Kernel Machines with Shape Constraints*

Pierre-Cyril Aubin-Frankowski[1], Zoltán Szabó[2]

[1]INRIA, PSL Research University, [2]Department of Statistics, LSE

Goal

- Impose shape constraints
  - for rich function classes (RKHS),
  - in a hard fashion (e.g. *all* points of an interval),
  - with modularity in the shape constraints (D below),
  - with theoretical guarantees (not discussed here).

Shape constraints

$$0 \le Df(\mathbf{x}) \quad \forall \mathbf{x}.$$

- Examples:
  - non-negativity: $0 \le f(x)$, monotonicity: $0 \le f'(x)$, convexity: $0 \le f''(x)$,
  - monotonicity w.r.t. partial orderings:

$$0 \le \partial^{\mathbf{e}_j} f(\mathbf{x}), \quad (\forall j \in [d]), \quad 0 \le \partial^{\mathbf{e}_d} f(\mathbf{x}) \le \ldots \le \partial^{\mathbf{e}_1} f(\mathbf{x}),$$

  - supermodularity: $0 \le \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}$ $(\forall i \neq j \in [d])$.
- Applications: economics (utility function, demand function, production function), statistics (quantile function, pdf), finance (option pricing), RL (value function), supply chain models and game theory.

Rich function class: kernel, RKHS

$k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is a kernel if

$$k(x, y) = \langle \varphi(x), \varphi(y) \rangle_{\mathcal{F}}.$$

Examples:

$$k_p(\mathbf{x}, \mathbf{y}) = (\langle \mathbf{x}, \mathbf{y} \rangle + c)^p, \qquad k_G(\mathbf{x}, \mathbf{y}) = e^{-\gamma \|\mathbf{x}-\mathbf{y}\|_2^2},$$
$$k_L(\mathbf{x}, \mathbf{y}) = e^{-\gamma \|\mathbf{x}-\mathbf{y}\|_1}, \qquad k_e(\mathbf{x}, \mathbf{y}) = e^{\gamma \langle \mathbf{x}, \mathbf{y} \rangle}.$$

