arXiv:2412.06204v2 [cs.CV] 2 May 2025

You KAN Do It in a Single Shot: Plug-and-Play Methods with Single-Instance Priors

Yanqi Cheng[1], Carola-Bibiane Schönlieb[1], Angelica I Aviles-Rivero[2]

[1]Department of Applied Mathematics and Theoretical Physics, University of Cambridge.
[2]Yau Mathematical Sciences Center, Tsinghua University.

Contributing authors: yc443@cam.ac.uk; cbs31@cam.ac.uk; aviles-rivero@tsinghua.edu.cn;

Abstract

The use of Plug-and-Play (PnP) methods has become a central approach for solving inverse problems, with denoisers serving as regularising priors that guide optimisation towards a clean solution. In this work, we introduce KAN-PnP, an optimisation framework that incorporates Kolmogorov-Arnold Networks (KANs) as denoisers within the Plug-and-Play (PnP) paradigm. KAN-PnP is specifically designed to solve inverse problems with single-instance priors, where only a single noisy observation is available, eliminating the need for large datasets typically required by traditional denoising methods. We show that KANs, based on the Kolmogorov-Arnold representation theorem, serve effectively as priors in such settings, providing a robust approach to denoising. We prove that the KAN denoiser is Lipschitz continuous, ensuring stability and convergence in optimisation algorithms like PnP-ADMM, even in the context of single-shot learning. Additionally, we provide theoretical guarantees for KAN-PnP, demonstrating its convergence under key conditions: the convexity of the data fidelity term, Lipschitz continuity of the denoiser, and boundedness of the regularisation functional. These conditions are crucial for stable and reliable optimisation. Our experimental results show, on super-resolution and joint optimisation, that KAN-PnP outperforms exiting methods, delivering superior performance in single-shot learning with minimal data. The method exhibits strong convergence properties, achieving high accuracy with fewer iterations.

Keywords: Inverse Problems, Plug-and-Play Methods, Deep Denoiser Priors

1 Introduction

Inverse problems are widely studied in scientific computing and imaging, where the goal is to recover an unknown signal $x \in \mathbb{R}^N$ from indirect and often noisy measurements $y \in \mathbb{R}^M$. These measurements are typically modelled as:

$$y = A(x) + \epsilon, \tag{1}$$

where $A : \mathbb{R}^N \to \mathbb{R}^M$ is a forward operator representing the measurement process, and ϵ accounts for noise or perturbations in the system. Depending on the application, A may be linear, such as in magnetic resonance imaging (MRI) [1] or computed tomography (CT) [2], or nonlinear, such as in phase retrieval or optical imaging.

Due to challenges like noise contamination, incomplete data (e.g., undersampling), or the ill-posed nature [3–5] of the inverse operator A^{-1}, direct inversion is generally unstable or infeasible. To address this, regularisation methods are employed to incorporate prior knowledge about the desired solution x. Mathematically, this leads to solving a variational problem:

$$\min_{x \in \mathbb{R}^N} D(x) + \lambda R(x), \tag{2}$$

where $D(x)$ is the data-fidelity term that ensures consistency with the measurements y, $R(x)$ is the regularisation term encoding prior knowledge about x, and λ balances the two terms.

The proximal operator is a key mathematical tool [6] in solving regularised problems of the form given in (2). For a function $R(x)$, its proximal operator is defined as:

$$\mathrm{Prox}_{\gamma R}(v) = \arg\min_{x \in \mathbb{R}^N} \left\{ R(x) + \frac{1}{2\gamma} \|x - v\|^2 \right\}, \tag{3}$$

where $\gamma > 0$ is a step size parameter. Intuitively, the proximal operator can be interpreted as a denoising step that produces a regularised estimate of v by minimising the sum of $R(x)$ and a quadratic penalisation term.

Many iterative optimisation algorithms, such as proximal gradient methods and the alternating direction method of multipliers (ADMM) [7], leverage the proximal operator to solve problems with composite objectives. For instance, in ADMM, the proximal operator of the regularisation term $R(x)$ is applied iteratively to refine the solution.

