Regulating Stochastic Uncertainty: Covariance Control and Applications

Yongxin Chen & Tryphon Georgiou

based on joint work with

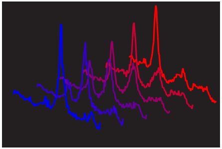
Michele Pavon

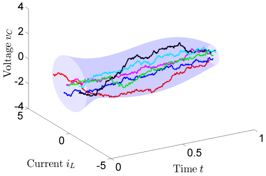
Gatech & UC, Irvine

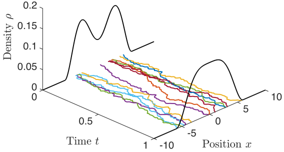
58^{th} IEEE CDC, 2019

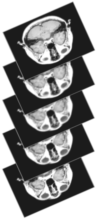
distribution flows and uncertainty

modeling
interpolation
regulation





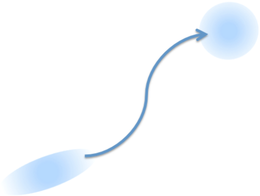


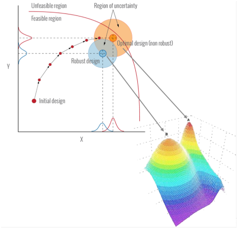


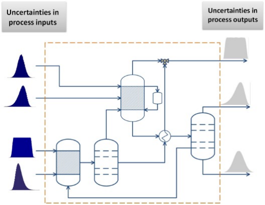
motivation

optimally drive a system from one uncertain state to another

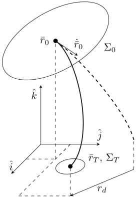
- quality control
- chemical industry
- missile guidance
- spacecraft landing





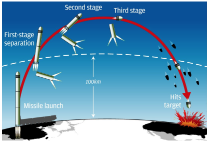


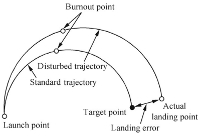
landing



Tsiotras et al 18

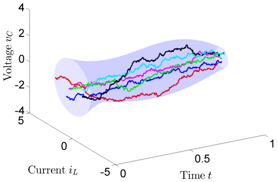
guidance

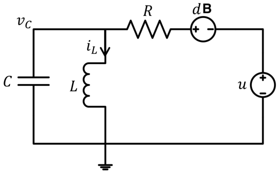




Nyquist-Johnson noise driven oscillator

$$
\begin{aligned}
L di_L(t) &= v_C(t)dt \\
RC dv_C(t) &= -v_C(t)dt - Ri_L(t)dt + u(t)dt + d\mathbf{B}(t)
\end{aligned}
$$





$\binom{v_C}{i_L} \sim \rho_0$ at $t = 0$, $\binom{v_C}{i_L} \sim \rho_1$ at $t = 1$

distribution flows and uncertainty

Monge (1781)
Kantorovich (1940's)

Schrödinger (1932)

Mass Transport $\leftrightarrow$ Schrödinger bridges

Stochastic control
Covariance control

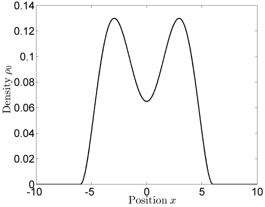
"Covariance analysis permeates almost all of systems theory."

Hotz & Skelton, 1987

optimal mass transport

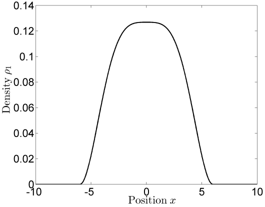
- Monge 1781
- Kantorovich 1942
- Brenier, McCann, Gangbo, Villani, Rachev, Figalli, ...

