Gradient Flow Algorithms for Density Propagation in Stochastic Systems

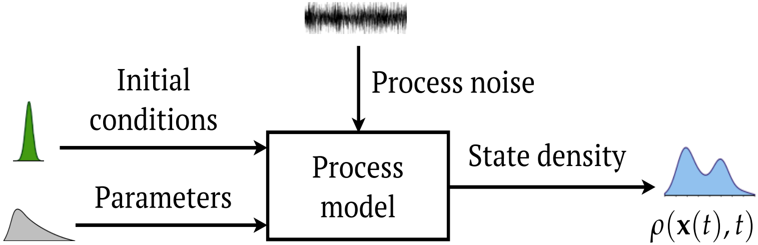
Kenneth Caluya

Joint work with Abhishek Halder

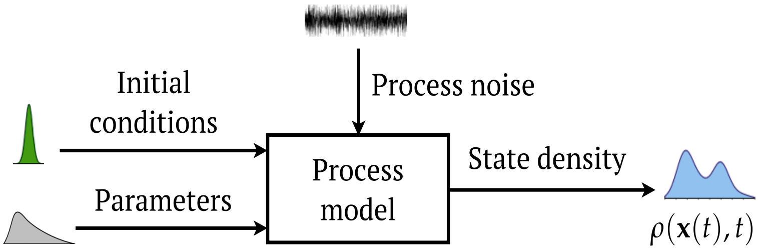
University of California, Santa Cruz

CITRIS/CPAR Control Theory and Automation Symposium
2nd NorCal Control Workshop, UC Berkeley, April 26, 2019

Problem: Density Propagation



Problem: Density Propagation



State Dynamics:

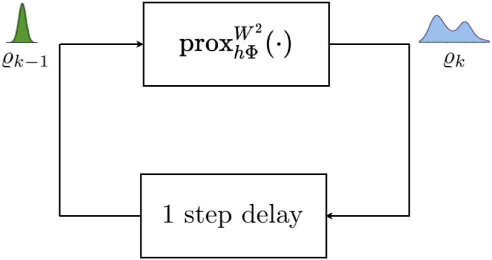
$$\mathrm{d}\boldsymbol{x} = -\nabla\psi(\boldsymbol{x}, t)\,\mathrm{d}t + \sqrt{2\beta^{-1}}\,\mathrm{d}\boldsymbol{w}, \;\; \boldsymbol{x}_0 \sim \rho_0, \;\; \mathrm{d}\boldsymbol{w}(t) \sim \mathcal{N}(0, \boldsymbol{I}\mathrm{d}t), \;\; \boldsymbol{x} \in \mathbb{R}^n$$

Probability Density Function (PDF) Dynamics:

$$\frac{\partial\rho}{\partial t} = \mathcal{L}\rho := \nabla\cdot(\nabla\psi\rho) + \beta^{-1}\Delta\rho, \quad \rho(\boldsymbol{x}, 0) = \rho_0$$

What's new?

Main Idea: Solve $\dfrac{\partial \rho}{\partial t} = \mathcal{L}\rho,\ \rho(\boldsymbol{x}, 0) = \rho_0$ **as gradient flow in** $\mathcal{P}_2(\mathbb{R}^n)$



Proximal Operator: $\rho_k = \mathrm{prox}_{h\Phi}^{W^2}(\rho_{k-1}) := \underset{\rho \in \mathcal{P}_2(\mathbb{R}^n)}{\arg\inf} \left\{ \dfrac{1}{2} W^2(\rho, \rho_{k-1}) + h\,\Phi(\rho) \right\}$

Optimal Transport Cost: $W^2(\rho, \rho_{k-1}) := \displaystyle\inf_{\pi \in \Pi(\rho, \rho_{k-1})} \int_{\mathbb{R}^n \times \mathbb{R}^n} c(\boldsymbol{x}, \boldsymbol{y}) \mathrm{d}\pi(\boldsymbol{x}, \boldsymbol{y})$

Free Energy Functional: $\Phi(\rho) := \displaystyle\int_{\mathbb{R}^n} \psi \rho \, \mathrm{d}\boldsymbol{x} + \beta^{-1} \int_{\mathbb{R}^n} \rho \log \rho \, \mathrm{d}\boldsymbol{x}$