Traditionally, $R(x)$ is chosen to be a mathematically explicit regularisation function, such as total variation, wavelet sparsity, or nuclear norm. However, modern approaches have shifted toward *implicit regularisation*, where the proximal operator is replaced with a learned or pre-defined denoiser $H_\sigma(x)$. This substitution leads to the Plug-and-Play (PnP) framework, introduced by Venkatakrishnan et al. [8], which exploits the equivalence between regularised denoising and the proximal operator. In the PnP framework, the optimisation problem is reformulated using a denoiser $H_\sigma(x)$

as a surrogate for $\text{Prox}_{\gamma R}$. For example, the ADMM iteration in PnP is given by:

$$\begin{aligned} x^{k+1} &= H_\sigma(z^k - u^k), \\ z^{k+1} &= \text{Prox}_{\mu^{-1}D}(x^{k+1} + u^k), \\ u^{k+1} &= u^k + x^{k+1} - z^{k+1}. \end{aligned} \tag{4}$$

where σ controls the denoising strength of H_σ, and μ is a parameter related to the data-fidelity term.

Building on this foundation, PnP has evolved significantly since its introduction by [8]. Early approaches demonstrated its effectiveness across diverse imaging tasks [9], while subsequent studies [10–13] refined the framework by enhancing stability, convergence, and applicability. Additionally, works such as [11, 14–17] extended its reach to more complex imaging scenarios. A notable milestone was the development of TFPnP (Tuning-Free Plug-and-Play) [18], which eliminated the need for parameter tuning, simplifying practical deployment.

The core of the Plug-and-Play (PnP) framework lies in the denoiser, H_σ, which serves as an implicit regulariser for the reconstruction process. This denoiser is typically trained using a large amount of data, allowing it to learn rich priors that generalise well to diverse instances of the underlying task. PnP frameworks have been developed utilising a variety of denoisers. Classical methods like BM3D [19] and TV [20] remain popular due to their robustness and minimal reliance on training data. In the meantime, the rise of deep learning-based denoisers [9, 16, 21] has significantly advanced the field by capturing complex data patterns, such as DnCNN [22] or U-Net [23] achieve their performance by leveraging extensive datasets to capture the statistical properties of natural images or specific domains. Recent works on Weakly Convex Ridge Regularisers (WCRR-NNs) [24] proposes learning explicit regularisation functionals $R(x)$ such that the inverse problem remains (weakly) convex. These methods operate in a variational setting and typically require access to a dataset for training. However, this reliance on large-scale training datasets poses a significant challenge in domains where data availability is limited.

A key question thus arises: *how can PnP frameworks be adapted to solve inverse problems effectively when only a single instance or minimal data is available?* Addressing this question requires rethinking the fundamental reliance of PnP methods on large-scale training datasets and instead developing approaches that can utilise the limited data at hand efficiently. The recent application of Deep Image Prior (DIP) [25] within the Plug-and-Play (PnP) framework [26] represents a noteworthy development. However, in practice, DIP-PnP still depends on training the model with multiple instances. Our focus, in contrast, is on minimising the data required for prior training in the PnP setting—specifically, training on a single instance. To the best of our knowledge, the only existing work that aligns with ours is that of Cheng et al. [27], which introduces the single-shot paradigm within the Plug-and-Play context. We highlight key differences between their approach and ours: Firstly, their prior is based on implicit neural representations, whereas ours leverages the Kolmogorov-Arnold representation theorem. Moreover, unlike their work, we provide theoretical results on both the denoiser's properties and the convergence of the entire scheme.

Constributions. We introduce a novel optimisation framework called KAN-PnP, which leverages Kolmogorov-Arnold Networks (KANs) as denoisers within the Plug-and-Play (PnP) optimisation paradigm. This method addresses the challenge of solving inverse problems in the context of single-instance priors, where only a single noisy observation is available. Our approach offers a significant departure from traditional denoising methods that rely on large datasets, providing a robust solution for cases with minimal data. The key contributions of this work are as follows:

- We present the first framework to incorporate Kolmogorov-Arnold Networks (KANs) [28] as denoisers in the PnP paradigm, focusing on solving inverse problems with single-instance priors:
  - We demonstrate how KANs, based on the Kolmogorov-Arnold representation theorem, can effectively serve as priors in settings where only one noisy observation is available, bypassing the need for large datasets typically required by traditional denoising techniques.
  - We prove that the KAN denoiser is Lipschitz continuous, ensuring its stability and guaranteeing convergence in optimisation algorithms like PnP-ADMM (Theorem 1).
  - The framework is designed to handle single-shot learning, where the denoiser is adapted to a specific image instance, allowing it to perform well in real-world scenarios with limited data.
- We establish the convergence of the KAN-PnP algorithm, proving that it converges to a stationary point under certain conditions (Theorem 2): I) Convexity of the Data Fidelity Term: Ensures the well-posedness of the optimisation problem and the solvability of the data fidelity subproblem. II) Lipschitz Continuity of the Denoiser: Guarantees the stability of the denoising step, ensuring that the method behaves well during optimisation. III) Boundedness of the Regularisation Functional: Prevents unbounded behaviour in the regularisation term, ensuring a stable optimisation process.
- We validate the effectiveness of KAN-PnP through extensive experiments on various inverse problems. We demonstrate that KAN-PnP significantly outperforms existing methods, when trained with a single instance observation. Moreover, we show that our method shows excellent convergence properties, reliably approaching a stationary point with minimal iterations. Finally, We compare KAN-PnP with state-of-the-art denoising methods and show that it provides superior performance for single and multi-operator inverse problems including super-resolution and joint optimisation.

