Jackpot: Approximating Uncertainty Domains with Adversarial Manifolds

Nathanaël Munier, Emmanuel Soubies, Pierre Weiss

Jackpot: Approximating Uncertainty Domains with Adversarial Manifolds

Nathanaël Munier[1] Emmanuel Soubies[2] Pierre Weiss[2]

October 18, 2024

Contents

Abstract

Given a forward mapping $\Phi : \mathbb{R}^N \to \mathbb{R}^M$, the region $\{x \in \mathbb{R}^N, \|\Phi(x) - y\|_2 \leq \varepsilon\}$, where $y \in \mathbb{R}^M$ is a given vector and $\varepsilon \geq 0$ is a perturbation amplitude, represents the set of all possible inputs x that could have produced the measurement y within an acceptable error margin. This set reflects the inherent uncertainty or indeterminacy in recovering the true input x solely from the noisy observation y, which is a key challenge in inverse problems. In this work, we develop a numerical algorithm called Jackpot (Jacobian Kernel Projection Optimization) which approximates this set with a low-dimensional *adversarial manifold*. The proposed algorithm leverages automatic differentation, allowing it to handle complex, high dimensional mappings such as those found when dealing with dynamical systems or neural networks. We demonstrate the effectiveness of our algorithm on various challenging large-scale, non-linear problems including parameter identification in dynamical systems and blind image deblurring. The algorithm is integrated within the Python package deepinv.

1 Introduction

Let $\Phi : \mathbb{R}^N \to \mathbb{R}^M$ denote an arbitrary C^1 mapping and $x^\star \in \mathbb{R}^N$ denote a point. Our main objective in this paper is to design a numerical algorithm to describe the set $\mathcal{S}^\varepsilon$ defined by

$$\begin{aligned} \mathcal{S}^\varepsilon &\stackrel{\text{def}}{=} \left\{x \in \mathbb{R}^N, \|\Phi(x) - \Phi(x^\star)\|_2 \leq \varepsilon\right\} \\ &= \Phi^{-1}\left(\mathcal{B}\left(\Phi(x^\star), \varepsilon\right)\right), \end{aligned} \tag{1}$$

for some $\varepsilon \geq 0$, where $\|\cdot\|_2$ denotes the standard ℓ^2-norm and $\mathcal{B}(y, r)$ denotes an ℓ^2-ball centered at y of radius r. As will be explained later, many important practical problems can be framed in this manner.

If the mapping Φ is linear, then $\mathcal{S}^\varepsilon$ is an ellipsoid that can be characterized using linear algebra tools. In the general nonlinear case, the set $\mathcal{S}^\varepsilon$ is the sublevel set of a near arbitrary function. Indeed, an arbitrary closed set $\mathcal{E}$ can be written as the 0-level set of the function $\Phi(x) = \text{dist}(x, \mathcal{E})^2$. Given the level of generality, it is unreasonable to expect solving the problem properly for all mappings Φ. Describing $\mathcal{S}^\varepsilon$ is more challenging than finding all the global minimizers of a near arbitrary non convex function.

Therefore, we aim for a more modest goal: describe $\mathcal{S}^\varepsilon$ locally around the point $x^\star$. Additionally, we seek to approximate this set by another one of relatively low intrinsic dimensionality. This serves two purposes: i) to design tractable algorithms and ii) to provide an output that can be easily visualized. Our main objective in this paper can therefore be summarized as follows.

Main objective Given a mapping $\Phi : \mathbb{R}^N \to \mathbb{R}^M$ and a point $x^\star \in \mathbb{R}^N$, find a D-dimensional manifold $\mathcal{M}_\delta^\varepsilon$ that approximates the set

$$\mathcal{S}_\delta^\varepsilon \stackrel{\text{def}}{=} \{x \in \mathcal{B}(x^\star, \delta),\ \|\Phi(x) - \Phi(x^\star)\|_2 \leq \varepsilon\}, \tag{2}$$

where $\delta \geq 0$ bounds the diameter of the approximating manifold. The set $\mathcal{S}_\delta^\varepsilon$ can be encountered under different names, depending on the scientific field: *equivalence domain* [21], *low misfit region* [15], *solution set* [31] or *uncertainty region* [16]. Throughout this paper, we will call it *uncertainty region*.

We are particularly interested in complex models Φ (e.g. dynamical systems, neural networks), which can lead to problems involving high-dimensional (latent) spaces. Throughout the paper, we assume that the Jacobian of Φ can be accessed either through analytical derivation or by using automatic differentation algorithms.

1.1 Application examples

Many problems involve identifying parameters of physical, biomedical, or chemical systems from indirect observations. The primary applications we have in mind are related to the field of inverse problems. In this field, the set $\mathcal{S}^\varepsilon$ describes what can be identified in a system and what cannot. Let us illustrate this point with a few examples studied in the numerical experiments.

1.1.1 Identifying dynamical systems parameters

Assume that some dynamical system $u : \mathbb{R}_+ \to \mathbb{R}^{P \times D}$ of P particles in $\mathbb{R}^D$ is governed by a first or second order equation:

$$\dot{u}(t) = f(u(t), x) \qquad \text{or} \qquad \ddot{u}(t) = f(u(t), x)$$

where $u(t) = (u_1(t), \ldots, u_P(t))$ denotes the particles positions and $f : \mathbb{R}^{P \times D} \times \mathbb{R}^N \to \mathbb{R}^{P \times D}$ is a force or velocity field parameterized by the vector $x \in \mathbb{R}^N$. After measuring the particles positions at multiple time points, we would like to recover the unknown parameter x that describes the dynamical system.

An example that we will be studied later is the solar system, where $u(t)$ represents the positions of $P = N$ planets at time t, f are forces given by Newton's second law of attraction and $x \in \mathbb{R}^N$ is the mass of the different planets. We assume that a measurement system (e.g. telescope) returns estimates of the positions u at some times, leading to a measurement vector $(y_1, \ldots, y_K)$ of the form $y_k = u(t_k) + b_k$ for sampling times $t_1 < \cdots < t_K$ and unknown measurement errors b_k.

The scientific question we would like to answer is: *can we recover the mass of the planets from the partial observation of their positions?* To this end, we can set $\Phi : x \mapsto (u(t_1), \ldots, u(t_K))$ and find a first guess $x^\star$ of the mass using a least square regression:

$$x^\star = \operatorname*{argmin}_{x \in \mathbb{R}^N} \frac{1}{2} \|\Phi(x) - y\|_2^2.$$

Describing the set $\mathcal{S}_\delta^\varepsilon$ provides a much finer description of the uncertainties than the single estimate $x^\star$. As we will demonstrate, our algorithm reveals that predicting the mass of Pluto is highly challenging. This observation aligns with historical estimates: initially, Pluto's mass was thought to be twice that of Earth, but in 2015, the New Horizons probe measured it to be just 1/455th of Earth's mass.

1.1.2 Blind inverse problems

Assume that we acquire measurements y of the form

$$y = \mathcal{P}(A(\theta)(x)),$$

where $\mathcal{P} : \mathbb{R}^M \to \mathbb{R}^M$ represents a perturbation (e.g. noise, quantization) and $\theta \in \mathbb{R}^N$ is a parameter that describes the state of an acquisition system $A(\theta) : \mathbb{R}^{N'} \to \mathbb{R}^M$. This parameter could represent an unknown point spread function in optics or uncertain projection angles in tomography for instance. The goal in blind inverse problems is to recover both x and θ from the measurements y.

One approach we will explore in the numerical section involves constructing a learned reconstruction mapping $R(A(\theta), y)$, which, given a forward operator $A(\theta)$ and a measurement vector y, outputs an estimate $\hat{x}(\theta) = R(A(\theta), y)$ of x. Then, a simple approach to estimate θ is to minimize the discrepancy:

$$\operatorname*{argmin}_{\theta \in \mathbb{R}^N} \frac{1}{2} \left\| A(\theta)(\hat{x}(\theta)) - y \right\|_2^2,$$

ensuring that the pair $(A(\theta), \hat{x}(\theta))$ is coherent with the observation y.

By setting $\Phi(\theta) = A(\theta)(\hat{x}(\theta))$ and exploring $\mathcal{S}_\delta^\varepsilon$ around a local minimizer $\theta^\star$, we can identify all pairs coherent with the data. This is a way to describe some uncertainty properties of the blind inverse problem.

1.1.3 Bayesian posterior exploration

Similar to the previous example, assume that a system yields measurements y of the form

$$y = \mathcal{P}(A(x)),$$

where $A : \mathbb{R}^N \to \mathbb{R}^M$ is a known operator describing the acquisition device. Whenever A is not invertible, recovering x from y requires regularization. An elegant approach to formalize this principle uses Bayesian reasoning: we assume that x is the realization of some random vector $\mathbb{x}$ with probability distribution function $p_{\mathbb{x}}$. This makes it possible to construct a posterior distribution

$$p_{\mathbb{x}|\mathbb{y}}(x|y) = \frac{p_{\mathbb{y}|\mathbb{x}}(y|x) \cdot p_{\mathbb{x}}(x)}{p_{\mathbb{y}}(y)}$$

and Bayesian estimators such as the Maximum A Posteriori (MAP) estimate:

$$\begin{aligned}\hat{x}_{\mathrm{MAP}}(y) &\stackrel{\text{def}}{=} \operatorname*{argmax}_{x\in\mathbb{R}^N} p_{\mathbb{x}|\mathbb{y}}(x|y) \\ &= \operatorname*{argmin}_{x\in\mathbb{R}^N} -\log\left(p_{\mathbb{y}|\mathbb{x}}(y|x)\right) - \log\left(p_{\mathbb{x}}(x)\right).\end{aligned}$$

The prior $p_{\mathbb{x}}$ can be handcrafted or learned with large databases. The latter strategy oftentimes provides surprisingly good results, which are rapidly replacing older strategies [44, 43, 53]. In scientific applications, it however becomes critical to certify the results, i.e. describe what information can be safely regarded as valid and the one which is uncertain.

One approach to achieving this involves describing the high probability region

$$H^\alpha = \{x \in \mathbb{R}^N,\ \log p_{\mathbb{x}|\mathbb{y}}(x|y) \geq \alpha\} \tag{3}$$

for some probability threshold α. However, the prior $p_{\mathbb{x}}(x)$ can often only be accessed indirectly through its gradient [28, 58]. In that situation, another approach is to compute a local minimizer $x^\star$ of the posterior distribution and to explore the set

$$\mathcal{S}_\delta^\varepsilon = \left\{x \in \mathcal{B}(x^\star, \delta), \|\nabla \log p_{\mathbb{x}|\mathbb{y}}(x|y)\|_2 \leq \varepsilon\right\}. \tag{4}$$

This problem fits our formalism with $\Phi = \nabla \log p_{\mathbb{x}|\mathbb{y}}(\cdot|y)$ and $\Phi(x^\star) = 0$, since $x^\star$ is a minimizer. When $x^\star = \hat{x}_{\mathrm{MAP}}$, we have $\mathcal{S}_\delta^\varepsilon \subseteq H^\alpha$ for some parameter $\varepsilon > 0$ that depends on α and sufficiently small $\delta > 0$. However the reverse inclusion is not true in general.

1.2 Related works

As illustrated by these examples, describing uncertainty regions can be useful in various fields such as system identification, inverse and blind inverse problems, data assimilation. It can also be regarded as a type of sensibility analysis or uncertainty quantification. Let us review some approaches available in the literature. We will use the nomenclature of the identifiability theory [59], which classifies the study as *structural* in the noiseless setting ($\varepsilon = 0$) and *practical* when noise is considered ($\varepsilon > 0$).

1.2.1 Structural identifiability

Using our notation, *global structural identifiability* means that $\mathcal{S}^0 = \{x^\star\}$ while *local structural identifiability* means that there exists $\delta > 0$ such that $\mathcal{S}_\delta^0 = \{x^\star\}$. The literature devoted to those two notions of *identifiability* is vast. Under mild regularity assumptions on Φ, a necessary and sufficient condition for local structural identifiability is given by $\ker(J_\Phi(x^\star)) = \{0\}$, where $J_\Phi(x^\star) \in \mathbb{R}^{M\times N}$ is the Jacobian matrix of Φ at $x^\star$ [52]. When Φ is linear, this condition reads $\ker(\Phi) = \{0\}$ and implies global structural identifiability. Beyond the linear case, there exist rich theories regarding the (local) identifiability of bilinear

inverse problems [38, 33], phase retrieval problems [30, 22], partial differential equations [57] or neural networks [46, 9], to name a few.

When $x^\star$ is non-identifiable, the analysis and characterization of the non-trivial set $\mathcal{S}^0$ locally around $x^\star$ has also been the object of many works. Existing methods include computing the uncertainty region $\mathcal{S}^0$ in explicit form using differential algebra [6], numerical algebraic geometry (e.g. diffalg [26], daisy [7]) or Lie group theory [40]. However, these algebraic methods are limited to the analysis of rational functions and become computationally prohibitive as the number of parameters increases.

1.2.2 Practical identifiability

Practical identifiability, which guarantees stability to noise on the observations, is the main topic of this paper. Alternative names are sensitivity analysis or uncertainty quantification, which are closely related. For the local characterization of the uncertainty region $\mathcal{S}^\varepsilon$, various approaches exist.

Coordinate-based methods Coordinate-based methods examine the model's sensitivity along each direction of the canonical basis. In this setting, the function $x \mapsto \exp\left(-\|\Phi(x) - \Phi(x^\star)\|_2\right)$ is often called likelihood.

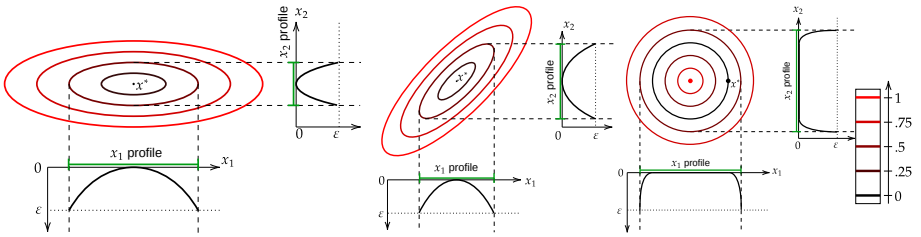
Profile likelihood methods [56, 50] compute one-dimensional profiles. The profiles are evaluated by fixing one coordinate and optimizing over the remaining ones:

$$\min_{\substack{x \in \mathbb{R}^N \\ x_i = x_i^\star + t_j}} \|\Phi(x) - \Phi(x^\star)\|_2.$$

This is done for different deviations $t_j \in \mathbb{R}$ and all coordinates i. The directions i, where the likelihood is high are uncertain. An illustration with simple likelihood functions is given in Fig. 1. Notice that in practice, only endpoints of the confidence intervals are computed [18]. A weakness of these approaches is that they act component-wise. Hence, in Fig. 1 examples (a), (b), we see that the profiles change a lot depending on the coordinate system, showing that the profile directions should be chosen with care. In Fig. 1 (c) for instance, the profiles are locally flat in both directions, while the uncertainty is only one-dimensional along the unit circle.

