Convex and Nonconvex Sublinear Regression with Application to Data-driven Learning of Reach Sets

Shadi Haddad and Abhishek Halder

Department of Applied Mathematics

University of California, Santa Cruz

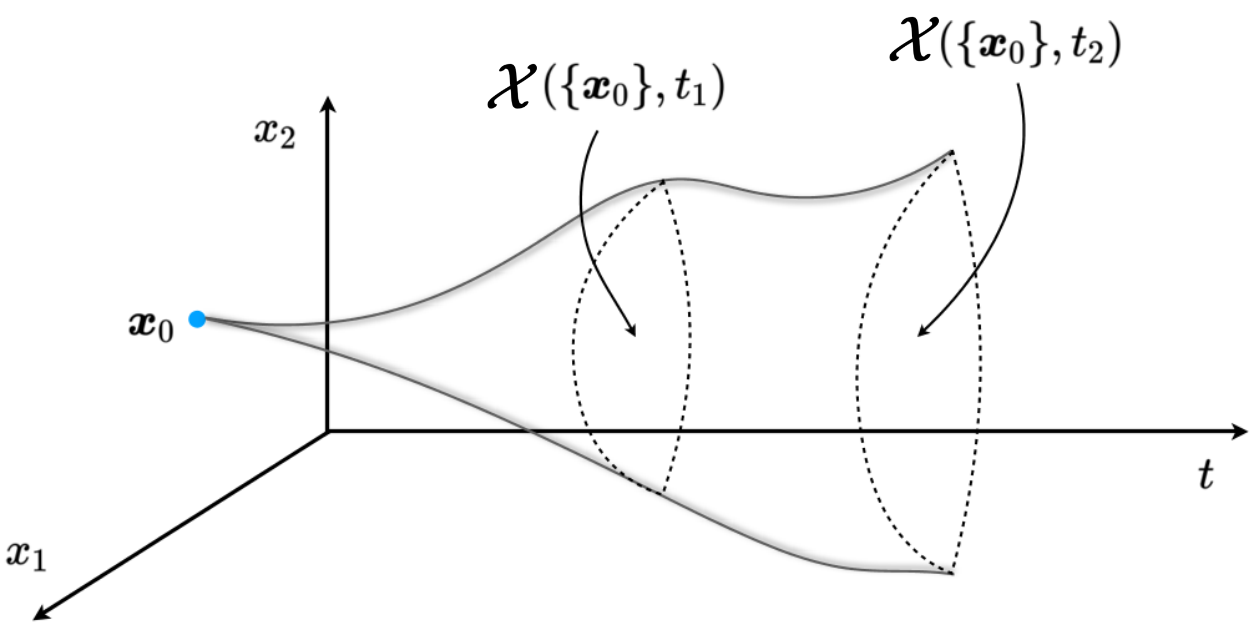
2023 American Control Conference, San Diego, June 2, 2023

Reach set

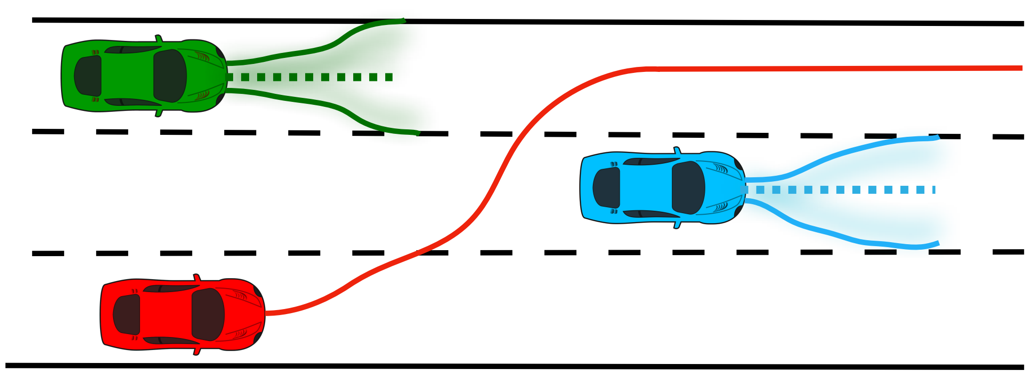
$$\mathcal{X}_t := \{\boldsymbol{x}(t) \in \mathbb{R}^d \mid \dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})\boldsymbol{u},$$

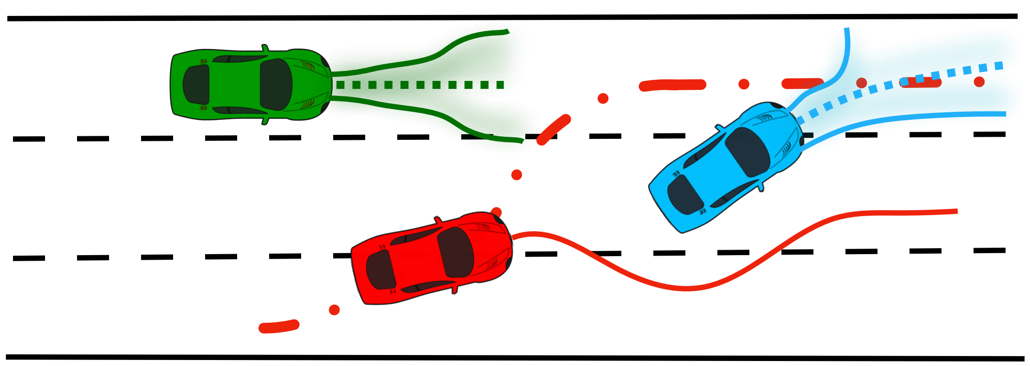
$$\boldsymbol{x}(t=0) = \boldsymbol{x}_0, \boldsymbol{u}(s) \in \mathcal{U}(s)$$