Move mass from distribution ρ_0 to distribution ρ_1



ρ_0

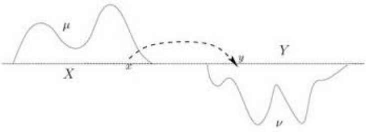
$y = T(x)$



ρ_1

Optimal mass transport (OMT)

Gaspard Monge 1781

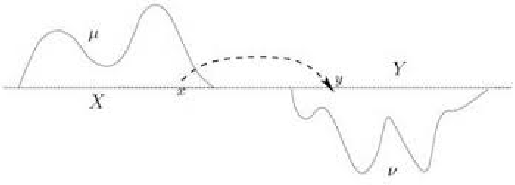


Leonid Kantorovich, work 1940's

Nobel 1975

OMT – Monge's formulation

Le mémoire sur les déblais et les remblais
Gaspard Monge 1781



$$W_2(\mu,\nu)^2 := \min_{T} \int \underbrace{\|x - \underbrace{T(x)}_{y}\|^2}_{\text{cost } c(x,y)} d\mu(x)$$

where $T\#\mu = \nu$

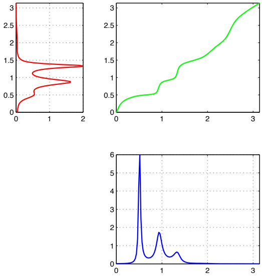
OMT – Kantorovich's formulation

$$W_2(\mu,\nu)^2 = \inf_{\pi\in\Pi(\rho_0,\rho_1|\mathbb{R}\times\mathbb{R})} \iint \underbrace{\|x-y\|^2}_{\text{cost } c(x,y)} d\pi(x,y)$$

$\Pi(\mu,\nu)$: "couplings"

$\int_y \pi(dx,dy) = \rho_0(x)dx = d\mu(x)$
$\int_x \pi(dx,dy) = \rho_1(y)dy = d\nu(y)$

linear program



OMT – fluid dynamic & stochastic control formulation

fluid dynamics

$$\inf_{(\rho,u)} \int_{\mathbb{R}^n} \int_0^1 \|u(x,t)\|^2 \rho(x,t)\,dt\,dx$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (u\rho) = 0 \quad \text{and} \quad \rho(x,0) = \rho_0(x), \;\; \rho(y,1) = \rho_1(y)$$

stochastic control

$$\inf_u \mathbb{E}_\rho \left\{ \int_0^1 \|u(x,t)\|^2 dt \right\}$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (u\rho) = 0 \quad \text{and} \quad \rho(x,0) = \rho_0(x), \;\; \rho(y,1) = \rho_1(y)$$

Benamou and Brenier

OMT – linking Kantorovich to fluid formulations (BB)

In $\mathbb{R}^n$: over paths $\mathcal{X}_{xy} = \{\mathbf{x} \in C^1 \mid \mathbf{x}(0) = x, \mathbf{x}(1) = y\}$,

$$\|x - y\|^2 = \inf_{\mathbf{x} \in \mathcal{X}_{xy}} \int_0^1 \|\dot{\mathbf{x}}\|^2 dt$$

Inf attained at constant speed geodesic $\mathbf{x}^*(t) = (1 - t)x + ty$

$$= \inf_{P_{xy}} \mathbb{E}_{P_{xy}} \left\{ \int_0^1 \|\dot{\mathbf{x}}(t)\|^2 dt \right\}$$

over $P_{xy} \in \Pi(\delta_x, \delta_y \mid C^1)$: prob. measures on C^1 with marginals δ_x, δ_y

connections to Schrödinger bridges

Optimal mass transport:

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

Schrödinger bridges:

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

Schrödinger's Bridge problem (SB)

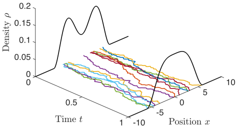
Erwin Schrödinger

Work in 1926, Nobel 1935

Bridges 1931/32

Schrödinger Bridge

- Cloud of N independent Brownian particles (N large)
- $\rho_0(x)$ and $\rho_1(y)$ at $t = 0$ and $t = 1$
- but not compatible with transition mechanism



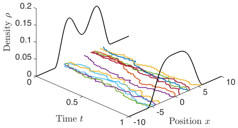
$$\rho_1(y) \neq \int_0^1 p(t_0, x; t_1, y)\rho_0(x)dx,$$

with $p(s, y; t, x)$ the Gaussian kernel

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?"