A possible approach to avoid this drawback, is to find meaningful directions using the Fisher information matrix, defined by $J_\Phi(x^\star)^T \cdot J_\Phi(x^\star)$. The eigenvectors associated to the lowest singular values indicate the directions which are hardest to identify. This idea was used by [14] to identify combinations of coordinates which are jointly identifiable and draw profiles likelihood for these subsets only. A similar mechanism is at the core of our approach, though used in a different manner.

Monte Carlo sampling methods One way to express our problem is to recover $\Phi^{-1}(\mathcal{B}(\Phi(x^\star), \varepsilon))$, the pre-image by Φ of a ball centered on $\Phi(x^\star)$. A simple way to do so is to deploy a Monte Carlo sampling method [42]. A few



(a) $\Phi = A$ (b) $\Phi = UAU^T$ (c) $\Phi(x) = \|x\|_2^2$

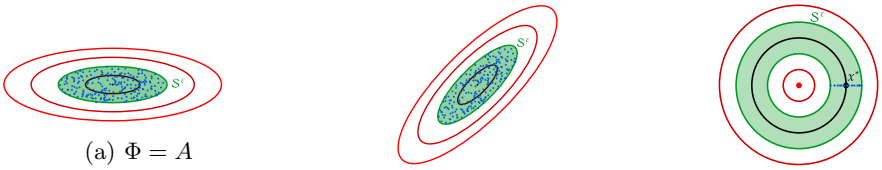
Figure 1: Illustration of the profile likelihood method. The red-black curves are the level sets of the discrepancy $\|\Phi(\cdot) - \Phi(x^\star)\|_2$. The discrepancy profiles along each axis are plotted in black, together with the confidence intervals in green. In (a), (b), we fixed $A = \mathrm{diag}(.2, 1)$, an unitary matrix U and $x^\star = (0, 0)$. In (c), we fixed $x^\star = (1, 0)$.

samples y_k for $1 \leq k \leq K$ are drawn at random within $\mathcal{B}(\Phi(x^\star), \varepsilon)$. Then, the antecedents can be computed as

$$\hat{x}_k \overset{\text{def}}{=} \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \|\Phi(x) - y_k\|_2^2. \tag{5}$$

The discrete set $(\hat{x}_k)$ provides a snapshot of $\mathcal{S}^\varepsilon$. An illustration is provided in Fig. 2 with our basic examples.

Statistical analyses of the samples $(\hat{x}_k)$ can then be performed to quantify the uncertainty. These include, for instance, variance or covariance estimates, bootstrapping [24], sensitivity-based methods [23, 4], or randomize-then-optimize strategies [5].



(a) $\Phi = A$ (b) $\Phi = UAU^T$ (c) $\Phi(x) = \|x\|_2^2$

Figure 2: Illustration of Monte Carlo sampling on the same examples as in Fig. 1. The set $\mathcal{S}^\varepsilon$ is represented in green and the sampling points are in blue.

As opposed to the profile likelihood approach, Monte Carlo sampling accounts for relationships between coordinates. Yet, it suffers from multiple weaknesses.

- It can be computationally expensive given that each sample $\hat{x}_k$ requires solving an optimization problem and that the number of sampling points required for statistical analyses grows exponentially with the dimension of the problem [16, 17].

- For degenerate problems, where Φ is non injective, the optimization problem may possess multiple minimizers. In those situations, sampling becomes inefficient. In Fig. 2 (c) for instance, we started the optimization algorithm by setting $x_{\text{init}} = x^\star$ as an initialization point. As a result, the sampling approach only yields a 1D subset of $\mathcal{S}^\varepsilon$, that does not capture the complete annulus.

- Uniform sampling in $\mathcal{B}(\Phi(x^\star), \varepsilon)$ might not translate to a uniform sampling in $\mathcal{S}^\varepsilon$, potentially over-representing some regions and under-representing others. This issue will be striking in our numerical experiments.

Some of these challenges can be addressed with advanced diffusion models, which have recently gained significant attention in artificial intelligence research. These methods were initially developed for generative modeling tasks [25] and have since been adapted for various inverse problems [12, 36, 11]. Ongoing research focuses on certifying the accuracy of posterior sampling [54] and improving the efficiency of sample generation. However, given the computational cost of these approaches, we have chosen not to use them for comparisons.

Geometric analysis Finally, the last class of methods, the one we endorse, aims to compute geometrically the structure of the set $\mathcal{S}^\varepsilon$. [3] proved that the boundary of $\mathcal{S}^\varepsilon$ is a $N-1$-dimensional manifold and proposed a numerical method to compute its boundary. It relies on the construction of a complete vector field tangential to the level sets of $\|\Phi(x) - \Phi(x^\star)\|_2$. Then, the determination of the boundary of $\mathcal{S}^\varepsilon$ amounts to solve an ODE. This approach is limited to low-dimensional problems such as the two-dimensional non-polynomial and was applied only to three-dimensional polynomial by [3]. Instead of considering the whole space, [49] proposed to determine a trajectory referred to as *minimally disruptive curve* or *talweg* within the structural or practical non-identifiable sets $\mathcal{S}^0$ and $\mathcal{S}^\varepsilon$. This trajectory is a one-dimensional sub-manifold approximation of $\mathcal{S}^\varepsilon$ with the smallest discrepancy values. However, it is unclear how to extend this approach to higher dimensions. Another approach given by [55] consists in successively eliminating one parameter at a time to finally obtain a partial low-dimensional boundary of the uncertainty region. This method is not applicable to high-dimensional problems.

For further details and references related to these approaches, we refer the reader to the comprehensive reviews [41, 59, 35].

1.3 Main contribution

We aim at advancing the field of local practical identifiability by designing the Jackpot algorithm. It is applicable to large dimensional problems and yields a manifold of arbitrary (but preferably small) dimension D. This manifold is constructed so as to provide a good approximation of the ε-uncertainty region in Hausdorff distance, regardless of the non-linear and non-algebraic nature of the model. A simple illustration of its output is given in Fig. 5.

The proposed algorithm is based on automatic differentiation, allowing us to tackle high-dimensional problems. In particular, we illustrate its behavior in problems involving dynamical systems and neural networks. It is being integrated within the Python package deepinv.

2 The Jackpot algorithm and its guarantees

Our main goal is to describe $\mathcal{S}_\delta^\varepsilon = \{x \in \mathcal{B}(x^\star, \delta),\ \|\Phi(x) - \Phi(x^\star)\|_2 \leq \varepsilon\}$, the local $\varepsilon^2/2$-sublevel set of the function

$$F(x) \stackrel{\text{def}}{=} \frac{1}{2}\|\Phi(x) - \Phi(x^\star)\|_2^2. \tag{6}$$

In what follows, to measure how well we approximate the set $\mathcal{S}_\delta^\varepsilon$, we will use Kolmogorov and Hausdorff distances which we define next.

Definition 1 (D-width or Kolmogorov distance)**.** Let $\mathcal{G}(D, N)$ denote the set of D- dimensional vector subspaces in dimension N (the Grassmanian). The Kolmogorov distance δ_D of a set $\mathcal{S} \subset \mathbb{R}^N$ is defined by:

$$\delta_D(\mathcal{S}) \stackrel{\text{def}}{=} \inf_{\mathcal{V} \in \mathcal{G}(D,N)} \sup_{x \in \mathcal{S}} \inf_{v \in \mathcal{V}} \|x - v\|_2. \tag{7}$$

It measures how well a set $\mathcal{S}$ can be approximated by a D-dimensional subspace. Let us also recall the notion of Hausdorff distance.

Definition 2 (Hausdorff distance)**.** For two subsets $\mathcal{X}, \mathcal{Y}$ of a metric space $(\mathcal{M}, d)$, the Hausdorff distance is defined as

$$d_{\mathcal{H}}(\mathcal{X}, \mathcal{Y}) \stackrel{\text{def}}{=} \max\left\{\sup_{x \in \mathcal{X}} d(x, \mathcal{Y}),\ \sup_{y \in \mathcal{Y}} d(\mathcal{X}, y)\right\} \tag{8}$$

$$= \inf_{\eta \geq 0} \left\{\mathcal{X} \subseteq \mathcal{Y} + \mathcal{B}(0, \eta) \text{ and } \mathcal{Y} \subseteq \mathcal{X} + \mathcal{B}(0, \eta)\right\}. \tag{9}$$

This distance intuitively measures how much a set should be dilated to include the other one and vice versa.

In this section, we start by proposing a linear manifold approximation $\widetilde{\mathcal{M}}_\delta^\varepsilon$ obtained from a quadratic approximation of F at $x^\star$. This allows us to present the main approximation results in a simplified format. We then turn to the construction of a more precise nonlinear manifold approximation $\mathcal{M}_\delta^\varepsilon$. We describe its approximation guarantees, and a practical algorithm to compute it.

2.1 Linear approximation

Under a $\mathcal{C}^2$ regularity assumption, we can approximate F by a quadratic function. Most terms in the Taylor expansion vanish since $x^\star$ is a minimizer and for any $h \in \mathbb{R}^N$, we obtain:

$$F(x^\star + h) = \frac{1}{2}\|J_\Phi(x^\star) \cdot h\|_2^2 + o(\|h\|_2^2). \tag{10}$$

This motivates us to consider the quadratic approximation of F at $x^\star$:

$$\widetilde{F}(x) \stackrel{\text{def}}{=} \frac{1}{2}\|J_\Phi(x^\star)\cdot(x-x^\star)\|_2^2. \tag{11}$$

The uncertainty region can be described as the $\varepsilon^2/2$-sublevel set of this function, denoted $\widetilde{\mathcal{S}}^\varepsilon = \left\{x\in\mathbb{R}^N,\ \tilde{F}(x)\leq \varepsilon^2/2\right\}$. Similarly to the definition of $\mathcal{S}_\delta^\varepsilon$, we also set

$$\widetilde{\mathcal{S}}_\delta^\varepsilon = \left\{x\in\mathcal{B}(x^\star,\delta),\ \tilde{F}(x)\leq\frac{\varepsilon^2}{2}\right\}. \tag{12}$$

2.1.1 Approximation result

In the linear case, the best approximation of $\widetilde{\mathcal{S}}^\varepsilon$ by a D-dimensional subspace is explicit using the singular vectors of the Jacobian matrix of Φ. We can indeed decompose the Jacobian as $J_\Phi(x) = U(x)\Sigma(x)V^T(x)$, where $U(x)\in\mathbb{R}^{M\times M}$ and $V(x)\in\mathbb{R}^{N\times N}$ are orthogonal matrices and $\Sigma(x)\in\mathbb{R}^{M\times N}$ is a diagonal rectangular matrix with nonnegative and decaying entries. We let $J^\star \stackrel{\text{def}}{=} J_\Phi(x^\star) = U^\star\Sigma^\star V^{\star,T}$, where $U^\star = [u_1^\star,\dots,u_M^\star]$, $V^\star = [v_1^\star,\dots,v_N^\star]$, $\mathrm{diag}(\Sigma^\star) = [\sigma_1^\star,\dots,\sigma_{R^\star}^\star,0,\dots,0]$, and $R^\star$ is the rank of $J^\star$. By convention, we let $\sigma_{R^\star+1}^\star,\sigma_{R^\star+2}^\star,\dots,\sigma_N^\star = 0$. The following description is standard.

Proposition 3. *(From [32]) The sublevel set $\widetilde{\mathcal{S}}^\varepsilon$ is a (degenerate) hyper-ellipsoid with axes aligned with the vectors $(v_n^\star)$. The length of the ellipsoid along the semi-axis $v_n^\star$ is given by $\frac{\varepsilon}{\sigma_n^\star}$ for $n\in\{1,\dots,R^\star\}$ and it is infinite for the other axes.*

The following result is rather standard in approximation theory, see e.g. [47, Thm. 1.3]. First the uncertainty region is approximated by an affine subspace $\mathcal{L}$ and then by a troncation of this subset which gives the approximation manifold $\widetilde{\mathcal{M}}_\delta^\varepsilon$ of $\widetilde{\mathcal{S}}_\delta^\varepsilon$. Guarantees on the approximation distances are also provided and the proof is provided in Appendix A.1.1.

Proposition 4. *Firstly, the affine $D-$dimensional affine space that best approximates $\widetilde{\mathcal{S}}^\varepsilon$ for the Kolmogorov distance is given by $\mathcal{T}_{x^\star} \stackrel{\text{def}}{=} x^\star + \mathcal{V}_D^\star$ with*

$$\mathcal{V}_D^\star \stackrel{\text{def}}{=} \mathrm{span}\left(v_{N-D+1}^\star,\dots,v_N^\star\right). \tag{13}$$

The Kolmogorov distance is given by $\delta_D(\widetilde{\mathcal{S}}^\varepsilon - x^\star) = \frac{\varepsilon}{\sigma_{N-D}^\star}$.

Secondly, for any $\delta>0$, we define the set

$$\widetilde{\mathcal{M}}_\delta^\varepsilon \stackrel{\text{def}}{=} \left\{x\in\mathcal{T}_{x^\star};\left\|\Sigma^\star V^{\star,T}(x-x^\star)\right\|_2^2\leq\varepsilon^2 \text{ and } \|x-x^\star\|_2\leq\delta\right\}. \tag{14}$$

It satisfies the following equality:

$$d_{\mathcal{H}}(\widetilde{\mathcal{M}}_\delta^\varepsilon,\widetilde{\mathcal{S}}_\delta^\varepsilon) = \min\left(\frac{\varepsilon}{\sigma_{N-D}^\star},\ \delta\right). \tag{15}$$

This linearized approximate set $\widetilde{\mathcal{M}}_\delta^\varepsilon$ was already described in [16].

2.1.2 Lowest singular vectors computation

The results above show that the singular values and singular vectors of the Jacobian $J^\star = J_\Phi(x^\star)$ play a critical role for the description of the set $\widetilde{\mathcal{S}}^\varepsilon$. To compute them, two operation are required: i) evaluating the Jacobian matrix or matrix-vector products with it and ii) computing the singular value decomposition.

