The Most Likely Evolution of Diffusing and Vanishing Particles in the Spirit of Erwin Schrödinger:

Constructing bridges with unbalanced marginals

Tryphon Georgiou
University of California, Irvine

joint work with

Yongxin Chen and **Michele Pavon**
Georgia Tech — University of Padova

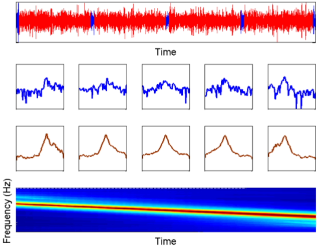
Dynamics and Discretization: PDEs, Sampling, and Optimization
Simons Institute, Berkeley
Oct 25 – Oct 29, 2021

Outline of the talk

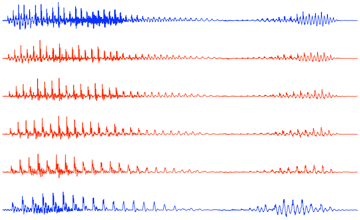
- Some context/motivation: interpolation of distributions (bridges)
- Optimal mass transport and Schrödinger's bridge problem
- Diffusing and vanishing Particles in the spirit of Schrödinger
  - Bridges between unbalanced marginals

Interpolation of distributions

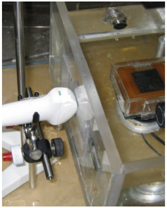
aka Morphing



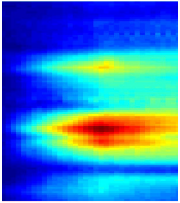
Doppler frequency tracking



Voice morphing

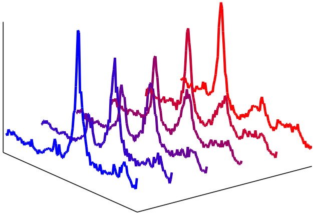


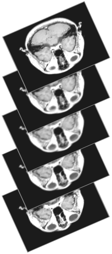
⇒

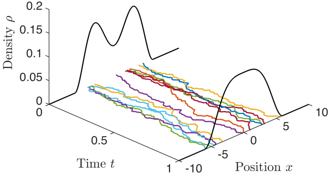


Noninvasive temperature sensing - temperature field

Interpolation of distributions – unbalanced marginals

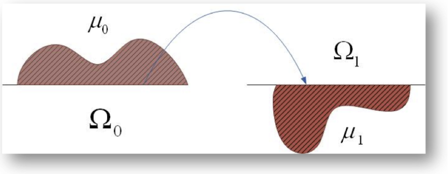






Optimal mass transport

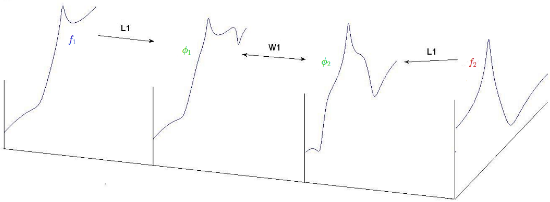
G. Monge (1871)



L. Kantorovich (1942)

and then McCann, Ganbgo, Brenier, Benamou, Ambrosio,. . . (1990's on)
Rachev-Ruschendorf, Villani, . . .

Unbalanced marginals - before 2010



$\mu_0(\Omega_0) \neq \mu_1(\Omega_1)$

$$d_{\mathrm{mixed},\kappa}(\mu_0,\mu_1) = \inf_{\hat\mu_0,\hat\mu_1} d_{\mathrm{W}}(\hat\mu_0,\hat\mu_1) + \kappa \sum_{i=0}^{1} \|\hat\mu_i - \mu_i\|_{\mathrm{TV}}$$
$$= \sup_f \left\{ \int f d(\mu_0 - \mu_1) \mid \|f\|_{\mathrm{Lip}} \leq 1, \|f\|_\infty \leq \kappa \right\}$$

$\hat\mu_0, \hat\mu_1$: noise-free measures

$\mu_i - \hat\mu_i$: noise components

e.g.., see G-Karlsson-Takyar 2009

Unbalanced marginals - post 2010

$$\inf_{\rho,v,\tilde\rho_1} \int_0^1 \int_{\mathbb{R}^m} \rho(t,x)\|v\|^2\, dx\, dt + \alpha \int_{\mathbb{R}^m} (\rho_1(x) - \tilde\rho_1(x))^2\, dx,$$