Gradient Flow

Gradient Flow in $\mathbb{R}^n$

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = -\nabla\varphi(\boldsymbol{x}), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0$$

Recursion:

$$\begin{aligned}\boldsymbol{x}_k &= \boldsymbol{x}_{k-1} - h\nabla\varphi(\boldsymbol{x}_k)\\ &= \underset{\boldsymbol{x}\in\mathbb{R}^n}{\arg\min}\left\{\frac{1}{2}\|\boldsymbol{x}-\boldsymbol{x}_{k-1}\|_2^2 + h\varphi(\boldsymbol{x})\right\}\\ &=: \mathrm{prox}_{h\varphi}^{\|\cdot\|_2}(\boldsymbol{x}_{k-1})\end{aligned}$$

Convergence:

$$\boldsymbol{x}_k \to \boldsymbol{x}(t = kh) \quad \text{as} \quad h \downarrow 0$$

Gradient Flow

Gradient Flow in $\mathbb{R}^n$

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = -\nabla\varphi(\boldsymbol{x}), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0$$

Recursion:

$$\begin{aligned}\boldsymbol{x}_k &= \boldsymbol{x}_{k-1} - h\nabla\varphi(\boldsymbol{x}_k)\\ &= \underset{\boldsymbol{x}\in\mathbb{R}^n}{\arg\min}\left\{\frac{1}{2}\|\boldsymbol{x}-\boldsymbol{x}_{k-1}\|_2^2 + h\varphi(\boldsymbol{x})\right\}\\ &=: \mathrm{prox}_{h\varphi}^{\|\cdot\|_2}(\boldsymbol{x}_{k-1})\end{aligned}$$

Convergence:

$$\boldsymbol{x}_k \to \boldsymbol{x}(t = kh) \quad \text{as} \quad h \downarrow 0$$

Gradient Flow in $\mathcal{P}_2(\mathbb{R}^n)$

$$\frac{\partial\rho}{\partial t} = -\nabla^W\Phi(\rho), \quad \rho(\boldsymbol{x}, 0) = \rho_0$$

Recursion:

$$\begin{aligned}\rho_k &= \rho(\cdot, t = kh)\\ &= \underset{\rho\in\mathcal{P}_2(\mathbb{R}^n)}{\arg\min}\left\{\frac{1}{2}W^2(\rho,\rho_{k-1}) + h\Phi(\rho)\right\}\\ &=: \mathrm{prox}_{h\Phi}^{W^2}(\rho_{k-1})\end{aligned}$$

Convergence:

$$\rho_k \to \rho(\cdot, t = kh) \quad \text{as} \quad h \downarrow 0$$

Algorithm: Gradient Ascent on the Dual Space

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\nabla \psi \rho) + \beta^{-1} \Delta \rho$$

$\Updownarrow$ **Proximal Recursion**

$$\rho_k = \rho(\boldsymbol{x}, t = kh) = \underset{\rho \in \mathcal{P}_2(\mathbb{R}^n)}{\arg\inf} \left\{ \frac{1}{2} W^2(\rho, \rho_{k-1}) + h\, \Phi(\rho) \right\}$$

Algorithm: Gradient Ascent on the Dual Space

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\nabla \psi \rho) + \beta^{-1} \Delta \rho$$

$\Updownarrow$ **Proximal Recursion**

$$\rho_k = \rho(\boldsymbol{x}, t = kh) = \underset{\rho \in \mathcal{P}_2(\mathbb{R}^n)}{\arg\inf} \left\{ \frac{1}{2} W^2(\rho, \rho_{k-1}) + h\, \Phi(\rho) \right\}$$

$\Downarrow$ **Discrete Primal Formulation**

$$\boldsymbol{\varrho}_k = \arg\min_{\boldsymbol{\varrho}} \left\{ \min_{\boldsymbol{M} \in \Pi(\boldsymbol{\varrho}_{k-1}, \boldsymbol{\varrho})} \frac{1}{2} \langle \boldsymbol{C}_k, \boldsymbol{M} \rangle + h \, \langle \boldsymbol{\psi}_{k-1} + \beta^{-1} \log \boldsymbol{\varrho}, \boldsymbol{\varrho} \rangle \right\}$$

Algorithm: Gradient Ascent on the Dual Space

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\nabla \psi \rho) + \beta^{-1} \Delta \rho$$

