On the Contraction Coefficient of the Schrödinger Bridge for Stochastic Linear Systems

Alexis M.H. Teter

Department of Applied Mathematics
University of California, Santa Cruz

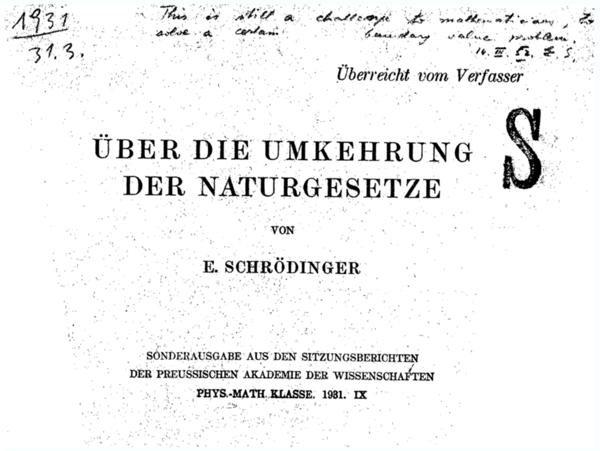
Joint work with

Abhishek Halder
Iowa State University

Yongxin Chen
Georgia Institute of Technology

2024 American Control Conference, Toronto, July 12, 2024

What is a Schrödinger Bridge Problem (SBP)



1931
31.3.

This is still a challenge to mathematicians, to solve a certain boundary value problem.
14. III. 52. E. S.

Überreicht vom Verfasser

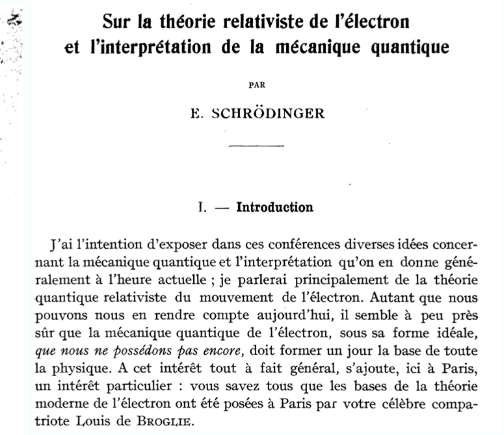
S

ÜBER DIE UMKEHRUNG
DER NATURGESETZE

VON

E. SCHRÖDINGER

SONDERAUSGABE AUS DEN SITZUNGSBERICHTEN
DER PREUSSISCHEN AKADEMIE DER WISSENSCHAFTEN
PHYS.-MATH. KLASSE. 1931. IX



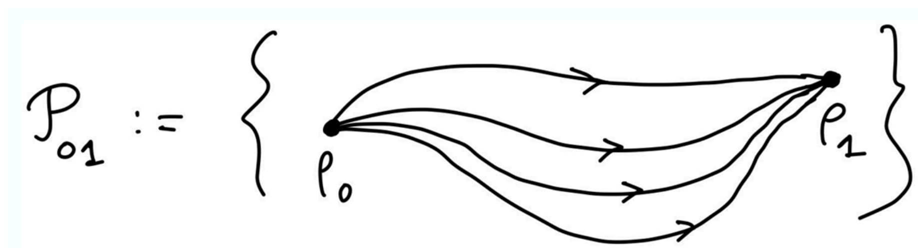
Sur la théorie relativiste de l'électron
et l'interprétation de la mécanique quantique

