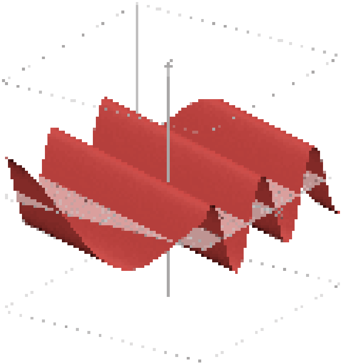
Dimension reduction using Poincaré inequalities

Alexandre PASCO

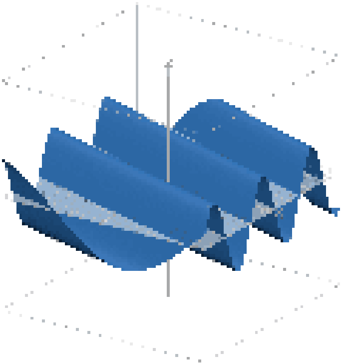
Centrale Nantes, Nantes Université, France
In collaboration with Anthony NOUY.

Introduction

$$u(x) = e^{x_1+3x_2}$$

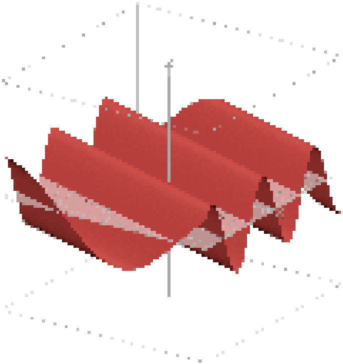


$$v(x) = e^{x_1+3x_2+0.05x_1x_2}$$

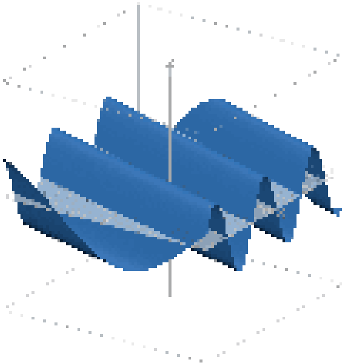


Introduction

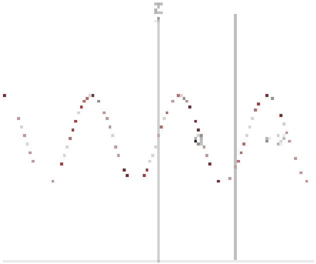
$u(x) = e^{x_1+3x_2}$



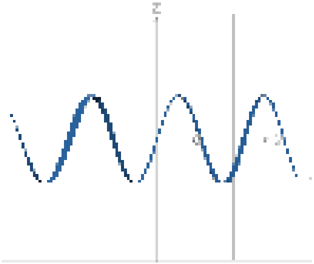
$v(x) = e^{x_1+3x_2+0.05x_1x_2}$



$$u(x) = e^{x_1+3x_2}$$



$$v(x) = e^{x_1+3x_2+0.05x_1x_2}$$



Introduction

- Goal: Approximate $u : \mathbb{R}^d \to \mathbb{R} \in \mathcal{C}^1$ with $d \gg 1$, i.e. minimize

$$\mathcal{E}(\tilde{u}) := \mathbb{E}\left[(u(\mathbf{X}) - \tilde{u}(\mathbf{X}))^2\right]$$

  where $\mathbf{X}$ has probability density $\mu_{\mathbf{X}}$.

- Model class: ridge functions

$$\tilde{u}(x) = f(G^T x)$$

  with $G \in \mathbb{R}^{d\times m}$ and $f : \mathbb{R}^m \to \mathbb{R}$

- Minimizing $\mathcal{E}(\tilde{u})$ means solving

$$\inf_{G \in \mathbb{R}^{d\times m}} \inf_{f:\mathbb{R}^m \to \mathbb{R}} \mathbb{E}\left[(u(\mathbf{X}) - f(G^T\mathbf{X}))^2\right]$$

- For $G \in \mathbb{R}^{d\times m}$ of rank k, we can decompose $G = QR$ with $Q \in \mathbb{R}^{d\times k}$ such that $Q^TQ = I_k$ and $R \in \mathbb{R}^{k\times m}$.

- For some $f : \mathbb{R}^m \to \mathbb{R}$ we can write

$$f(G^Tx) = \tilde{f}(Q^Tx) \quad \text{where} \quad \tilde{f} = f \circ R^T.$$

$\to$ Only range(G) matters ! Now assume $m < d$ and $G^TG = I_m$.