$$\text{for all } 0 \leq s \leq t\}$$



Safety critical applications such as motion planning & collision warning systems

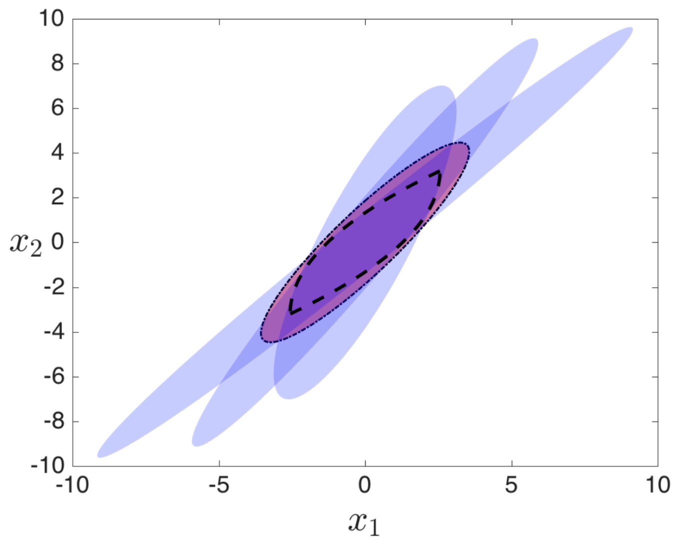




Existing algorithms for reach set computation

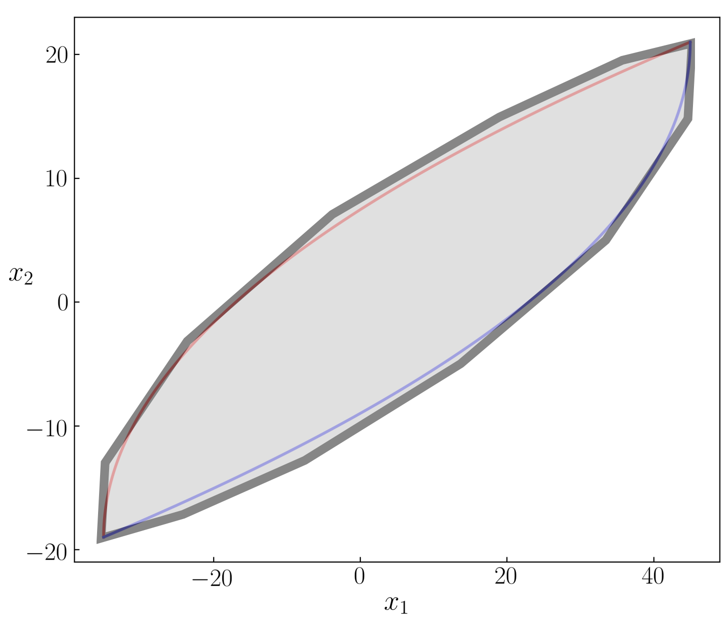
Parametric

Ellipsoidal over-approximation



Elipsoidal toolbox
[Kurzhansky et al., 2006]

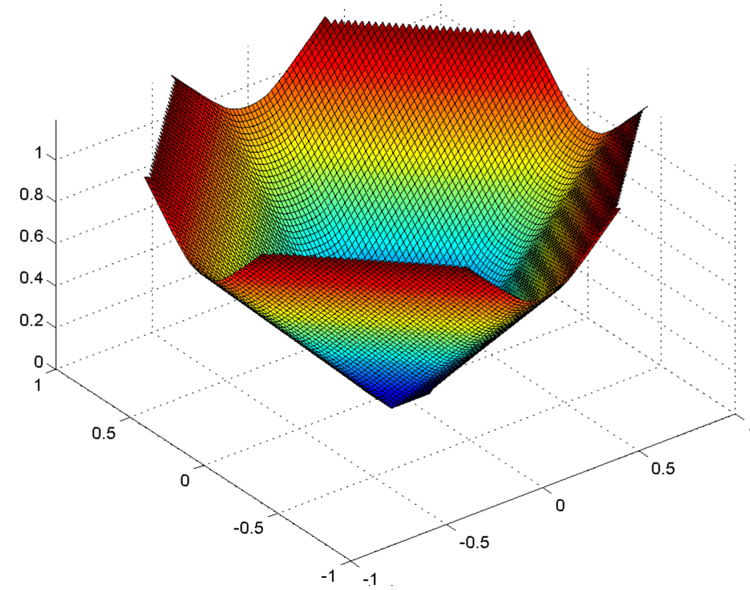
Zonotopic over-approximation

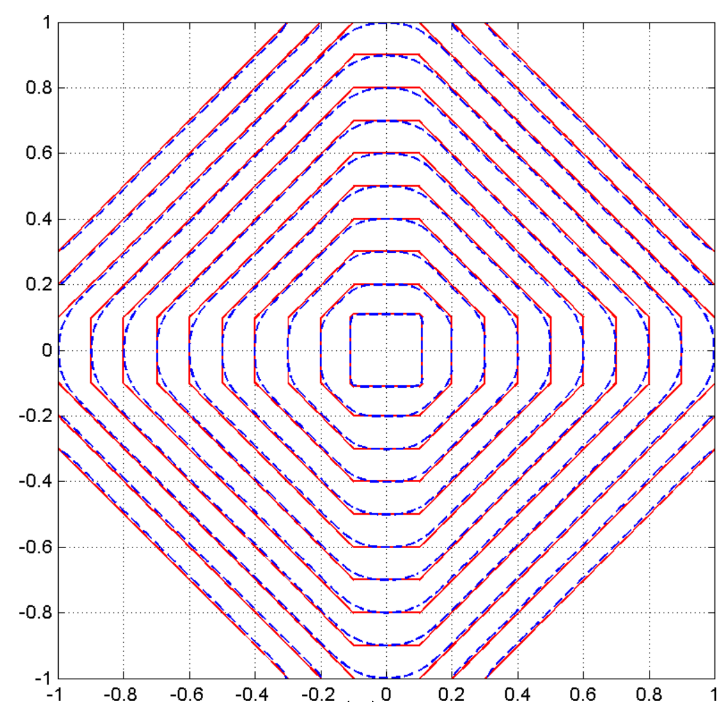


CORA toolbox
[Althoff et al., 2015]