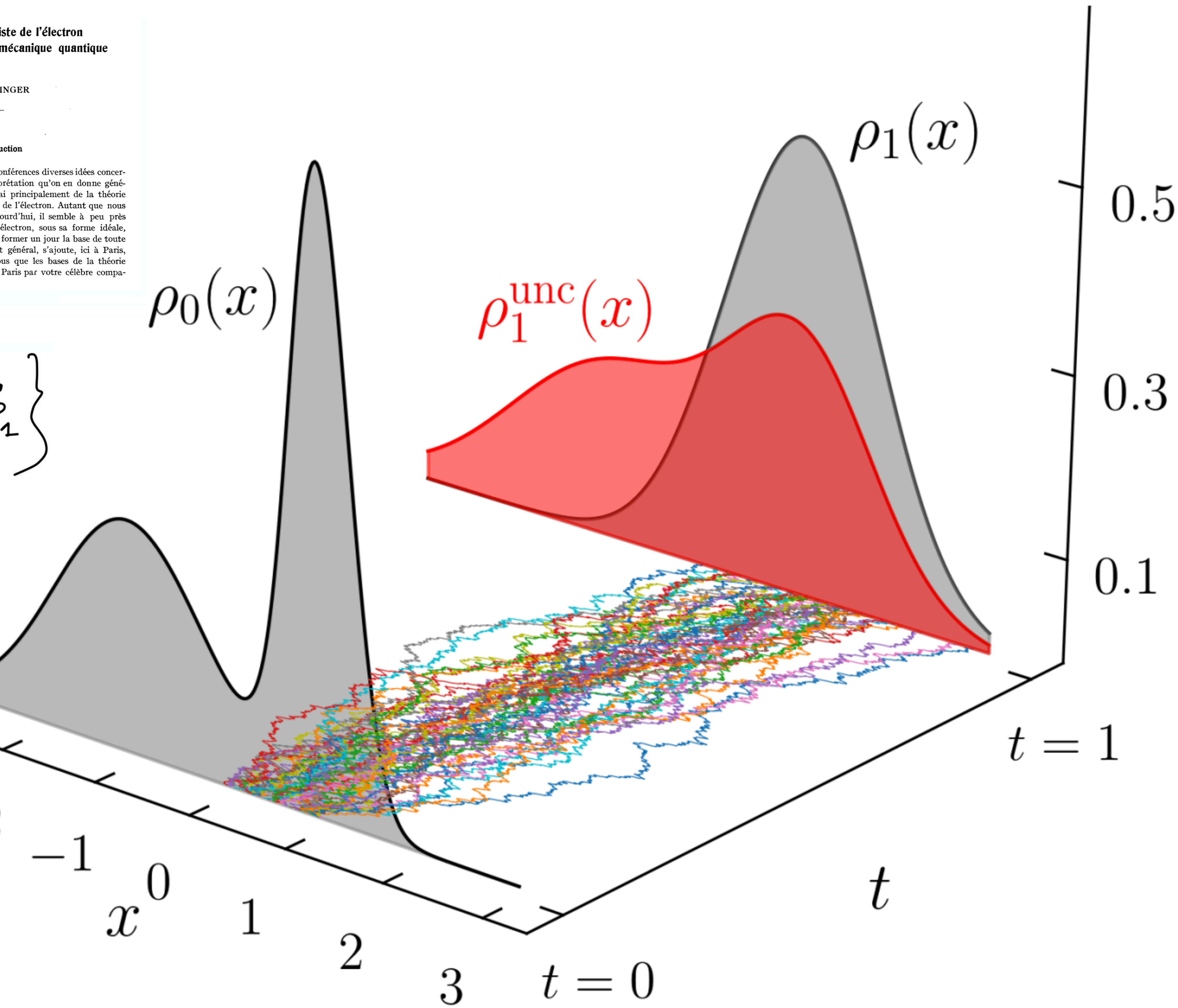
PAR

E. SCHRÖDINGER

I. — **Introduction**

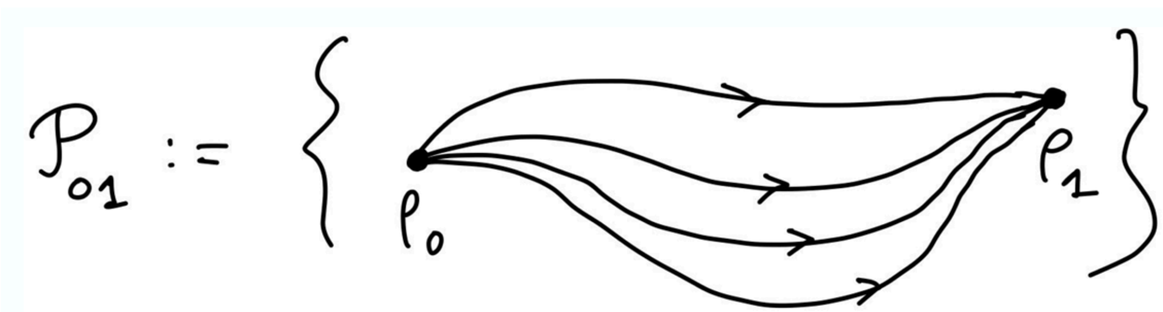
J'ai l'intention d'exposer dans ces conférences diverses idées concernant la mécanique quantique et l'interprétation qu'on en donne généralement à l'heure actuelle ; je parlerai principalement de la théorie quantique relativiste du mouvement de l'électron. Autant que nous pouvons nous en rendre compte aujourd'hui, il semble à peu près sûr que la mécanique quantique de l'électron, sous sa forme idéale, *que nous ne possédons pas encore*, doit former un jour la base de toute la physique. A cet intérêt tout à fait général, s'ajoute, ici à Paris, un intérêt particulier : vous savez tous que les bases de la théorie moderne de l'électron ont été posées à Paris par votre célèbre compatriote Louis de Broglie.

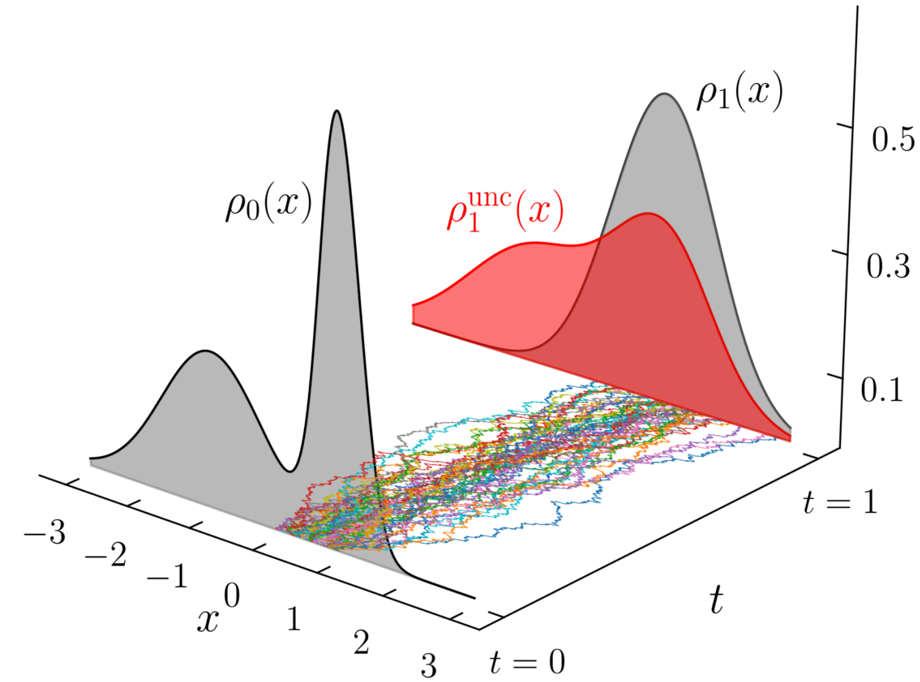




Most likely evolution between 2 distributional snapshots

Classical SBP





$$\underset{(\rho,\boldsymbol{v})\in\mathcal{P}_{01}\times\mathcal{V}}{\operatorname{arginf}} \int_{t_0}^{t_1}\int_{\mathbb{R}^n} \frac{1}{2}|\boldsymbol{v}|^2\rho(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t$$

$$\frac{\partial\rho}{\partial t} + \nabla_{\boldsymbol{x}}\cdot(\rho\boldsymbol{v}) = \varepsilon\Delta_{\boldsymbol{x}}\rho, \quad \varepsilon > 0,$$

$$\rho(\boldsymbol{x}, t = t_0) = \rho_0, \quad \rho(\boldsymbol{x}, t = t_1) = \rho_1,$$