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho v) = 0, \;\; \rho(0,\cdot) = \rho_0(\cdot), \;\; \rho(1,\cdot) = \tilde\rho_1(\cdot) \text{ (not necessarily } = \rho_1(\cdot)).$$

$$\inf_{\rho,v,s} \int_0^1 \int_{\mathbb{R}^m} \left\{\rho(t,x)\|v\|^2 + \alpha s(t,x)^2\right\} dx\, dt,$$

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho v) = s, \;\; \rho(0,\cdot) = \rho_0(\cdot), \;\; \rho(1,\cdot) = \rho_1(\cdot).$$

$$\inf_{\rho,v,r} \int_0^1 \int_{\mathbb{R}^m} \left\{\rho(t,x)\|v\|^2 + \alpha \frac{s^2}{\rho(t,x)}\right\} dx\, dt$$

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho v) = s, \;\; \rho(0,\cdot) = \rho_0(\cdot), \;\; \rho(1,\cdot) = \rho_1(\cdot).$$

see Liero-Mielke-Savaré (arxiv.org/pdf/1508.07941),
Peyré-Cuturi 2020, also Chen-G-Tannenbaum 2018

Optimal Mass Transport regularization: Schrödinger's Bridge Problem (SBP)

Balanced marginals for now

A problem in large-deviations that leads to:

$$\inf_{(\rho,v)} \int_{\mathbb{R}^n} \int_0^1 \rho(t,x) \|v(t,x)\|^2 dt dx,$$
$$\frac{\partial \rho}{\partial t} + \nabla \cdot (v\rho) = \frac{1}{2}\Delta\rho$$
$$\rho(0,x) = \rho_0(x), \quad \rho(1,y) = \rho_1(y)$$

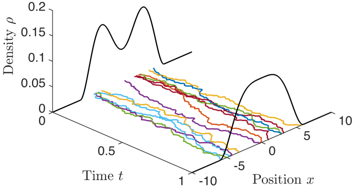
And a fluid-dynamic, time-symmetric, formulation:

$$\inf_{(\rho,v)} \int_{\mathbb{R}^n} \int_0^1 \left[\|v(t,x)\|^2 + \|\frac{1}{2}\nabla \log \rho(t,x)\|^2 \right] \rho(t,x) dt dx,$$
$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) = 0,$$
$$\rho(0,x) = \rho_0(x), \quad \rho(1,y) = \rho_1(y).$$

Blaquière, Dai Pra, Pavon-Wakolbinger, Filliger-Hongler-Streit, Mikami, Thieulien, Leonard, Chen-G-Pavon

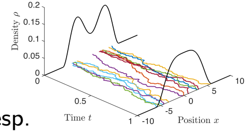
Schrödinger's Bridge Problem (SBP)

Erwin Schrödinger
Schrödinger bridges 1931/32



$\sim$ Nelson's stochastic mechanics

Schrödinger's Bridge Problem (SBP)



Consider:

- Cloud of N independent Brownian particles (N large)
- empirical distr. $\rho_0(x)dx$ and $\rho_1(y)dy$ at $t = 0$ and $t = 1$, resp.
- ρ_0 and ρ_1 not compatible with transition mechanism

$$\rho_1(y) \neq \int_0^1 p(0, x, 1, y)\rho_0(x)dx,$$

where

$$p(s, y, t, x) = [2\pi(t-s)]^{-\frac{n}{2}} \exp\left[-\frac{|x-y|^2}{2(t-s)}\right], \quad s < t$$

Particles have been transported in an unlikely way

Schrödinger (1931): **Of the many unlikely ways in which this could have happened, which one is the most likely?**

Schrödinger's Bridge Problem (SBP)

Large deviations formulation

$$\min_P H(P|R) = \min_P E_P\left[\log \frac{dP}{dR}\right]$$

over $P \in \{$distributions on paths with marginals $\rho_0, \rho_1\}$;
$H(\cdot|\cdot)$ is the relative entropy
R reference Wiener measure

Föllmer 1988: **SBP is a large deviations problem of the empirical distribution on paths $\equiv$ maximum entropy problem via Sanov's thm**

Connection to stochastic control & OMT:
For prior the law of a diffusion: $dX = vdt + dB$, Girsanov's thm:

$$E_Q\left[\log \frac{dQ}{dR}\right] = E_Q\left[\frac{1}{2}\int_0^1 \|v\|^2 ds\right]$$