SBP – maximum likelihood



SB problem $\Rightarrow$ minimize $\mathrm{KL}(Q, W) = E_Q\left[\log \frac{dQ}{dW}\right]$

W Wiener measure on paths (prior)
Q sought distribution on paths consistent with ρ_0, ρ_1

SBP – Girsanov thm – stochastic control

For Q the law of X_t, $dX_t = v_t dt + d\mathbf{B}_t$

$$KL(Q, W) = E\{\frac{1}{2}\int \|v\|^2 dt\}$$

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

Blaquière, Dai Pra

Schrödinger's insight on SBP

the density factors into

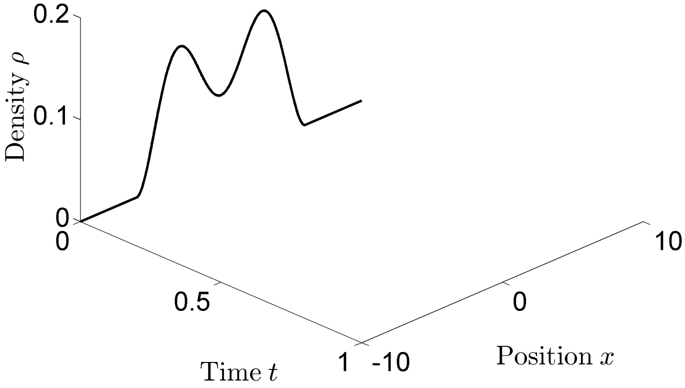
$$\rho(x,t) = \varphi(x,t)\hat{\varphi}(x,t)$$

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

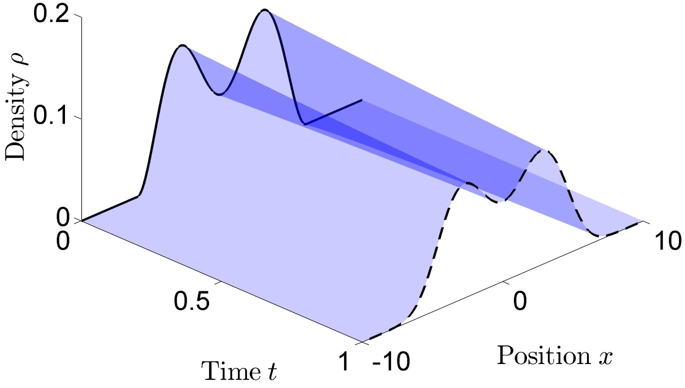
$$\begin{aligned}
\varphi(x,t) &= \int p(t,x,1,y)\varphi(y,1)dy, \quad \varphi(x,0)\hat{\varphi}(x,0) = \rho_0(x) \\
\hat{\varphi}(x,t) &= \int p(0,y,t,x)\hat{\varphi}(y,0)dy, \quad \varphi(x,1)\hat{\varphi}(x,1) = \rho_1(x).
\end{aligned}$$