On the function f

On the function f

- For a given $G \in \mathbb{R}^{d\times m}$ let $\mathbf{Z} = G^T\mathbf{X}$, then the solution of

$$\inf_{f:\mathbb{R}^m\to\mathbb{R}} \mathbb{E}\left[(u(\mathbf{X}) - f(\mathbf{Z}))^2\right]$$

is given by $f_G : z \mapsto \mathbb{E}\left[u(\mathbf{X})|\mathbf{Z} = z\right]$

- We have an explicit expression for f_G, for any $z \in \mathbb{R}^m$,

$$f_G(z) = \mathbb{E}\left[u(Gz + (I_d - GG^T)\mathbf{X})\right],$$

where $I_d - GG^T$ is the orthogonal projector onto range$(G)^\perp$.

On the function f

- For a given $G \in \mathbb{R}^{d\times m}$ we can re-write the initial problem as

$$\inf_{\tilde{u}=f\circ G^T} \mathcal{E}(\tilde{u}) = \inf_{\substack{G\in\mathbb{R}^{d\times m}\\ G^TG=I_m}} \mathbb{E}\left[(u(\mathbf{X}) - \mathbb{E}\left[u(\mathbf{X})|G^T\mathbf{X}\right])^2\right]$$

- In practice we compute some $\tilde{f}_G \neq f_G$ by regression.

Poincaré Inequality

Poincaré Inequality: the classic

- Let $U \subset \mathbb{R}^n$ open and bounded with $\mathcal{C}^1$ boundary.
- There exists $C(U) < \infty$ such that

$$C(U) := \left(\inf_{h \in H_0^1(U)} \frac{\int_U \|\nabla h\|^2}{\int_U h^2} \right)^{-1}$$

- For any $h \in H_0^1(U)$, the following Poincaré Inequality holds

$$\int_U h^2 \leq C(U) \int_U \|\nabla h\|^2.$$

Poincaré Inequality: the probabilistic

- Let $\mathbf{Y} \in \mathbb{R}^n$ with probability density γ.

- Let $C(\gamma) \in \mathbb{R} \cup \{+\infty\}$ the smallest constant such that

$$\mathbb{E}\left[h(\mathbf{Y})^2\right] \leq C(\gamma)\mathbb{E}\left[\|\nabla h(\mathbf{Y})\|^2\right]$$

  for any $h : \mathbb{R}^d \to \mathbb{R}$ such that $h \in \mathcal{C}^1(\mathbb{R}^d)$ and $\mathbb{E}\left[h(\mathbf{Y})\right] = 0$,

- We can bound $C(\gamma)$ for some classical γ.

Poincaré Inequality: bounds for the constant

- If $\gamma = \mathcal{N}(0,1)$ then $C(\gamma) = 1$.
- If $\gamma = \mathcal{N}(m,\Sigma)$ then $C(\gamma) = \lambda_{\min}(\Sigma)^{-1}$.
- If γ is α-uniformly log-concave then $C(\gamma) = \alpha^{-1}$, where

$$\gamma(dy) = \exp(-V(y))dx \quad \text{with} \quad \alpha = \inf_{y\in\mathbb{R}^n} \lambda_{\min}(V''(y)).$$

- If $\gamma \propto \mathcal{U}(\mathcal{Y})$ with $\mathcal{Y}$ bounded convex then $C(\gamma) \leq \frac{\mathrm{diam}(\mathcal{Y})^2}{\pi^2}\frac{\sup\gamma}{\inf\gamma}$.

$\rightarrow$ For those examples $C(\gamma)$ is independent on the dimension n. In general this is not the case !

Poincaré Inequality and G

Poincaré Inequality and G

- Recall that $\mathbf{Z} = G^T\mathbf{X}$ and $f_G(\mathbf{Z}) = \mathbb{E}\left[u(\mathbf{X})|\mathbf{Z}\right]$

$$\mathbb{E}\left[(u(\mathbf{X}) - f_G(\mathbf{Z}))^2\right] = \mathbb{E}\left[\mathbb{E}\left[(u(\mathbf{X}) - f_G(\mathbf{Z}))^2|\mathbf{Z}\right]\right]$$

- Decompose $\mathbf{X} = G\mathbf{Z} + (I_d - GG^T)\mathbf{X}$, thus we can write

$$\mathbb{E}\left[(u(\mathbf{X}) - f_G(\mathbf{Z}))^2|\mathbf{Z}\right] = \mathbb{E}\left[h_{\mathbf{Z}}(\mathbf{X})^2|\mathbf{Z}\right]$$

where $h_{\mathbf{Z}}(\mathbf{X}) := u(G\mathbf{Z} + (I_d - GG^T)\mathbf{X}) - f_G(\mathbf{Z})$.