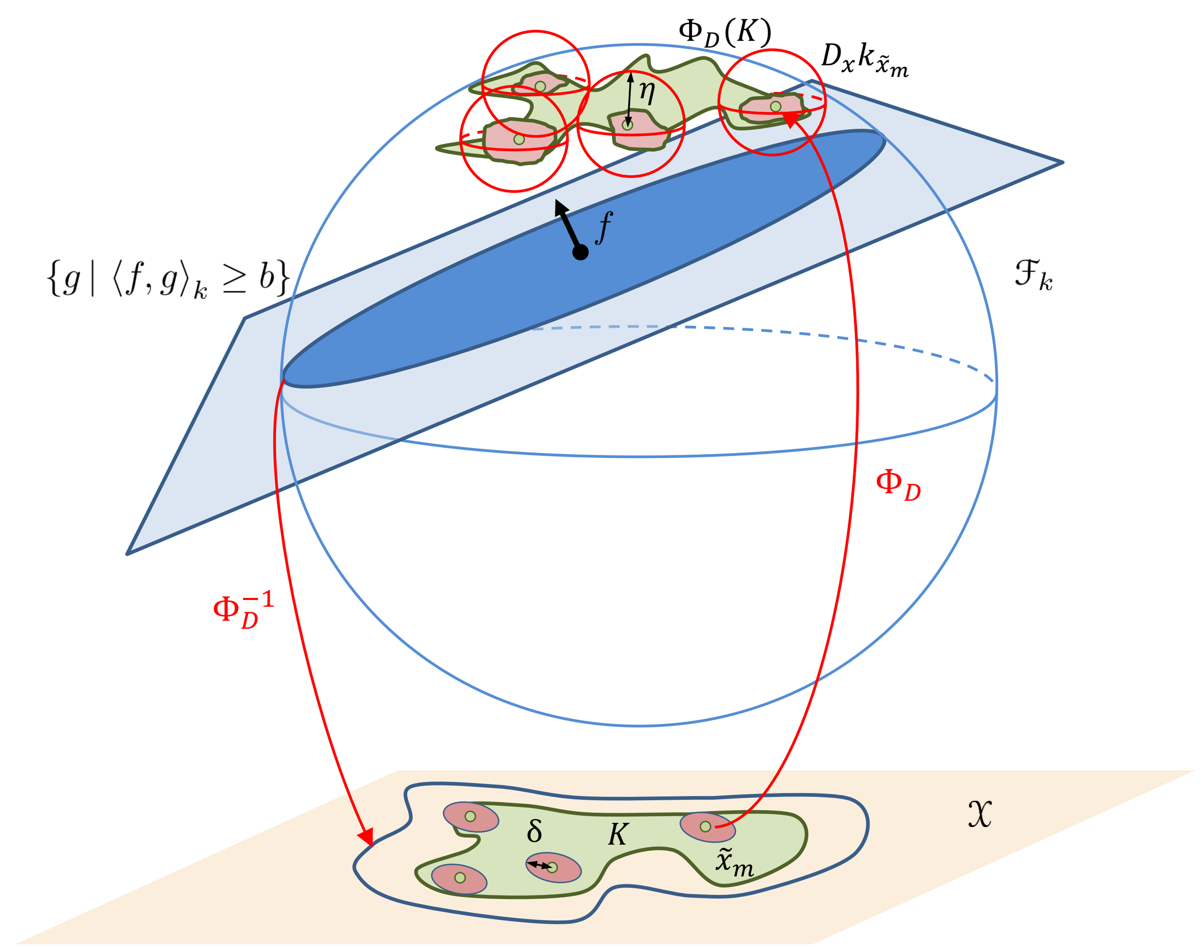
RKHS: $\mathcal{F}_k = \overline{\text{span}}\{k(\cdot, \mathbf{x}) : \mathbf{x} \in \mathcal{X}\} \overset{1:1}{\longleftrightarrow} k$.
Included ($\mathcal{F}_k$): Fourier analysis, polynomials, splines, ...

Idea

SOC-tightening [1, 2]:

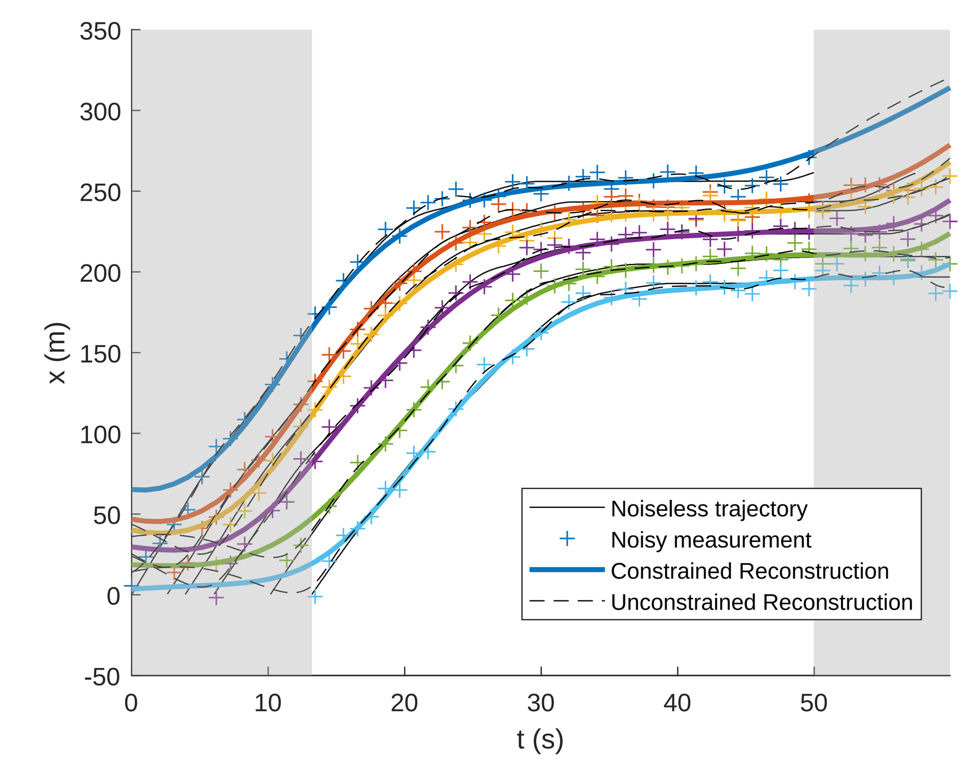
$$0 \le Df(\mathbf{x}), \quad \forall \mathbf{x} \in \mathcal{K}$$
$$\eta_m \|f\|_{\mathcal{F}_k} \le Df(\tilde{\mathbf{x}}_m), \quad \forall m \in [M]$$

