Well-Posed KL-Regularized Control via Wasserstein and Kalman–Wasserstein KL Divergences

Viktor Stein

Technische Universität München

TUM School of Computation, Information and Technology

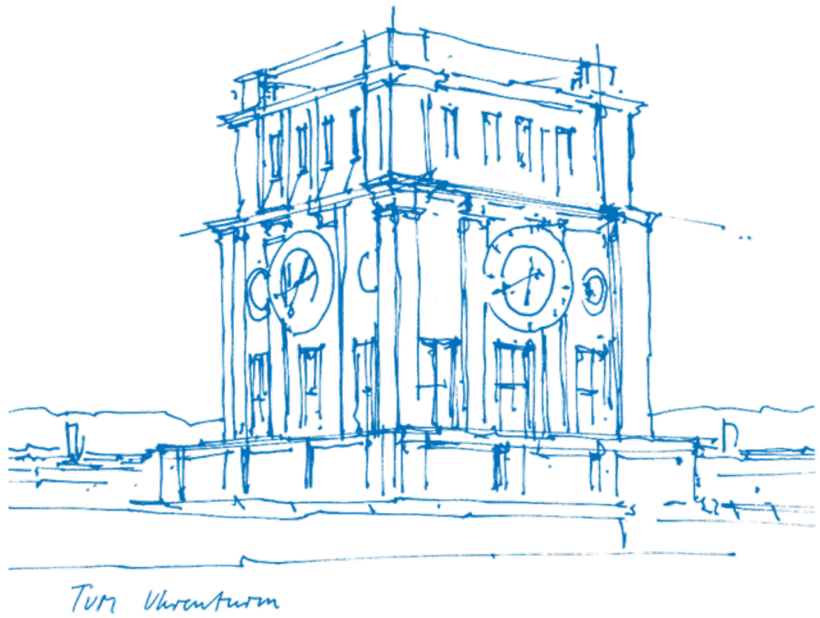
Research Group Applied Numerical Analysis

& Munich Center for Machine Learning

The majority of the work was conducted while at the

Institute for Mathematics, Technische Universität Berlin

& Berlin Mathematical School (BMS)



ICML 2026

Acknowledgements

This is joint work with

Adwait Datar, TUHH

Nihat Ay, TUHH & Santa Fe Institute

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. This project has received funding from the European Research Council (ERC) under the European Union's Horizon Europe research and innovation programme (grant agreement No. 101198055, project acronym NEITALG).

Take home message

- KL divergence used as regularizer in control problems
- **Problem:** KL degenerates under support mismatch / low noise
- **Idea:** Modify dynamical formulation of KL $\rightsquigarrow$ (Kalman-)Wasserstein KL, which respects the state space geometry

The dynamical formulation of the KL divergence

The Kullback-Leibler (KL) divergence between two probability measures $\mu, \nu \in \mathcal{P}(M)$ with $\mu \ll \nu$ is

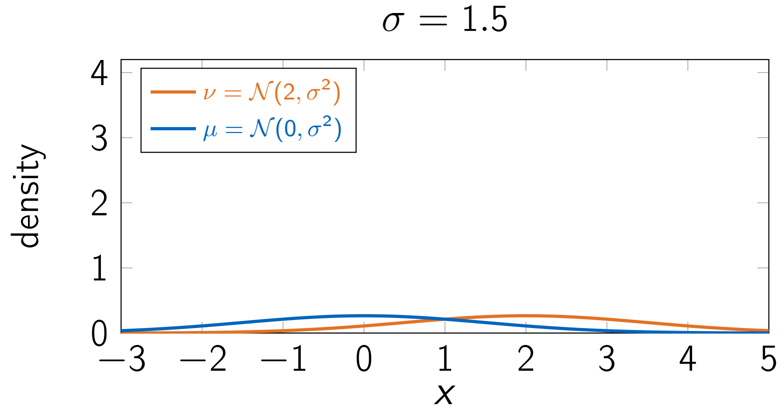
$$\mathrm{KL}(\mu \mid \nu) := \int_M \ln\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right) \mathrm{d}\mu = \int_0^1 t \|\dot{\gamma}(t)\|^2_{\gamma(t)} \, \mathrm{d}t, \qquad (\star)$$

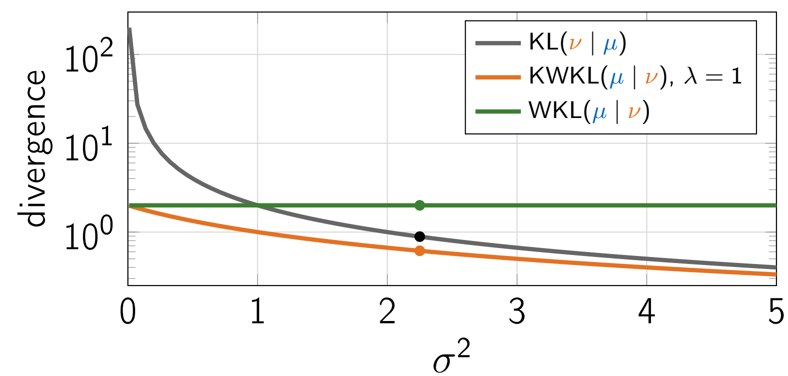
$\gamma =$ Fisher-Rao dual geodesic wrt the mixture connection with $\gamma(0) = \nu, \ \gamma(1) = \mu$.