Nonparametric

Zero sub-level set of the viscosity solution of HJB PDE

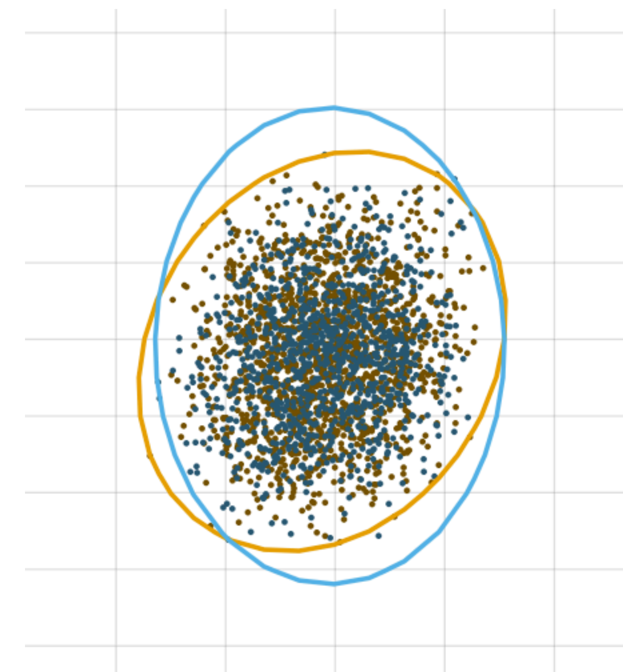


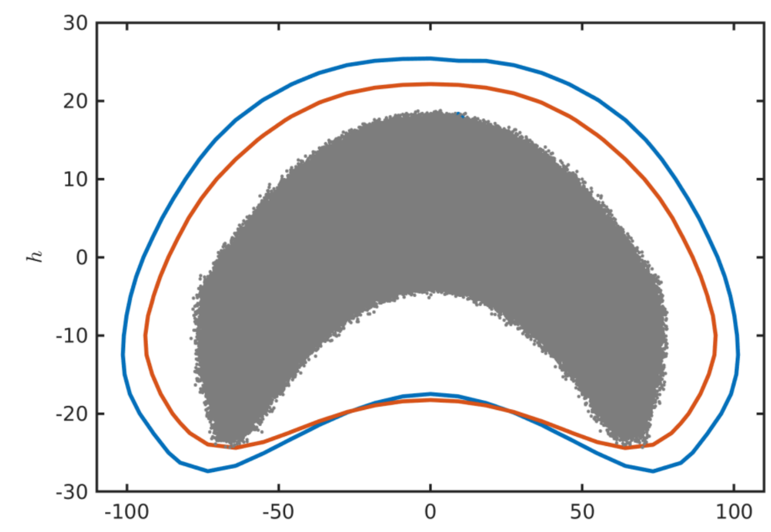


Level set toolbox
[Mitchell et al., 2008]

Semiparametric

Sample-based statistical learning



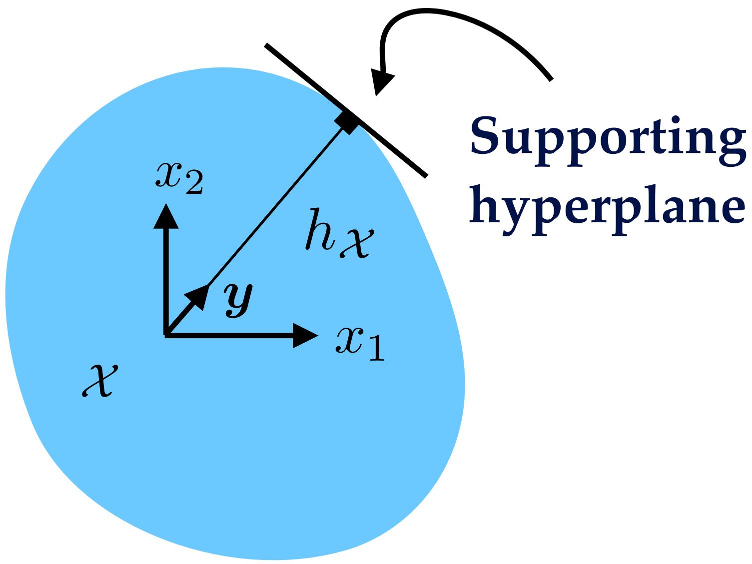


[Devonport and Arcak, 2020]

Our approach: support function representation learning

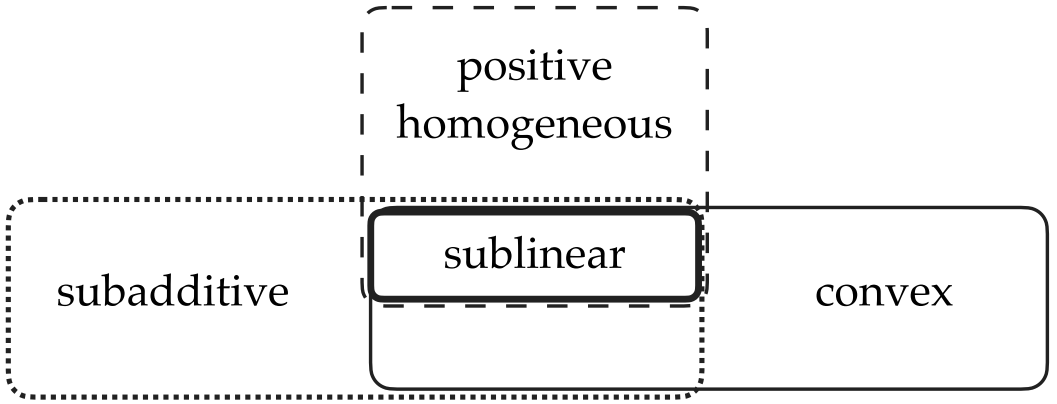
Support function:

$$h_{\mathcal{X}(\mathcal{X}_0,t)}(\boldsymbol{y}) := \sup_{\boldsymbol{x}\in\mathcal{X}} \left\{ \langle \boldsymbol{x}, \boldsymbol{y} \rangle \mid \boldsymbol{y} \in \mathbb{S}^{d-1} \right\}$$



- **Support function is positive homogeneous of degree 1:**

$$h_{\mathcal{X}}(a\boldsymbol{y}) = a h_{\mathcal{X}}(\boldsymbol{y}), \quad a \in \mathbb{R}_{>0}, \quad \forall \boldsymbol{y} \in \mathbb{S}^{d-1}$$

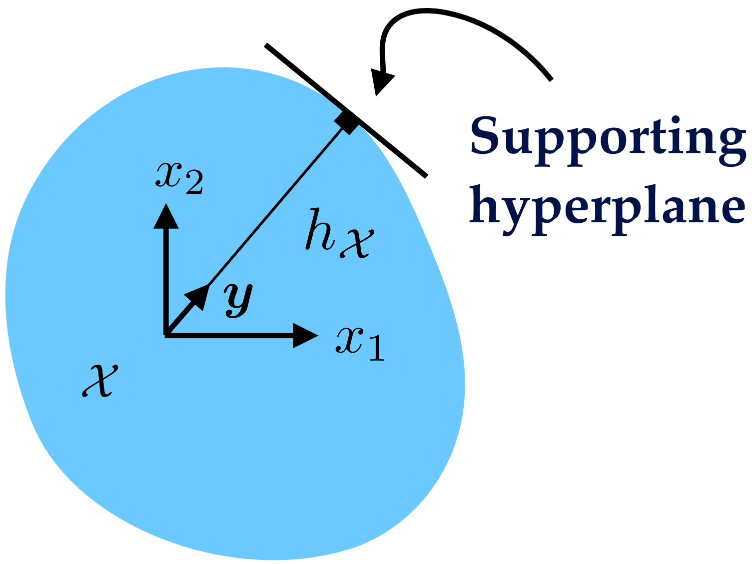


- **Convex and positive homogeneous $\Longleftrightarrow$ Sublinear**

Our approach: support function representation learning

Support function:

$$h_{\mathcal{X}(\mathcal{X}_0,t)}(\boldsymbol{y}) := \sup_{\boldsymbol{x}\in\mathcal{X}} \left\{ \langle \boldsymbol{x}, \boldsymbol{y} \rangle \;\middle|\; \boldsymbol{y} \in \mathbb{S}^{d-1} \right\}$$