Controlled sample path dynamics

$$\mathrm{d}\boldsymbol{x} = \boldsymbol{v}(\boldsymbol{x},t)\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}\boldsymbol{w}(t)$$

Solution to the Classical SBP

The pair $(\rho_\varepsilon^{\text{opt}}, \boldsymbol{v}_\varepsilon^{\text{opt}})$ solves the coupled PDEs

Value function

$$\frac{\partial \psi_\varepsilon}{\partial t} + \frac{1}{2}\left|\nabla_x \psi_\varepsilon\right|^2 + \varepsilon \Delta_x \psi_\varepsilon = 0,$$

$$\frac{\partial \rho_\varepsilon^{\text{opt}}}{\partial t} + \nabla_x \cdot \left(\rho_\varepsilon^{\text{opt}} \nabla_x \psi_\varepsilon\right) = \varepsilon \Delta_x \rho_\varepsilon^{\text{opt}}$$

Hopf-Cole transform

$$\varphi_\varepsilon := \exp\left(\frac{\psi_\varepsilon}{2\varepsilon}\right), \quad \widehat{\varphi}_\varepsilon := \rho_\varepsilon^{\text{opt}} \exp\left(-\frac{\psi_\varepsilon}{2\varepsilon}\right)$$

Schrödinger factors

Optimally controlled joint state PDF: $\rho_\varepsilon^{\text{opt}}(\cdot, t) = \widehat{\varphi}_\varepsilon(\cdot, t)\varphi_\varepsilon(\cdot, t)$

Optimal control: $\boldsymbol{v}_\varepsilon^{\text{opt}}(\cdot, t) = 2\varepsilon \nabla_{\boldsymbol{x}} \log \varphi_\varepsilon(\cdot, t)$

Schrödinger factors

Algorithm

Fixed point recursion over pair $\left(\varphi_{\varepsilon,1}, \widehat{\varphi}_{\varepsilon,0}\right)$

$$
\begin{tikzcd}[column sep=huge, row sep=huge]
\widehat{\varphi}_{\varepsilon,0}(\cdot) \arrow[r, "\int \mathcal{L}_{\text{forward}} \widehat{\varphi}"] & \widehat{\varphi}_{\varepsilon}(\cdot, t = t_1) \arrow[d, "\rho_1(\cdot) / \widehat{\varphi}_{\varepsilon}(\cdot, t = t_1)"] \\
\varphi_{\varepsilon}(\cdot, t = t_0) \arrow[u, "\rho_0(\cdot) / \varphi_{\varepsilon}(\cdot, t = t_0)"] & \varphi_{\varepsilon,1}(\cdot) \arrow[l, "\int \mathcal{L}_{\text{backward}} \varphi"]
\end{tikzcd}
$$

Schrödinger system

Markov kernel

$$\rho_0(\boldsymbol{x}) = \widehat{\varphi}_{\varepsilon,0}(\boldsymbol{x}) \int_{\mathbb{R}^n} k(t_0, \boldsymbol{x}, t_1, \boldsymbol{y}) \varphi_{\varepsilon,1}(\boldsymbol{y}) \mathrm{d}\boldsymbol{y}$$

$$\rho_1(\boldsymbol{x}) = \varphi_{\varepsilon,1}(\boldsymbol{x}) \int_{\mathbb{R}^n} k(t_0, \boldsymbol{y}, t_1, \boldsymbol{x}) \widehat{\varphi}_{\varepsilon,0}(\boldsymbol{y}) \mathrm{d}\boldsymbol{y}$$

Coupled nonlinear integral equations

Contraction Coefficient γ in Classical SBP

Let

$$\tilde{\alpha}_{\mathrm{B}} := \max_{\boldsymbol{x}_0 \in \mathcal{X}_0, \boldsymbol{x}_1 \in \mathcal{X}_1} |\boldsymbol{x}_0 - \boldsymbol{x}_1|^2 \quad \text{and} \quad \tilde{\beta}_{\mathrm{B}} := \min_{\boldsymbol{x}_0 \in \mathcal{X}_0, \boldsymbol{x}_1 \in \mathcal{X}_1} |\boldsymbol{x}_0 - \boldsymbol{x}_1|^2$$

Then

$$\boxed{\gamma_{\mathrm{B}} = \tanh^2\left(\frac{\tilde{\alpha}_{\mathrm{B}} - \tilde{\beta}_{\mathrm{B}}}{8\varepsilon}\right) \in (0,1)}$$

Chen, Georgiou, Pavon, ***SIAM J. Applied Math, 2016***

Linear SBP

$$\underset{(\rho,\boldsymbol{v})\in\mathcal{P}_{01}\times\mathcal{V}}{\arg\inf} \quad \int_{t_0}^{t_1}\int_{\mathbb{R}^n} \frac{1}{2}|\boldsymbol{v}|^2\rho(\boldsymbol{x},t)\,d\boldsymbol{x}\,dt$$

$$\frac{\partial\rho}{\partial t} + \nabla_{\boldsymbol{x}}\cdot(\rho(\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{v})) = \varepsilon\langle\text{Hess}, \boldsymbol{B}(t)\boldsymbol{B}(t)^\top\rho\rangle$$

resp. compact supports $\mathcal{X}_0, \mathcal{X}_1$

$$\rho(\boldsymbol{x},t=t_0) = \rho_0, \quad \rho(\boldsymbol{x},t=t_1) = \rho_1$$