Jacobian computation Computing matrix-vector products with the Jacobian can be achieved efficiently using automatic differentiation algorithms. In Pytorch for instance, this can be achieved through the *torch.jvp* and *torch.vjp* functions. If the dimension N or M is sufficiently small, then the whole Jacobian matrix $J^\star$ can be evaluated and stored. If the dimensions are too large, then automatic differentation algorithms make it possible to compute left and right matrix-vector products, i.e. products of the form $J^\star v$ or $u^T J^\star$ for arbitrary directions $u \in \mathbb{R}^M$ and $v \in \mathbb{R}^N$.

Singular vectors computation The computation of singular pairs of $J^\star$ or the eigenpairs of $J^{\star,T} J^\star$ are two equivalent problems. The latter is known as a symmetric eigenvalue problem [45]. Several algorithms, called eigensolvers [39], have been developed for this purpose. If the matrix is small enough, traditional decomposition methods based on the QR algorithm can be used.

In high dimension, we adopted the Locally Optimal Block Preconditioned Conjugate Gradient (LOBPCG) method [34, 13]. It is *matrix-free*: it only requires matrix-vector products and does not process the entire matrix. The computation is achieved by optimizing the generalized Rayleigh quotient. For $X \in \mathbb{R}^{N\times D}$, this quotient is defined as

$$f(X) = \mathrm{Tr}\left(\left(X^T X\right)^{-1} (J^\star X)^T J^\star X\right). \tag{16}$$

As such, finding the D leftmost singular vectors of $J^\star$ consists in finding the minimizer of

$$\operatorname*{argmin}_{X\in \mathrm{St}(N,D)} f(X),$$

where $\mathrm{St}(N,D) \stackrel{\text{def}}{=} \left\{X \in \mathbb{R}^{N\times D} \;;\; X^T X = \mathrm{Id}_D\right\}$ is the set of $N \times D$ orthonormal matrices also called the Stiefel manifold. The LOBPCG algorithm can be interpreted as a manifold gradient descent [1] with a carefully designed step-size based on the previous iterates.

2.2 Nonlinear approximation

We now turn to the approximation of the sublevel sets of F without resorting to a quadratic approximation: we want to find a D-dimensional manifold $\mathcal{M}_\delta^\varepsilon$ that approximates $\mathcal{S}_\delta^\varepsilon$ well.

2.2.1 The Jackpot manifold

The linear subspace $\mathcal{T}_{x^\star} = x^\star + \mathcal{V}_D^\star$ can be a poor approximation of $\mathcal{S}^\varepsilon$ when Φ is nonlinear. However, it is a natural tangent space for the set $\mathcal{M}_\delta^\varepsilon$ at $x^\star$. The idea behind our algorithm consists in bending this tangent space, to better fit the uncertainty region $\mathcal{S}_\delta^\varepsilon$. A similar idea was proposed by W. Rheinboldt for polynomials in [51]. However, the algorithm was limited to three dimensions and could not visualize the uncertainty region, as it was constrained to the case where $\varepsilon = 0$.

In mathematical terms, we want to find a one-to-one map γ that sends points of $\mathcal{T}_{x^\star}$ onto $\mathcal{S}^\varepsilon$. This can be achieved thanks to the following optimization problem:

$$\gamma(z) \stackrel{\text{def}}{=} \underset{V_D^{\star,T}(x-x^\star)=z}{\operatorname{argmin}} \frac{1}{2}\left\|\Phi(x) - \Phi(x^\star)\right\|_2^2, \qquad (\mathcal{P}_z)$$

where

$$V_D^\star \stackrel{\text{def}}{=} [v_{N-D+1}^\star, \cdots, v_{N-1}^\star, v_N^\star] \in \mathbb{R}^{N\times D}. \qquad (17)$$

The constraint $V_D^{\star,T}(x-x^\star) = z$ means that the projection of $\gamma(z)$ on the tangent space $\mathcal{T}_{x^\star}$ has coordinate z. It ensures the injectivity of the map γ. The argmin notation refers to a local minimizer near $x^\star$. The main tools appearing in this description are depicted in Fig. 3.

A simple illustration To illustrate this principle, we consider the mapping $\Phi(x) = \frac{1}{2}\|x\|_2^2$ and set $x^\star = e_1$, where $(e_1, \ldots, e_N)$ is the canonical basis. It is a basic example from a broader context named algebraic implicit curves parameterization. For this specific function, the set $\mathcal{S}^0$ is a sphere of radius 1 centered at 0. The tangent space is the hyperplane orthogonal to e_1 passing through $x^\star$. Simple computation yields $\gamma(z) = \sqrt{1-\|z\|_2^2}e_1 + \sum_{n=2}^N z_n e_n$. We see that the proposed parameterization allows us to recover the half sphere, which is illustrated in Fig. 3. This example also shows the importance of considering local minimizers, since the problem $(\mathcal{P}_z)$ admits two global minimizers.

Main definition The mapping γ provides a good candidate manifold that approximates the uncertainty region $\mathcal{S}_\delta^\varepsilon$. It writes as follows.

Definition 5 (The Jackpot manifold)**.** For $\delta > 0$, we define the set

$$\mathcal{M}_\delta^\varepsilon \stackrel{\text{def}}{=} \left\{\gamma(z)\ ;\ z \in \mathbb{R}^D, \|\Phi(\gamma(z)) - \Phi(x^\star)\|_2 \le \varepsilon \text{ and } \|\gamma(z) - x^\star\|_2 \le \delta\right\}. \quad (18)$$

As shown below, this set is well defined provided that δ is small enough. The proof is provided in Appendix A.2.1

Theorem 6 (Manifold approximation is well defined)**.** *Let $\Phi : \mathbb{R}^N \to \mathbb{R}^M$ be a $\mathcal{C}^1$ map, $x^\star \in \mathbb{R}^N$, and $D \ge \dim\ker J_\Phi(x^\star)$ denote an integer.*

Then there exists a neighborhood $\mathcal{U} \subseteq \mathbb{R}^D$ such that $(\mathcal{P}_z)$ admits a unique solution $\gamma(z)$ for $z \in \mathcal{U}$. Moreover, the mapping $\gamma : \mathcal{U} \to \mathbb{R}^N$ verifies $\gamma(0_D) = x^\star$ and it has the following properties

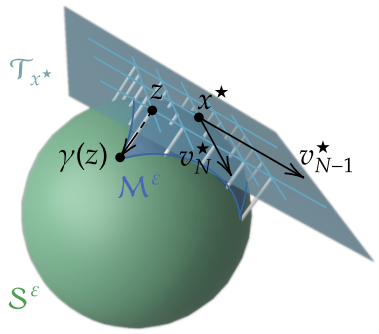


Figure 3: Summary of Jackpot (JACobian Kernel Projection by OpTimization). In this example, we set $\Phi(x) = \|x\|_2^2$, $\varepsilon = 0$ and $x^\star = e_1$. In this setting, the uncertainty region $\mathcal{S}^\varepsilon$ is just a unit sphere, and the affine space $\mathcal{T}_{x^\star}$ is the tangent space to $\mathcal{S}^\varepsilon$ at $x^\star$. The uncertainty region is obtained by projecting the tangent plane onto $\mathcal{S}^\varepsilon$.

- *If Φ is of class $\mathcal{C}^2$, then γ is of class $\mathcal{C}^1$.*
- *If Φ is of class $\mathcal{C}^1$ and J_Φ is locally Lipschitz and definable, then γ is a locally Lipschitz definable function.*

Remark 7. The second condition in Theorem 6 is weaker than the first one and allows us to handle a larger class of operators. The notion of definable function [8] generalizes the notion of (implicitly defined) semi-algebraic function. It encompasses all functions commonly considered in engineering. We refer the reader to [8, Appendix A.2] for a precise definition.

Our approximation manifold $\mathcal{M}_\delta^\varepsilon$ is now well defined. In what follows we will refer to γ as the *parameterization* of the manifold $\mathcal{M}_\delta^\varepsilon$.

2.2.2 Approximation guarantees

In this paragraph, we show in Theorem 10 that $\mathcal{M}_\delta^\varepsilon$ provides a good approximation of $\mathcal{S}_\delta^\varepsilon$ under certain regularity assumptions.

Assumption 8 (A simple approximability condition)**.** Let

$$\mathcal{P}_z = \left\{ x \in \mathcal{S}_\delta^\varepsilon, V_D^{\star,T}(x - x^\star) = z \right\}$$

denote the slice of $\mathcal{S}_\delta^\varepsilon$ which has coordinate z on $\mathcal{T}_{x^\star}$. We assume that

$$\sup_{\substack{z \in \mathbb{R}^D \\ \|\gamma(z) - x^\star\|_2 \leq \delta}} \operatorname{diam}(\mathcal{P}_z) \leq \eta. \tag{19}$$

This assumption is illustrated in Fig. 4. Notice that the value $\eta > 0$ can be large even if the set is thin, when the set is bent a lot with respect to $\mathcal{T}_{x^\star}$. For a $\mathcal{C}^2$ mapping Φ, the thickness η can be controlled by regularity properties as shown in the following proposition, the proof of which is deferred to Appendix A.2.2.

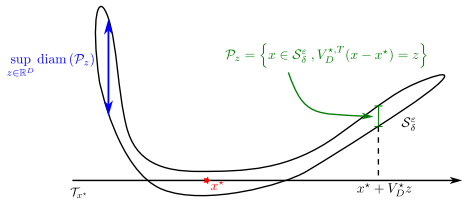


Figure 4: An illustration of Assumption 8. The slices $\mathcal{P}_z$ of $\mathcal{S}_\delta^\varepsilon$ have a diameter uniformly bounded by η.

Proposition 9 (Uniformly bounded curvature)**.** *Assume that Φ is of class $\mathcal{C}^2$, that $D \geq \dim\ker(J_\Phi(x^\star)) = N - R^\star$ and that the Hessian tensor H_Φ satisfies*

$$\|H_\Phi\|_{2,\delta} \stackrel{\text{def}}{=} \sup_{x,x',x''\in\mathcal{B}(x^\star,\delta)} \frac{\|H_\Phi(x)(x',x'')\|_2}{\|x'\|_2\,\|x''\|_2} < \frac{\sigma^\star_{N-D}}{\delta}$$

where $\sigma^\star_{N-D}$ is the $(N-D)$-th singular value of the Jacobian $J_\Phi(x^\star)$. Then Assumption 8 is verified with

$$\eta = \frac{2\varepsilon}{\sigma^\star_{N-D} - \delta\|H_\Phi\|_{2,\delta}}. \tag{20}$$

We are now in a position to show how well $\mathcal{M}_\delta^\varepsilon$ approximates $\mathcal{S}_\delta^\varepsilon$ in terms of the Hausdorff distance defined in (2). The proof is provided in Appendix A.2.2.

Theorem 10 (Approximation guarantees)**.** *Set a point $x^\star \in \mathbb{R}^N$, a map Φ of class $\mathcal{C}^2$, a threshold $\varepsilon > 0$ and $D \geq \dim\ker J_\Phi(x^\star)$. Then $\mathcal{M}_\delta^\varepsilon$ defines a manifold with parameterization γ for sufficiently small $\delta > 0$.*

Moreover, under Assumption 8, $\mathcal{M}_\delta^\varepsilon$ is a good approximation of $\mathcal{S}_\delta^\varepsilon$ in the sense that:

$$d_{\mathcal{H}}(\mathcal{M}_\delta^\varepsilon,\ \mathcal{S}_\delta^\varepsilon) \leq \eta, \tag{21}$$

where $d_{\mathcal{H}}$ is defined in (2)*. More precisely, we have*

$$\mathcal{M}_\delta^\varepsilon \subseteq \mathcal{S}_\delta^\varepsilon \subseteq \mathcal{M}_\delta^\varepsilon + \mathcal{B}(0,\eta). \tag{22}$$

Remark 11. A few remarks are in order:

- In the linear case, since the Hessian is null, any radius $\delta > 0$ suits as long as the dimension D satisfies $D \geq \dim\ker J_\Phi(x^\star)$.
- An upper-bound on the Hessian norm $\|H_\Phi\|_{2,\delta}$ is

  $$\sup_{x\in\mathcal{B}(x^\star,\delta)} \left\|\left(\sigma_{\max}\left(H_{\Phi_m}(x)\right)\right)_{m\leq M}\right\|_2$$

  with Φ_m being the m-th component of Φ.

2.2.3 Numerical computation

From a numerical perspective, we have to discretize the manifold $\mathcal{M}_\delta^\varepsilon$. This can be done by evaluating $(\mathcal{P}_z)$ on a discrete subset $Z = \{z_k\}_{0\leq k\leq K}$ of $\mathbb{R}^D$. We let $\mathcal{M}_\delta^\varepsilon(Z)$ denote the discretized manifold:

$$\mathcal{M}_\delta^\varepsilon(Z) \stackrel{\text{def}}{=} \mathcal{M}_\delta^\varepsilon \cap \{\gamma(z); z \in Z\}\,. \tag{23}$$

The main problem now reduces to evaluating γ for grid points Z satisfying the constraints. The principle is to propagate a front starting from $z = 0_D$ with a grid-search algorithm and to solve $(\mathcal{P}_z)$ at each z. A pseudo-code is given in Algorithm 1. We detail the two steps below.

Grid search In this work, we simply discretize $\mathbb{R}^D$ with a Cartesian grid. Other choices such as the honeycomb pattern are possible, but have not been explored. Fixing a grid length $s > 0$, we define the following Cartesian grid

$$Z_s \stackrel{\text{def}}{=} \left(s\mathbb{Z}^D\right) \cap \mathcal{B}_\infty(0_D, \delta) = s\left\{-\left\lfloor\frac{\delta}{s}\right\rfloor, \ldots, \left\lfloor\frac{\delta}{s}\right\rfloor\right\}^D. \tag{24}$$

Since the initial point $\gamma(0) = x^\star$ is known, the grid search should start with the point $0_D \in Z$. Solving $(\mathcal{P}_z)$ can be achieved with first order methods, which require a starting point. Since γ is a smooth mapping, a good approximation of $\gamma(z)$ is $\gamma(z')$ for some z' where γ has already been computed. This leads to consider a grid search method starting from $0_D \in Z$ and passing from neighbor to neighbor. The Breadth-First Search (BFS) algorithm is well suited for this task. We add a graph structure $G = (Z, E)$ (where E is the set of edges) on the grid in order to apply this algorithm. Given this structure, the natural choice of initial guess for the optimization at a non already-explored node is thus given by the closest already computed neighbor. BFS method also makes it possible to stop the search in the neighborhood of a vertex whenever one of the two criteria defining (23) are violated.

Optimization algorithm Each step of Algorithm 1 requires solving the optimization problem $(\mathcal{P}_z)$. Obtaining the following result is based on inserting the linear constraint in the objective function. Its proof is provided in Appendix A.3.1.