2 Proposed Method

We remain to the reader that we consider the following class of optimisation problems from the Plug-and-Play perspective:

$$\min_{x,z} f(x) + \lambda R(H_\sigma(z)), \quad \text{subject to} \quad x = z \tag{5}$$

where $f(x)$ is the data fidelity term, λ is a regularisation parameter, and $R(x)$ is the regularisation functional implicitly defined through a network denoiser $H_\sigma(x)$. This

framework generalises a wide range of inverse problems. We note that (5) should be understood as a conceptual variational formulation underpinning the Plug-and-Play scheme. While $R(H_\sigma(z))$ is not used explicitly in the algorithm, it allows us to interpret the denoising step as a surrogate proximal operator of an implicit regulariser R.

A critical component of the Plug-and-Play methods is the choice of the denoiser. In traditional denoising techniques, methods such as non-local means (NLM) [29], BM3D [19], and deep learning-based approaches like DnCNN [22] have been widely used. These methods, although effective, *often rely on large datasets for training, which limits their application in scenarios where only a single noisy observation is available*. In this section, we detail a novel framework called KAN-PnP, which uses a single-instance prior.

2.1 KAN as the Ideal Single-Instance Prior

In our framework, *we focus on a single-shot learning setting*, where we only have a single noisy observation of the underlying clean signal. This presents a challenge, as traditional denoisers may struggle in such settings, where the noise characteristics are not fixed across samples. To address this, we propose a novel prior drawing from Kolmogorov-Arnold Networks (KANs) as the denoiser. KANs are a powerful class of neural networks based on the Kolmogorov-Arnold representation theorem, which ensures that any continuous function can be approximated by a finite composition of smooth, univariate functions. This ability to smoothly approximate complex mappings makes KANs particularly well-suited for denoising tasks, as they naturally preserve the underlying structure of the signal while removing noise. KANs introduce a form of structured inductive bias through kernel-based interpolation over learnable anchors, which induces local smoothness and adaptivity — properties that are well-aligned with the statistical structure of images. Unlike MLP-based INRs, which are globally connected and prone to spectral bias (favoring low-frequency content), KAN's piecewise smooth interpolation better captures high-frequency details and edges, which are essential for image restoration tasks (imaging inverse problems). In the single-shot setting, where no external supervision is available, this inductive bias acts as a strong prior that regularises the learning dynamics and helps avoid overfitting to noise. Kolmogorov-Arnold Networks (KANs) are a class of neural networks based on the Kolmogorov-Arnold representation theorem [30–32], which guarantees that any continuous function on a multidimensional domain can be approximated by a finite composition of continuous functions. KAN is defined as follows. Let $x \in \mathbb{R}^d$ be the input vector, and let $y \in \mathbb{R}$ be the output. A KAN approximates a continuous function $f(x)$ by composing several layers of univariate basis functions. Specifically, a KAN can be expressed as:

$$f(x) = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} \phi_{ij}(x_i),$$

where $\phi_{ij}(x_i)$ are univariate functions (typically chosen to be continuous and smooth, such as B-splines or polynomial functions). Moreover, α_{ij} are the weights (parameters) of the network, and x_i represents the components of the input vector x. Also, m and n are the number of hidden layers and basis functions, respectively. Each hidden

layer $\phi_{ij}(x_i)$ is constructed from a basis function that is chosen to provide smooth approximations, and the overall network combines these functions to model the desired output.

KAN-Plug-and-Play Method. In the context of Plug-and-Play Methods, and in particular in this work PnP-ADMM, the denoiser serves as the regularising prior, which guides the optimisation process toward a clean solution. One critical property of the denoiser is Lipschitz continuity [12], which ensures stability of the denoiser. This property is essential for the convergence of iterative algorithms like PnP-ADMM, where stability and well-behaved priors are necessary for guaranteeing the algorithm's success.