Controlled sample path dynamics

$$\mathrm{d}\boldsymbol{x}(t) = (\boldsymbol{A}(t)\boldsymbol{x}(t)+\boldsymbol{B}(t)\boldsymbol{v}(\boldsymbol{x},t))\mathrm{d}t + \sqrt{2\varepsilon}\boldsymbol{B}(t)\mathrm{d}\boldsymbol{w}(t)$$

State transition matrix $\boldsymbol{\Phi}_{t\tau} := \boldsymbol{\Phi}(t,\tau) \quad \forall t_0\leq\tau\leq t\leq t_1$

Assume controllability: $\boldsymbol{M}_{10} := \int_{t_0}^{t_1}\boldsymbol{\Phi}_{t_1\tau}\boldsymbol{B}(\tau)\boldsymbol{B}^\top(\tau)\boldsymbol{\Phi}_{t_1\tau}^\top\mathrm{d}\tau \succ \boldsymbol{0}$

Structure of the Solution for Linear SBP

Optimally controlled joint state PDF: $\rho_{\varepsilon}^{\text{opt}}(\cdot, t) = \widehat{\varphi}_{\varepsilon}(\cdot, t)\varphi_{\varepsilon}(\cdot, t)$

Optimal control: $\boldsymbol{v}_{\varepsilon}^{\text{opt}}(\cdot, t) = 2\varepsilon \boldsymbol{B}(t)^{\top}\nabla_{\boldsymbol{x}} \log \varphi_{\varepsilon}(\cdot, t)$

Schrödinger factors

Define: $\widehat{\varphi}_{\varepsilon,0}(\cdot) := \widehat{\varphi}_{\varepsilon}(\cdot, t = t_0), \quad \varphi_{\varepsilon,1}(\cdot) := \varphi_{\varepsilon}(\cdot, t = t_1)$

Schrödinger system

Markov kernel

$$\rho_0(\boldsymbol{x}) = \widehat{\varphi}_{\varepsilon,0}(\boldsymbol{x}) \int_{\mathbb{R}^n} \underbrace{k(t_0, \boldsymbol{x}, t_1, \boldsymbol{y})}_{\text{Markov kernel}} \varphi_{\varepsilon,1}(\boldsymbol{y}) \mathrm{d}\boldsymbol{y}$$

$$\rho_1(\boldsymbol{x}) = \varphi_{\varepsilon,1}(\boldsymbol{x}) \int_{\mathbb{R}^n} k(t_0, \boldsymbol{y}, t_1, \boldsymbol{x}) \widehat{\varphi}_{\varepsilon,0}(\boldsymbol{y}) \mathrm{d}\boldsymbol{y}$$

Coupled nonlinear integral equations

Here

$$k(t_0, \boldsymbol{x}_0, t_1, \boldsymbol{x}_1) := \frac{\exp\left(-\frac{\left(\boldsymbol{\Phi}_{t_1 t_0}\boldsymbol{x}_0 - \boldsymbol{x}_1\right)^{\top} \boldsymbol{M}_{10}^{-1} \left(\boldsymbol{\Phi}_{t_1 t_0}\boldsymbol{x}_0 - \boldsymbol{x}_1\right)}{4\varepsilon}\right)}{\sqrt{(4\pi\varepsilon)^n \det(\boldsymbol{M}_{10})}}$$

Contractive Fixed Point Algorithm

Fixed point recursion over pair $(\varphi_{\varepsilon,1}, \widehat{\varphi}_{\varepsilon,0})$

$$\begin{tikzcd}[column sep=huge, row sep=huge]
\widehat{\varphi}_{\varepsilon,0}(\cdot) \arrow[r, "\int", "\text{IVP forward in time}"'] & \widehat{\varphi}_{\varepsilon}(\cdot, t = t_1) \arrow[d, "\rho_1(\cdot)/\widehat{\varphi}_{\varepsilon}(\cdot, t = t_1)"] \\
\varphi_{\varepsilon}(\cdot, t = t_0) \arrow[u, "\rho_0(\cdot)/\varphi_{\varepsilon}(\cdot, t = t_0)"] & \varphi_{\varepsilon,1}(\cdot) \arrow[l, "\text{IVP backward in time}"', "\int"]
\end{tikzcd}$$

Guaranteed linear convergence with contraction rate $\kappa \in (0,1)$

But exact rate depends on problem data $(\mathcal{X}_0, \mathcal{X}_1, \varepsilon, \boldsymbol{A}(t), \boldsymbol{B}(t))$

Worst case contraction coefficient $\gamma := \sup\limits_{\text{Linear SBPs with fixed } (\mathcal{X}_0, \mathcal{X}_1, \varepsilon, \boldsymbol{A}(t), \boldsymbol{B}(t))} \kappa$

Contraction Coefficient γ in Linear SBP

Thm. (informal)