with suitable $\eta_m > 0$ and $\{\tilde{\mathbf{x}}_m\}_{m \in [M]} \subset \mathcal{K}$ covering.
Intuition of η_m:



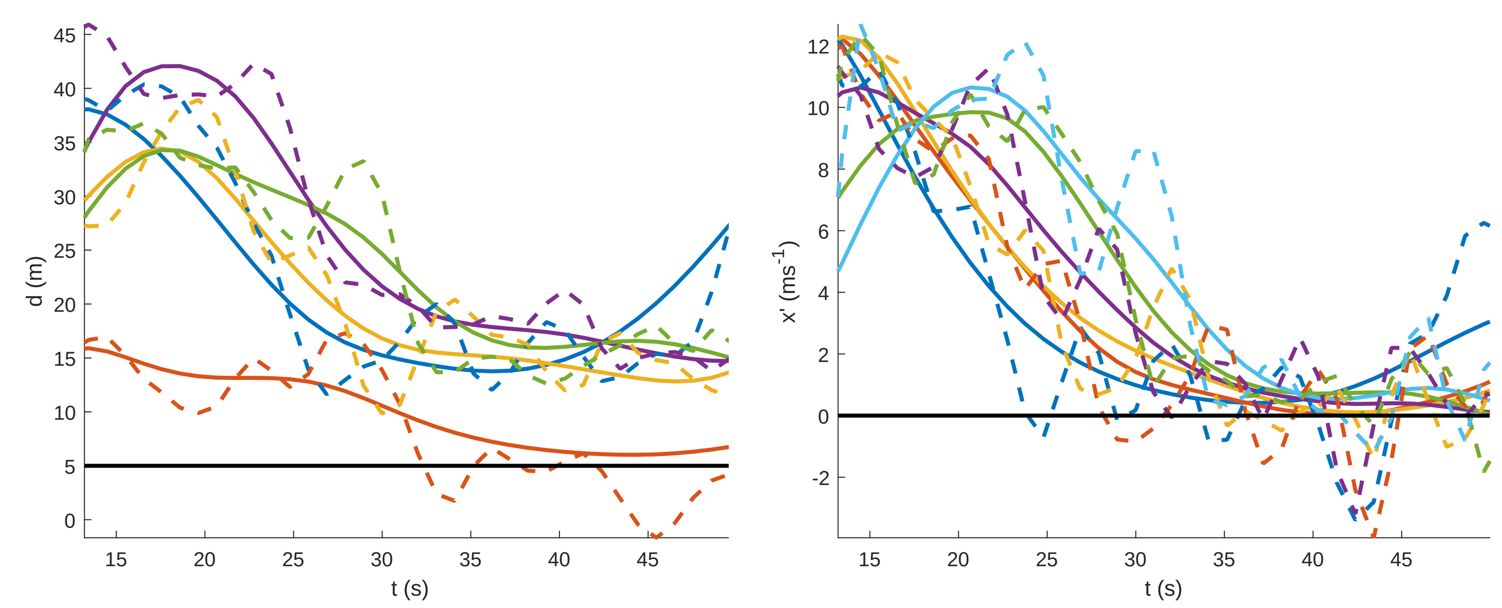
Convoy localization with traffic jam

- D: no overtaking, min speed & inter-vehicular distance.
- Setting: $Q = 6$, $d_{\min} = 5m$, $v_{\min} = 0$.