Proposition 12. *Let $\Pi_\perp \stackrel{\text{def}}{=} \mathrm{Id}_N - V_D^\star V_D^{\star,T}$ be the orthogonal projection on the subspace $(\mathrm{Im}V_D^\star)^\perp$ and consider the following optimization problem for $z \in \mathbb{R}^D$*

$$\eta_\perp(z) \stackrel{\text{def}}{=} \operatorname*{argmin}_{x\in\mathbb{R}^N} \frac{1}{2}\left\|\Phi(x^\star + V_D^\star z + \Pi_\perp x) - \Phi(x^\star)\right\|_2^2. \tag{$\mathcal{P}'_z$}$$

Then there exists a local neighbor $\mathcal{U} \subseteq \mathbb{R}^D$ of 0_D such that the solution γ of $(\mathcal{P}_z)$ verifies for $z \in \mathcal{U}$,

$$\gamma(z) = x^\star + V_D^\star z + \Pi_\perp \eta_\perp(z). \tag{25}$$

Algorithm 1 BFS parameterization of $\mathcal{M}_\delta^\varepsilon(Z)$

Inputs of the model:
- $\Phi : \mathbb{R}^N \to \mathbb{R}^M$ a $\mathcal{C}^1$ mapping.
- $x^\star \in \mathbb{R}^N$ an estimation of input parameter.
- $1 \leq D < N$ the manifold dimension.
- $V_D^\star \in \mathbb{R}^{N\times D}$ the D lowest right singular vectors of $J_\Phi(x^\star)$.

Inputs of the paramaterization:
- $\delta > 0$ a radius around $x^\star$ (to deal with $\|x - x^\star\|_2 \leq \delta$).
- $\varepsilon > 0$ a discrepancy threshold (to deal with $\|\Phi(x) - \Phi(x^\star)\|_2 \leq \varepsilon$).
- $G = (Z, E)$ a connected graph with $0_D \in Z \subseteq \mathbb{R}^D$.

Initialization:
$\eta_\perp(0_D) \leftarrow 0_N$.
$\gamma(0_D) \leftarrow x^\star$.
repeat
  $(z, z_{\text{prev}}) \leftarrow$ the couple (non-evaluated, evaluated) neighboring vertices with smallest distance.
  $\eta_\perp(z) \leftarrow$ solution of $(\mathcal{P}'_z)$ initialized at $\eta_\perp(z_{\text{prev}})$ with an L-BFGS method.
  $\gamma(z) \leftarrow x^\star + V_D^\star z + (\mathrm{Id}_N - V_D^\star V_D^{\star,T})\eta_\perp(z)$.
  if $\|\Phi(\gamma(z)) - \Phi(x^\star)\|_2 > \varepsilon$ **or** $\|\gamma(z) - x^\star\|_2 > \delta$ **then**
    Remove vertex z from Z and all its edges from E.
  end if
until No couple (z, z_{prev}) exists
Return $\{\gamma(z)\}_{z\in Z}$

Moreover, if Φ is of class $\mathcal{C}^2$, a gradient descent algorithm applied to $(\mathcal{P}'_z)$ converges linearly.

As before, the argmin notation refers to a local minimizer near 0_N.

To further accelerate the computation of $\gamma(z)$ we propose to use a quasi-Newton method on $(\mathcal{P}'_z)$. More specifically we suggest to use a L-BFGS method. It accelerates the convergence by using a preconditionning based on an estimation of the Hessian using the gradients at each step. We summarize the whole process and give a pseudo-code in Algorithm 1.

Approximation result One can extend the approximation result of the previous section to this discrete framework. The proof of the following corollary is provided in Appendix A.3.2.

Corollary 13. *Under the same assumptions as in Theorem 10, and using a Cartesian grid $Z = Z_s$ defined in* (24)*, the discretized manifold provides a good approximation of the uncertainty region. Specifically, we have*

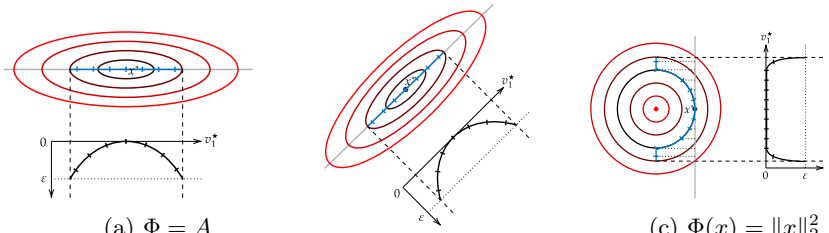
$$d_{\mathcal{H}}(\mathcal{M}_\delta^\varepsilon(Z), \mathcal{S}_\delta^\varepsilon) \leq \eta + L\sqrt{D}\frac{s}{2}, \tag{26}$$

where $L > 0$ is the Lipschitz constant of the local solution γ to $(\mathcal{P}_z)$, $d_{\mathcal{H}}$ is the Hausdorff distance as defined in Definition 2, and $\eta > 0$ is the threshold defined in Theorem 10.

2.2.4 Discussion

Before turning to the numerical experiments, we describe the strengths and weaknesses of the approach in order to highlight the problems for which it is best suited.

Behavior on the toy examples To begin with, let us illustrate how the algorithm behaves on the three introductory examples of Fig. 1. We see that the algorithm outputs one dimensional sets (in blue), which correctly approximate the uncertainty regions in all three cases. The method suffers neither from the limitation of the choice of coordinates as were profile likelihood methods, nor from the instability of Monte Carlo-based methods when the Jacobian of Φ is degenerate.



(a) $\Phi = A$ (b) $\Phi = UAU^T$ (c) $\Phi(x) = \|x\|_2^2$

Figure 5: Illustration of the proposed Jackpot method with $D = 1$ on the same examples as in Fig. 1. The low dimensional manifold $\mathcal{M}_\delta^\varepsilon$ is represented with blue lines. In contrast to the profile likelihood method, the profile is computed along the most suitable direction. In addition, the set $\mathcal{M}_\delta^\varepsilon$ is the best possible one-dimensional approximation of the 2 dimensional ellipses in (a) and (b). In example (c), the approximating set $\mathcal{M}_\delta^\varepsilon$ precisely coincides with the unit circle (i.e., $\mathcal{S}_\delta^\varepsilon$), except on the extremities.

Adapting to large dimensions Since the LOBPCG algorithm is matrix-free and combined with an automatic differentiation algorithm, our algorithm is scalable to large model (up to millions of parameters). The approximation being of smallest dimension D, it can be represented quite easily at least for dimensions $D = 1, 2, 3$. Depending on the geometry of the problem, our method gives a full exploration of the uncertainty region or only a partial one as shown in Fig. 6.

When is it most useful? From our theoretical analysis in Section 2.2.2, one can see that the low-dimensional manifold $\mathcal{M}_\delta^\varepsilon$ is a good approximation of $\mathcal{S}_\delta^\varepsilon$

provided that D is sufficient large relatively to the number of "low" singular values of $J_\Phi(x^\star)$. Yet, for large-scale problems, it is not possible to set D too large due to computational limitations. This leads us to classify problems in three categories depending on the Jacobian singular spectrum, as illustrated in Fig. 6.

- When the Jacobian is well-conditioned ("flat" spectrum, Fig. 6.a), we have practical local identifiability and the Jackpot method will detect it.
- When the Jacobian is badly-conditioned, we distinguish two cases.
  - If the spectrum admits few low singular values with a sharp transition (Fig. 6.b), $\mathcal{S}_\delta^\varepsilon$ can be well approximated by a low-dimensional manifold and the Jackpot method will compute it.
  - Otherwise (e.g., Fig. 6.c), $\mathcal{S}_\delta^\varepsilon$ cannot be well approximated by a low-dimensional manifold. In this situation, the manifold $\mathcal{M}_\delta^\varepsilon$ computed by Jackpot (with "small" D) will only partially represent $\mathcal{S}_\delta^\varepsilon$.

Notice that in the last case, we cannot expect to "sketch" the uncertainty set efficiently due to its intrinsically high dimensional nature.

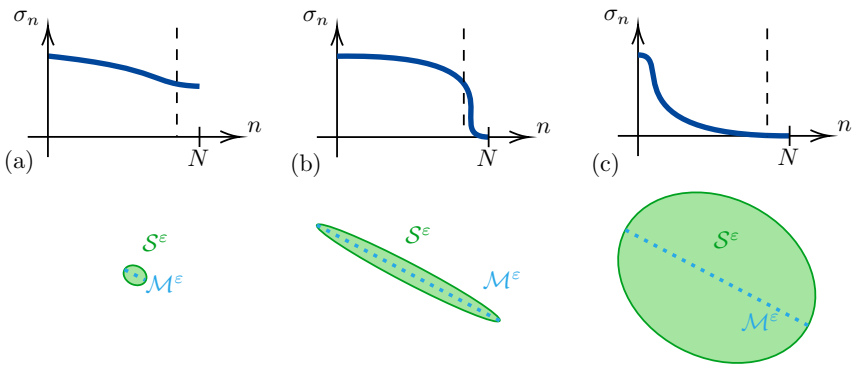


Figure 6: Relation between Jacobian singular spectrum and manifold approximation of the uncertainty region $\mathcal{S}^\varepsilon$. (a) No small singular values: the problem is identifiable. (b) Small number of small singular values. (c) Large number of small singular values. The set $\mathcal{S}^\varepsilon$ can not be approximated by a low dimensional manifold.

Adversarial versus natural perturbations Our goal in this paper is to design a manifold that approximates $\mathcal{S}^\varepsilon = \Phi^{-1}(\mathcal{B}(\Phi(x^\star), \varepsilon))$ in Hausdorff distance. We call the resulting set an *adversarial manifold*. The directions $(v^\star_{N-D+1}, \ldots, v^\star_N)$ represent the perturbation directions that most significantly affect the inverse map Φ^{-1}. These directions also capture the local geometry of $\mathcal{S}^\varepsilon$ most effectively, as seen in Theorem 10.

In contrast, Monte Carlo sampling methods serve a different purpose: they simulate "natural" perturbations, which are typically observed in real-world scenarios. Given a random vector $Y = \Phi(x)+B$, where B follows some probability distribution μ, they capture the pushforward distribution $\Phi_\sharp^{-1}(\mu(\cdot - \Phi(x)))$. Points derived from the Jackpot method, however, may possess very low likelihood and would almost never appear through Monte Carlo sampling. This distinction becomes clearer in our numerical experiments.

Adversarial and natural perturbations offer distinct advantages, depending on the nature of uncertainty one aims to describe. Adversarial perturbations [2] are critical in scenarios where incorrect decisions may lead to severe consequences, as they identify the worst-case distortions. On the other hand, natural perturbations provide an averaged view of uncertainty, reflecting more common or benign variations.

The limitations of locality Our approximation and our analysis is only local. For instance, in Fig. 3, only half of the circle can be parameterized by our method. Hence the method should only be used locally. Solutions such as a periodic reassessment of the tangent plane could be considered, but we did not explore this rather computationally heavy approach in this paper.

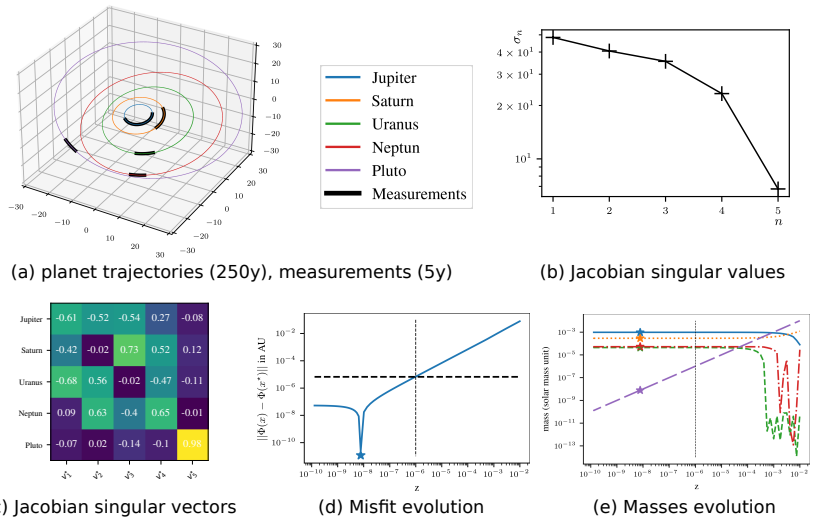
3 Numerical experiments

In this section, we provide a few numerical examples on complex problems and compare it with some alternatives in the main three cases of practical identifiability described in Section 1.1.

3.1 Measuring masses in the Solar system

In this first example, we address the problem of identifying the masses of five planets of the solar system (Jupiter, Saturn, Uranus, Neptun, and Pluto), from their approximate positions at a few time points. We assume that the planets are moving according to Newton's law of gravitation. This corresponds to the problem described in section 1.1.1.

Notation We sort the planets with respect to their distance to the Sun. We let $1 \leq n \leq N$ denote the index of the n-th outer planet and associate the index 0 to the Sun. We let $u_n(t) \in \mathbb{R}^3$ denote the positions in AU of the n-th planet at time t, in an heliocentric frame. Therefore $u_0(t) = 0$ for all time. We let w_n denote the weight of the n-th planet in the solar mass unit, that is $w_0 = 1$. Moreover, we let $u(t) = (u_1(t), \dots, u_N(t)) \in \mathbb{R}^{N\times 3}$ denote the vector of positions and $w = (w_1, \dots, w_N)$ denote the vector of masses. The time is expressed in years. The time $t = 0$ corresponds to January 1, 2000. Initial positions and speeds are taken from [29].



(a) planet trajectories (250y), measurements (5y)

(b) Jacobian singular values

(c) Jacobian singular vectors

(d) Misfit evolution

(e) Masses evolution

Figure 7: Solar system experiment with Jackpot. After measuring the approximate planet's positions for 5 years, we wish to recover their masses. The spectrum of the Jacobian (b) indicates that a direction is particularly uncertain. Looking at the Jacobian singular vectors (c), we see that it corresponds to Pluto's mass. In (d) and (e), we see the evolution of the misfit and of the masses along the 1D set $\mathcal{M}_\delta^\varepsilon$, and observe that all masses are stable (relative to their mass), except Pluto's which can vary by 4 orders of magnitude.