- We have $\mathbb{E}\left[h_{\mathbf{Z}}(\mathbf{X})|\mathbf{Z}\right] = 0$ and

$$\nabla h_{\mathbf{Z}}(\mathbf{X}) = (I_d - GG^T)\nabla u(\mathbf{X}).$$

Poincaré Inequality and G

- Applying the Poincaré Inequality gives

$$\mathbb{E}\left[h_{\mathbf{Z}}(\mathbf{X})^2|\mathbf{Z}\right] \leq C(\mu_{\mathbf{X}|\mathbf{Z}})\mathbb{E}\left[\|(I_d - GG^T)\nabla u(\mathbf{X})\|^2|\mathbf{Z}\right]$$

- Assuming that $C(G,\mu_{\mathbf{X}}) := \sup_{\mathbf{Z}} C(\mu_{\mathbf{X}|\mathbf{Z}}) < +\infty$ we obtain

$$\mathbb{E}\left[(u(\mathbf{X}) - f_G(\mathbf{Z}))^2\right] \leq C(G,\mu_{\mathbf{X}})\mathbb{E}\left[\|(I_d - GG^T)\nabla u(\mathbf{X})\|^2\right]$$

Poincaré Inequality and G

- Let $G_\perp \in \mathbb{R}^{d\times(d-m)}$ s.t. $G_\perp^T G_\perp = I_{d-m}$ and $G_\perp^T G = 0$ then

$$\begin{aligned}
\|(I_d - GG^T)\nabla u(\mathbf{X})\|^2 &= \|G_\perp^T \nabla u(\mathbf{X})\|^2 \\
\|(I_d - GG^T)\nabla u(\mathbf{X})\|^2 &= \mathsf{Trace}(G_\perp^T \nabla u(\mathbf{X}) \nabla u(\mathbf{X})^T G_\perp) \\
\mathbb{E}\left[\|(I_d - GG^T)\nabla u(\mathbf{X})\|^2\right] &= \mathsf{Trace}(G_\perp^T H G_\perp)
\end{aligned}$$

  where

$$H := \mathbb{E}\left[\nabla u(\mathbf{X}) \nabla u(\mathbf{X})\right] \in \mathbb{R}^{d\times d}$$

- We are now with

$$\mathbb{E}\left[(u(\mathbf{X}) - f_G(\mathbf{Z}))^2\right] \leq C(G, \mu_{\mathbf{X}})\mathsf{Trace}(G_\perp^T H G_\perp)$$

Poincaré Inequality and G

- We are now with

$$\mathbb{E}\left[(u(\mathbf{X}) - f_G(\mathbf{Z}))^2\right] \leq C(G, \mu_{\mathbf{X}})\mathsf{Trace}(G_\perp^T H G_\perp)$$

- Trace term minimized by taking $G_\perp$ as the $d-m$ last eigen vectors of H, i.e. taking G as the first m ones, which gives

$$\mathsf{Trace}(G_\perp^T H G_\perp) = \lambda_{m+1} + \cdots + \lambda_d$$

- Ok, but what about $C(G, \mu_{\mathbf{X}})$? Just take the previous examples and everything goes well ?

Poincaré Inequality and G: bounding conditional constant

- Recall that $C(G, \mu_{\mathbf{X}}) := \sup_{\mathbf{Z}} C(\mu_{\mathbf{X}|\mathbf{Z}})$.
- What property do we preserve when taking a "slice" of $\mu_{\mathbf{X}}$?
- If $\mu_{\mathbf{X}}$ is gaussian, α-uniformly log-concave, or proportional to $\mathcal{U}(\mathcal{X})$ with $\mathcal{X}$ bounded convex, then so is $\mu_{\mathbf{X}|G^T\mathbf{X}}$ and

$$C(\mu_{\mathbf{X}|G^T\mathbf{X}}) \leq C(\mu_{\mathbf{X}})$$

  for any orthonormal G.