Theorem 1 (Lipschitz Continuity of the KAN Denoiser)**.** *Let $H(x)$ represent the KAN denoiser defined as a composition of layers:*

$$H(x) = \Phi_L \circ \Phi_{L-1} \circ \cdots \circ \Phi_1(x), \tag{6}$$

where each layer Φ_l is a matrix of univariate B-spline functions $\phi_{l,j,i}(x)$ with bounded derivatives. Then $H(x)$ is Lipschitz continuous, i.e., there exists a constant $L > 0$ such that:

$$\|H(x) - H(y)\| \leq L\|x - y\|, \quad \forall x, y \in \mathbb{R}^n. \tag{7}$$

Proof. We begin by analysing the Lipschitz continuity of each individual layer in the KAN architecture. Recall that the activation value at layer $l + 1$ is computed as:

$$x_{l+1,j} = \sum_{i=1}^{n_l} \phi_{l,j,i}(x_{l,i}), \tag{8}$$

where $x_{l,i}$ represents the output of neuron i in layer l, and $\phi_{l,j,i}$ denotes the univariate B-spline function that connects neuron i in layer l to neuron j in layer $l + 1$. The pre-activation values are aggregated using a summation operation, and each $\phi_{l,j,i}$ is differentiable and piecewise smooth.

Let Φ_l denote the mapping from x_l to x_{l+1}. Each univariate spline $\phi_{l,j,i}(x)$ is parametrised using B-splines, whose maximum derivative depends on the number of control points and the smoothness order k. Assuming the splines are constructed with G grid points and bounded coefficients, the maximum derivative $L_{\phi_{l,j,i}}$ can be explicitly bounded as:

$$L_{\phi_{l,j,i}} \leq CG^{k-1}, \tag{9}$$

where C is a constant dependent on the initialisation scale and spline parameters.

Since the post-activation value at neuron j in layer $l + 1$ is a summation of the outputs of splines, we can bound the Lipschitz constant of each neuron as:

$$L_{\Phi_l,j} = \sum_{i=1}^{n_l} L_{\phi_{l,j,i}}. \tag{10}$$

The Lipschitz constant for the entire layer Φ_l is therefore:

$$L_{\Phi_l} \leq \max_j \sum_{i=1}^{n_l} CG^{k-1}. \tag{11}$$

The overall denoiser $H(x)$ is a composition of L such layers:

$$H(x) = \Phi_L \circ \Phi_{L-1} \circ \cdots \circ \Phi_1(x). \tag{12}$$

By the chain rule for Lipschitz continuity, the Lipschitz constant of the composite function is given by the product of the Lipschitz constants of the individual layers:

$$L_H = \prod_{l=1}^{L} L_{\Phi_l}. \tag{13}$$

Substituting the bound for L_{Φ_l} into this expression, we have:

$$L_H \leq \prod_{l=1}^{L} \left(\max_j \sum_{i=1}^{n_l} CG^{k-1} \right). \tag{14}$$

Since each layer Φ_l has a finite Lipschitz constant L_{Φ_l}, their composition $H(x)$ is also Lipschitz continuous with a finite constant L_H. This completes the proof. □

On the Theorem 1 and its Implications. While Lipschitz continuity of neural networks is well-known, we emphasise that our result is tailored to the Kolmogorov-Arnold Network (KAN) architecture, which differs significantly from standard ReLU-based networks. The use of smooth univariate B-spline basis functions requires a more explicit analysis to ensure global Lipschitz continuity, especially in the context of single-instance priors, where no data-driven regularisation is available. Although the derived Lipschitz bound in Theorem 1 may be conservative, it offers a principled way to control the smoothness and stability of the KAN denoiser via architectural parameters (e.g., number of grid points G, spline smoothness k, and layer width). This enables practitioners to design networks with desirable properties, including potential non-expansiveness ($L_H \leq 1$), which is beneficial for convergence in Plug-and-Play schemes.

2.2 Key Assumptions for KAN-PnP

Having established that the KAN denoiser H_σ is Lipschitz continuous, we now outline the key assumptions required to ensure the convergence of the Plug-and-Play framework. In particular, we focus on PnP-ADMM framework in the context of single-shot learning setting. Unlike traditional settings where the denoiser is trained on a large dataset, single-shot learning involves tailoring H_σ to a single image instance. This adaptation introduces unique challenges in satisfying the assumptions, which we address below:

• **Convexity of the Data Fidelity Term.** The data fidelity term $f(x)$ must be convex, proper, and lower semi-continuous. In this work, we use the squared ℓ_2-norm, defined as:

$$f(x) = \|Ax - y\|_2^2, \tag{15}$$

which is a standard choice in inverse problems due to its simplicity and convexity. This ensures that the subproblem for z^{k+1} in the ADMM iterations is well-posed and solvable. The use of $f(x)$ remains valid regardless of whether H_σ is trained on a large dataset or a single image.