Modeling The solar system can be modeled through the second order dynamical system:

$$\begin{cases} u_0(t) = 0 \\ \ddot{u}_n(t) \stackrel{\text{def}}{=} -\displaystyle\sum_{n'=0}^{N} \frac{G w_{n'}}{\|u_n(t) - u_{n'}(t)\|_2^3}(u_n(t) - u_{n'}(t)), \quad 1 \le n \le N. \end{cases} \tag{27}$$

To discretize this system, we use a Runge-Kutta scheme of order 4 [10]. This yields a forward mapping Φ of the form:

$$\Phi : w \mapsto (u(t_k))_{0 \le k \le K}, \tag{28}$$

where $t_k = k\Delta t$ denote uniformly spaced sampling points. The measurements are taken on a period of 5 years within intervals of $\Delta t = 7$ days, leading to $K = 260$ positions for each planet. Notice that 5 Earth-years do not cover a complete revolution for the outer planets. For instance, Pluto's revolution lasts about 250 years. This is illustrated on Fig. 7 (a), where the black lines illustrate the measurements. We add white Gaussian noise to the measurements with a standard deviation $\varepsilon = 1000$ km ($\simeq 7 \cdot 10^{-6}$ AU, where AU is the Earth-Sun distance). This yields a measurements vector $y = \Phi(\bar{w}) + b$, where $\bar{w}$ denotes the true planets' weights and b is the random perturbation.

Uncertainty with Jackpot To estimate the masses w, we first solve the nonlinear problem

$$w^\star = \underset{w\in\mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|\Phi(w) - y\|_2^2 \tag{29}$$

using a gradient descent to high accuracy. The gradient is evaluated through automatic differentiation.

This gives a fairly good estimate of Pluto's mass, overestimating the true estimate by 4%. Then, we run algorithm 1. The singular values of the Jacobian of Φ are displayed in Fig. 7 b). As can be seen, the last singular value is nearly ten times smaller than the others. Taking a close look at the singular vectors in Fig. 7 c), shows that the last singular vector is entirely concentrated on Pluto's mass, indicating that the estimation of this planet's mass is dubious.

We then compute a $D=1$-dimensional manifold approximation $\mathcal{M}_\delta^\varepsilon$ of the uncertainty region with the Jackpot method. As can be seen in Fig. 7, we can vary the range of Pluto's mass within the range $[10^{-10}, 10^{-6}]$, while staying consistent with the data. In that range, the other planets masses nearly do not change (at least relative to their weight). This illustrates a huge uncertainty on the mass of Pluto when evaluating it with 5 years of planetary observations only.

Comparisons with other approaches In Fig. 8 we use profile likelihood on this problem to estimate the uncertainty. This problem is particularly well suited to this approach, since only Pluto's mass is uncertain. In this case, we see that all profiles are concentrated on the true planets mass, apart for Pluto. There, the profile likelihood method recovers the same uncertainty interval as the one provided by Jackpot. The only difference is that Jackpot was able to automatically find this direction, without having to explore all of them. Also notice that profile-likelihood methods would likely bring partial information on the uncertainty if Pluto was discarded from the set. Indeed, the top-left 4x4 block in Fig. 7 c) indicates strong correlation effects between the different weights.

We do not show the results of the Monte Carlo sampling method since it fails for this example. Indeed, the sampling process only provides points that are very close to $x^\star$, and the method fails to recover the uncertainty set. This phenomenon is similar to the one observed in Fig. 2 (c): the regression problem to recover the samples x^i is very flat along the direction of Pluto's mass.

3.2 Blind deblurring

In this section, we explore the problem described in section 1.1.2. Observing a biological object with a microscope yields diffraction limited (i.e. blurry) images. When the impulse response, or Point Spread Function (PSF) of the system is unknown, improving the image sharpness requires solving a blind deblurring problem. In this section, we showcase how Jackpot makes it possible to evaluate uncertainty on the recovered blur kernels and images.

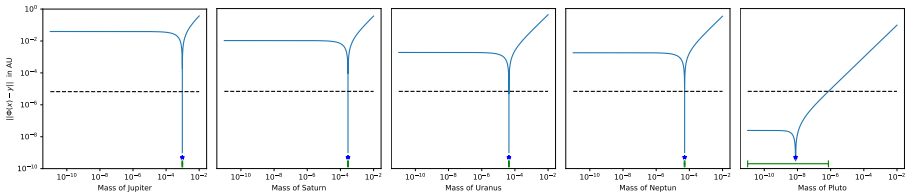


Figure 8: Profile likelihood on the solar system problem. The blue stars represent the true values of the parameters. Green segments represent confident interval. Black dashed lines represent the error threshold ε. Profile likelihood detects the Pluto's mass non-identifiability.

Modeling We consider the classical image formation model

$$y = k(\theta) * x + b, \tag{30}$$

where $x \in \mathbb{R}^M$ denotes the observed sample, $*$ denotes the convolution product, $b \in \mathbb{R}^M$ is a noise vector, and $k : \mathbb{R}^N \to \mathbb{R}^M$ represents the point spread function (PSF) of the system. It depends on parameters $\theta \in \mathbb{R}^N$. The scalar theory of diffraction [19], informs us that the PSF $k(\theta)$ can be characterized by its pupil function; a two-dimensional function supported on a disk whose radius depends on the numerical aperture of the objective and the wavelength of the collected light. A common practice is to expand this pupil function on a truncated Zernike polynomial basis [19]. With this model, only a few Zernike polynomials are sufficient to describe a rich class of realistic PSF that includes typical aberrations. Hence, in our model (30), $\theta \in \mathbb{R}^N$ denotes the vector of Zernike coefficients. We fix $N = 8$ in this work, corresponding to the following optical aberrations: defocus, vertical and oblique astigmatism, trefoil and coma and primary spherical. The blind inverse problem consists in estimating both θ and x from y.

To that end, we follow the approach proposed by [20]. We first train a reconstruction mapping $R(\theta, y)$ which computes an estimate $\hat{x}(\theta)$ such that the pair $(\theta, \hat{x}(\theta))$ is coherent with the data y. In this experiment, R is a deep plug-and-play image restoration method (DPIR). It corresponds to a half-quadratic splitting algorithm where the proximal step is replaced by a DRUNet denoiser with pre-trained weights [60]. Equipped with this (non-blind) reconstruction mapping R, we want to find the pair $(\theta, \hat{x}(\theta))$ that best fits the data y. That is, we want to minimize $\|k(\theta) * R(\theta, y) - y\|_2^2$. Using a Bayesian formalism, this means that we are trying to find the maximum a posteriori pair $(\hat{\theta}, \hat{x}(\hat{\theta}))$. This blind deblurring problem can be cast in our framework by using the following forward mapping

$$\Phi : \theta \mapsto k(\theta) * R(\theta, y). \tag{31}$$

In the first row of Fig. 9 we display the input image (a) as well as the noiseless measurements (b) we generated for this experiment.

Uncertainty with Jackpot Given y, we first estimate the parameters $\theta^\star$ by solving

$$\theta^\star = \operatorname*{argmin}_{\theta \in \mathbb{R}^N} \frac{1}{2} \|\Phi(\theta) - y\|_2^2 \tag{32}$$

using a gradient descent with automatic differentiation. In Fig. 9 (c) and (d) we show the associated PSF $k(\theta^\star)$ and deconvolved image $\hat{x}(\theta^\star) = R(\theta^\star, y)$. Note that the image seems significantly better resolved than the observation, which suggests that reconstruction mapping R is constructed carefully.

Then, we run Algorithm 1. The singular values of the Jacobian of Φ and the two last singular vectors are displayed in Fig. 9 (e) and (f), respectively. As opposed to the previous experiment, we see no clear gap in the amplitude of the last singular values. The singular vectors are clearly mixing the different Zernike coefficients, suggesting that the profile likelihood methods will struggle indicating the main directions of uncertainty.

We then compute a $D=2$-dimensional manifold approximation $\mathcal{M}_\delta^\varepsilon$ of the uncertainty region with the Jackpot method. We obtain a parameterization of the Zernike coefficients vector $\theta(z)$ for $z \in \mathbb{R}^2$ with $\theta(0) = \theta^\star$. In Fig. 9 (g), we report the SNR between the outputs $\Phi(\theta(z))$ and $\Phi(\theta^\star)$. The level line of 40dBs is displayed in green.

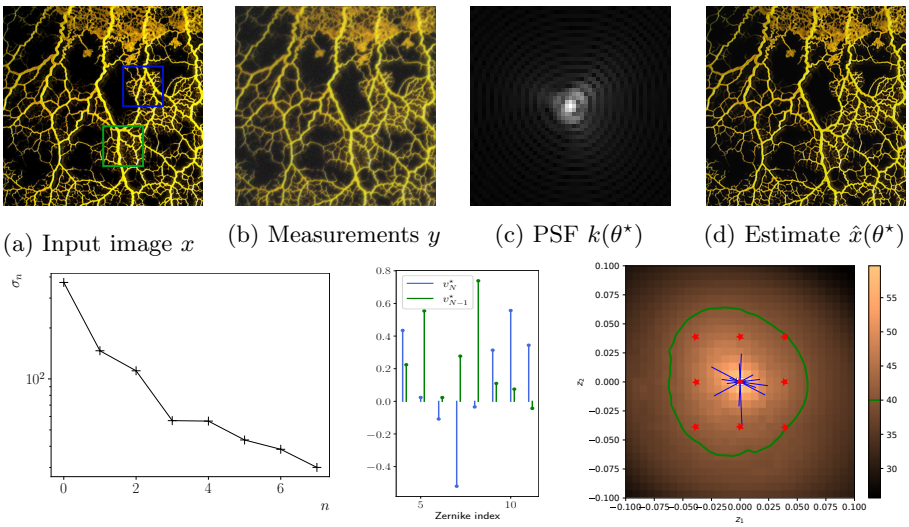
In Fig. 11 we show how the profile likelihood behaves on this problem. The confidence intervals obtained with the profile likelihood method are also displayed as blue segments in Fig. 9 (g). They have been projected on the tangent plane $\mathcal{T}_{\theta^\star}$ for comparison. As can be seen from this plot, the profile likelihood method significantly underestimates the uncertainty domain, since the blue segments tips are far from the level line boundaries.

In Fig. 10, we provide some zoomed regions of the deblurred images in (a) and of the corresponding PSF in (b) for different parameters $\theta(z) \in \mathcal{M}_\delta^\varepsilon$. The corresponding parameters z are shown as red stars in Fig. 9 (g). We see that it is possible to significantly deform the PSFs with a negligible modification of the measurements. We can also observe on Fig.s 10 (a) and (b) that some structures of the sample can be hallucinated without affecting the measurements too much. Overall, this shows that the studied blind deblurring process is unstable, though it produces sharp and nice looking images for all parameters.

Finally, let us mention that the Monte Carlo sampling method strongly underestimates uncertainty regions as illustrated in Fig. 12. The reason is related to the fact that we wish to recover a relatively low dimensional manifold (less than 8), but add noise on a high dimensional space of color images ($200 \times 200 \times 3$).

3.3 Posterior exploration for an image deblurring problem

In contrast to the previous section, we tackle regular inverse problems where the forward operator A is known perfectly.



(a) Input image x (b) Measurements y (c) PSF $k(\theta^\star)$ (d) Estimate $\hat{x}(\theta^\star)$

(e) Jacobian singular values (f) Singular vectors (g) Output SNR in dB

Figure 9: Blind delurring experiment. The reconstructed image $\hat{x}(\theta^\star) = R(\theta^\star, y)$ in (d) seems significantly sharper than the observation y. There is no clear gap for the lowest singular values in (e), suggesting that the uncertainty domain has a rather large intrinsic dimension. In (f), the Zernike coefficients of the two last singular vectors of $J^\star$ are mixed, meaning that the main uncertainty directions are not aligned with pure optical abberations. In (g), we evaluated the function $\mathrm{SNR}(\Phi(\theta), \Phi(\theta^\star))$ for $\theta \in \mathcal{M}_\delta^\varepsilon$. The level line of 40dB, corresponding to a high fidelity is displayed as a green curve. We also projected the confidence intervals obtained with the profile likelihood method as blue segments.

Formalism We explore the problem described in Section 1.1.3. Given $y = Ax + b$, with $b \sim \mathcal{N}(0, \sigma^2 \mathrm{Id})$, we construct the MAP estimator

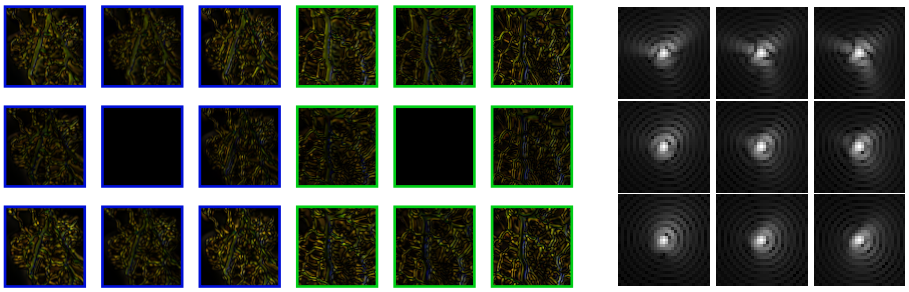
$$\begin{aligned}\hat{x}_{\mathrm{MAP}} &\stackrel{\mathrm{def}}{=} \operatorname*{argmin}_{x \in \mathbb{R}^N} -\log p_{\mathtt{x}|\mathtt{y}}(x|y) \\ &= \operatorname*{argmin}_{x \in \mathbb{R}^N} \frac{1}{2\sigma^2} \|Ax - y\|_2^2 - \log p_{\mathtt{x}}(x).\end{aligned}$$

By first order optimality conditions, $\hat{x}_{\mathrm{MAP}}$ satisfies

$$\nabla \log p_{\mathtt{x}|\mathtt{y}}(\hat{x}_{\mathrm{MAP}}|y) = \frac{1}{\sigma^2} A^T (A\hat{x}_{\mathrm{MAP}} - y) - \nabla \log p_{\mathtt{x}}(\hat{x}_{\mathrm{MAP}}) = 0. \tag{33}$$

The term $\nabla \log p_{\mathtt{x}}(\hat{x}_{\mathrm{MAP}})$ can be estimated using a learned denoiser and Tweedie formula [25]. In this paper, we use the denoiser proposed by [27], which enables to compute both the log likelihood and its gradient.

Following Section 1.1.3, we set $\Phi = -\nabla \log p_{\mathtt{x}|\mathtt{y}}(\cdot|y)$. The uncertainty region $\mathcal{S}_\delta^\varepsilon$ now can be interpreted as the set of images with smallest gradient amplitude.



(a) Zoom on blue and green regions in Fig. 9 (a)

(b) PSF $k(\theta(z))$

Figure 10: *Blind deblurring identifiability* - Grid sampling of the 2-dimensional manifold $\mathcal{M}_\delta^\varepsilon$. (a-b) Grid sampling of image differences $|\hat{x}(\theta(z)) - x^\star|$. (c) PSF grid-sampling $k(\theta(z))$.