Sinkhorn iteration: Cuturi 2013, Georgiou-Pavon 2015

SBP schematic

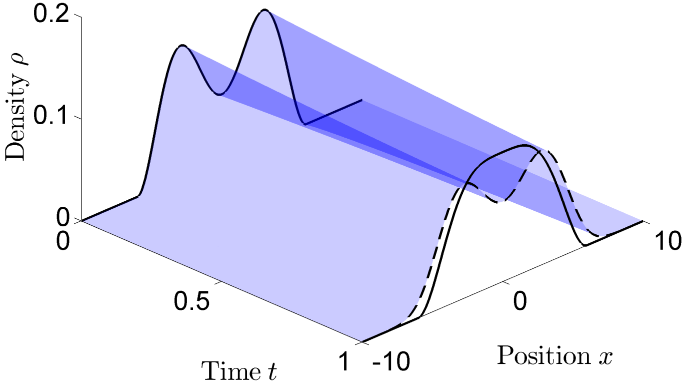


SBP schematic



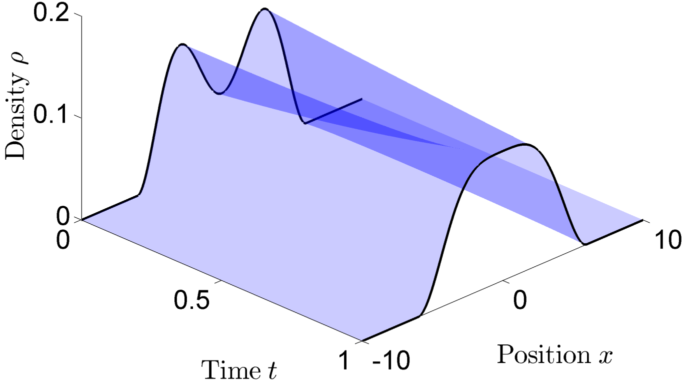
$\frac{\partial \rho}{\partial t} = \frac{1}{2}\Delta\rho$

SBP schematic

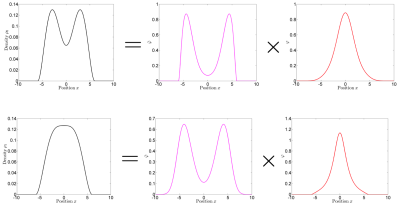


$\frac{\partial \rho}{\partial t} + \nabla \cdot (u\rho) = \frac{1}{2}\Delta\rho$

SBP schematic



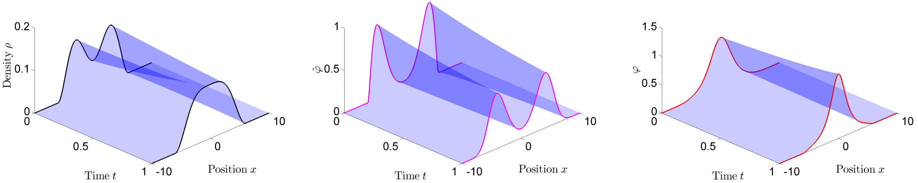
Schrödinger system



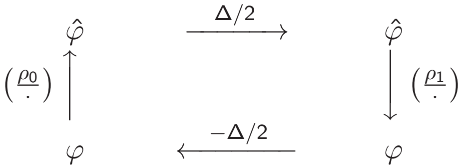
$$-\frac{\partial \varphi}{\partial t}(t,x) = \frac{1}{2}\Delta \varphi(t,x)$$

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

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



Existence & uniqueness (Sinkhorn scaling)



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

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

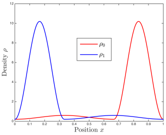
$$\varphi(0,x)\hat{\varphi}(0,x) = \rho_0(x)$$
$$\varphi(1,x)\hat{\varphi}(1,x) = \rho_1(x)$$

iteration is contractive in the Hilbert metric!

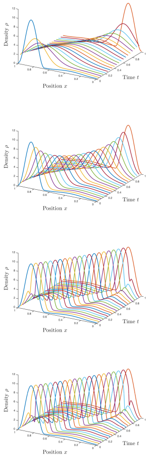
$$d_H(p,q) = \log \frac{M(p,q)}{m(p,q)}$$

$$\begin{aligned} M(p,q) &:= \inf\{\lambda \mid p \leq \lambda q\} \\ m(p,q) &:= \sup\{\lambda \mid \lambda q \leq p\} \end{aligned}$$

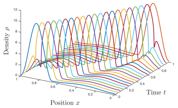
Chen etal, "... computational approach using Hilbert metric," *SIAM Appl. Math.*
Cuturi , "lightspeed computation..."



Marginal distributions



$\rho_t + \nabla \cdot \rho v = \epsilon \Delta \rho$, varying ϵ



OMT interpolation:

$\rho_t + \nabla \cdot \rho v = 0$

Schrödinger bridges to stochastic control

Linear stochastic process $\mathcal{W}$

$$dx(t) = Ax(t)dt + Bdw(t), \quad x(0) \sim \rho_0(x)$$

Controlled process $\mathcal{Q}$

$$dx(t) = Ax(t)dt + Bu(t,x)dt + Bdw(t), \quad x(0) \sim \rho_0(x)$$

u causal, finite energy

KL divergence between $\mathcal{Q}$ and $\mathcal{W}$ in the interval $t \in [0,1]$

$$\mathrm{KL}(\mathcal{Q},\mathcal{W}) = \frac{1}{2}\mathbb{E}\left\{\int_0^1 |u(t)|^2 dt\right\}$$