- **Sublinear functions** $\overset{\text{Isomorphism}}{\Longleftrightarrow}$ **Finite dimensional compact convex sets**

Set operations $\Longleftrightarrow$ Support function operations

- **The Legendre-Fenchel conjugate of indicator function**

Uniquely determines a compact set, up to closure of convexification

$$h_{\mathcal{X}}(\boldsymbol{y}) = h_{\overline{\text{conv}}(\mathcal{X})}(\boldsymbol{y}) \quad \forall \boldsymbol{y} \in \mathbb{S}^{d-1}$$

Set operations $\Longleftrightarrow$ Support function operations

Set operands	Support function operands
• Membership	Inequality
• Intersection	Infimal convolution
• Affine transformation	Composition
• Convergence in Hausdorff topology	Pointwise convergence
• p-sum	p-norm
• Minkowski sum	Sum
• Union	Pointwise maximum

Set operations $\Longleftrightarrow$ Support function operations

Geometric functionals

- **Hausdorff distance metric** $$\delta_{\mathrm{H}}(\mathcal{P},\mathcal{Q}) = \sup_{\boldsymbol{y}\in\mathbb{S}^{d-1}} |h_{\mathcal{P}}(\boldsymbol{y}) - h_{\mathcal{Q}}(\boldsymbol{y})|$$

- **Width of $\mathcal{X}\subset\mathbb{R}^d$ in the direction of $\boldsymbol{y}\in\mathbb{S}^{d-1}$** $$\mathrm{width}_{\mathcal{X}}(\boldsymbol{y}) := h_{\mathcal{X}}(\boldsymbol{y}) + h_{\mathcal{X}}(-\boldsymbol{y})$$

- **Diameter of $\mathcal{X}\subset\mathbb{R}^d$** $$\mathrm{diam}_{\mathcal{X}} := \max_{\boldsymbol{y}\in\mathbb{S}^{d-1}} \left(h_{\mathcal{X}}(\boldsymbol{y}) + h_{\mathcal{X}}(-\boldsymbol{y})\right)$$

- **The Polar dual $\mathcal{X}^{\circ}$ of $\mathcal{X}\subset\mathbb{R}^d$** $$\mathcal{X}^{\circ} = \left\{\boldsymbol{y}\in\mathbb{R}^d \mid h_{\mathcal{X}}(\boldsymbol{y}) \leq 1\right\}$$

- **Intersection oracle** [Haddad, Halder 2022] $$\min_{\boldsymbol{y}\in\mathbb{S}^{n-1}} \{h_{\mathcal{P}}(\boldsymbol{y}) + h_{\mathcal{Q}}(-\boldsymbol{y})\} \geq (<)0$$ $$\Longleftrightarrow \mathcal{P}\cap\mathcal{Q} \neq (=)\varnothing$$

Regression Algorithm

Deterministic

Vector of i.i.d samples with zero mean

Available data

$$\{\widehat{\boldsymbol{x}}_j\}_{j=1}^{n_x} = \{\boldsymbol{x}_j + \boldsymbol{\nu}_j\}_{j=1}^{n_x}, \quad \boldsymbol{x}_j \in \mathcal{X}, \quad \boldsymbol{\nu} \in \mathbb{R}^d$$

Training data

Support function data

$$\widehat{h}_{\mathcal{X}}(\boldsymbol{y}_i) = \sup_{\widehat{\boldsymbol{x}} \in \{\widehat{\boldsymbol{x}}_j\}_{j=1}^{n_x}} \langle \boldsymbol{y}_i, \widehat{\boldsymbol{x}} \rangle, \quad \boldsymbol{y}_i \in \mathbb{S}^{d-1}, \quad \forall i \in [\![n_y]\!] \longrightarrow \{(\boldsymbol{y}_i, \widehat{h}_{\mathcal{X}}(\boldsymbol{y}_i))\}_{i=1}^{n_y}$$

n_x **: given number samples from** $\mathcal{X}$

n_y **: arbitrary number of samples from unit sphere**

We seek a sublinear function that "well fits" the training data

Quadratic Programming (QP)

Infinite dimensional least square

Minimizer

$$\underset{\{h_{n_x}:\mathbb{R}^d\to\mathbb{R} \mid h_{n_x}(\cdot)\text{ is sublinear }\}}{\arg\inf} \sum_{i=1}^{n_y} \left(\widehat{h}_{\mathcal{X}}(\boldsymbol{y}_i) - h_{n_x}(\boldsymbol{y}_i)\right)^2$$

Finite dimensional convex QP

Subgradients

$$\underset{\boldsymbol{g}_1,\cdots,\boldsymbol{g}_{n_y}\in\mathbb{R}^d,\ \boldsymbol{h}\in\mathbb{R}^{n_y}}{\arg\min} \sum_{i=1}^{n_y} \left(\widehat{h}_{\mathcal{X}}(\boldsymbol{y}_i) - h_i\right)^2$$

$$\text{subject to } h_j \geq h_i + \langle \boldsymbol{g}_i, \boldsymbol{y}_j - \boldsymbol{y}_i\rangle, \quad \forall (i,j) \in [\![n_y]\!] \times [\![n_y]\!]$$

Piecewise linear (PWL) estimate

$$h^{\text{PWL}}(\cdot) = \max_{i=1,\cdots,n_y} \left\{ \widehat{h}_{\mathcal{X}}(\boldsymbol{y}_i) + \left\langle \boldsymbol{g}_i^{\text{opt}}, \cdot - \boldsymbol{y}_i \right\rangle \right\}$$