State transition matrix   Controllability Gramian

Let

$$\tilde{\alpha}_{\mathrm{L}} := \max_{\boldsymbol{x}_0 \in \mathcal{X}_0, \boldsymbol{x}_1 \in \mathcal{X}_1} (\boldsymbol{\Phi}_{t_1 t_0} \boldsymbol{x}_0 - \boldsymbol{x}_1)^\top \boldsymbol{M}_{10}^{-1} (\boldsymbol{\Phi}_{t_1 t_0} \boldsymbol{x}_0 - \boldsymbol{x}_1)$$

$$\tilde{\beta}_{\mathrm{L}} := \min_{\boldsymbol{x}_0 \in \mathcal{X}_0, \boldsymbol{x}_1 \in \mathcal{X}_1} (\boldsymbol{\Phi}_{t_1 t_0} \boldsymbol{x}_0 - \boldsymbol{x}_1)^\top \boldsymbol{M}_{10}^{-1} (\boldsymbol{\Phi}_{t_1 t_0} \boldsymbol{x}_0 - \boldsymbol{x}_1)$$

Then

$$\gamma_{\mathrm{L}} = \tanh^2 \left(\frac{\tilde{\alpha}_{\mathrm{L}} - \tilde{\beta}_{\mathrm{L}}}{8\varepsilon} \right)$$

Control-theoretic Interpretation for γ_{L}

$$\tilde{\alpha}_{\mathrm{L}} := \max_{\boldsymbol{x}_0 \in \mathcal{X}_0, \boldsymbol{x}_1 \in \mathcal{X}_1} (\boldsymbol{\Phi}_{t_1 t_0}\boldsymbol{x}_0 - \boldsymbol{x}_1)^{\top} \boldsymbol{M}_{10}^{-1} (\boldsymbol{\Phi}_{t_1 t_0}\boldsymbol{x}_0 - \boldsymbol{x}_1)$$

$$\tilde{\beta}_{\mathrm{L}} := \min_{\boldsymbol{x}_0 \in \mathcal{X}_0, \boldsymbol{x}_1 \in \mathcal{X}_1} (\boldsymbol{\Phi}_{t_1 t_0}\boldsymbol{x}_0 - \boldsymbol{x}_1)^{\top} \boldsymbol{M}_{10}^{-1} (\boldsymbol{\Phi}_{t_1 t_0}\boldsymbol{x}_0 - \boldsymbol{x}_1)$$

$$\begin{aligned}
&\underset{\boldsymbol{v}}{\text{minimum}} \int_{t_0}^{t_1} \frac{1}{2} |\boldsymbol{v}|^2 dt \\
&\text{subject to} \quad \dot{\boldsymbol{x}} = \boldsymbol{A}(t)\boldsymbol{x} + \boldsymbol{B}(t)\boldsymbol{v} \\
&\qquad\qquad\quad \boldsymbol{x}(t = t_0) = \boldsymbol{x}_0, \boldsymbol{x}(t = t_1) = \boldsymbol{x}_1
\end{aligned}$$

Minimum cost for deterministic OCP

Control-theoretic Interpretation for γ_{L}

Range of optimal state transfer cost

$$\gamma_{\mathrm{L}} = \tanh^2\left(\frac{\tilde{\alpha}_{\mathrm{L}} - \tilde{\beta}_{\mathrm{L}}}{8\varepsilon}\right)$$

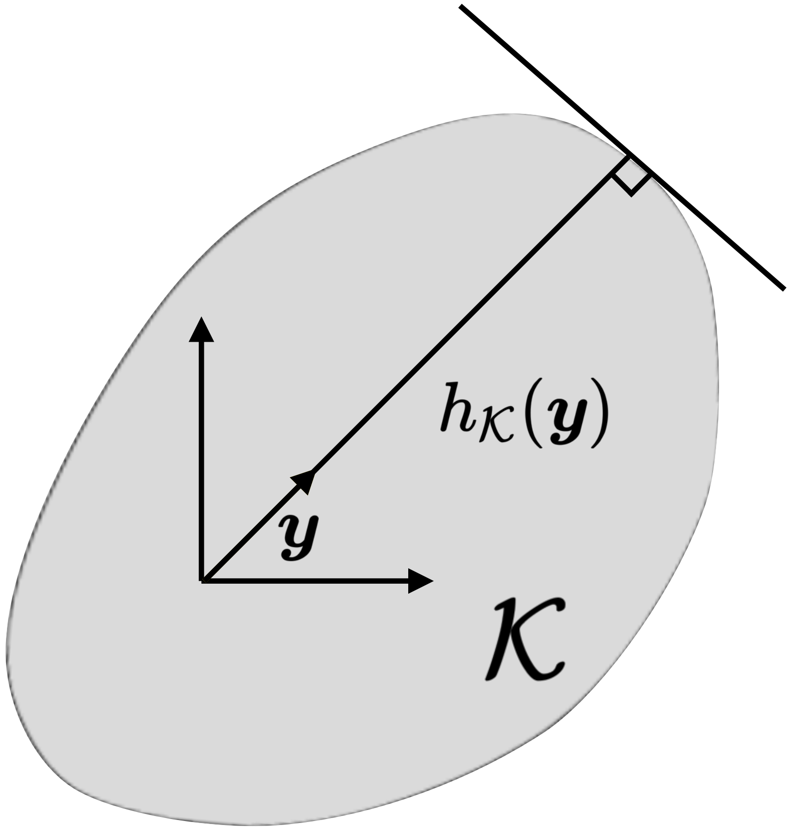
Process noise

Intuitive:

$$\boxed{\varepsilon \uparrow \quad \Rightarrow \quad \gamma_{\mathrm{L}} \downarrow}$$

$$\boxed{\tilde{\alpha}_{\mathrm{L}} - \tilde{\beta}_{\mathrm{L}} \uparrow \quad \Rightarrow \quad \gamma_{\mathrm{L}} \uparrow}$$