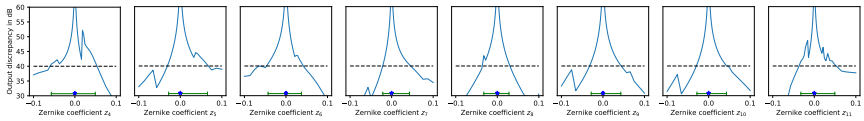


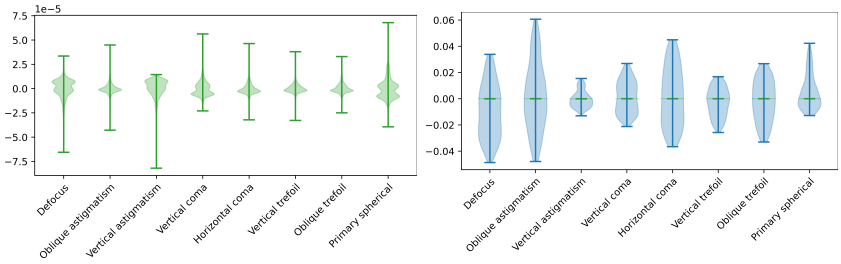
Figure 11: Profile likelihood on blind inverse problem. The blue stars represent the true values of the parameters. The green segments represent confidence intervals. The two first coefficients seem less identifiable than the other from this analysis.

We compute $\hat{x}_{\text{MAP}}$ using a gradient descent run for many iterations to reach a near 0 gradient.

Jackpot results We treat the 200×200 color image shown on Fig. 13d. The forward operator A is defined as a convolution with the PSF shown on Fig. 13a. Its discrete Fourier transform is shown on Fig. 13b. It possesses only two zeros (green crosses). Considering the Fourier transform symmetries for real signals, it means that the kernel is 2D for each color channel and therefore $\dim(\ker(A)) = 6$. This deblurring problem can be considered as mildly ill-posed.

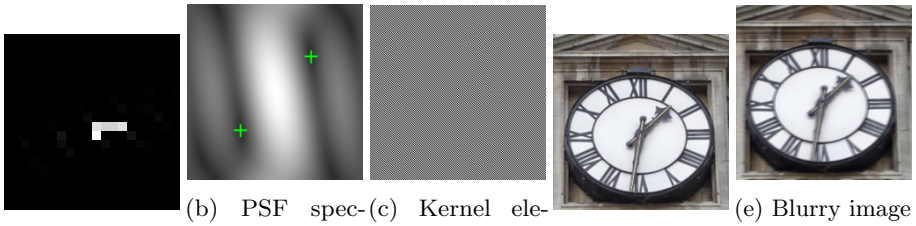
The resulting blurry image and recovered images are shown on Fig. 13e and 14a respectively. The five last singular values of the Jacobian matrix $J^\star$ are displayed in Fig. 14d. Surprisingly, they all possess roughly the same amplitude. We could expect that the only directions of uncertainty are unlocalized patterns such as the one in Fig. 13c. Looking closer at the associated singular vectors in Fig. 14e, we see that they all consist of localized and highly oscillatory patterns. The reason why the uncertainty patterns can be localized is that we do not constrain $Ax = y$, but just promote it through the data term $\frac{1}{2\sigma^2}\|Ax - y\|_2^2$, leaving space for more uncertainty.

In this examples, profile-likelihood methods are unusable due to the too large number of parameters ($3 \times 200 \times 200$). We propose a comparison with Monte



(a) Monte Carlo sampling (b) Jackpot

Figure 12: Violinplots of the Zernike coefficients obtained using a Monte Carlo sampling with 1000 points (left) and the Jackpot algorithm (right) for the blind deblurring experiment. Observe that the amplitudes obtained with Monte Carlo sampling are about 10^3 times lower than those obtained with Jackpot, illustrating the large discrepancy between natural and adversarial perturbations.



(a) PSF (b) PSF spectrum (c) Kernel element (d) Reference (e) Blurry image y

Figure 13: The deblurring problem. To illustrate Jackpot, we designed a specific PSF with a highly oscillatory 2D kernel (null-space). An element is shown in (c).

Carlo sampling in Fig. 15. After drawing 1000 samples, we display maximal pixel-wise distance to $x^\star$ both for the Monte Carlo and the Jackpot methods. As can be seen, the Monte Carlo method significantly underestimates the size of the uncertainty region. This again illustrates the difference between the effect of natural (Monte Carlo) and adversarial (Jackpot) perturbations.

A Proofs

A.1 Linear Approximation

A.1.1 Proof of Proposition 4

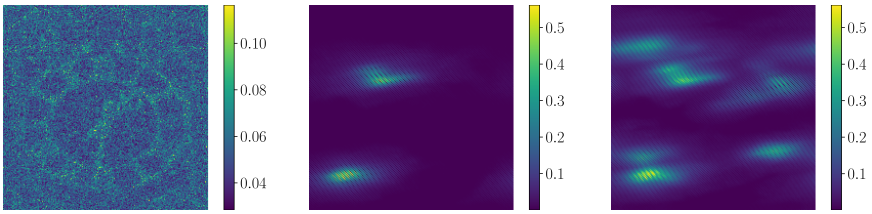
Defining the ellipsoid $\mathcal{E} \overset{\text{def}}{=} \{x \in \mathbb{R}^N \; ; \; \left\|\Sigma^\star V^{\star,T}(x - x^\star)\right\|_2^2 \leq \varepsilon^2\}$ we can reformulate

$$\widetilde{\mathcal{S}}_\delta^\varepsilon = \mathcal{E} \cap \mathcal{B}(x^\star, \delta) \text{ and } \widetilde{\mathcal{M}}_\delta^\varepsilon = \mathcal{E} \cap \mathcal{B}(x^\star, \delta) \cap \mathcal{T}_{x^\star}. \tag{34}$$

(a) Recovered $x^\star$ (b) Jackpot $v_N^\star$ (c) Jackpot $v_{N-1}^\star$ (d) Singular values

(e) The last 5 leftmost eigenvectors of $J^\star$ for the deblurring inverse problem.

Figure 14: Recovered image $x^\star$, 2 images recovered by Jackpot, the singular spectrum and 5 leftmost eigenpairs of $J^\star$.



(a) Monte Carlo sampling (b) Jackpot with $D = 2$ (c) Jackpot with $D = 10$

Figure 15: Comparison between Monte Carlo sampling and Jackpot for image deblurring. The pixel-wise maximum distance with $x^\star$ is displayed for 1000 Monte Carlo points in (a) and for 20 Jackpot points and $D = 2$ in (b) and 100 Jackpot points and $D = 10$ in (c). Notice the different amplitude ranges.

As $\widetilde{\mathcal{M}}_\delta^\varepsilon = \widetilde{\mathcal{S}}_\delta^\varepsilon \cap \mathcal{T}_{x^\star} \subseteq \widetilde{\mathcal{S}}_\delta^\varepsilon$, the Hausdorff distance is given by $d_{\mathcal{H}}(\widetilde{\mathcal{M}}_\delta^\varepsilon, \widetilde{\mathcal{S}}_\delta^\varepsilon) = \sup_{x \in \widetilde{\mathcal{S}}_\delta^\varepsilon} d(x, \widetilde{\mathcal{M}}_\delta^\varepsilon)$. Without considering the constraint $x \in \mathcal{B}(x^\star, \delta)$, we get the same for $\mathcal{E}$ and $\mathcal{E} \cap \mathcal{T}_{x^\star}$

$$d_{\mathcal{H}}(\mathcal{E} \cap \mathcal{T}_{x^\star}, \mathcal{E}) = \sup_{x \in \mathcal{E}} d(x, \mathcal{E} \cap \mathcal{T}_{x^\star}). \tag{35}$$

Now, we can consider the following equivalent optimization problem

$$\sup_{x \in \mathcal{E}} \frac{1}{2} \left\| (\mathrm{Id}_N - \Pi_{\mathcal{V}_D^\star})(x - x^\star) \right\|_2^2. \tag{36}$$

as $d(x, \mathcal{E} \cap \mathcal{T}_{x^\star})^2 = \left\| (\mathrm{Id}_N - \Pi_{\mathcal{V}_D^\star})(x - x^\star) \right\|_2^2$ where $\mathcal{V}_D^\star$ is defined in (13). The associated Lagrangian is given by

$$\mathcal{L}(\lambda, x) = \frac{1}{2} \left\| (\mathrm{Id}_N - \Pi_{\mathcal{V}_D^\star})(x - x^\star) \right\|_2^2 - \frac{\lambda}{2} \left(\left\| \Sigma^\star V^{\star,T}(x - x^\star) \right\|_2^2 - \varepsilon^2 \right). \tag{37}$$

And its gradient is

$$\nabla \mathcal{L}(\lambda, x) = \begin{pmatrix} \sum_{n=1}^{N} s_n^2 (\sigma_n^\star)^2 - \varepsilon^2 \\ \sum_{n=1}^{N-D} s_n v_n^\star - \lambda \sum_{n=1}^{N} (\sigma_n^\star)^2 s_n v_n^\star \end{pmatrix} \tag{38}$$

where we denote the coordinates $x - x^\star = \sum_{n=1}^{N} s_n v_n^\star$. The optimum is achieved where the gradient vanishes, that is

$$\sum_{n=1}^{N} s_n^2 (\sigma_n^\star)^2 = \varepsilon^2 \;;\; \begin{cases} s_n = 0, & \forall n \geq N - D + 1 \\ s_n(1 + \lambda \sigma_n^\star) = 0, & \forall n \leq N - D \end{cases}.$$

From this, only one s_n is non zero. Among the $N - D$ local optima, only the following is global as long as $\sigma_{N-D}^\star < \sigma_{N-D-1}^\star$ (otherwise, it is still a global optimum but not unique)

$$s_n = 0, \; \forall n \neq N - D \;;\; s_{N-D} = \frac{\varepsilon}{\sigma_{N-D}^\star} \text{ and } \lambda = \frac{1}{\sigma_{N-D}^\star}. \tag{39}$$

Finally let's add the constraint $\|x - x^\star\|_2 \leq \delta$. If $\delta > \frac{\varepsilon}{\sigma_{N-D}^\star}$, then the supremum is achieved at $x^\star + \frac{\varepsilon}{\sigma_N^\star} v_{N-D}^\star$. Otherwise, $\delta < \frac{\varepsilon}{\sigma_{N-D}^\star}$ and the supremum is achieved at $x^\star + \delta v_{N-D}^\star$. In both cases, the Hausdorff distance verifies (15).

A.2 Nonlinear Approximation

All proof are provided in a more general setting with the parametrization

$$\gamma(z) \stackrel{\text{def}}{=} \operatorname*{argmin}_{A(x - x^\star) = z} \frac{1}{2} \|\Phi(x) - \Phi(x^\star)\|_2^2, \tag{$\mathcal{P}_A$}$$

where $A \in \mathbb{R}^{D \times N}$ with $1 \leq D \leq N$ satisfies the following assumption.

Assumption 14. The matrix A verifies $AA^T = \mathrm{Id}_D$ and the matrix $\Pi_{\mathrm{im} A^T} + \Pi_{\ker A} \cdot J_\Phi(x^\star)^T J_\Phi(x^\star)$ is invertible. We denote the projections operators on the image and kernel of A as $\Pi_{\mathrm{im} A^T} \stackrel{\text{def}}{=} A^T A$ and $\Pi_{\ker A} \stackrel{\text{def}}{=} \mathrm{Id}_N - A^T A$ respectively.

The results provided in the main text are specific instances of the results reported below with $A = V_D^{\star,T}$ for which Assumption 14 holds.

Lemma 15. *Assumption 14 holds true for $A = V_D^{\star,T}$.*

Proof. With $A = V_D^{\star,T}$, the matrix

$$\Pi_{\mathrm{im} A^T} + \Pi_{\ker A} \cdot J_\Phi(x^\star)^T J_\Phi(x^\star) = V_D^\star \, \mathrm{Diag}\left(\sigma_1^2, \cdots, \sigma_{N-D}^2, 1, \cdots, 1\right) V_D^{\star,T}$$

is invertible by construction of $V_D^{\star,T}$. □

A.2.1 Proof of Theorem 6

Theorem 16. *Let $\Phi : \mathbb{R}^N \to \mathbb{R}^M$ be a $\mathcal{C}^1$ map, $x^\star \in \mathbb{R}^N$, and $A \in \mathbb{R}^{D\times N}$ satisfying Assumption 14. Then, there exists an open neighborhood $\mathcal{U} \subseteq \mathbb{R}^D$ such that $(\mathcal{P}_A)$ admits a unique solution $\gamma : \mathcal{U} \to \mathbb{R}^N$ verifying $\gamma(0_D) = x^\star$. Moreover, it has the following properties*

- *If Φ is of class $\mathcal{C}^2$, then γ is of class $\mathcal{C}^1$. Moreover, its Jacobian is given by*

$$J_\gamma(z) = \left[\Pi_{\mathrm{im}A^T} + \Pi_{\ker A} \cdot \nabla^2 F(\gamma(z))\right]^{-1} A^T. \tag{40}$$

- *If Φ is of class $\mathcal{C}^1$ and J_Φ is locally Lipschitz and definable, then γ is a locally Lipschitz definable function. Moreover, we can express a conservative Jacobian for γ as*

$$\tilde{J}_\gamma : z \rightrightarrows \left\{ \left[\Pi_{\mathrm{im}A^T} + \Pi_{\ker A} \cdot B\right]^{-1} A^T \ ; \ B \in \tilde{J}_{\nabla F}(\gamma(z)) \right\}. \tag{41}$$

where $F(x) = \frac{1}{2}\left\|\Phi(x) - \Phi(x^\star)\right\|_2^2$ denotes the objective function of $(\mathcal{P}_A)$.

Proof. Starting from the Karush-Kuhn-Tucker (KKT) conditions of Problem $(\mathcal{P}_A)$, the idea of the proof is to apply an implicit function theorem on the dual formulation to get existence of solutions for both on primal and dual problems. We moreover derive an explicit expression of the (conservative) Jacobian of γ.