$\Updownarrow$ **Proximal Recursion**

$$\rho_k = \rho(\boldsymbol{x}, t = kh) = \underset{\rho \in \mathcal{P}_2(\mathbb{R}^n)}{\arg\inf} \left\{ \frac{1}{2} W^2(\rho, \rho_{k-1}) + h\, \Phi(\rho) \right\}$$

$\Downarrow$ **Discrete Primal Formulation**

$$\boldsymbol{\varrho}_k = \arg\min_{\boldsymbol{\varrho}} \left\{ \min_{\boldsymbol{M} \in \Pi(\boldsymbol{\varrho}_{k-1}, \boldsymbol{\varrho})} \frac{1}{2} \langle \boldsymbol{C}_k, \boldsymbol{M} \rangle + h \left\langle \boldsymbol{\psi}_{k-1} + \beta^{-1} \log \boldsymbol{\varrho}, \boldsymbol{\varrho} \right\rangle \right\}$$

$\Downarrow$ **Entropic Regularization**

$$\boldsymbol{\varrho}_k = \arg\min_{\boldsymbol{\varrho}} \left\{ \min_{\boldsymbol{M} \in \Pi(\boldsymbol{\varrho}_{k-1}, \boldsymbol{\varrho})} \frac{1}{2} \langle \boldsymbol{C}_k, \boldsymbol{M} \rangle + \epsilon H(\boldsymbol{M}) + h \left\langle \boldsymbol{\psi}_{k-1} + \beta^{-1} \log \boldsymbol{\varrho}, \boldsymbol{\varrho} \right\rangle \right\}$$

Algorithm: Gradient Ascent on the Dual Space

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\nabla \psi \rho) + \beta^{-1} \Delta \rho$$

$\Updownarrow$ **Proximal Recursion**

$$\rho_k = \rho(\boldsymbol{x}, t = kh) = \underset{\rho \in \mathcal{P}_2(\mathbb{R}^n)}{\arg\inf} \left\{ \frac{1}{2} W^2(\rho, \rho_{k-1}) + h\, \Phi(\rho) \right\}$$

$\Downarrow$ **Discrete Primal Formulation**

$$\boldsymbol{\varrho}_k = \arg\min_{\boldsymbol{\varrho}} \left\{ \min_{\boldsymbol{M} \in \Pi(\boldsymbol{\varrho}_{k-1}, \boldsymbol{\varrho})} \frac{1}{2} \langle \boldsymbol{C}_k, \boldsymbol{M} \rangle + h \left\langle \boldsymbol{\psi}_{k-1} + \beta^{-1} \log \boldsymbol{\varrho}, \boldsymbol{\varrho} \right\rangle \right\}$$

$\Downarrow$ **Entropic Regularization**

$$\boldsymbol{\varrho}_k = \arg\min_{\boldsymbol{\varrho}} \left\{ \min_{\boldsymbol{M} \in \Pi(\boldsymbol{\varrho}_{k-1}, \boldsymbol{\varrho})} \frac{1}{2} \langle \boldsymbol{C}_k, \boldsymbol{M} \rangle + \epsilon H(\boldsymbol{M}) + h \left\langle \boldsymbol{\psi}_{k-1} + \beta^{-1} \log \boldsymbol{\varrho}, \boldsymbol{\varrho} \right\rangle \right\}$$

$\Updownarrow$ **Dualization**

$$\boldsymbol{\lambda}_0^{\mathrm{opt}}, \boldsymbol{\lambda}_1^{\mathrm{opt}} = \underset{\boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1 \geq 0}{\arg\max} \Bigg\{ \langle \boldsymbol{\lambda}_0, \boldsymbol{\varrho}_{k-1} \rangle - F^{\star}(-\boldsymbol{\lambda}_1) - \frac{\epsilon}{h} \left(\exp(\boldsymbol{\lambda}_0^{\top} h/\epsilon) \exp(-\boldsymbol{C}_k/2\epsilon) \exp(\boldsymbol{\lambda}_1 h/\epsilon) \right) \Bigg\}$$

Fixed Point Recursion

$$\boldsymbol{y} = e^{\frac{\boldsymbol{\lambda}_0^*}{\epsilon}} h \Big\downarrow \qquad \Big\downarrow \boldsymbol{z} = e^{\frac{\boldsymbol{\lambda}_1^*}{\epsilon}} h$$