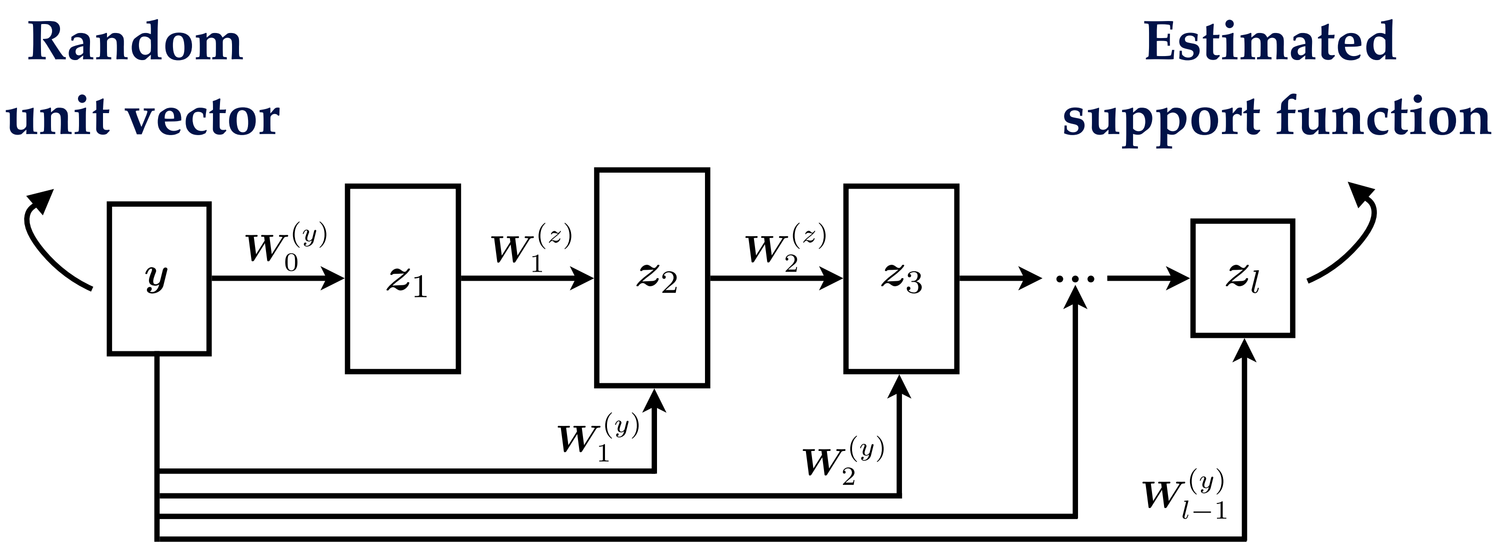
Theorem . The minimizer of (1) converge to the true support function as $n_y, n_x \to \infty$.

$$\boxed{h_{n_x}(\cdot) \longrightarrow h_{\mathcal{X}}(\cdot), \ \text{ as } \ n_y, n_x \to \infty}$$

Input Sublinear Neural Network (ISNN)

Input convex neural network (ICNN) [Amos, 2017]

The Output z_l is convex if $\boldsymbol{W}^{(y)}_{1:\ell} \geq 0$, and $\boldsymbol{\sigma}(.)$ is convex and non-decreasing.



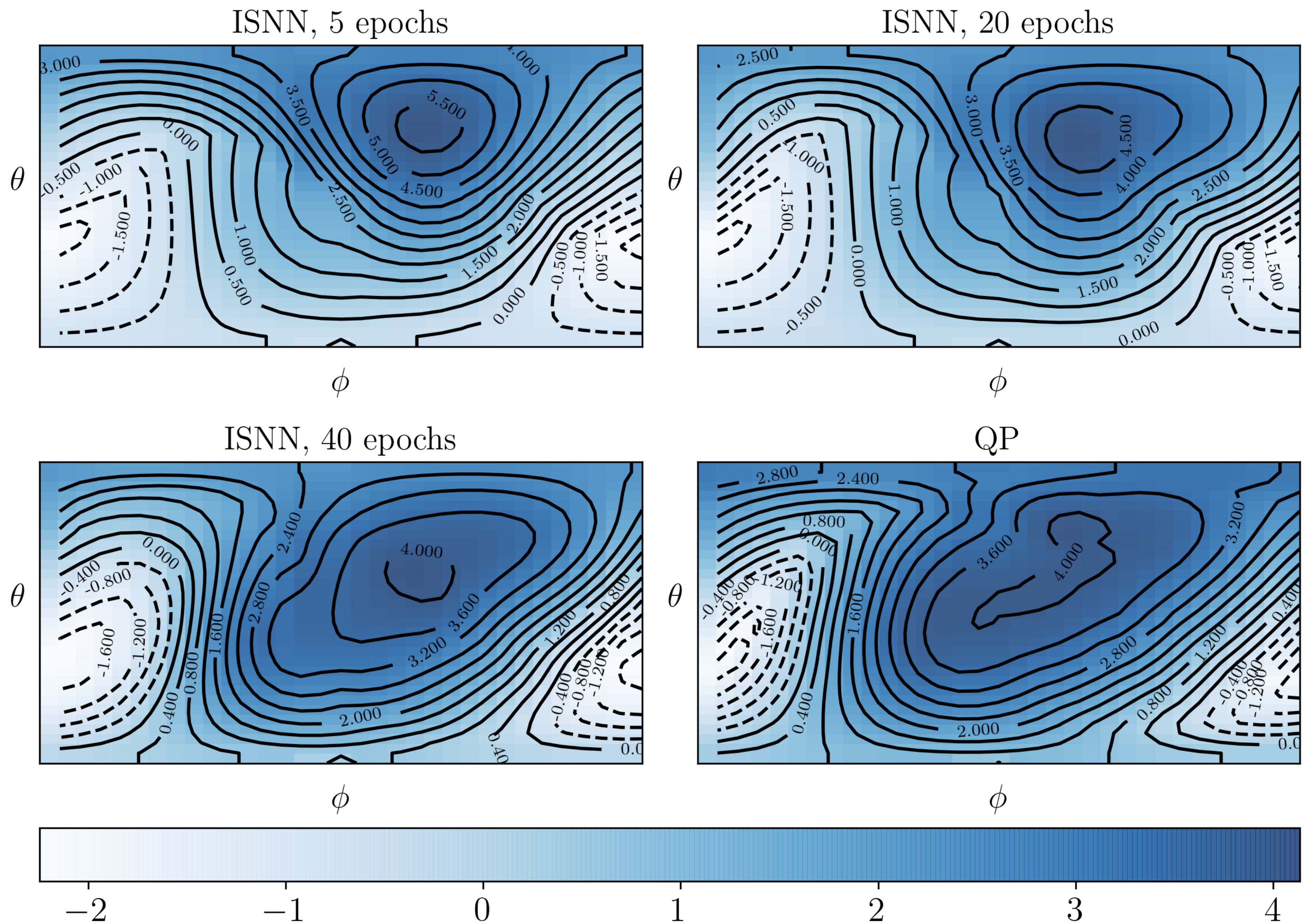
$$
\begin{aligned}
\boldsymbol{z}_1 &= \boldsymbol{\sigma}\Big(\boldsymbol{W}_1^{(y)} \boldsymbol{y} + \boldsymbol{b}_1\Big), \\
\boldsymbol{z}_{k+1} &= \boldsymbol{\sigma}\Big(\boldsymbol{W}_{k+1}^{(z)} \boldsymbol{z}_k + \boldsymbol{W}_{k+1}^{(y)} \boldsymbol{y} + \boldsymbol{b}_{k+1}\Big), \\
\boldsymbol{z}_\ell &= \boldsymbol{W}_\ell^{(z)} \boldsymbol{z}_{\ell-1} + \boldsymbol{W}_\ell^{(y)} \boldsymbol{y} + \boldsymbol{b}_\ell
\end{aligned}
$$