From constraint optimization theory, the Lagrangian function associated to $(\mathcal{P}_A)$ reads

$$L_z(\lambda, x) \overset{\text{def}}{=} F(x) + \lambda^T \left(A(x - x^\star) - z\right).$$

where $\lambda \in \mathbb{R}^D$ is the vector of Lagrange multipliers and $F(x) = \frac{1}{2}\left\|\Phi(x) - \Phi(x^\star)\right\|_2^2$ is the objective function. As Φ is of class $\mathcal{C}^1$, and given that the constraint is linear, the KKT conditions reduce to

$$\nabla_{L_z}(\lambda, x) = \begin{bmatrix} A(x - x^\star) - z \\ \nabla F(x) + A^T\lambda \end{bmatrix} = 0_{D+N},$$

for $z \in \mathbb{R}^D$. In particular, multiplying the second line of ∇_{L_z} by A yields that $\lambda(z) = -A\nabla F(x)$ since $AA^T = \mathrm{Id}_D$.

Hence, finding a solution (λ, x) of the KKT conditions of $(\mathcal{P}_A)$ is equivalent to solving the implicit system

$$G(z, x) \overset{\text{def}}{=} \begin{bmatrix} A(x - x^\star) - z \\ (\mathrm{Id}_N - A^T A)\nabla F(x) \end{bmatrix} = 0_{D+N}.$$

Let's notice that G has values in $\mathbb{R}^{D+N}$ but its image spans only the N dimensional subspace $\mathbb{R}^D \times \mathrm{Ker}A$. Indeed, the matrix $(\mathrm{Id}_N - A^T A)$ is the projection onto the subspace $\ker A = (\mathrm{im}A^T)^\perp$. Thus one can reduce the output dimension of G to make it surjective while keeping the same zeros level sets. This is done

by left multiplying with $[A^T, \mathrm{Id}_N]$ which embeds the first term $A(x - x^\star) - z$ of G onto $\mathrm{im}A^T$ while keeping the second term in $(\mathrm{im}A^T)^\perp$. One obtains

$$H(z, x) \overset{\text{def}}{=} A^T(A(x - x^\star) - z) + (\mathrm{Id}_N - A^T A)\nabla F(x).$$

Our problem is now equivalent to the following implicit equation

$$H(z, \gamma(z)) = 0_N, \tag{42}$$

for $z \in \mathbb{R}^D$ and with $H : \mathbb{R}^D \times \mathbb{R}^N \to \mathbb{R}^N$ defined before.

Given that a solution of (43) verifies the KKT conditions, it thus is also a solution of the primal equation $(\mathcal{P}_A)$. We now use two forms of implicit function theorem to get existence of a solution of (43) and thus existence of a mapping γ for $(\mathcal{P}_A)$.

- $\mathcal{C}^2$ *case:* From implicit function theorem [37, Thm. 5.15], the only assumption required is the invertibility condition on the second sub-matrix $\mathcal{D}_2 H(0_D, x^\star)$ of the differential $\mathcal{D}H = [\mathcal{D}_1 H, \mathcal{D}_2 H]$ of H at $(0_D, x^\star)$. With some computations, one exactly recovers the second condition in Assumption 14. Then as a consequence, the Jacobian of γ is given by

  $$J_\gamma(z) = -\mathcal{D}_2 H(z, \gamma(z))^{-1} \mathcal{D}_1 H(z, \gamma(z))$$

  from which we recover the formula (41).

- $\mathcal{C}^1$ *case:* We use the Lipschitz definable version of implicit function theorem from [8, Thm. 1]. As J_Φ is locally Lipschitz and definable, so is F and thus H. With $\gamma(0_D) = x^\star$, one recovers $H(0_D, \gamma(0_D)) = 0_N$. A conservative Jacobian of H can be written as

  $$\tilde{J}_H : (z, x) \rightrightarrows \left\{ \left[-A^T, \Pi_{\mathrm{im}A^T} + \Pi_{\ker A} \cdot B \right] \ ; \ B \in \tilde{J}_{\nabla F}(\gamma(z)) \right\}. \tag{43}$$

  Since $\nabla F(x) = J_\Phi(x)^T(\Phi(x) - \Phi(x^\star))$, by path-differentiation rules, one derives the formula

  $$\tilde{J}_{\nabla F}(x) = \left\{ B \ (\Phi(x) - \Phi(x^\star)) + J_\Phi(x)^T J_\Phi(x) \ ; \ B \in \tilde{J}_{J_\Phi^T}(x) \right\}$$

  and evaluating in $x = x^\star$ provides, since the first term vanishes,

  $$\tilde{J}_{\nabla F}(x^\star) = \left\{ J_\Phi(x^\star)^T J_\Phi(x^\star) \right\}. \tag{44}$$

  Finally inserting (45) in (44) allows us to express

  $$\tilde{J}_H(0_D, x^\star) = \left\{ \left[-A^T, \Pi_{\mathrm{im}A^T} + \Pi_{\ker A} \cdot J_\Phi(x^\star)^T J_\Phi(x^\star) \right] \right\}.$$

  This last equation implies that the assumptions of [8, Thm. 1] are verified: $\tilde{J}_H(0_D, x^\star)$ is a singleton and thus convex and its second sub-matrix is invertible from Assumption 14. This completes the proof.

□

A.2.2 Proof of Proposition 9 and Theorem 10

Assumption 8, Proposition 9 and Theorem 10 are specific instances of Assumption 17, Proposition 18 and Theorem 19 below with $A = V_D^{\star,T}$.

Assumption 17 (Controlled thickness)**.** Let

$$\mathcal{P}_z = \{x \in \mathcal{S}_\delta^\varepsilon, A(x - x^\star) = z\}$$

denote the slice of $\mathcal{S}_\delta^\varepsilon$ which has coordinate z on $\mathcal{T}_{x^\star}$. We assume that

$$\sup_{z:\|\gamma(z)-x^\star\|_2 \leq \delta} \operatorname{diam}(\mathcal{P}_z) \leq \eta. \tag{45}$$

Proposition 18 (Uniformly bounded curvature)**.** *Let Φ be of class $\mathcal{C}^2$ and assume that the Hessian tensor H_Φ satisfies*

$$\|H_\Phi\|_{2,\delta} \stackrel{\text{def}}{=} \sup_{x,x',x'' \in \mathcal{B}(x^\star,\delta)} \frac{\|H_\Phi(x)(x',x'')\|_2}{\|x'\|_2 \|x''\|_2} < \frac{\sigma_{\min}}{\delta}$$

where σ_{min} is the lowest singular value of the Jacobian $J_\Phi(x^\star)$ restricted to the kernel of A. Then Assumption 17 is verified with

$$\eta = \frac{2\varepsilon}{\sigma_{min} - \delta\|H_\Phi\|_{2,\delta}}. \tag{46}$$

Proof. Setting $z \in \mathcal{B}(0_D, \delta)$, let's upper bound the diameter of the intersection of $\mathcal{S}_\delta^\varepsilon$ with the affine hyperplane $A(\cdot - x^\star) = z$. When this slice is empty there is nothing to do, and when it is not, it only remains to bound the norm $\|x_1 - x_2\|_2$ for any couple $x_1, x_2 \in \mathcal{S}_\delta^\varepsilon$ verifying $A(x_1 - x^\star) = A(x_2 - x^\star) = z$.

To that end, one can remark that

$$\|\Phi(x_1) - \Phi(x_2)\|_2 \leq \|\Phi(x_1) - \Phi(x^\star)\|_2 + \|\Phi(x^\star) - \Phi(x_2)\|_2 \leq 2\varepsilon$$

as x_1 and x_2 belong to $\mathcal{S}_\delta^\varepsilon$. The remaining of the proof consists in lower bounding the left-hand side by a term of the form $\mu \|x_1 - x_2\|_2$ with some $\mu > 0$.

Applying Taylor expansion formula on Φ and J_Φ, one gets that there exists a point c lying on the open segment between x_1 and x_2, (i.e., $c = \alpha x_1 + (1-\alpha)x_2$ for $\alpha \in (0,1)$), and a point c' lying on the open segment between c and $x^\star$ (i.e., $c' = \alpha c + (1-\alpha)x^\star$ for $\alpha \in (0,1)$) such that

$$\begin{aligned}\Phi(x_1) - \Phi(x_2) &= J_\Phi(c)(x_1 - x_2) \\ &= J_\Phi(x^\star)(x_1 - x_2) + H_\Phi(c')(c - x^\star)(x_1 - x_2).\end{aligned}$$

Then, from the inverse triangular inequality, we get

$$\|\Phi(x_1) - \Phi(x_2)\|_2 \geq \left| \|J_\Phi(x^\star)(x_1 - x_2)\|_2 - \|H_\Phi(c')(c - x^\star)(x_1 - x_2)\|_2 \right|. \tag{47}$$

The first right term of (48) can be handled noticing that the difference $x_1 - x_2$ lies in the kernel of A since $A(x_1 - x_2) = A(x_1 - x^\star) + A(x^\star - x_2) = z - z = 0$.

Thus by definition of the minimal singular value of $J_\Phi(x^\star)$ restricted to the kernel of A, denoted $\sigma_{\min}$, it follows that

$$\|J_\Phi(x^\star)(x_1 - x_2)\|_2 \geq \sigma_{\min} \|x_1 - x_2\|_2 .$$

Moreover, the second right term of (48) can be bounded using Hessian norm. Indeed, since $x^\star$, x_1 and x_2 belong to the ball $\mathcal{B}(x^\star, \delta)$, the points c and c' are also in $\mathcal{B}(x^\star, \delta)$ by construction. It follows that

$$\|H_\Phi(c')(c - x^\star)(x_1 - x_2)\|_2 \leq \|H_\Phi\|_{2,\delta} \cdot \delta \cdot \|x_1 - x_2\|_2 .$$

And under the condition $\|H_\Phi\|_{2,\delta} \cdot \delta < \sigma_{\min}$, equation (48) becomes

$$\|\Phi(x_1) - \Phi(x_2)\|_2 \geq (\sigma_{\min} - \|H_\Phi\|_{2,\delta} \cdot \delta) \|x_1 - x_2\|_2$$

as intended. □

Theorem 19. *Let $\Phi : \mathbb{R}^N \to \mathbb{R}^M$ be a function of class $\mathcal{C}^2$, $x^\star \in \mathbb{R}^N$ be a point, $A \in \mathbb{R}^{D \times N}$ satisfying Assumption 14, and let $\varepsilon > 0$ be a threshold. Then $\mathcal{M}_\delta^\varepsilon$ defines a manifold with parameterization γ in $(\mathcal{P}_A)$ for sufficiently small $\delta > 0$. Moreover, under Assumption 17, the manifold $\mathcal{M}_\delta^\varepsilon$ is a good approximation of $\mathcal{S}_\delta^\varepsilon$ in the sense that:*

$$d_{\mathcal{H}}(\mathcal{M}_\delta^\varepsilon,\ \mathcal{S}_\delta^\varepsilon) < \eta, \tag{48}$$

where the Hausdorff distance $d_{\mathcal{H}}$ is defined in (2). More precisely, we have the following inclusions:

$$\mathcal{M}_\delta^\varepsilon \subseteq \mathcal{S}_\delta^\varepsilon \subseteq \mathcal{M}_\delta^\varepsilon + \mathcal{B}(0, \eta). \tag{49}$$

Proof. First of all, from Theorem 16, the set $\mathcal{M}_\delta^\varepsilon$ is well defined for some $\delta > 0$ and by definition verifies the first inclusion $\mathcal{M}_\delta^\varepsilon \subseteq \mathcal{S}_\delta^\varepsilon$. Furthermore, using a standard result on smooth manifolds from [37], which is summarized in Proposition 20 below, the image of γ locally defines an embedded manifold since $J_\gamma(0)$ is injective as the product of an invertible matrix with A^T which is injective by assumption.

To prove the second inclusion, given a point $\tilde{x} \in \mathcal{S}_\delta^\varepsilon$, we need to find a point $x' \in \mathcal{M}_\delta^\varepsilon$ such that $\|\tilde{x} - x'\|_2 \leq \eta$. A natural candidate is $x' \stackrel{\text{def}}{=} \gamma(\tilde{z})$ with $\tilde{z} \stackrel{\text{def}}{=} A(\tilde{x} - x^\star)$. Indeed, given that both $\tilde{x}$ and x' are in $\mathcal{S}_\delta^\varepsilon$ and (by definition) verify $A(\tilde{x} - x^\star) = A(x' - x^\star) = \tilde{z}$, Assumption 17 implies that $\|\tilde{x} - x'\|_2 \leq \eta$. Moreover, we have

$$\|\Phi(\gamma(\tilde{z})) - \Phi(x^\star)\|_2 \leq \|\Phi(\tilde{x}) - \Phi(x^\star)\|_2 \leq \varepsilon, \tag{50}$$

and thus $x' = \gamma(\tilde{z}) \in \mathcal{M}_\delta^\varepsilon$. Note that the first inequality in (51) comes from the fact that $\gamma(\tilde{z})$ solves $(\mathcal{P}_A)$ while $\tilde{x}$ is an admissible point. The second one is due to the fact that $\tilde{x} \in \mathcal{S}_\delta^\varepsilon$. This completes the proof. □

Proposition 20 (Injective Jacobian implies a local structure of manifold)**.** *Let $F : \mathbb{R}^N \to \mathbb{R}^M$ be a function of class $\mathcal{C}^1$ and $x^\star \in \mathbb{R}^N$. Assume that $J_F(x^\star)$ is injective. Then there exists a neighborhood $\mathcal{U} \subset \mathbb{R}^N$ of $x^\star$ such that $F(\mathcal{U})$ is an embedding submanifold of $\mathbb{R}^N$.*

Proof. First of all, J_F is locally injective near $x^\star$ because of the lower semi-continuity of the rank function. Indeed, the set of points on which J_F is injective can be reformulated as the inverse image of the open set $\mathbb{R}^*$ under the continuous map $\det(J_F^T J_F)$ and thus is open.

From rank theorem [37, Thm. 5.13], up to diffeomorphisms, F is locally equal to its Jacobian which is locally injective, so is F. We actually get that F is locally an immersion.

From [37, Lem. 5.34], since F is locally an immersion, F is also locally an embedding on a neighborhood $\mathcal{U}$.

Finally from [37, Thm. 5.31], $F(\mathcal{U})$ is an embedded submanifold as the image of a smooth embedding, which concludes the proof. □

A.3 Numerical computation

A.3.1 Proof of Proposition 12

Proof. Proof of Proposition 12. In $(\mathcal{P}_z)$, splitting the vector $x - x^\star \in \mathbb{R}^N$ within the space decomposition $\mathbb{R}^N = \mathrm{Im}V_D^\star \oplus (\mathrm{Im}V_D^\star)^\perp$ yields the reformulation $(\mathcal{P'}_z)$. Here the constraint $V_D^{\star,T}(x - x^\star) = z$ is included in the objective function that we denote

$$F_z : x \mapsto \frac{1}{2} \left\| \Phi(r(z,x)) - \Phi(x^\star) \right\|_2^2 \tag{51}$$

where $r(z,x) \overset{\text{def}}{=} x^\star + V_D^\star z + \Pi_\perp x$.