• **Lipschitz Continuity of the Denoiser.** As demonstrated in the previous section, the KAN denoiser H_σ is Lipschitz continuous with a constant L_H. This guarantees the stability of the denoising step, making it suitable as a surrogate for the proximal operator in the PnP framework. In single-shot learning, ensuring Lipschitz continuity requires architectural constraints within the KAN framework. The use of bounded spline parameters and regularisation within the network ensures controlled outputs, even when the denoiser is adapted to a single image.

• **Boundedness of the Regularisation Functional.** The regulariser $R(x)$ associated with the KAN denoiser is defined implicitly via the proximal operator formulation:

$$H_\sigma(x) = \arg\min_{z \in \mathbb{R}^n} \left\{ R(z) + \frac{1}{2\sigma^2} \|z - x\|^2 \right\}. \tag{16}$$

This implies that $R(x)$ is bounded below due to the quadratic penalty term, which prevents $R(x)$ from decreasing. However, in the single-shot setting, the behavior of $R(x)$ depends on how H_σ is constructed and adapted. To ensure boundedness:

» Architectural smoothness (e.g., using B-splines with bounded derivatives) in the KAN denoiser suppresses non-coercive behaviour.

» The quadratic term in the proximal operator formulation helps dominate $R(z)$ for large z, further ensuring boundedness.

We remark that (16) defines H_σ as the proximal operator of a regularisation functional R. However, such an R exists only if H_σ is a firmly nonexpansive operator — a condition not guaranteed in the learned denoising setting. In our framework, we treat this formulation as a conceptual model aligned with recent works (e.g., [10]), and do not assume the explicit existence or convexity of R. Our convergence results rely on the Lipschitz continuity of H_σ, convexity of the data fidelity term, and boundedness of the surrogate regularisation.

• **Differentiability of the Implicit Regulariser.** The implicit regulariser $R(x)$ derived from H_σ must have a Lipschitz continuous gradient. This condition can be satisfied if $H_\sigma(x)$ is differentiable and smooth, which follows from the construction of the KAN denoiser using B-splines. In single-shot learning, we ensure the smoothness of H_σ, avoiding overfitting artifacts or discontinuities introduced during adaptation to the single image. We emphasise that while (16) conceptually associates the denoiser H_σ with a regularisation functional R, we do not assume that R is differentiable. However, since the KAN denoiser H_σ is constructed from smooth B-spline basis functions, it is differentiable. This differentiability enables practical use of updates in PnP-ADMM, and contributes to stable convergence behavior. Nonetheless, the differentiability of H_σ does not imply the existence or differentiability of an underlying functional R.

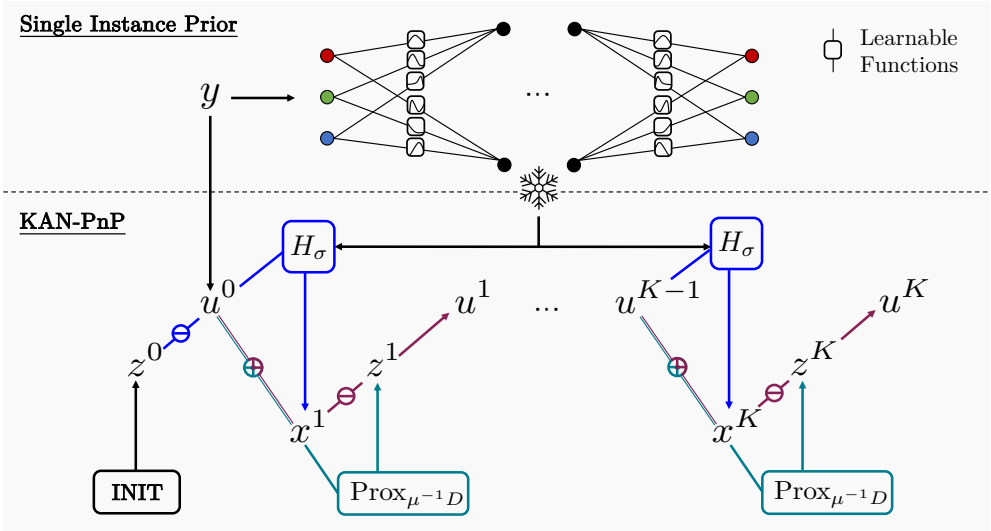

Fig. 1 The proposed KAN-PnP framework, comprising the training of a single-instance prior and its integration into an iterative optimisation algorithm.