Schrödinger's Bridge Problem (SBP)

Stochastic control formulation & structure of solutions

- Girsanov's thm gives:

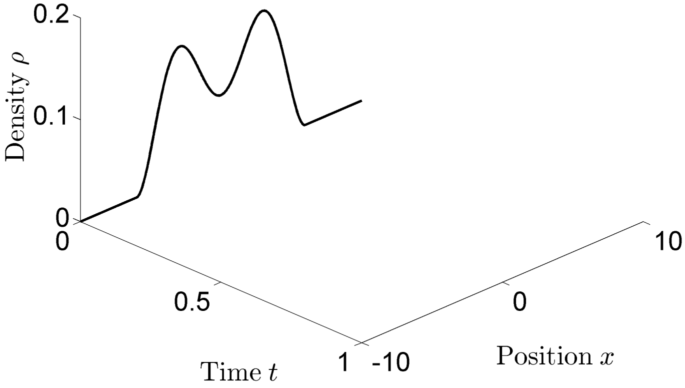
$$\inf_{(\rho,v)} \int_{\mathbb{R}^n} \int_0^1 \|v(t,x)\|^2 \rho(t,x) dt dx,$$
$$\frac{\partial \rho}{\partial t} + \nabla \cdot (v\rho) = \frac{1}{2}\Delta\rho$$
$$\rho(0,x) = \rho_0(x), \quad \rho(1,y) = \rho_1(y)$$

- min $H(P|R) \Rightarrow \rho(t,x) = \varphi(t,x)\hat{\varphi}(t,x)$ (t-time marginal of P)

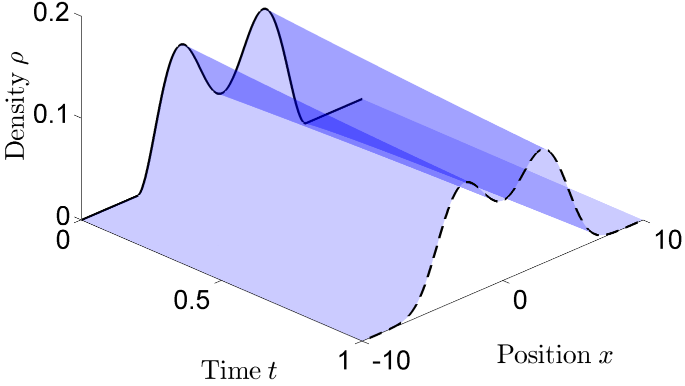
where φ and $\hat{\varphi}$ solve the Schrödinger's system:

$$\varphi(t,x) = \int p(t,x,1,y)\varphi(1,y)dy, \quad \varphi(0,x)\hat{\varphi}(x,0) = \rho_0(x)$$
$$\hat{\varphi}(t,x) = \int p(0,y,t,x)\hat{\varphi}(0,y)dy, \quad \varphi(1,x)\hat{\varphi}(1,x) = \rho_1(x).$$