Theorem. The output of the network is sublinear with respect to the input vector if

$\boldsymbol{W}^{(y)}_{1:\ell} \geq 0, \quad \boldsymbol{b}^{(y)}_{1:\ell} = \boldsymbol{0}, \quad \boldsymbol{\sigma}(.)$ **sublinear and non-decreasing**

Example: Dubin's car

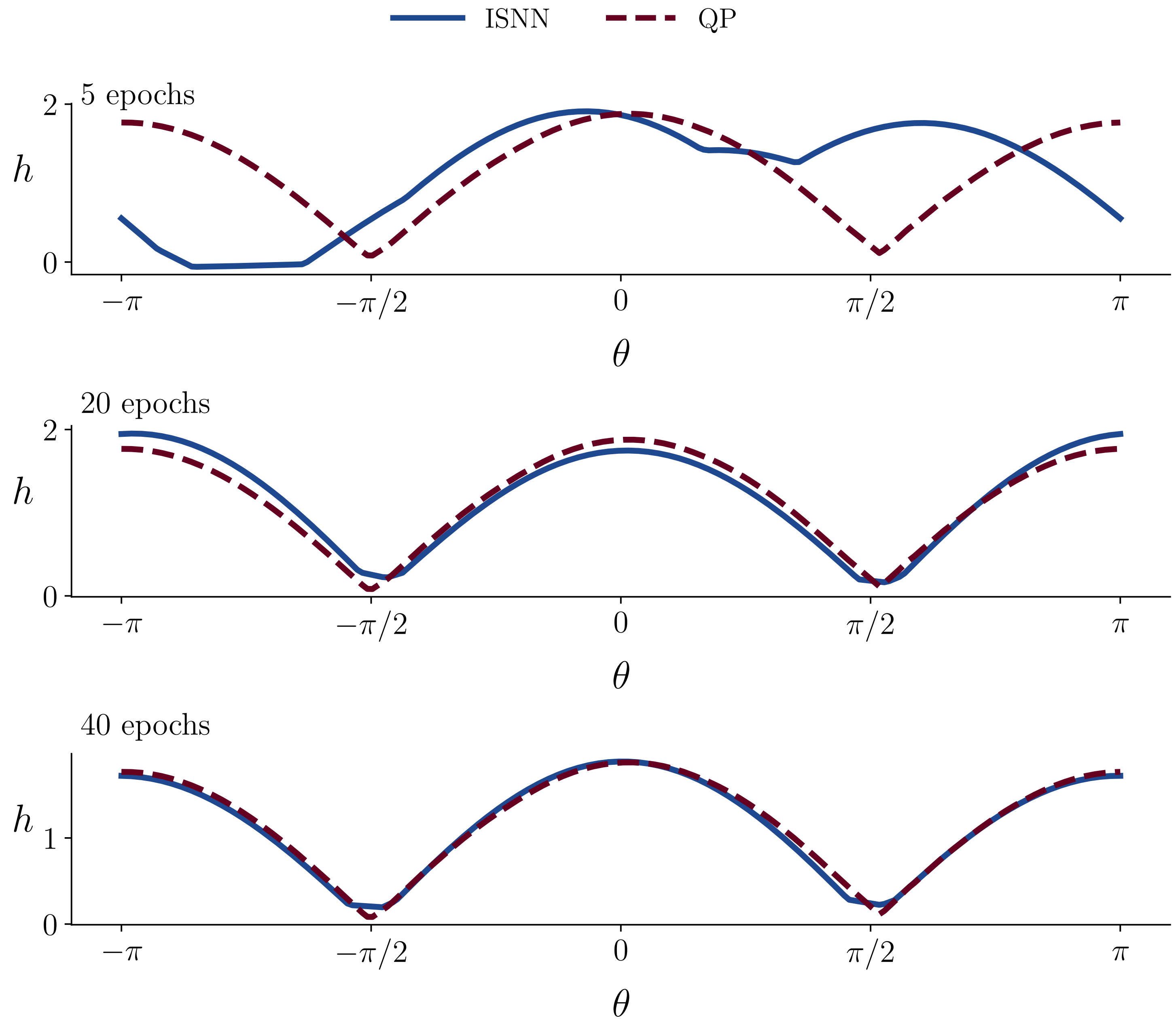
ISNN → QP as # epochs increases



Contour plots of the estimated support function at t = 2 s

$$\dot{x}_1 = v\cos x_3, \quad \dot{x}_2 = v\sin x_3, \quad \dot{x}_3 = u, \quad u(t) \in \mathcal{U} := [-30^\circ, 90^\circ]\mathrm{deg/s}^2 \qquad \mathcal{X}_0 = \{\mathbf{0}\}$$

Example: kinematic bicycle model

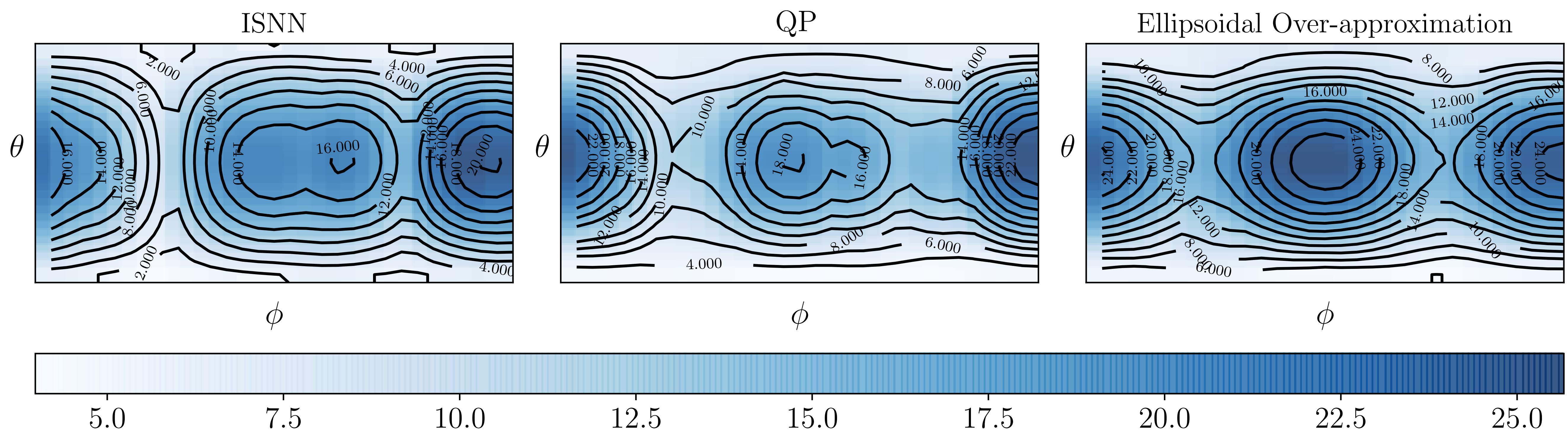


The estimated support function of the projection of the reach set onto the position coordinates.