Poincaré Inequality and G: final bound

- Assuming that $\sup_G \sup_{\mathbf{X}} C(\mu_{\mathbf{X}|G^T\mathbf{X}}) = C < +\infty$, then

$$\mathbb{E}\left[(u(\mathbf{X}) - f_G(\mathbf{Z}))^2\right] \leq C\,\mathsf{Trace}(G_\perp^T H G_\perp)$$

- Choosing the m first eigen vectors for G leads to

$$\min_{f:\mathbb{R}^d\to\mathbb{R}} \mathbb{E}\left[(u(\mathbf{X}) - f(G^T\mathbf{X}))^2\right] \leq C \sum_{k=m+1}^{d} \lambda_k.$$

with $\lambda_1 \geq \cdots \geq \lambda_d \geq 0$ the eigen values of $H = \mathbb{E}\left[\nabla u \nabla u^T\right]$.

- Remark: this bound is in general not sharp.

Towards nonlinear dimension reduction

Towards nonlinear dimension reduction

- Approximate $u(\mathbf{X})$ by $f(\mathbf{Z})$ with $\mathbf{Z} = g(\mathbf{X})$ where $g \in \mathcal{G}_m \subset \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^m)$ may be nonlinear.

- If $J_g(\mathbf{X})$ has full rank then $\mu_{\mathbf{X}|\mathbf{Z}}$ where $\mathbf{Z} = g(\mathbf{X})$ is supported on the m-dimensional submanifold $g^{-1}(\mathbf{Z}) \subset \mathbb{R}^d$.

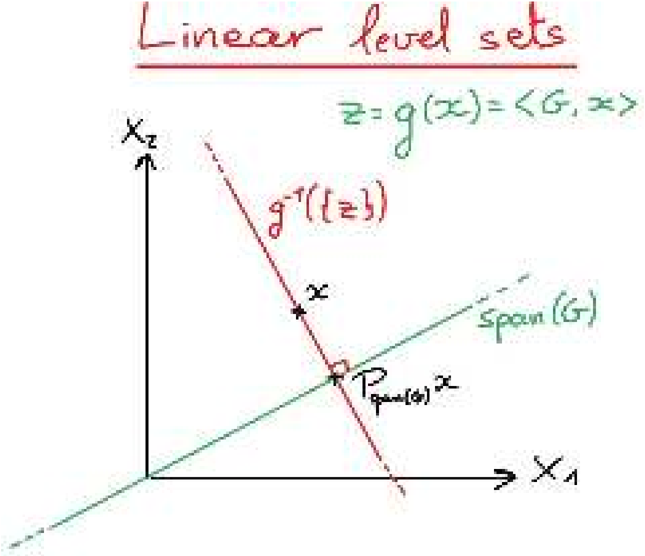
$+$ Allows for lower m, i.e. better dimension reduction.

\- The support of $\mu_{\mathbf{X}|\mathbf{Z}}$ not affine anymore, but a Riemannian submanifold of $\mathbb{R}^d$.

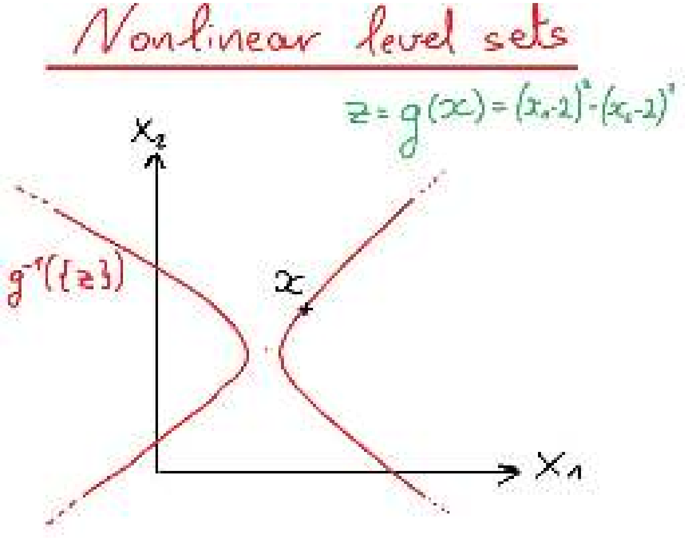
\- Cannot say much on the Poincaré Constant $C(\mu_{\mathbf{X}|\mathbf{Z}})$.

\- The Poincaré Inequality involves $\mathbb{E}\left[\|\nabla u_{|g^{-1}(\mathbf{Z})}(\mathbf{X})\|^2\right]$, which is not minimized by an eigen-value problem anymore...