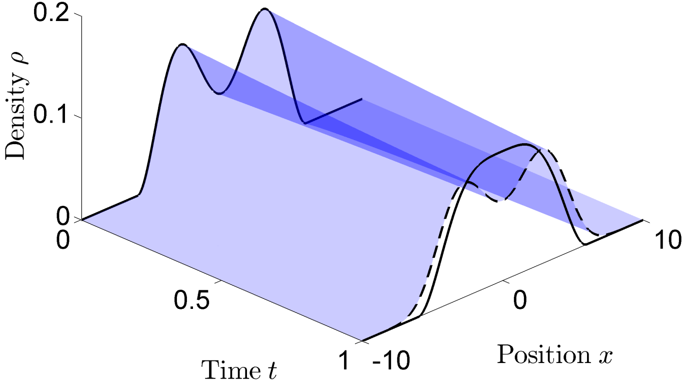
SBP schematic - marginal



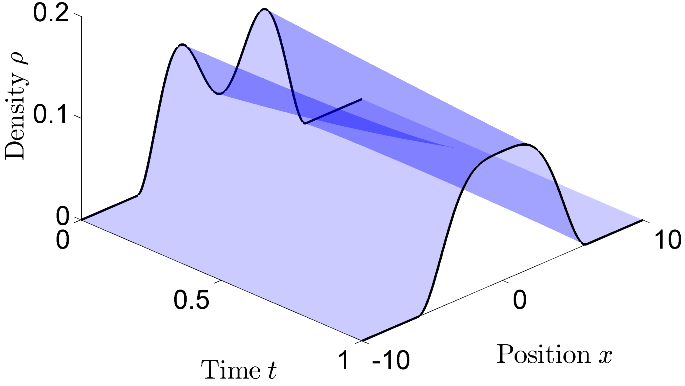
SBP schematic - prior



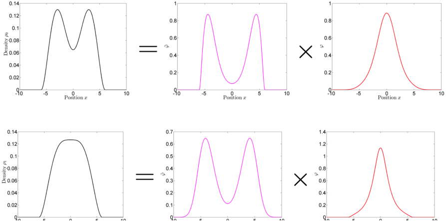
SBP schematic - prior vs. mismatched end-point marginal



SBP schematic - Schrödinger bridge



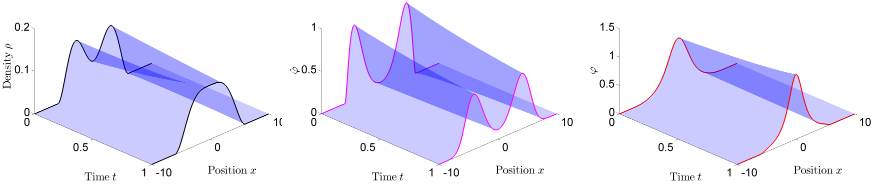
Schrödinger system



$$-\frac{\partial\varphi}{\partial t}(t,x) = \frac{1}{2}\Delta\varphi(t,x)$$

$$\frac{\partial\hat{\varphi}}{\partial t}(t,x) = \frac{1}{2}\Delta\hat{\varphi}(t,x)$$

$$\varphi(0,x)\hat{\varphi}(0,x) = \rho_0(x)$$
$$\varphi(1,x)\hat{\varphi}(1,x) = \rho_1(x)$$



For $dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \ (\star)$

$a(t, X) := \sigma(t, X)\sigma(t, X)' > 0$

Notation:

R : "prior" law of $(\star)$ on paths

R_t, R_{st} : marginals at times t, and jointly t, s

$R^{xy}(\cdot)$ law conditioned on $X_0 = x$, $X_1 = y$

disintegration of measure $R(\cdot) = \int_x \int_y R^{xy}(\cdot) R_{01}(dxdy)$

SBP: Find

$$P^\star = \arg\min_P \{H(P|R) \mid P_0 = \rho_0, \ P_1 = \rho_1\}$$

$$H(P|R) = H(P_{01}|R_{01}) + \int H(P^{xy}|R^{xy}) P_{01}(dxdy)$$

Static SBP: Find

$$P_{01}^\star = \arg\min_{P_{01}} \{H(P_{01}|R_{01}) \mid P_0 = \rho_0, \ P_1 = \rho_1\}$$

Relation static-dynamic SBP:

$$P^\star(\cdot) = \int_x \int_y R^{xy}(\cdot) P^\star_{01}(dxdy)$$

Solution:

Under mild/natural assumptions, $\exists$ f, g so that:
$P^\star_{01} = f(X_0)g(X_1)R_{01}$. These are solutions of the **Schrödinger system**

$$\begin{aligned} \frac{d\rho_0}{dR_0}(x) &= f(x)R(g(X_1) \mid X_0 = x), \\ \frac{d\rho_1}{dR_1}(y) &= g(y)R(f(X_0) \mid X_1 = y). \end{aligned}$$

$P^\star_{01} = f(X_0)g(X_1)R_{01} \Leftrightarrow P^\star = f(X_0)g(X_1)R.$

Solution:

for $\hat{\varphi}(0, x) := f(x)R_0(x)$, $\varphi(1, y) := g(y)$

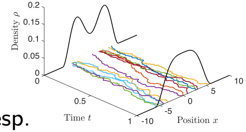
$$\begin{aligned}
\partial_t \hat{\varphi} &= -\nabla \cdot (b\hat{\varphi}) + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 (a_{ij}\hat{\varphi})}{\partial x_i \partial x_j} \\
\partial_t \varphi &= -b \cdot \nabla \varphi - \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}\frac{\partial^2 \varphi}{\partial x_i \partial x_j} \\
\rho_0 &= \varphi(0, \cdot)\hat{\varphi}(0, \cdot) \\
\rho_1 &= \varphi(1, \cdot)\hat{\varphi}(1, \cdot).
\end{aligned}$$