$$\dot{x}_1 = x_3 \cos(x_4 + \beta), \quad \dot{x}_2 = x_3 \sin(x_4 + \beta),$$
$$\dot{x}_3 = u_1, \quad \dot{x}_4 = x_3 \sin(\beta)/1.5$$

$$(u_1, u_2) \in \mathcal{U} := [-1, 1]\mathrm{m/s}^2 \times [-10^\circ, 10^\circ]$$

Example: quadrotor



Contour plots of the estimated support function at t = 2 s.

$$\dot{\boldsymbol{x}} = \boldsymbol{A}_{\mathrm{cl}}(t)\boldsymbol{x} + \boldsymbol{B}_{\mathrm{cl}}\boldsymbol{\eta} + \boldsymbol{G}\boldsymbol{w}$$

$$\boldsymbol{A}_{\mathrm{cl}} := \boldsymbol{A} + \boldsymbol{B}\boldsymbol{K}(t)$$

$$\boldsymbol{B}_{\mathrm{cl}} := \boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^{\top}$$

$$\boldsymbol{\eta}(t) \in \mathcal{E}(\boldsymbol{v}(t), \boldsymbol{V}(t))$$

From Riccati ODE

$$\boldsymbol{x}_0 \sim \mathcal{N}(\mathbf{0}_{12\times 1}, 2\boldsymbol{I}_{12})$$

$$\boldsymbol{w}(t) \sim \mathcal{N}\left((\cos t, \sin t, \cos t)^{\top}, 0.01\boldsymbol{I}_3\right)$$

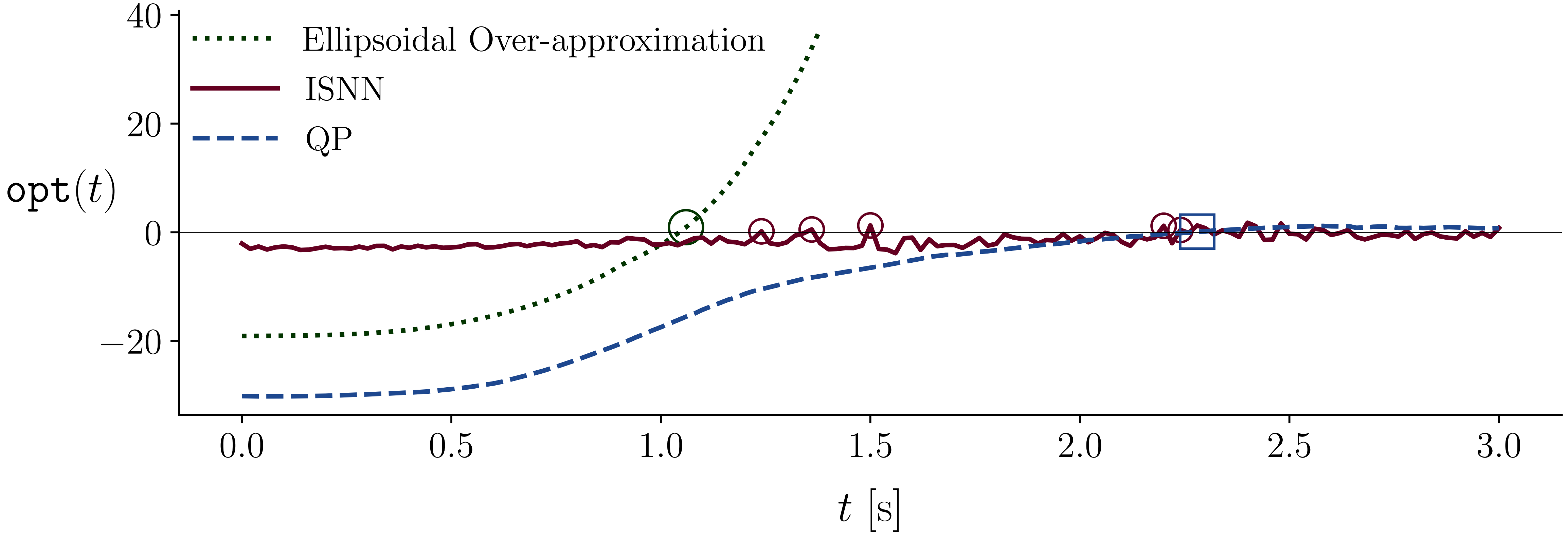
Ellipsoidal over-approximation [Kurzhanskiy, 2006]

$$\mathcal{X}_{\mathrm{LTV}}(\boldsymbol{x}_0, t) \subseteq \widehat{\mathcal{X}}^{(N)}_{\mathrm{LTV}}(\boldsymbol{X}_0, t) := \bigcap_{i=1}^{N} \mathcal{E}(\boldsymbol{x}_c(t), \boldsymbol{X}_i(t))$$

Example: quadrotor

Collision detection between the quadrotors A and B

$$\mathrm{opt}(t) = \min_{\boldsymbol{y}\in\mathbb{S}^{n-1}} \left\{ h_{\mathrm{Proj}\left(\mathcal{X}_t^A\right)}(\boldsymbol{y}) + h_{\mathrm{Proj}(\mathcal{X}_t^{\mathrm{B}})}(-\boldsymbol{y}) \right\}$$