Dai Pra
Chen etal

control of densities

$$dx(t) = Ax(t)dt + Bu(t,x)dt + Bdw(t)$$
$$x(0) \sim \rho_0(x)$$

(A, B) controllable

Find u causal, finite-energy control such that the cost

$$J(u) = \mathbb{E}\left\{\int_0^1 |u(t)|^2 dt\right\}$$

is minimized and

$$x(1) \sim \rho_1(x)$$

Control of densities

unconstrained $\tilde{\mathcal{U}} := \{u \mid \text{ causal, finite-energy } \}$

constrained $\mathcal{U} := \{u \mid x(1) \sim \rho_1(x)\} \subset \tilde{\mathcal{U}}$

1. Choose g, let $\tilde{J}(u) = \mathbb{E}\left\{\int_0^1 \|u(t)\|^2 dt + g(x(1))\right\}$

   $\mathrm{argmin}_{u\in\mathcal{U}} J(u) = \mathrm{argmin}_{u\in\mathcal{U}} \tilde{J}(u)$

2. Compute unconstrained optimal control $u^\star = \mathrm{argmin}_{u\in\tilde{\mathcal{U}}} \tilde{J}(u)$
3. Compute distribution $x^\star(1) \sim \rho_1^\star(x)$

Approach: study $g \mapsto \rho_1^\star$

Choose g

If $\rho_1^\star = \rho_1$ then $u^\star \in \mathcal{U}$

It follows $u^\star = \mathrm{argmin}_{u\in\mathcal{U}} \tilde{J}(u)$

Linear-Quadratic Brigdes = covariance control

$\rho_0 \sim N(0, \Sigma_0)$, $\rho_1 \sim N(0, \Sigma_1)$ means can be controlled separately

1. $\tilde{J}(u) = \mathbb{E}\left\{\int_0^1 \|u(t)\|^2\, dt + x(1)'\Pi(1)x(1)\right\}$
2. Compute unconstrained optimal control $u^*(t) = -B'\Pi(t)x(t)$
   $\dot{\Pi}(t) = -A'\Pi(t) - \Pi(t)A + \Pi(t)BB'\Pi(t)$
3. Compute covariance
   $\dot{\Sigma}(t) = (A - BB'\Pi(t))\,\Sigma(t) + \Sigma(t)\,(A - BB'\Pi(t))' + BB'$

Induces a map

$$\Pi(1) \mapsto \Sigma(1), \quad S_n \to S_n^+$$

Choose $\Pi(1)$

If $\Sigma(1) = \Sigma_1$

then $u(t) = -B'\Pi(t)x(t)$

Linear-Quadratic Bridges = covariance control

Optimal control

$$u(t,x) = -B^T \Pi(t) x$$

coupled Riccati equations

$$\begin{aligned} -\dot{\Pi}(t) &= A^T \Pi(t) + \Pi(t) A - \Pi(t) B B^T \Pi(t) \\ -\dot{\mathrm{H}}(t) &= A^T \mathrm{H}(t) + \mathrm{H}(t) A + \mathrm{H}(t) B B^T \mathrm{H}(t) \\ \Sigma_0^{-1} &= \Pi(0) + \mathrm{H}(0), \quad \Sigma_1^{-1} = \Pi(1) + \mathrm{H}(1) \end{aligned}$$

Controlled process

$$dx(t) = (A - BB^T \Pi(t)) x(t) dt + B dw(t)$$

Chen et al 15

relation to OMT/SBP

relation to Schrödinger's $\phi, \hat{\phi}$'s:

$$\phi(t,x) = \text{coef} \times \exp(-\|x\|^2_{H(t)^{-1}})$$
$$\hat{\phi}(t,x) = \text{coef} \times \exp(-\|x\|^2_{\Pi(t)^{-1}})$$