Then, $P_t^\star = \rho(t, \cdot) = \phi(t, \cdot)\hat{\phi}(t, \cdot)$ (t-time marginal) of the law of

$dX_t = (b(t, X_t) + a(t, X_t)\nabla \log \varphi(t, X_t))dt + \sigma(t, X_t)dW_t$

Schrödinger's Bridge with losses
most likely evolution of diffusing and vanishing particles

Consider:



- Cloud of N "tracer" particles (N large)
- empirical distr. $\rho_0(x)dx$ and $\rho_1(y)dy$ at $t = 0$ and $t = 1$, resp.
- ρ_0 and ρ_1 not compatible with transition mechanism

$$\rho_1(y) \neq \int_0^1 p(t_0, x, t_1, y)\rho_0(x)dx,$$

Besides having been transported in an unlikely way,
the particles remain in suspension for a duration of time, and thus,
at $t = 1$ a random portion of the particles have been lost (sunk), and $\int \rho_1 < \int \rho_0$

Question - in the spirit of Schrödinger:

What is the most likely evolution that accounts for losses?

Stochastic transport with losses – Prior:

$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \; (\star\star)$

with killing rate $V(t, x)$

State space: $\mathcal{X} = \mathbb{R}^n \cup \{\mathfrak{c}\}$ with $\mathfrak{c}$ a "coffin state"

Paths $\mathbf{\Omega} = D([0,1], \mathcal{X})$ càdlàg
(X_t on $\mathbb{R}^n$ with killing) $\equiv$ ($\mathbf{X}_t$ on $\mathcal{X}$ with a law on $\mathcal{P}(\mathbf{\Omega})$)

p_0, p_1 natural augmentation of ρ_0, ρ_1 so that $p_0, p_1 \in \mathcal{P}(\mathcal{X})$
i.e., assuming $\int \rho_1 = 1$, **set** $p_0 = (\rho_0(\cdot), \;\; 0)$, and $p_1 = (\rho_1(\cdot), \;\; 1 - \int \rho_1)$

$$\mathbf{P}^\star := \arg \min_{\mathbf{P} \in \mathcal{P}(\mathbf{\Omega})} \left\{ H(\mathbf{P} \mid \mathbf{R}) \;\mid\; \mathbf{P}_0 = p_0, \mathbf{P}_1 = p_1 \right\}.$$

Schrödinger Bridge with losses

unbalanced SBP – $\int \rho_0 > \int \rho_1$

Prior: Fokker-Planck equation for a diffusion with killing rate $V(t,x)$

$$\partial_t R_t + \nabla \cdot (bR_t) + VR_t = \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 (a_{ij}R_t)}{\partial x_i \partial x_j}.$$

SB with losses: "new ϕ" $= (\phi(t,\cdot), \psi(t))$ on $\mathcal{X}$, same for "new ψ," via the **Schrödinger system:**

$$\partial_t \hat\varphi = -\nabla \cdot (b\hat\varphi) - V\hat\varphi + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 (a_{ij}\hat\varphi)}{\partial x_i \partial x_j}$$

$$\frac{d\hat\psi}{dt} = \int V\hat\varphi(t,x)dx$$

$$\partial_t \varphi = -b \cdot \nabla\varphi + V\varphi - \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}\frac{\partial^2 \varphi}{\partial x_i \partial x_j} - V\psi$$

$$\frac{d\psi}{dt} = 0$$

with b.c.

$$\rho_0 = \varphi(0,\cdot)\hat\varphi(0,\cdot)$$
$$\rho_1 = \varphi(1,\cdot)\hat\varphi(1,\cdot)$$
$$\hat\psi(0) = 0$$
$$\psi(1)\hat\psi(1) = 1 - \int \rho_1.$$

$\Rightarrow \mathbf{P}^\star = f(\mathbf{X}_0)g(\mathbf{X}_1)\mathbf{R}$

$\mathbf{P}^\star = (P_t^\star, q_t^\star)$, $\mathbf{R} = (R_t, s_t)$

Schrödinger Bridge with losses – dynamic formulation

$\mathbf{P}^\star$ is the law of a diffusion

$$dX_t = (b(t,X_t) + a(t,X_t)\nabla \log \varphi(t,X_t))dt + \sigma(t,X_t)dW_t$$

with killing rate $\psi V/\varphi$, and Fokker-Planck equation

$$\partial_t P_t + \nabla \cdot ((b + a\nabla \log \varphi)P_t) = \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 (a_{ij} P_t)}{\partial x_i \partial x_j} - \frac{\psi}{\varphi} V P_t.$$

mass $q(t)$ on $\mathfrak{c}$:
$\frac{dq_t}{dt} = \psi(t)\int V\hat{\varphi}(t,x)dx = \int \frac{\psi}{\varphi} V P_t dx$

