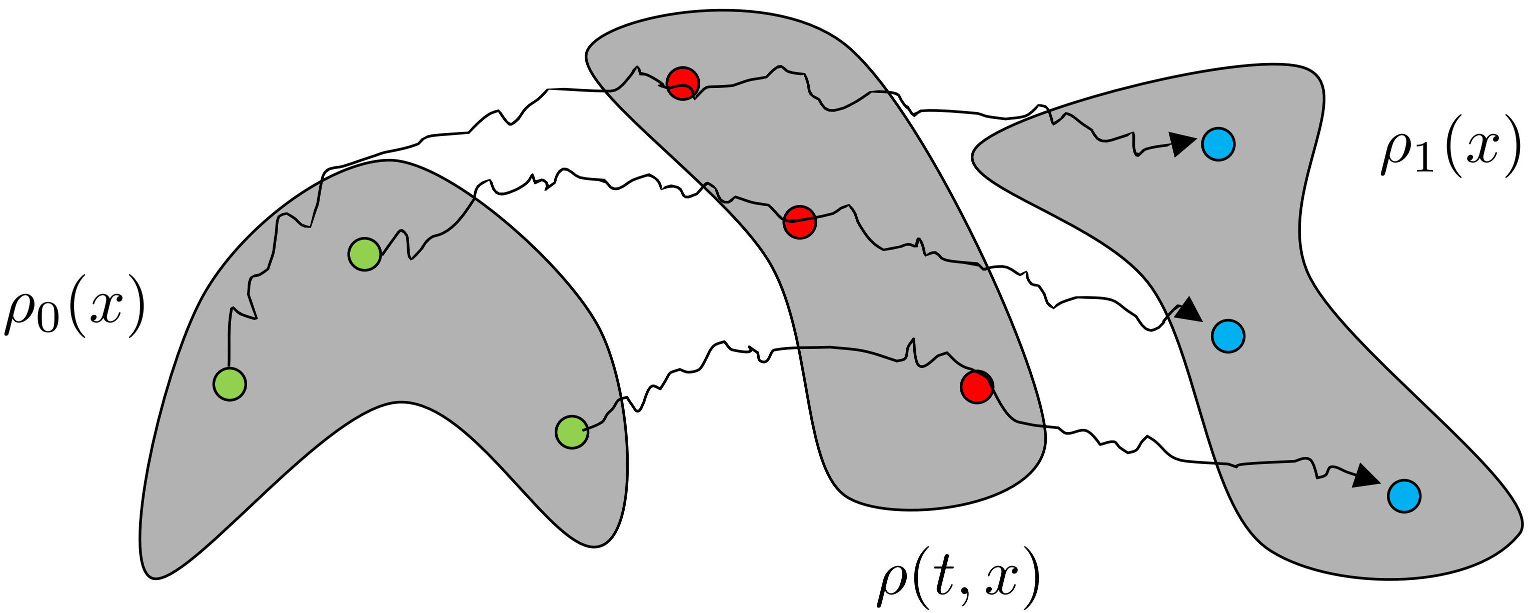
$$\rho(0,x) = \rho_0(x)$$

$$\rho(1,x) = \rho_1(x)$$

- Optimize flow field with Fokker-Planck equation constraints and initial and final distributions

Schrödinger bridge dynamics

$$dx_t = [v_0(t,x) + v(t,x)]dt + \epsilon_t dw_t$$



- Dynamics include prior on flow field and stochastic Brownian noise

Entropy regularized solution

$$\min_{(\rho,v)} \int_0^1 dt \int \frac{1}{2}\|v(t,x)\|^2 \rho(t,x) dx$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot ((v + v_0)\rho) - \frac{1}{2}\epsilon_t \nabla^2 \rho = 0$$