Comparison between QP and ISNN

- **QP offers a more robust result**

  ISNN → QP, as the num of epochs increases

- **ISNN computational time is considerably lower**

  Approximately 10 times lower, for the same cardinality

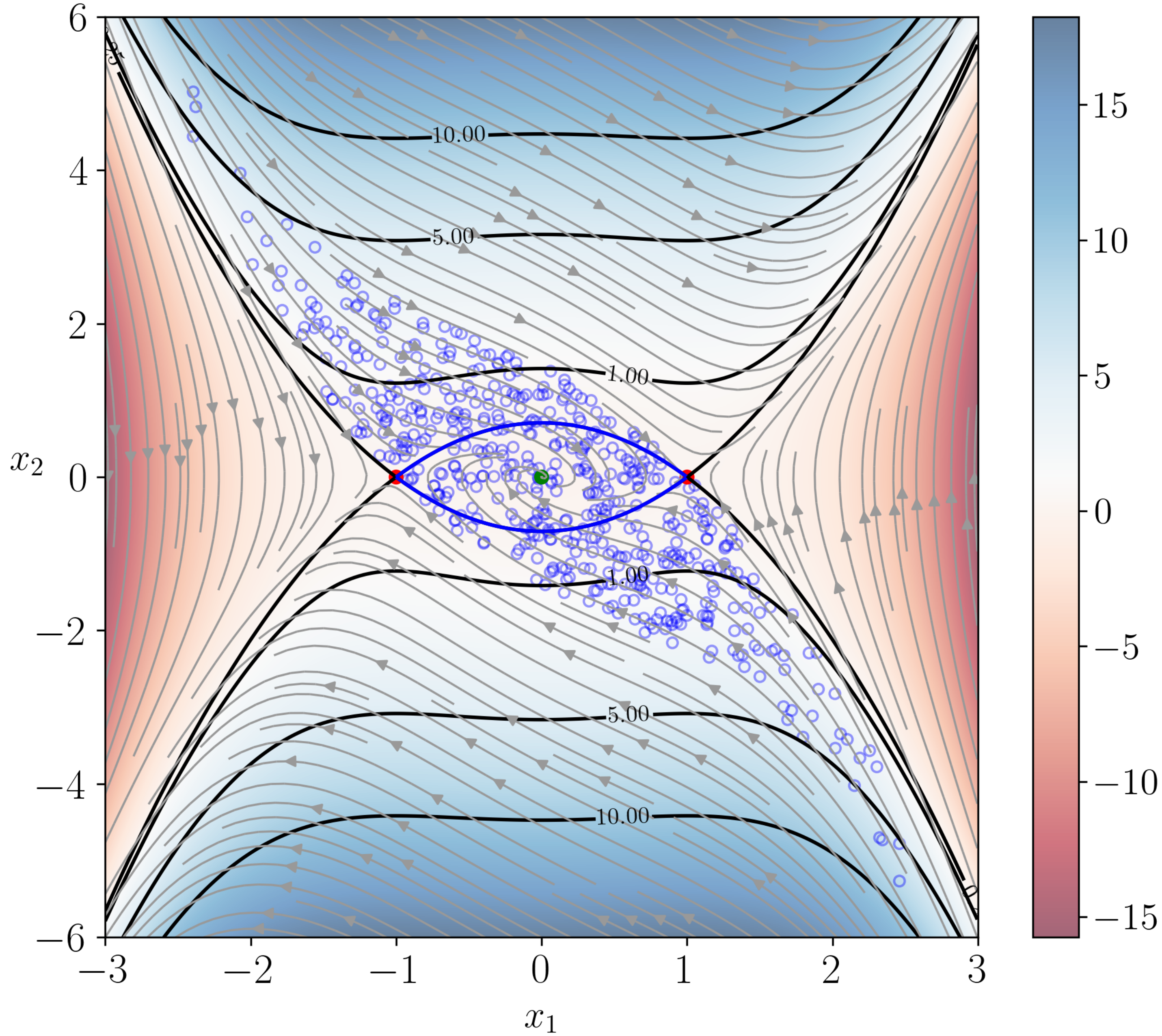
  The number of constraint of QP is quadratic in n_y

Instance	ISNN, 30 epochs	QP
1	6.76	60.78
2	6.66	60.52
3	6.88	63.54
4	6.68	67.02
5	6.63	66.55
6	7.09	66.92
7	6.63	73.82
8	6.66	74.56
9	6.66	71.91
10	6.98	69.60

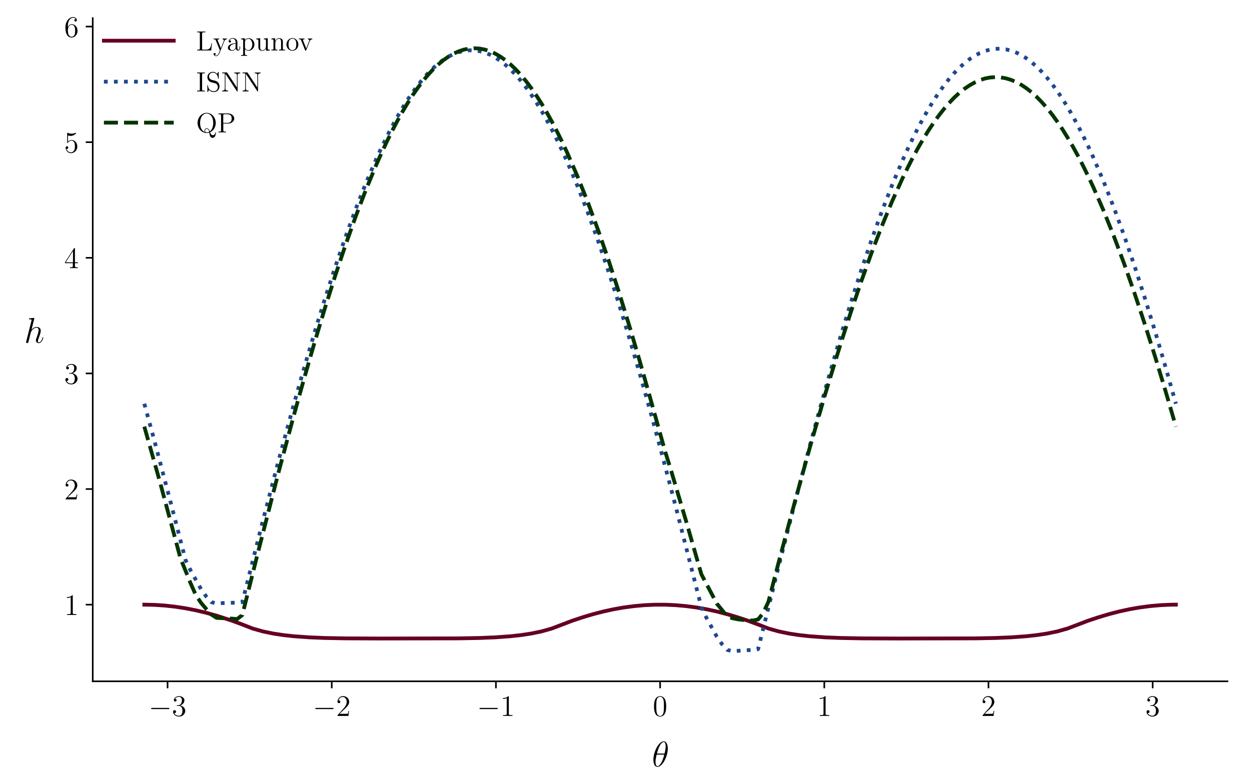
Computational times [s] for Dubin's car example

Future Work

- **Customized algorithm for solving the presented QP**
- **Learning the support function of arbitrary compact sets**



The Region of Attraction (ROA); Lyapunov vs sublinear regression



The estimated support function of ROA; Lyapunov, ISNN, QP

Thank You

shhaddad@ucsc.edu

ahaldder@ucsc.edu