Schrödinger Bridge with losses – fluid dynamic formulation

Contrast with original SB the added terms:

$$\min_{P_t(\cdot),u(t,\cdot)} \int_0^1 \int_{\mathbb{R}^n} [\frac{1}{2}\|u(t,x)\|^2 P_t + (\alpha \log \alpha - \alpha + 1) V P_t] dxdt$$

$$\partial_t P_t + \nabla \cdot ((b + \sigma u) P_t) + \alpha V P_t - \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 (a_{ij} P_t)}{\partial x_i \partial x_j} = 0$$

$$P_0 = \rho_0, \quad P_1 = \rho_1.$$

$u^\star(t,x) = \sigma(t,x)' \nabla \log \varphi(t,x)$
$\alpha^\star(t,x) = \frac{\psi(t)}{\varphi(t,x)}$

with marginals:
$P_t^\star(x) = \varphi(t,x)\hat{\varphi}(t,x)$ on $\mathbb{R}^n$
$q_t = \psi(t)\hat{\psi}(t)$ on $\mathfrak{c}$

SBP on Feynman-Kac reweighed processes

Earlier attempts to "model" losses – Nagasawa, Wakolbinger, Leonard, Chen-G-Pavon, . . .

Feynman-Kac reweighing of the prior

$$\hat{R} := \exp\left(-\int_0^1 V(t, X_t)dt\right) R \qquad \mapsto \qquad \hat{P}^\star = f(X_0)\exp\left(-\int_0^1 V(t, X_t)dt\right) g(X_1)R$$

via

$$\hat{P}^\star := \min_{P \in \mathcal{P}(\Omega)} \left\{ H(P \mid \hat{R}) \;\middle|\; P_0 = \rho_0,\ P_1 = \hat{\rho}_1 \right\},$$

with $\hat{\rho}_1$ normalized distribution of **survived particles**

- **upside:** simpler Schrödinger system
- **downside:** not physical & inconsistent with Schrödinger's dictum

ρ_0 distribution of all starting particles, ρ_1 surviving particles

starting distribution of survived partices in not knowable

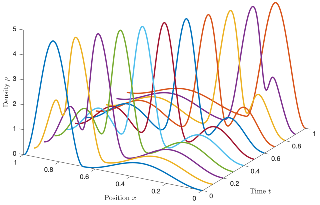
no mechanism to update V

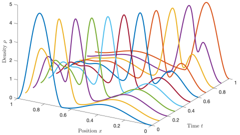
when ρ_1 consistent with prior model and losses in V, $\hat{P}^\star \neq \hat{R}$

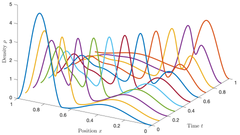
marginals of $\hat{R}$ and $\hat{P}^\star$ have constant mass

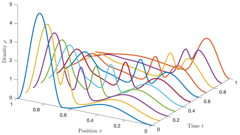
Numerical example

Prior:
$dX_t = \sigma dW_t$ with killing rate $V(t, x) = 1$.

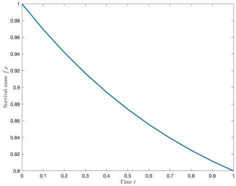


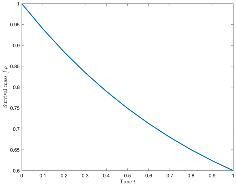


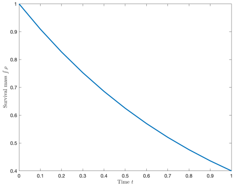




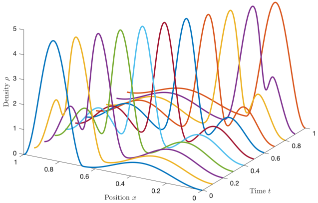
Numerical example



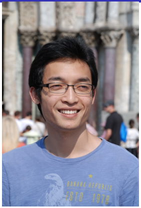




survived mass



reweighed process, regardless of end-point mass



Yongxin Chen

Michele Pavon

Thank you for your attention

CGP arxiv.org/abs/2108.02879