covariance control with state penalty

$$dx(t) = Ax(t)dt + Bu(t,x)dt + Bdw(t)$$

$$x(0) \sim N(0, \Sigma_0)$$

(A, B) controllable

Find u causal, finite-energy control such that the cost

$$J(u) = \mathbb{E}\left\{ \int_0^1 |u(t)|^2 + x(t)^T Qx(t)dt \right\}$$

is minimized and

$$x(1) \sim N(0, \Sigma_1)$$

covariance control with state penalty

Optimal control

$$u(t,x) = -B^T\Pi(t)x$$

coupled Riccati equations (closed-form solution difficult)

$$\begin{aligned} -\dot{\Pi}(t) &= A^T\Pi(t) + \Pi(t)A - \Pi(t)BB^T\Pi(t) + Q \\ -\dot{\mathrm{H}}(t) &= A^T\mathrm{H}(t) + \mathrm{H}(t)A + \mathrm{H}(t)BB^T\mathrm{H}(t) + Q \\ \Sigma_0^{-1} &= \Pi(0) + \mathrm{H}(0), \quad \Sigma_1^{-1} = \Pi(1) + \mathrm{H}(1) \end{aligned}$$

Controlled process

$$dx(t) = (A - BB^T\Pi(t))x(t)dt + Bdw(t)$$

Chen et al 17

covariance control with different input and noise channels

$$dx(t) = Ax(t)dt + Bu(t,x)dt + B_1 dw(t)$$

Optimal control

$$u(t,x) = -B^T \Pi(t) x$$

coupled Riccati equations (no closed-form solution)

$$\begin{aligned} -\dot{\Pi}(t) &= A^T \Pi(t) + \Pi(t) A - \Pi(t) B B^T \Pi(t) \\ -\dot{\mathrm{H}}(t) &= A^T \mathrm{H}(t) + \mathrm{H}(t) A + \mathrm{H}(t) B B^T \mathrm{H}(t) \\ &\quad + (\Pi + \mathrm{H})(B B^T - B_1 B_1^T)(\Pi + \mathrm{H}) \\ \Sigma_0^{-1} &= \Pi(0) + \mathrm{H}(0), \quad \Sigma_1^{-1} = \Pi(1) + \mathrm{H}(1) \end{aligned}$$

Controlled process

$$dx(t) = (A - B B^T \Pi(t)) x(t) dt + B_1 dw(t)$$

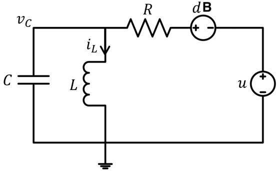
Chen et al 15

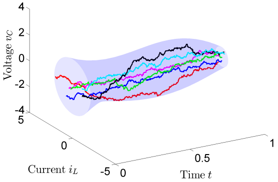
- thermodynamic systems, controlling collective response
- magnetization distribution in NMR spectroscopy,..

– Chen-Georgiou-Pavon, *J. Math. Phys.* 2015.

Nyquist-Johnson noise driven oscillator

$$\begin{aligned} L di_L(t) &= v_C(t)dt \\ RC dv_C(t) &= -v_C(t)dt - Ri_L(t)dt + u(t)dt + d\mathbf{B}(t) \end{aligned}$$





Stationary SBP

When can a given $\Sigma = \Sigma' > 0$ be a stationary state-covariance of

$$dx(t) = (A - BK)x(t)dt + B_1 dw(t)?$$

iff

$$\mathrm{rank}\begin{bmatrix} A\Sigma + \Sigma A' + B_1 B_1' & B \\ B & 0 \end{bmatrix} = \mathrm{rank}\begin{bmatrix} 0 & B \\ B & 0 \end{bmatrix}.$$

iff

$$0 = (A - BK)\Sigma + \Sigma(A - BK)' + B_1 B_1'$$

$$\underbrace{A\Sigma + \Sigma A' + B_1 B_1'}_{\Theta} - \underbrace{B K \Sigma}_{\Gamma G \Lambda} - \underbrace{(BK\Sigma)'}_{(\Gamma G \Lambda)'}$$