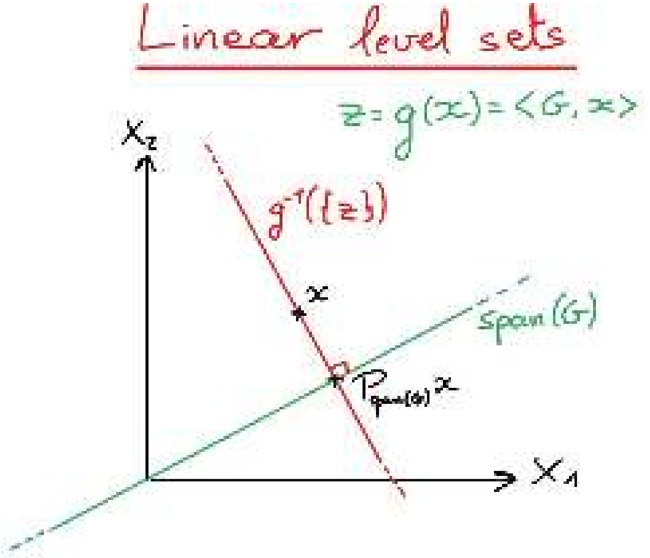
Towards nonlinear dimension reduction



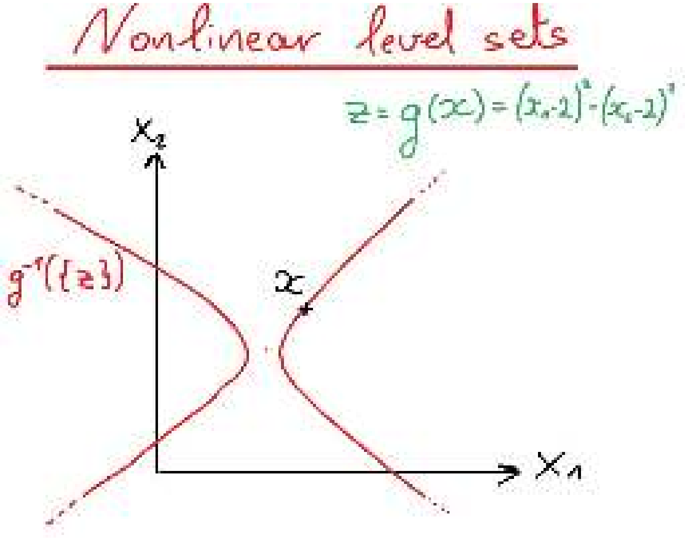
Towards nonlinear dimension reduction



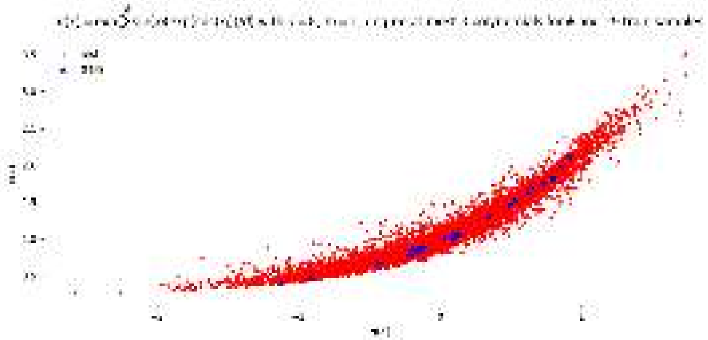
Towards nonlinear dimension reduction: level sets



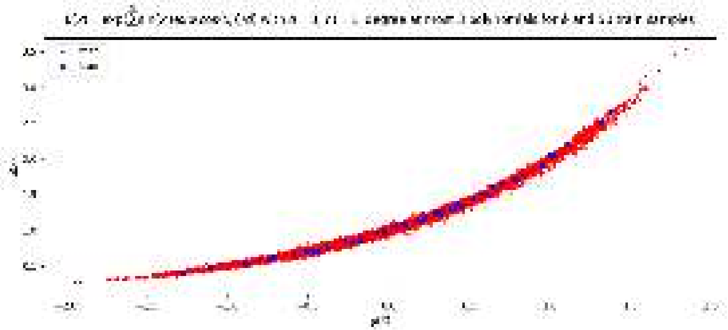
Towards nonlinear dimension reduction: level sets



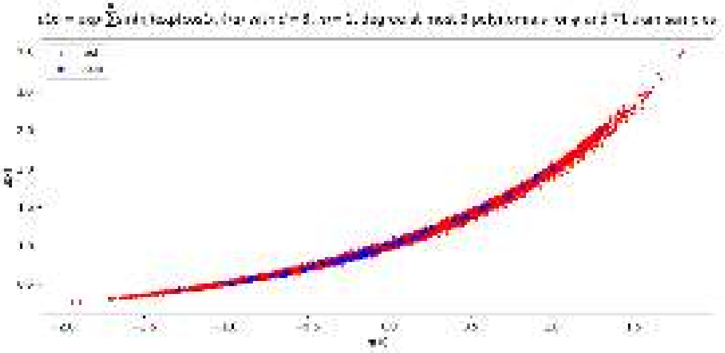
Towards nonlinear dimension reduction: numerics