KL degenerates when supports don't match:

$$\mathrm{KL}(\mathcal{N}(0, \sigma^2) \mid \mathcal{N}(2, \sigma^2)) = \frac{2}{\sigma^2} \xrightarrow{\sigma \to 0} \infty = \mathrm{KL}(\delta_x \mid \delta_y) \qquad \text{irregardless of } \|x - y\|_2$$

By contrast, $W_2(\delta_x, \delta_y) = \|x - y\|_2$.





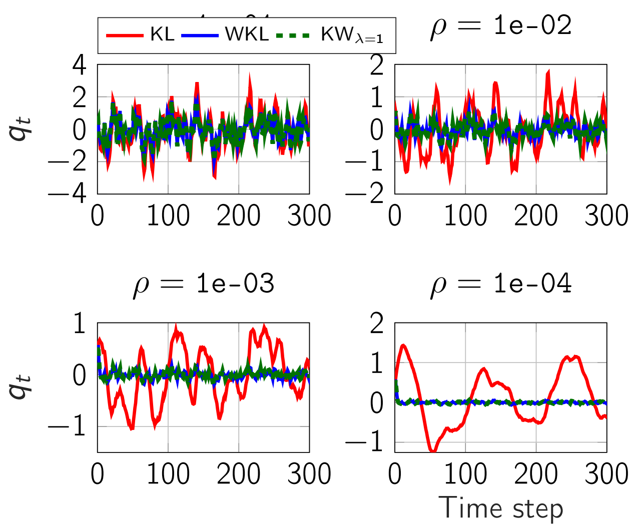
$\rightsquigarrow$ **Idea: Replace Fisher-Rao geometry by (Kalman-)Wasserstein geometry in ($\star$).**

Improving KL-regularized control

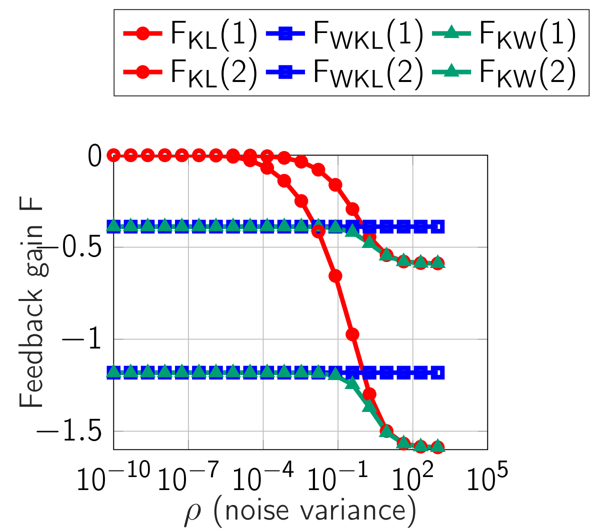
Consider the dynamics

$$x_{t+1} = Ax_t + Bu_t + w_t, \qquad u_t = Fx_t, \qquad w_t \sim \mathcal{N}(0, \rho), \qquad t \in \mathbb{N}\,.$$

Find optimal F by minimizing $J(F) := \mathbb{E}_{w_t \sim \mathcal{N}(0,\Sigma_w)}\left[\sum_{t=0}^{\infty} \gamma^t \left(\frac{1}{2} x_t^\top Q x_t + D(p_t \mid p_0)\right)\right]$, where $p_t := p\,(x_{t+1} \,|\, x_t, u_t) = \mathcal{N}(Ax_t + Bu_t, \Sigma_w)$ and $p_0 := p\,(x_{t+1} \,|\, x_t, u_t = 0) = \mathcal{N}(Ax_t, \Sigma_w)$.



Closed-loop trajectories of the double integrator. WKL- and KW-regularized controllers reduce oscillations.



Optimal feedback gains $F \in \mathbb{R}^{1\times 2}$ shown component-wise for regularized LQR.

Thank you for your attention!

`https://viktorajstein.github.io`

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. This project has received funding from the European Research Council (ERC) under the European Union's Horizon Europe research and innovation programme (grant agreement No. 101198055, project acronym NEITALG).