Note:
Skelton's talk: $\Theta - \Gamma G \Lambda - (\Gamma G \Lambda)'$

Stationary SBP

When Σ is feasible as a stationary state-covariance for

$$dx(t) = (A - BK)x(t)dt + B_1 dw(t)$$

minimize over admissible K's

$$J_{\text{power}}(u) := \mathbb{E}\{\|u\|^2\}$$

– a semidefinite program

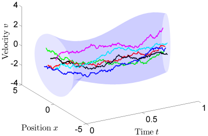
$$\begin{aligned}\mathbb{E}\{\|u\|^2\} &= K\Sigma K' \\ &= X\Sigma^{-1}X'\end{aligned}$$

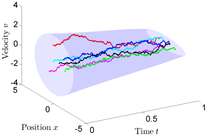
where $X = K\Sigma$

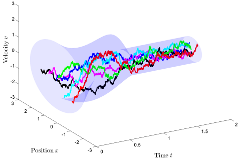
Note: $XY^{-1}X'$ just like $\frac{x^2}{y}$ is jointly convex

Inertial particles with stochastic excitation steered between marginals

$$\begin{aligned} dx(t) &= v(t)dt \\ dv(t) &= -u(t)dt + dw(t) \end{aligned}$$

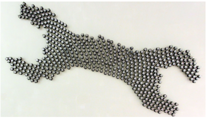






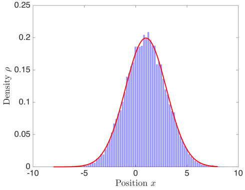
trajectories in phase space
transparent tube: "3σ region"

covariance control in mean-field games



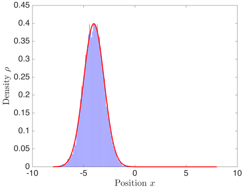
robot swarm
command
control

population dynamics
rational agents
incentive, game



initial distribution

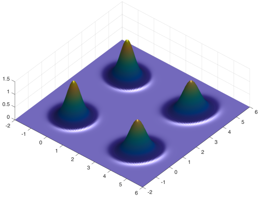
incentive

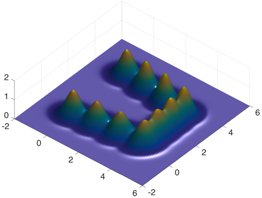


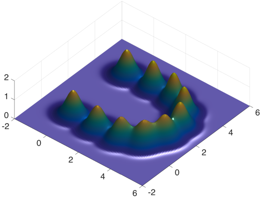
target distribution

Chen et al 17

Path covariance constraints – OMT/SB curves (splines)





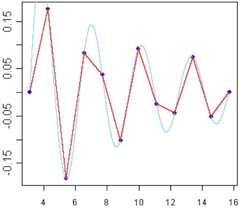


OMT interpolation

distributional (natural) splines

Cubic splines (in $\mathbb{R}^d$)

– Easy to compute
– Cubic polynomial on every subinterval



Holladay's Theorem (1957):
$(t_i, x_i)_{i=0,\ldots,N} \subset [0,1] \times \mathbb{R}^n$ time-space data
the variational problem

$$\inf_{\mathbf{x}} \int_0^1 |\ddot{\mathbf{x}}|^2 dt$$

$$\mathbf{x}(t_i) = x_i \quad i = 0, \ldots, N.$$

admits a unique solution - cubic interpolating spline

covariance (cubic) splines

contain uncertainty through specified midpoints
ρ_i, for $i = 0, \ldots, N$
or respective covariance specs

Chen-Conforti-Georgiou
Benamou-Gallouët-Viallard

Distributional (cubic) splines

– "transporting mass while minimizing squared acceleration"
– mass flows along any $\mathbf{x}$ in C^2 with square integrable second derivative

Distributional-Spline-Problem (DSP): Find

$$\inf_{\mathbf{x}_{t_i}\sharp P=\rho_i} \mathbb{E}_P\{\int_0^1 \| \underbrace{\ddot{\mathbf{x}}(t)}_{\text{acceleration}} \|^2 dt\}$$

with P a **probability measure** on path space.

when $\rho_i \sim \mathcal{N}(m_i, \sigma_i)$, the **quadratic cost function** ensures existence of a Gaussian optimal solution $\Rightarrow$ **Semidefinite program**

Thank you for your attention!