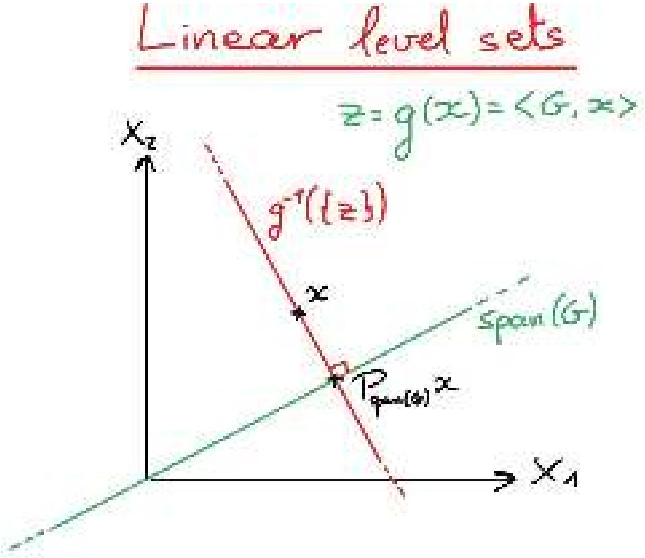
Towards nonlinear dimension reduction: numerics



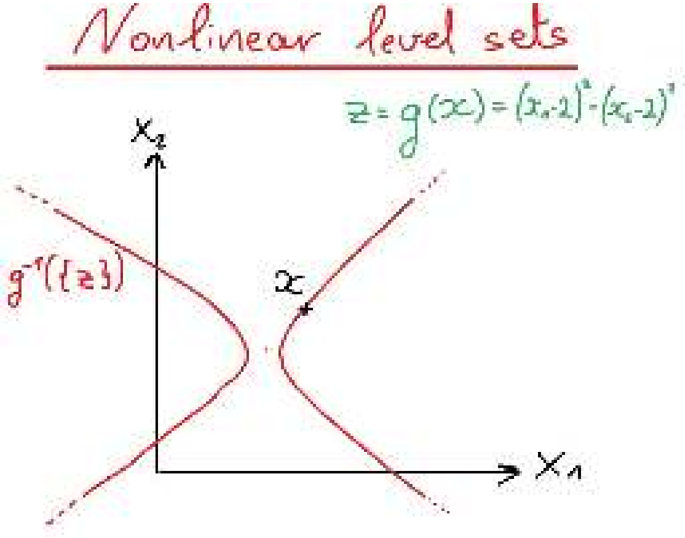
Towards nonlinear dimension reduction: numerics



Towards nonlinear dimension reduction



Towards nonlinear dimension reduction



Thank you for your attention !

References

References: Classical PI

- General PI: `https://stevensoojin.kim/blog/poincare-inequalities/`
- Bounds for the Poincaré constant: [Teixeira Parente et al., 2020]
- PI for dimension reduction:
  [Constantine et al., 2014, Zahm et al., 2020]
- PI for nonlinear dimension reduction:
  [Bigoni et al., 2022, Verdière et al., 2023]

Thank you !

References I

Bigoni, D., Marzouk, Y., Prieur, C., and Zahm, O. (2022).
Nonlinear dimension reduction for surrogate modeling using gradient information.
Information and Inference: A Journal of the IMA, 11(4):1597–1639.

Constantine, P. G., Dow, E., and Wang, Q. (2014).
Active Subspace Methods in Theory and Practice: Applications to Kriging Surfaces.
SIAM J. Sci. Comput., 36(4):A1500–A1524.

Teixeira Parente, M., Wallin, J., and Wohlmuth, B. (2020).
Generalized bounds for active subspaces.
Electron. J. Statist., 14(1).

References II

Verdière, R., Prieur, C., and Zahm, O. (2023).
Diffeomorphism-based feature learning using Poincaré inequalities on augmented input space.
Preprint.

Zahm, O., Constantine, P. G., Prieur, C., and Marzouk, Y. M. (2020).
Gradient-Based Dimension Reduction of Multivariate Vector-Valued Functions.
SIAM J. Sci. Comput., 42(1):A534–A558.