2.3 Convergence of KAN-PnP with Single-Instance Prior

In this section, we analyse the convergence of our KAN-PnP framework for solving the inverse problem, where the regulariser is replaced by the KAN denoiser H_σ. We recall that our problem is:

$$\min_{x,z} f(x) + \lambda R(H_\sigma(z)) \quad \text{subject to} \quad x = z,$$

where $f(x)$ is the data fidelity term and $R(H_\sigma(z))$ represents the implicit regularisation term provided by the KAN denoiser.

The KAN-PnP (see Figure 1) with ADMM proceeds by iteratively solving the following three subproblems:

$$\begin{aligned} \textbf{I. Denoising Step}: \; & x^{k+1} = H_\sigma(z^k - u^k), \\ \textbf{II. Data Fidelity Step}: \; & z^{k+1} = \arg\min_z \left(f(z) + \frac{\mu}{2}\|z - (x^{k+1} + u^k)\|^2 \right), \\ \textbf{III. Dual Update}: \; & u^{k+1} = u^k + x^{k+1} - z^{k+1}. \end{aligned} \tag{17}$$

In this section, we show that under the assumptions of Lipschitz continuity, convexity of $f(x)$, and boundedness of the regulariser $R(x)$, KAN-PnP algorithm converges to a stationary point.

Key Assumptions Recap. To ensure convergence of the KAN-PnP framework with the KAN denoiser, we rely on the following assumptions:

1. Convexity of $f(x)$, which ensures well-posedness of the data fidelity subproblem.

2. Lipschitz continuity of the KAN denoiser H ensures that the denoising step is stable and does not grow arbitrarily with input perturbations. While this does not imply non-expansiveness (i.e., 1-Lipschitz continuity), convergence can still be guaranteed under this weaker assumption, as discussed in [10].
3. Differentiability of the regulariser $R(x)$, if required for updates, which is satisfied under the smoothness of H_σ.

Theorem 2 (Fixed-Point Convergence of KAN-PnP)**.** *Under the assumptions:*
(i) $f : \mathbb{R}^n \to \mathbb{R}$ *is proper, closed, and convex,*
(ii) $H_\sigma : \mathbb{R}^n \to \mathbb{R}^n$ *is Lipschitz continuous with constant* L_H *(as established in Theorem 1),*
(iii) The penalty parameter $\mu > L_H$ *is constant across iterations,*
the PnP-ADMM iterates $\{(x^k, z^k, u^k)\}$ *defined in (17) converge to a fixed point* (x^*, z^*, u^*) *satisfying:*

$$x^* = H_\sigma(z^* - u^*), \quad z^* = \arg\min_z \left(f(z) + \frac{\mu}{2}\|z - (x^* + u^*)\|^2 \right), \quad x^* = z^*.$$

Proof Sketch.

We analyse the convergence of the KAN-PnP ADMM iterations using the fixed-point framework introduced in recent Plug-and-Play literature [10, 12]. This approach does not assume the existence of an explicit energy functional involving the denoiser, but instead analyses the behavior of the iterative updates as an operator composition acting on the optimisation variables. The denoising step $x^{k+1} = H_\sigma(z^k - u^k)$ applies a Lipschitz continuous operator H_σ with constant L_H (as guaranteed by Theorem 1). The z-update step:

$$z^{k+1} = \arg\min_z \left(f(z) + \frac{\mu}{2}\|z - (x^{k+1} + u^k)\|^2 \right),$$

is equivalent to the proximal operator of the convex function f, evaluated at $x^{k+1} + u^k$. Since f is convex and proper, the proximal operator $\mathrm{Prox}_{f/\mu}$ is firmly nonexpansive.