Support Function

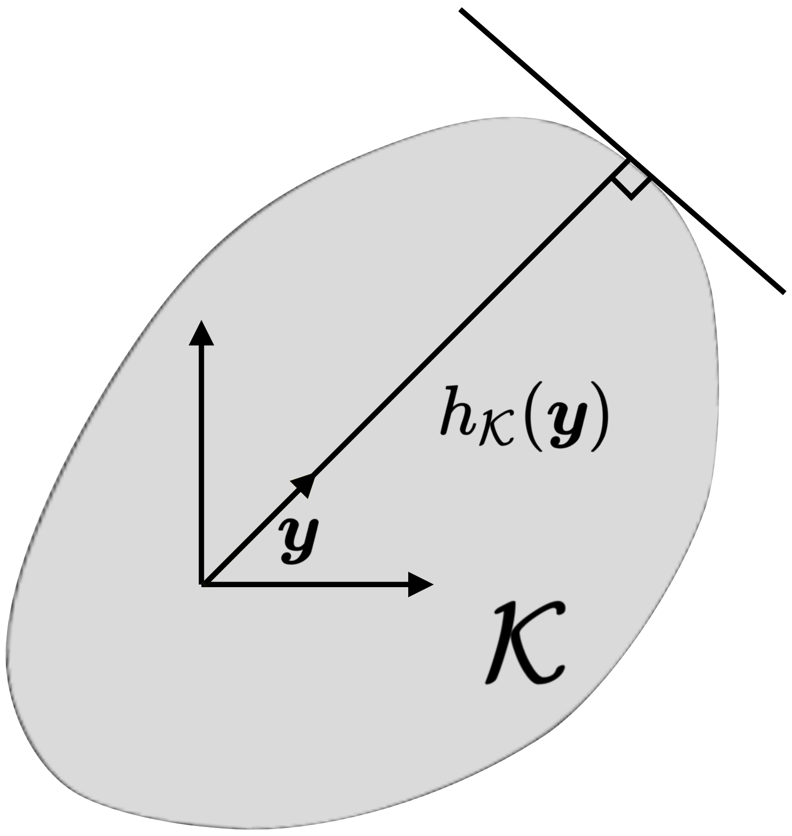


The support function $h_{\mathcal{K}}(\cdot)$ for closed convex set $\mathcal{K}$ is

$$h_{\mathcal{K}}(\boldsymbol{y}) := \sup_{\boldsymbol{x} \in \mathcal{K}} \langle \boldsymbol{y}, \boldsymbol{x} \rangle, \quad \boldsymbol{y} \in \mathbb{R}^n$$

e.g., distance from the origin to a supporting hyperplane of $\mathcal{K}$ with normal in direction of $\boldsymbol{y}$

Geometric Interpretation for γ_{L}



$$\gamma_{\mathrm{L}} = \tanh^2\left(\frac{\tilde{\alpha}_{\mathrm{L}} - \tilde{\beta}_{\mathrm{L}}}{8\varepsilon}\right)$$

$$\tilde{\alpha}_{\mathrm{L}} = \{\max_{\boldsymbol{y}\in\mathcal{S}^{n-1}} (h_{\mathcal{X}_0}(\boldsymbol{\Phi}_{t_1 t_0}^{\top}\boldsymbol{M}_{10}^{-1/2}\boldsymbol{y}) + h_{\mathcal{X}_1}(-\boldsymbol{M}_{10}^{-1/2}\boldsymbol{y}))\}^2$$

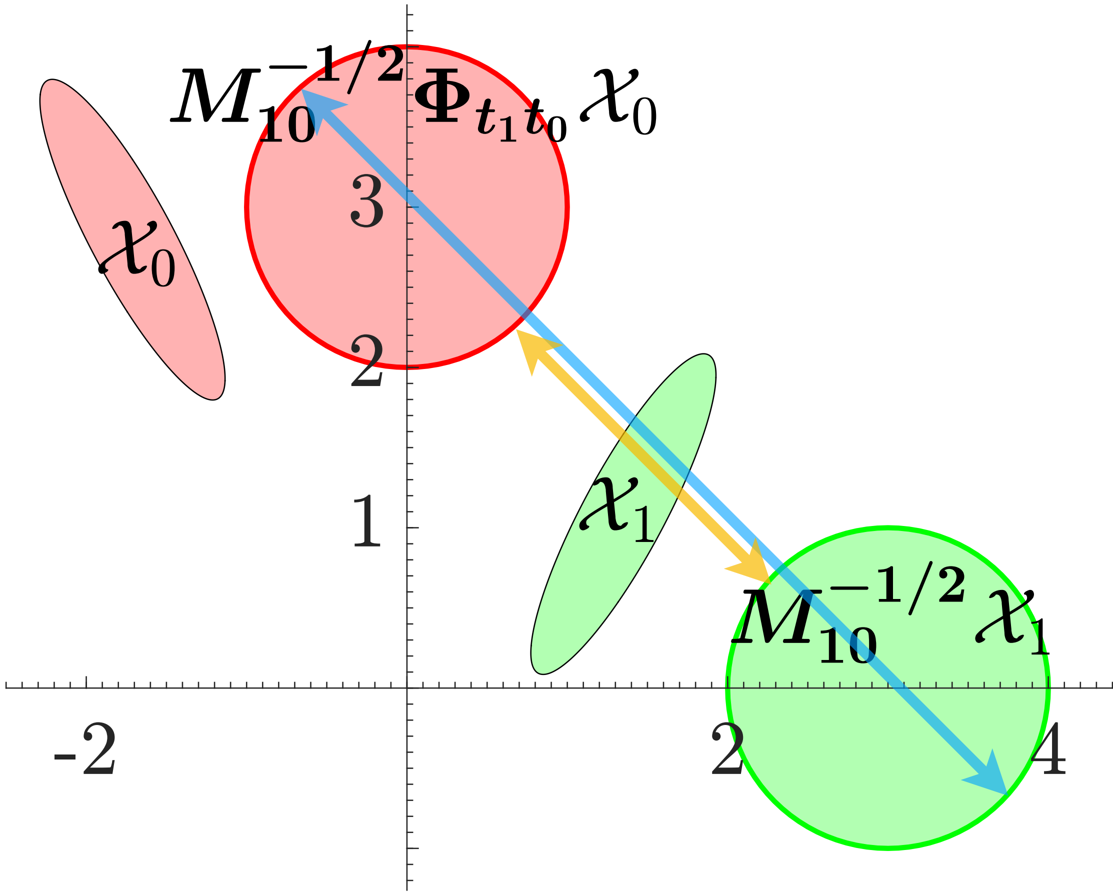
$$\tilde{\beta}_{\mathrm{L}} = \{\min_{\boldsymbol{y}\in\mathcal{S}^{n-1}} (h_{\mathcal{X}_0}(\boldsymbol{\Phi}_{t_1 t_0}^{\top}\boldsymbol{M}_{10}^{-1/2}\boldsymbol{y}) + h_{\mathcal{X}_1}(-\boldsymbol{M}_{10}^{-1/2}\boldsymbol{y}))\}^2$$