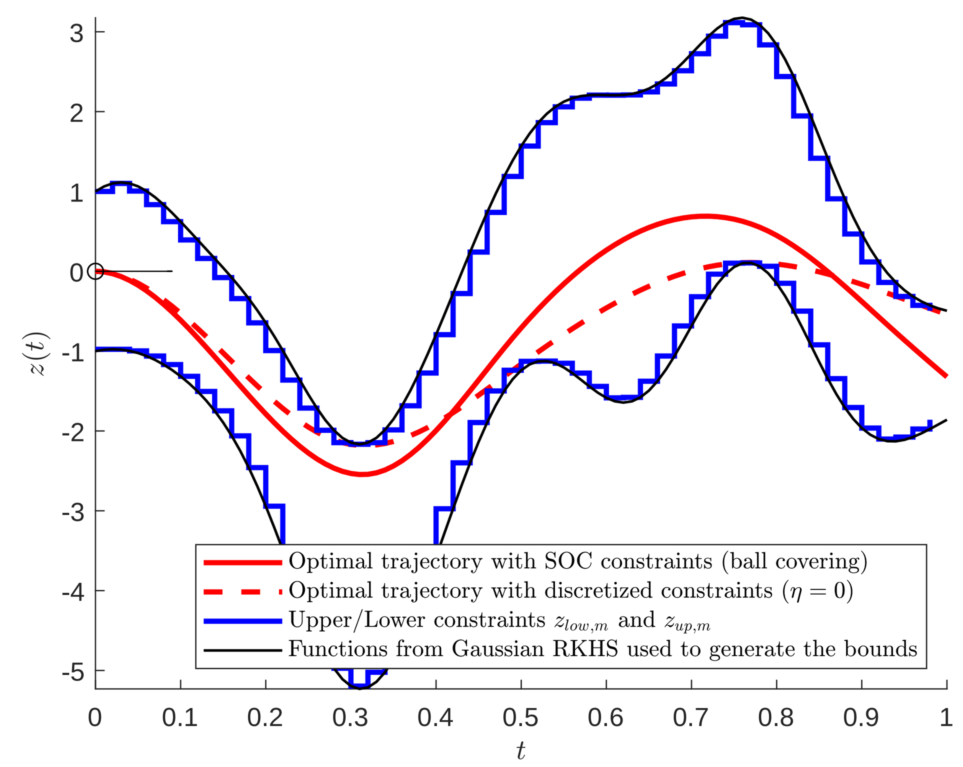
Convoy localization – continued

Pairwise distances: $t \mapsto f_q(t) - f_{q+1}(t)$ Speed: $t \mapsto f'_q(t)$



Safety-critical control

- D: avoid bumping into the ceiling & floor of the cave.



Further examples [1, 2]: quantile regression, shape optimization, econometrics, robotics.

References

[1] Pierre-Cyril Aubin-Frankowski and Zoltán Szabó. Hard shape-constrained kernel machines. In *NeurIPS*, pages 384–395, 2020.

[2] Pierre-Cyril Aubin-Frankowski and Zoltán Szabó. Handling hard affine SDP shape constraints in RKHSs. *JMLR (accepted)*, 2022. (http://arxiv.org/abs/2101.01519).

*Acknowledgements. ZSz was supported by the Europlace Institute of Finance and the Chair Stress Test RISK Management and Financial Steering, led by the French École Polytechnique and its Foundation and sponsored by BNP Paribas.