The dual update $u^{k+1} = u^k + x^{k+1} - z^{k+1}$ is a standard ADMM step enforcing consensus between x and z. When combining these three steps, the overall ADMM iteration can be interpreted as a fixed-point iteration of a composed operator acting on (x, z, u). Under the assumption that $\mu > L_H$, it follows from standard results in fixed-point theory that the composition of the Lipschitz operator H_σ and the firmly nonexpansive proximal operator defines an averaged operator with a contraction factor strictly less than one [10]. By the Banach fixed-point theorem, this guarantees that the sequence $\{(x^k, z^k, u^k)\}$ converges to a unique fixed point (x^*, z^*, u^*) satisfying:

$$x^* = H_\sigma(z^* - u^*), \quad z^* = \mathrm{Prox}_{f/\mu}(x^* + u^*), \quad x^* = z^*.$$

These conditions characterise a consistent fixed point of the algorithm, even though the objective function is implicit. This result justifies the convergence of the KAN-PnP method under mild and verifiable assumptions, without requiring knowledge of the underlying regulariser associated with H_σ. □

Remark. The convergence result in Theorem 2 adapts the fixed-point framework established in recent Plug-and-Play literature [10, 12]. While the overall structure follows these works, our analysis is specialised to the single-instance setting, where the denoiser H_σ is trained directly from a noisy observation. This introduces unique architectural properties (e.g., differentiability from B-splines) and excludes any access to external datasets or explicit regularisers. Our proof reflects this structure and confirms convergence under mild assumptions on the denoiser and fidelity term.

3 Experimental Results & Discussion

In this section, we evaluate the performance of our proposed KAN-PnP framework across challenging image reconstruction tasks, including super-resolution and joint optimisation for demosaicing and deconvolution. Our approach leverages a *single-instance prior*, wherein the denoiser is trained directly on the observed image itself, rather than on a dataset of similar images. The goal of these experiments is to assess the effectiveness of our approach compared to other single-instance denoisers and the conventional methods, including pretrained denoisers and total variation techniques within the Plug-and-Play framework. This training scheme is designed to exploit the unique characteristics of the target image, enhancing reconstruction quality without requiring extensive external datasets. This rationale motivates our comparison with traditional methods trained in a conventional manner.

For the experimental comparison, we followed the dual resizing strategies to preprocess the image. The images are sourced with Creative Commons Licenses which are resized to 512×384, and the selected data in [33, 34] were tested. The experiment were conducted on NVIDIA A100 GPU with 80GB RAM. The Plug-and-Play phase ensued with the ∇-Prox toolbox [35], adopting the default settings for all tasks.
Why Natural Images? While our method is applicable to diverse data types, we focus on natural images in this work for several reasons. First, natural image benchmarks (e.g., Set5, Set14) are widely used in foundational works on single-instance priors (e.g., DIP [25] and Cheng et al. [27]), enabling direct comparison with prior art. Second, these datasets offer well-understood structure and standardized metrics, which help isolate the methodological contribution of the learned prior. Third, since our aim is to investigate core behavior of single-shot Plug-and-Play priors — a relatively unexplored domain — natural images provide a valuable, interpretable setting for proof-of-concept evaluation.

• **Training Scheme.** The initial pre-training phase for the implicit neural coordinate network including the implicit neural representation (INR) and the implicit Kolmogorov-Arnold Network (KAN) denoising prior incorporated Gaussian noise with a standard deviation of 0.1. The single-shot denoising prior training spanned 100 iterations using a INR network composed of 2 hidden layers, each with 64 features, while the KAN network is composed of 3 hidden layers with $\{128, 32, 16\}$ features with grid

5. The learning rate was maintained at 0.001. Subsequently, image reconstruction was performed over 5 ADMM iterations, employing dynamically adjusted noise levels and a penalty parameter that followed a logarithmic descent across the steps.

- **Evaluation Protocol.** To ensure a robust comparison, we evaluated the Noise2Self pre-training approach [36] on three networks, each undergoing 100 iterations with a learning rate of 0.01. The FFDNet [37] model is designed with 8 hidden layers and 64 feature maps per layer, while the UNet [23] architecture featured a 4-level downsampling scheme and began with 32 feature maps in its initial convolutional layer. For consistency, the optimisation setup mirrored the configuration used in our proposed approach. We compare these results against both the Noise2Self approach and implicit neural MLP, which we denoted as INR, as well as our proposed implicit KAN method. The evaluation of the methods were measured by Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), where PSNR and SSIM scores signal are higher when reconstruction quality improved. **On Parameter Counts.** While the KAN and INR architectures differ in parameter count, we do not enforce equality since they encode fundamentally different inductive biases. KANs exploit local interpolation priors, while INRs rely on global MLP representations. For this reason, we evaluate models based on reconstruction quality rather than parameter parity.

3.1 KAN-PnP in High-Stakes Image Reconstruction

Our proposed KAN-PnP framework has been extensively evaluated in two challenging inverse problem settings: super-resolution and joint optimisation for demosaicing and deconvolution. Super-resolution tests the model's ability to enhance image detail across varying upscaling factors (2×, 4× and 8×), while the joint optimisation task emphasises its robustness and adaptability by tackling the intertwined challenges of demosaicing and deconvolution. The latter setting is particularly demanding due to the need to simultaneously address two operations—removing the effects of mosaic patterns while restoring the high-frequency details lost during image acquisition—requiring a balance of accuracy and computational.