$$\rho(0,x) = \rho_0(x)$$

$$\rho(1,x) = \rho_1(x)$$

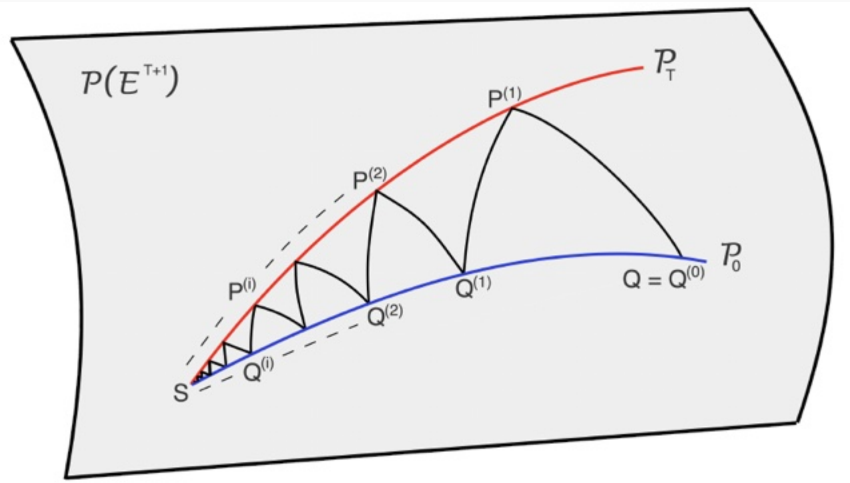
$\longleftrightarrow$

$$\min_{\rho(x,y)} \int dx\, dy\, \rho(x,y)\|y - x\|^2$$

$$+\varepsilon D_{KL}[\rho(x,y)\|\rho_0(x)\rho_1(y)]$$

- Schrödinger bridge is equivalent to entropy-regularized optimal transport

Iterative projection methods



$$\rho(0, x) = \rho_0(x)$$

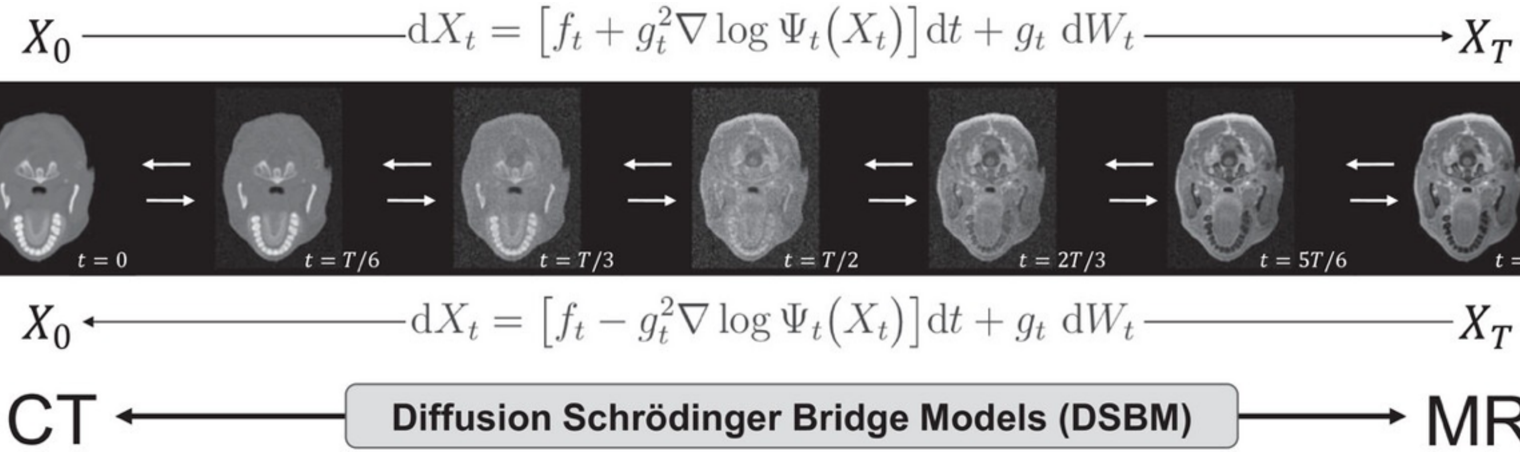
$$\rho(1, x) = \rho_1(x)$$

- Iterative Proportional Fitting (Sinkhorn) algorithm alternatively projects distribution onto initial and final marginal constraints (half-bridge solutions)

Applications

(a) Cat → Wild (b) Wild → Cat

De Bortoli et. al. (2024)



Li et. al. (2025)

Summary

- Unpaired data translation problem
- Optimal flow control to transport one distribution to another distribution
- Relationship between entropy-regularized optimal transport and Schrödinger bridges
- Active research on solution methods and new applications

References

Arnaud Doucet, NeurIPS 2024:

https://neurips.cc/virtual/2024/invited-talk/101133

Bert Kappen, stochastic optimal control notes:

http://www.snn.ru.nl/~bertk/

Chieh-hsin Lai, diffusion models monograph:

https://www.arxiv.org/abs/2510.21890