The main strategy to locally prove the linear convergence of gradient descent is to show a geometrical property on F_z. Since it is not locally strongly convex we use some *Polyak-Lojasiewicz (PL) condition*:

Definition 21. For a real number $\mu > 0$, a function f is said to verify the *μ-Lojasiewicz inequality* if for all x,

$$\frac{1}{2} \left\| \nabla f(x) \right\|_2^2 \geq \mu (f(x) - f(x^\star)). \tag{52}$$

The argumentation will proceed according to the following four steps, where the results are verified locally around $(z,x) = (0_D, 0_N)$:

1. The restricted map $\tilde{F}_z \overset{\text{def}}{=} F_z|_{(\mathrm{Im}V_D^\star)^\perp}$ on $(\mathrm{Im}V_D^\star)^\perp$ is strongly convex.

2. The map $\tilde{F}_z$ satisfies a PL inequality since any strongly convex map also satisfies a PL inequality [48, lemma].

3. Since F_z only depends on the orthogonal part $(\mathrm{Im}V_D^\star)^\perp$, i.e. $F_z(x) = F_z(x')$ whenever $\Pi_\perp x = \Pi_\perp x'$, the objective function F_z also verifies the same PL inequality.

4. Since F_z verifies a PL inequality and is also L-smooth for some $L > 0$ (since of class $\mathcal{C}^1$ on a neighbor set), this assures the linear convergence of the gradient descent [48, thm. 4].

The only remaining argument is to prove the first point : the local strong convexity of the objective function. The Hessian of F_z at x is

$$\begin{aligned}\nabla^2 F_z(x) =& \Pi_\perp J_\Phi^T(r(z,x)) J_\Phi(r(z,x)) \Pi_\perp \\ &+ \Pi_\perp H_\Phi^T(r(z,x))(\Phi(r(z,x)) - \Phi(x^\star)) \in \mathbb{R}^{N\times N}.\end{aligned}$$

Evaluating at $(z,x) = (0_D, 0_N)$ gives

$$H^\star \stackrel{\text{def}}{=} \Pi_\perp J^{\star,T} J^\star \Pi_\perp \in \mathbb{R}^{N\times N}.$$

Restricted to the subspace $(\mathrm{Im} V_D^\star)^\perp$, this matrix is a positive-definite matrix with eigenvalues equal to $\sigma_1^2 \geq \cdots \geq \sigma_{N-D}^2 > 0$ since $D > R^\star$. Moreover, this submatrix corresponds to the Hessian of the map $\tilde{F}_0$ at $x = 0_N$. By smoothness of Φ, the map $(z,x) \mapsto \nabla^2 \tilde{F}_z(x)$ is continuous and thus the matrices $\nabla^2 \tilde{F}_z(x)$ remain locally positive-definite matrix for (z,x) on a neigborhood $\mathcal{U} \times \mathcal{V}$ of $(0_D, 0_N)$. This property is a characterization of the strong convexity of the maps $\tilde{F}_z$ retricted to $\mathcal{V}$ for all $z \in \mathcal{U}$. □

A.3.2 Proof of Corollary 13

From Theorem 6, one knows that the parameterization function γ is locally Lipschitz. The key remaining argument is that any point z of the ball $\mathcal{B}_D(0,\delta)$ is at most at a distance of $\sqrt{D}s/2$ from a point z' of the grid Z_s. From Lipschitz consideration, $\|\gamma(z) - \gamma(z')\|_2 \leq L\sqrt{D}s/2$. This leads, using Theorem 19, the following inequalities

$$\begin{aligned}d_{\mathcal{H}}(\mathcal{S}_\delta^\varepsilon,\ \mathcal{M}_\delta^\varepsilon(Z)) &\leq d_{\mathcal{H}}(\mathcal{S}_\delta^\varepsilon,\ \mathcal{M}_\delta^\varepsilon) + d_{\mathcal{H}}(\mathcal{M}_\delta^\varepsilon,\ \mathcal{M}_\delta^\varepsilon(Z)) \\ &\leq \eta + L\sqrt{D}s/2,\end{aligned}$$

which completes the proof.

References

[1] P-A Absil, Robert Mahony, and Rodolphe Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2008.

[2] Vegard Antun, Francesco Renna, Clarice Poon, Ben Adcock, and Anders C Hansen. On instabilities of deep learning in image reconstruction and the potential costs of ai. *Proceedings of the National Academy of Sciences*, 117(48):30088–30095, 2020.

[3] Rafael Arutjunjan, Bjoern Malte Schaefer, and Clemens Kreutz. Constructing exact confidence regions on parameter manifolds of non-linear models. *arXiv preprint arXiv:2211.03421*, 2022.

[4] Johnathan M Bardsley. Mcmc-based image reconstruction with uncertainty quantification. *SIAM Journal on Scientific Computing*, 34(3):A1316–A1332, 2012.

[5] Johnathan M Bardsley, Antti Solonen, Heikki Haario, and Marko Laine. Randomize-then-optimize: A method for sampling from posterior distributions in nonlinear inverse problems. *SIAM Journal on Scientific Computing*, 36(4):A1895–A1910, 2014.

[6] Daniel J Bates, Jonathan D Hauenstein, and Nicolette Meshkat. Identifiability and numerical algebraic geometry. *Plos one*, 2019.

[7] Giuseppina Bellu, Maria Pia Saccomani, Stefania Audoly, and Leontina D'Angiò. Daisy: A new software tool to test global identifiability of biological and physiological systems. *Computer methods and programs in biomedicine*, 2007.

[8] Jérôme Bolte, Tam Le, Edouard Pauwels, and Tony Silveti-Falls. Nonsmooth implicit differentiation for machine-learning and optimization. *Advances in neural information processing systems*, 34:13537–13549, 2021.

[9] Joachim Bona-Pellissier, François Malgouyres, and François Bachoc. Local identifiability of deep relu neural networks: the theory. *Advances in Neural Information Processing Systems*, 35:27549–27562, 2022.

[10] John C. Butcher. *Numerical Methods for Ordinary Differential Equations*. John Wiley & Sons, 3rd edition, 2016.

[11] Jannis Chemseddine, Paul Hagemann, Christian Wald, and Gabriele Steidl. Conditional wasserstein distances with applications in bayesian ot flow matching. *arXiv preprint arXiv:2403.18705*, 2024.

[12] Hyungjin Chung, Jeongsol Kim, Michael T Mccann, Marc L Klasky, and Jong Chul Ye. Diffusion posterior sampling for general noisy inverse problems. *arXiv preprint arXiv:2209.14687*, 2022.

[13] Jed A Duersch, Meiyue Shao, Chao Yang, and Ming Gu. A robust and efficient implementation of lobpcg. *SIAM Journal on Scientific Computing*, 2018.

[14] Marisa C Eisenberg and Michael AL Hayashi. Determining identifiable parameter combinations using subset profiling. *Mathematical biosciences*, 2014.

[15] Juan L Fernandez Martinez, M Zulima Fernandez Muniz, and Michael J Tompkins. On the topography of the cost functional in linear and nonlinear inverse problems. *Geophysics*, 2012.

[16] Juan L Fernández-Martínez and Zulima Fernández-Muñiz. The curse of dimensionality in inverse problems. *Journal of Computational and Applied Mathematics*, 2020.

[17] Juan Luis Fernández-Martínez, Z Fernández-Muñiz, JLG Pallero, and Luis Mariano Pedruelo-González. From bayes to tarantola: new insights to understand uncertainty in inverse problems. *Journal of Applied Geophysics*, 98:62–72, 2013.

[18] Samuel M Fischer and Mark A Lewis. A robust and efficient algorithm to find profile likelihood confidence intervals. *Statistics and Computing*, 2021.

[19] Joseph W Goodman. *Introduction to Fourier optics*. Roberts and Company publishers, 2005.

[20] Alban Gossard and Pierre Weiss. Training adaptive reconstruction networks for blind inverse problems. *SIAM Journal on Imaging Sciences*, 17(2):1314–1346, 2024.

[21] Alexander V Grayver and Alexey V Kuvshinov. Exploring equivalence domain in nonlinear inverse problems using covariance matrix adaption evolution strategy (cmaes) and random sampling. *Geophysical Journal International*, 2016.

[22] Philipp Grohs, Sarah Koppensteiner, and Martin Rathmair. Phase retrieval: uniqueness and stability. *SIAM Review*, 62(2):301–350, 2020.

[23] Jon C Helton, Jay Dean Johnson, Cedric J Sallaberry, and Curt B Storlie. Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering & System Safety*, 2006.

[24] Stefan Hengl, Clemens Kreutz, Jens Timmer, and Thomas Maiwald. Data-based identifiability analysis of non-linear dynamical models. *bioinformatics*, 2007.

[25] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

[26] Evelyne Hubert. Essential components of an algebraic differential equation. *Journal of symbolic computation*, 1999.

[27] Samuel Hurault, Arthur Leclaire, and Nicolas Papadakis. Gradient step denoiser for convergent plug-and-play. *arXiv preprint arXiv:2110.03220*, 2021.

[28] Aapo Hyvärinen. Estimation of non-normalized statistical models by score matching. *J. Mach. Learn. Res.*, 6, dec 2005.

[29] IMCCE. position ephemeris, 2019.

[30] Kishore Jaganathan, Yonina C Eldar, and Babak Hassibi. Phase retrieval: An overview of recent developments. *Optical Compressive Imaging*, pages 279–312, 2016.

[31] Carine Jauberthie, Nathalie Verdière, and Louise Travé-Massuyès. Fault detection and identification relying on set-membership identifiability. *Annual Reviews in Control*, 2013.

[32] HK Johansen. A man/computer interpretation system for resistivity soundings over a horizontally strafified earth. *Geophysical Prospecting*, 1977.

[33] Michael Kech and Felix Krahmer. Optimal injectivity conditions for bilinear inverse problems with applications to identifiability of deconvolution problems. *SIAM Journal on Applied Algebra and Geometry*, 2017.

[34] Andrew V Knyazev. Toward the optimal preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method. *SIAM journal on scientific computing*, 2001.

[35] Nicholas N Lam, Paul D Docherty, and Rua Murray. Practical identifiability of parametrised models: A review of benefits and limitations of various approaches. *Mathematics and Computers in Simulation*, 2022.

[36] Rémi Laumont, Valentin De Bortoli, Andrés Almansa, Julie Delon, Alain Durmus, and Marcelo Pereyra. Bayesian imaging using plug & play priors: when langevin meets tweedie. *SIAM Journal on Imaging Sciences*, 15(2):701–737, 2022.

[37] John M. Lee. *Introduction to Smooth Manifolds*. Springer New York, 2012.

[38] Yanjun Li, Kiryung Lee, and Yoram Bresler. Identifiability in bilinear inverse problems with applications to subspace or sparsity-constrained blind gain and phase calibration. *IEEE Transactions on Information Theory*, 63(2):822–842, 2016.

[39] Erica Liu, Kelly Rivera, and Jordan Fox. Efficient eigensolvers and their applications. 2020.

[40] Benjamin Merkt, Jens Timmer, and Daniel Kaschek. Higher-order lie symmetries in identifiability and predictability analysis of dynamic models. *Physical Review E*, 92(1):012920, 2015.

[41] Hongyu Miao, Xiaohua Xia, Alan S Perelson, and Hulin Wu. On identifiability of nonlinear ode models and applications in viral dynamics. *SIAM review*, 2011.

[42] Klaus Mosegaard and Albert Tarantola. Monte carlo sampling of solutions to inverse problems. *Journal of Geophysical Research: Solid Earth*, 100(B7):12431–12447, 1995.

[43] Matthew J. Muckley and et al. Results of the 2020 fastmri challenge for machine learning mr image reconstruction. *IEEE Transactions on Medical Imaging*, 2021.

[44] Seungjun Nah, Sanghyun Son, Suyoung Lee, Radu Timofte, and Kyoung Mu Lee. Ntire 2021 challenge on image deblurring, 2021.

[45] Beresford N Parlett. *The symmetric eigenvalue problem*. SIAM, 1998.

[46] Mary Phuong and Christoph H Lampert. Functional vs. parametric equivalence of relu networks. In *International Conference on Learning Representations*, 2019.

[47] Allan Pinkus. *N-widths in Approximation Theory*, volume 7. Springer Science & Business Media, 2012.

[48] Boris T Polyak. Gradient methods for the minimisation of functionals. *USSR Computational Mathematics and Mathematical Physics*, 3(4):864–878, 1963.

[49] Dhruva V Raman, James Anderson, and Antonis Papachristodoulou. Delineating parameter unidentifiabilities in complex models. *Physical Review E*, 2017.

[50] Andreas Raue, Clemens Kreutz, Thomas Maiwald, Julie Bachmann, Marcel Schilling, Ursula Klingmüller, and Jens Timmer. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 2009.

[51] Werner C Rheinboldt. Manpak: A set of algorithms for computations on implicitly defined manifolds. *Computers & Mathematics with Applications*, 1996.

[52] Thomas J Rothenberg. Identification in parametric models. *Econometrica: Journal of the Econometric Society*, 1971.

[53] Emil Y. Sidky and Xiaochuan Pan. Report on the aapm deep-learning spectral ct grand challenge. *Medical physics*, 2024.

[54] David YW Thong, Charlesquin Kemajou Mbakam, and Marcelo Pereyra. Do bayesian imaging methods report trustworthy probabilities? *arXiv preprint arXiv:2405.08179*, 2024.

[55] Mark K Transtrum and Peng Qiu. Model reduction by manifold boundaries. *Physical review letters*, 2014.

[56] DJ Venzon and SH Moolgavkar. A method for computing profile-likelihood-based confidence intervals. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 37(1):87–94, 1988.

[57] Alejandro F Villaverde, Neil D Evans, Michael J Chappell, and Julio R Banga. Input-dependent structural identifiability of nonlinear systems. *IEEE Control Systems Letters*, 3(2):272–277, 2018.

[58] Pascal Vincent. A connection between score matching and denoising autoencoders. *Neural Comput.*, 23:1661–1674, jul 2011.

[59] Franz-Georg Wieland, Adrian L Hauber, Marcus Rosenblatt, Christian Tönsing, and Jens Timmer. On structural and practical identifiability. *Current Opinion in Systems Biology*, 2021.

[60] Kai Zhang, Yawei Li, Wangmeng Zuo, Lei Zhang, Luc Van Gool, and Radu Timofte. Plug-and-play image restoration with deep denoiser prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10):6360–6376, 2021.