Coupled Transcendental Equations in $\boldsymbol{y}$ and $\boldsymbol{z}$

$$\boldsymbol{\Gamma}_k = e^{\frac{-\boldsymbol{C}_k}{2\epsilon}} \longrightarrow, \quad \boldsymbol{\varrho}_{k-1} \longrightarrow, \quad \boldsymbol{\xi}_{k-1} = \frac{e^{-\beta \boldsymbol{\psi}_{k-1}}}{e} \longrightarrow$$

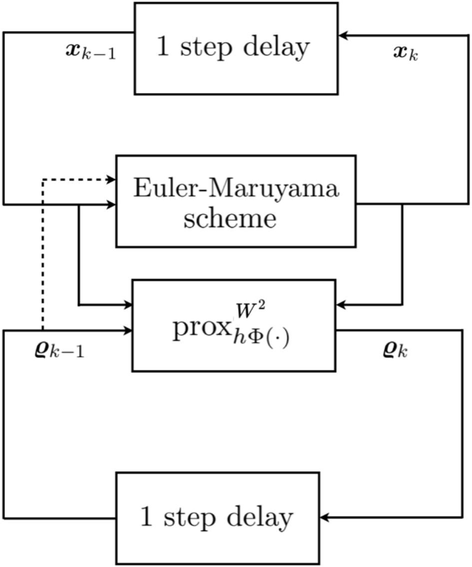
$$\boxed{\begin{aligned} \boldsymbol{y} \odot \boldsymbol{\Gamma}_k \boldsymbol{z} &= \boldsymbol{\varrho}_{k-1} \\ \boldsymbol{z} \odot \boldsymbol{\Gamma}_k^{\top} \boldsymbol{y} &= \boldsymbol{\xi}_{k-1} \odot \boldsymbol{z}^{-\beta\epsilon/2h} \end{aligned}} \longrightarrow \boldsymbol{\varrho}_k = \boldsymbol{z} \odot \boldsymbol{\Gamma}_k^{\top} \boldsymbol{y}$$

Theorem: Consider the recursion on the cone $\mathbb{R}^n_{\geq 0} \times \mathbb{R}^n_{\geq 0}$

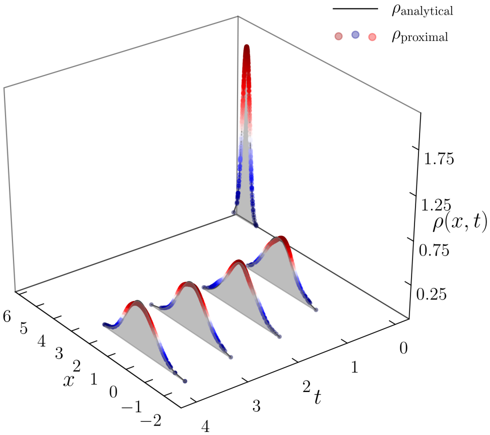
$$\boldsymbol{y} \odot (\boldsymbol{\Gamma}_k \boldsymbol{z}) = \boldsymbol{\varrho}_{k-1}, \quad \boldsymbol{z} \odot \left(\boldsymbol{\Gamma}_k{}^{\top} \boldsymbol{y}\right) = \boldsymbol{\xi}_{k-1} \odot \boldsymbol{z}^{-\frac{\beta\epsilon}{h}},$$

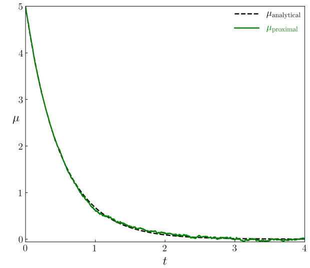
Then the solution $(\boldsymbol{y}^*, \boldsymbol{z}^*)$ gives the proximal update $\boldsymbol{\varrho}_k = \boldsymbol{z}^* \odot (\boldsymbol{\Gamma}_k{}^{\top} \boldsymbol{y}^*)$