Geometric Interpretation for γ_{L}

$$\gamma_{\mathrm{L}} = \tanh^2\left(\frac{\tilde{\alpha}_{\mathrm{L}} - \tilde{\beta}_{\mathrm{L}}}{8\varepsilon}\right)$$

$$\tilde{\alpha}_{\mathrm{L}} := \max_{\boldsymbol{x}_0 \in \boldsymbol{M}_{10}^{-1/2}\boldsymbol{\Phi}_{10}\mathcal{X}_0,\ \boldsymbol{x}_1 \in \boldsymbol{M}_{10}^{-1/2}\mathcal{X}_1} |\boldsymbol{x}_0 - \boldsymbol{x}_1|^2$$

$$\tilde{\beta}_{\mathrm{L}} := \min_{\boldsymbol{x}_0 \in \boldsymbol{M}_{10}^{-1/2}\boldsymbol{\Phi}_{10}\mathcal{X}_0,\ \boldsymbol{x}_1 \in \boldsymbol{M}_{10}^{-1/2}\mathcal{X}_1} |\boldsymbol{x}_0 - \boldsymbol{x}_1|^2$$



$\tilde{\alpha}_{\mathrm{L}}$ and $\tilde{\beta}_{\mathrm{L}}$ are the maximum and the minimum separation of the sets

$M_{10}^{-1/2}\boldsymbol{\Phi}_{t_1t_0}\mathcal{X}_0$ and $M_{10}^{-1/2}\mathcal{X}_1$

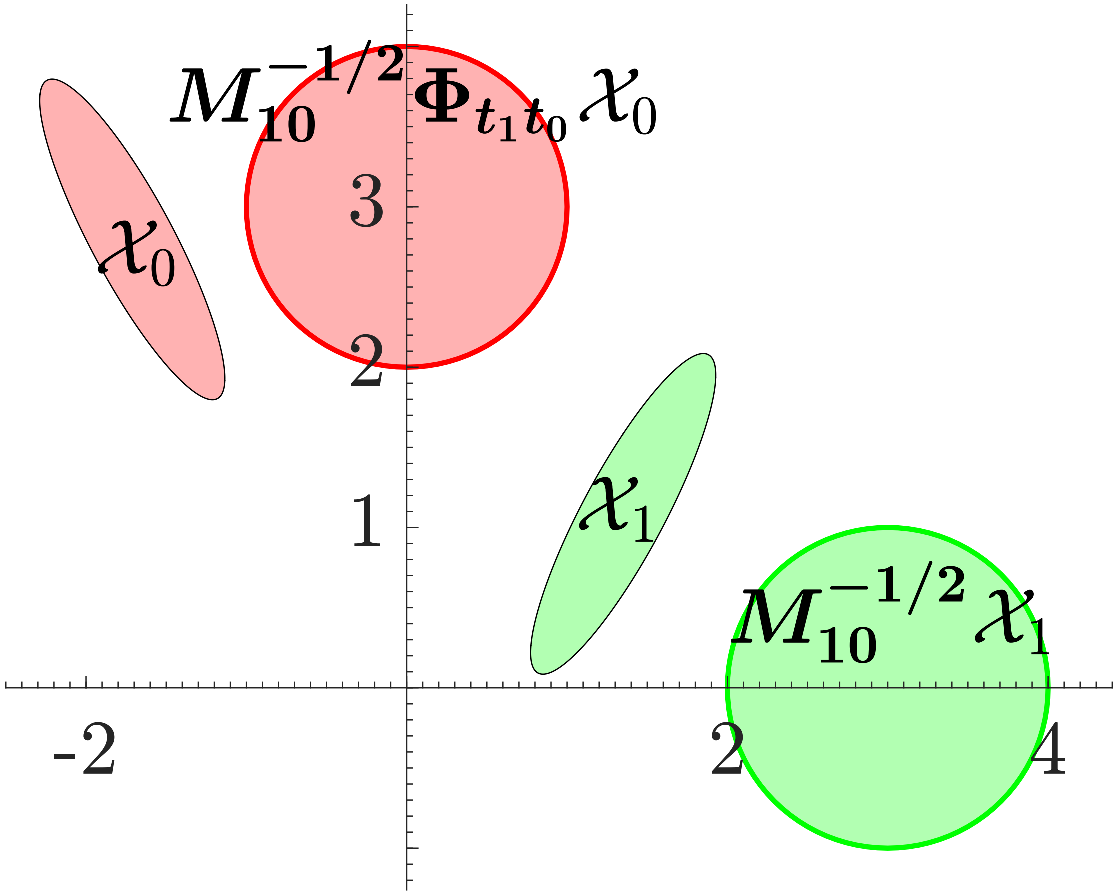
Applications to Preconditioning

Preconditioning to improve optimal transport algorithms
~ Kuang and Tabak, *SIAM J. Scientific Computing*, 2017