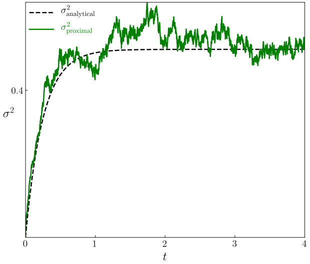
Algorithmic Setup



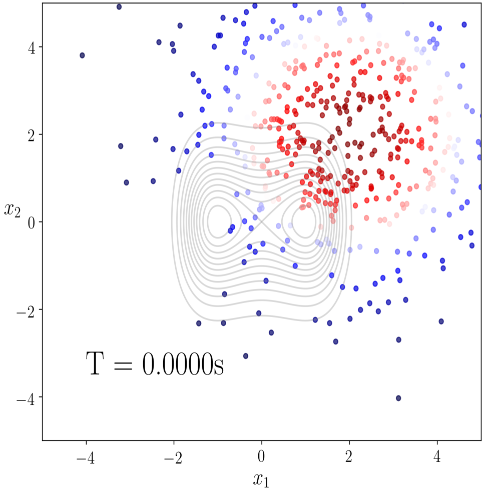
1D Linear Gaussian

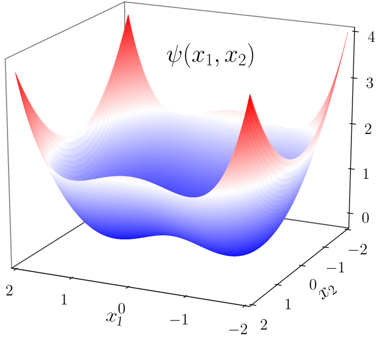






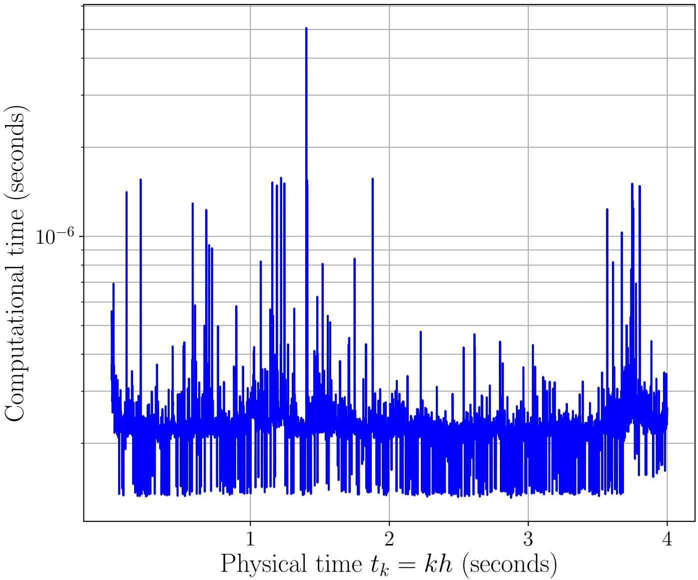
2D Nonlinear Non-Gaussian



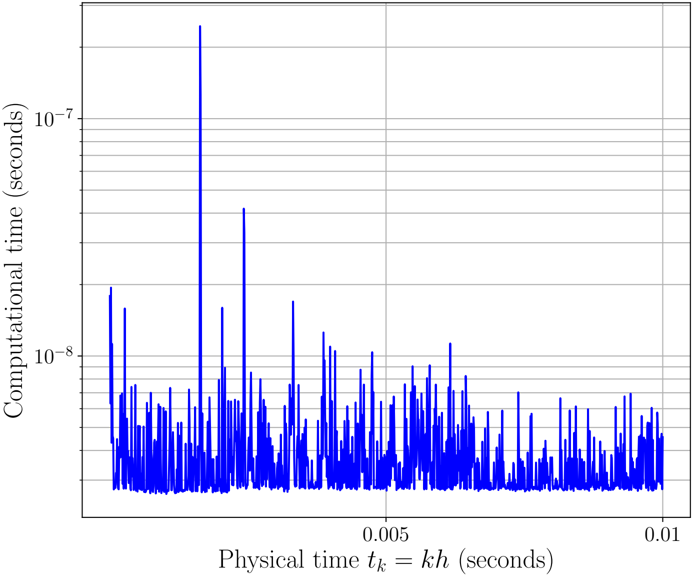


$$\rho_\infty \propto \exp(-\beta\psi)$$

Computational Time for 2D Nonlinear Non-Gaussian



Computational time for 6 State Satellite Motion

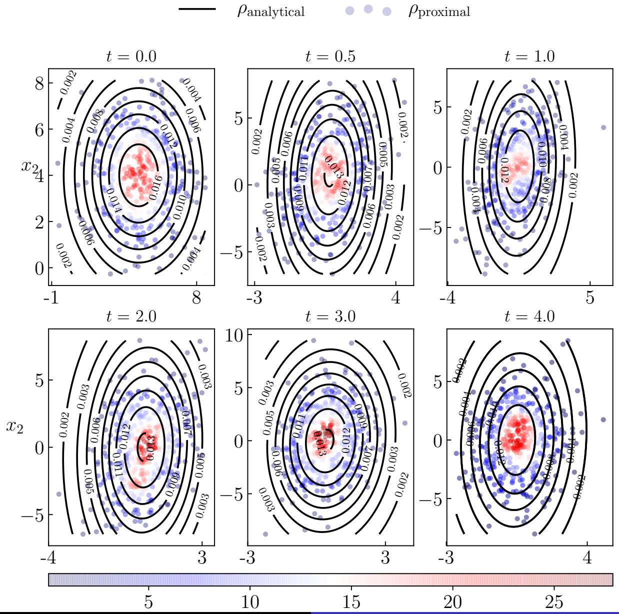


Summary

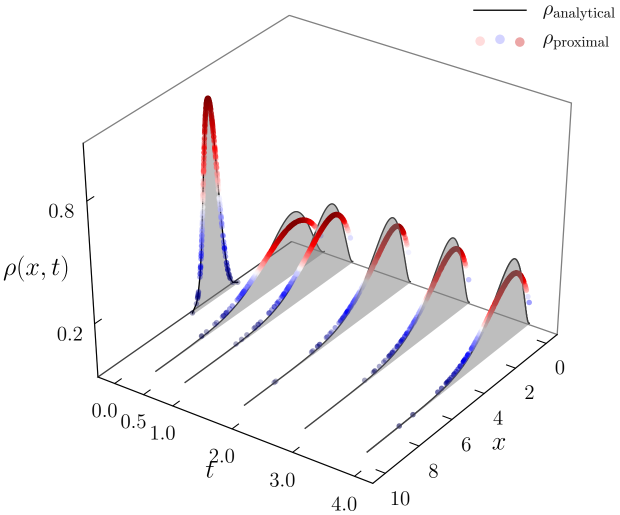
- New algorithm for density propagation
- No spatial discretization or function approximations
- Extremely fast runtime
- **Details:** `https://arxiv.org/pdf/1809.10844.pdf`

Thank You!

2D OU Process with Non-gradient Drift



Multiplicative Noise



Mixed Conservative-Dissipative Drift

Relative motion of a satellite in geocentric orbit:

$$\begin{pmatrix} \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}z \\ \mathrm{d}v_x \\ \mathrm{d}v_y \\ \mathrm{d}v_z \end{pmatrix} = \begin{pmatrix} v_x \\ v_y \\ v_z \\ -\dfrac{\mu x}{r^3} + (f_x)_{\text{pert}} - \gamma v_x \\ -\dfrac{\mu y}{r^3} + (f_y)_{\text{pert}} - \gamma v_y \\ -\dfrac{\mu z}{r^3} + (f_z)_{\text{pert}} - \gamma v_z \end{pmatrix} \mathrm{d}t + \sqrt{2\beta^{-1}\gamma} \begin{pmatrix} 0 \\ 0 \\ 0 \\ \mathrm{d}w_1 \\ \mathrm{d}w_2 \\ \mathrm{d}w_3 \end{pmatrix},$$

$$\begin{pmatrix} f_x \\ f_y \\ f_z \end{pmatrix}_{\text{pert}} = \begin{pmatrix} s\theta\, c\phi & c\theta\, c\phi & -s\phi \\ s\theta\, s\phi & c\theta\, s\phi & c\phi \\ c\theta & -s\theta & 0 \end{pmatrix} \begin{pmatrix} \dfrac{k}{2r^4}\left(3(s\theta)^2 - 1\right) \\ -\dfrac{k}{r^5} s\theta\, c\theta \\ 0 \end{pmatrix}, k := 3J_2R_{\mathrm{E}}^2, \mu = \text{constant}$$