Example: Linear SBP:

$$d\boldsymbol{x}(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \boldsymbol{x}(t)dt + \sqrt{2\varepsilon} \begin{bmatrix} 0 \\ 1 \end{bmatrix} d\boldsymbol{w}(t)$$

$$\boldsymbol{\Phi}_{t_1 t_0} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{M}_{10}^{-1} = \begin{bmatrix} 12 & -6 \\ -6 & 4 \end{bmatrix}.$$



No Preconditioning:

$$\tilde{\alpha}_{\mathrm{L}} = 2 + 2\sqrt{3} \qquad \tilde{\beta}_{\mathrm{L}} = -2 + 2\sqrt{3} \longrightarrow \gamma_{\mathrm{L}} = \tanh^2(1) \approx 0.580$$

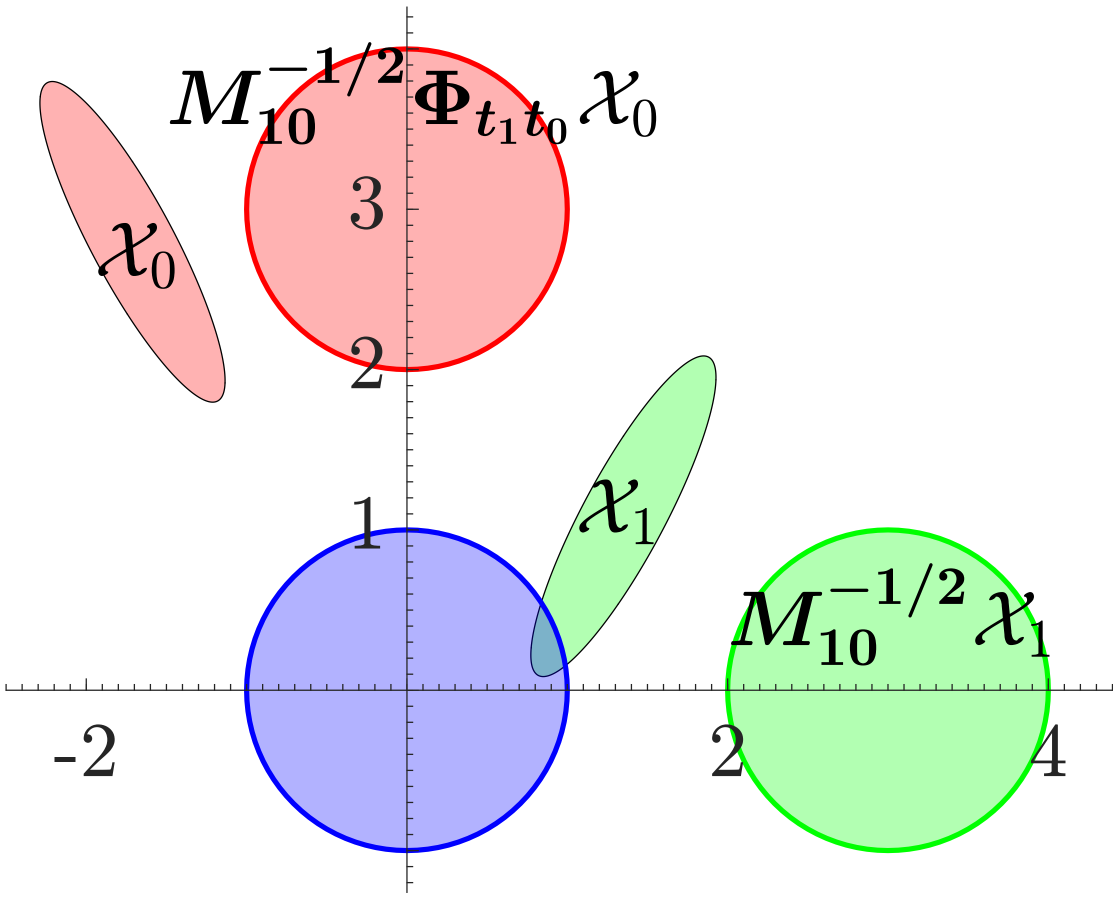
Applications to Preconditioning

Preconditioning to improve optimal transport algorithms
~ Kuang and Tabak, *SIAM J. Scientific Computing*, 2017

Example: Linear SBP:

$$d\boldsymbol{x}(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \boldsymbol{x}(t)dt + \sqrt{2\varepsilon} \begin{bmatrix} 0 \\ 1 \end{bmatrix} d\boldsymbol{w}(t)$$

$$\boldsymbol{\Phi}_{t_1 t_0} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{M}_{10}^{-1} = \begin{bmatrix} 12 & -6 \\ -6 & 4 \end{bmatrix}.$$



With Preconditioning:

$$\tilde{\alpha}_{\mathrm{L}}^{\mathrm{precond}} = 2, \ \tilde{\beta}_{\mathrm{L}}^{\mathrm{precond}} = 0 \quad \longrightarrow \quad \gamma_{\mathrm{L}}^{\mathrm{precond}} = \tanh^2(0.5) = 0.214$$

Avenues of Future Research

Preconditioning:

—> Demonstrate effectiveness of proposed methods

—> Inform the construction of new methods

Additional Generalizations of SBP:

—> Current: Quadratic state cost

—> Future: Keplerian state cost

Thank You

Acknowledgement: NSF 2112755