3.1.1 Yes, You KAN do Super-resolution Better!

The results in Table 1 underscore the significant advantages of KAN-PnP over a range of competing single-shot techniques, including INR, FFDNet, and UNet, across all scaling factors and datasets. KAN-PnP demonstrates a clear ability to generalise across diverse image types while maintaining robust performance under varying levels of degradation. Its superior adaptability, stability, and capacity to preserve intricate details highlight why it consistently outperforms other methods, particularly in challenging scenarios.

At 2x upscaling, KAN-PnP reports exceptional performance across all samples. For example, in the Dog instance, KAN-PnP achieves a PSNR of 25.82, outperforming FFDNet (24.02) and UNet (18.71) by notable margins. In *texture-heavy* samples like Squirrel and Koala, KAN-PnP achieves PSNR values of 27.07 and 25.00, respectively, further demonstrating its capacity to handle both structured and unstructured features. While INR is a close competitor in these cases, KAN-PnP consistently either

Table 1 The performance (PSNR(dB)) comparison of different methods on super-resolution (SR) tasks with 2×, 4×, and 8× upscaling factors.

Method	Dog			Fox			Giraffe		
	2×	4×	8×	2×	4×	8×	2×	4×	8×
FFDNet-PnP	24.02	21.98	14.88	25.10	23.19	6.51	4.29	13.20	17.44
UNet-PnP	18.71	14.34	6.86	19.82	14.38	14.99	17.81	14.12	6.31
INR-PnP	25.80	22.00	**15.40**	26.73	23.24	**16.21**	27.93	24.23	17.46
KAN-PnP	**25.82**	**22.05**	**15.40**	**27.06**	**23.29**	**16.21**	**27.94**	**24.24**	**17.47**

Method	Koala			Squirrel			Bird		
	2×	4×	8×	2×	4×	8×	2×	4×	8×
FFDNet-PnP	9.02	1.09	14.58	4.86	4.86	16.34	26.34	21.86	15.71
UNet-PnP	14.24	13.46	2.22	17.14	13.66	3.37	21.82	16.53	10.69
INR-PnP	24.97	20.51	**14.79**	**27.08**	**23.05**	**16.93**	**26.44**	22.06	15.74
KAN-PnP	**25.00**	**20.56**	**14.79**	27.07	**23.05**	**16.93**	26.43	**22.19**	**15.75**

matches or exceeds its performance, solidifying its position as a leading single-shot method. At 4x scaling, the differences between KAN-PnP and its competitors become even more pronounced. For example, on the Bird sample, KAN-PnP achieves a PSNR of 22.19, surpassing both FFDNet (21.86) and INR (22.06). UNet struggles significantly at this scale, achieving only 16.53, indicating its inability to handle moderate degradation effectively. Similarly, in instances like Fox and Giraffe, KAN-PnP maintains a distinct performance edge, with PSNR values of 23.29 and 24.24, respectively, compared to UNet's 14.38 and 14.12. These results emphasise KAN-PnP's ability to balance global structure recovery and fine detail preservation, a challenge that often undermines the performance of its competitors.

Extreme Degradation: At 8x scaling, the task of image reconstruction becomes particularly challenging, as the loss of high-frequency information is severe. While the margins between KAN-PnP and INR narrow in some datasets (e.g., Giraffe, where KAN achieves 17.47 compared to INR's 17.46), KAN-PnP demonstrates greater consistency across diverse datasets and excels in preserving fine details. On the Bird dataset at 8×, KAN-PnP achieves a PSNR of 15.75, outperforming FFDNet (15.71) and significantly exceeding UNet's 10.69. This consistency across datasets highlights the robustness of KAN-PnP in scenarios where competing methods often struggle to generalise or maintain stability. Additionally, KAN-PnP's architectural design allows it to preserve critical image features such as edges and textures, which are frequently lost in reconstructions by FFDNet and UNet. Particularly and unlike INR, KAN-PnP remains stable under extreme degradation (e.g., 8×) while excelling in texture-heavy and structured datasets alike.

Visual Results. We further support our numerical results through a set of visual outputs, which are displayed in Figures 2 and 3. KAN-PnP demonstrates remarkable detail preservation, particularly in the texture of the fox's fur and the sharpness of its

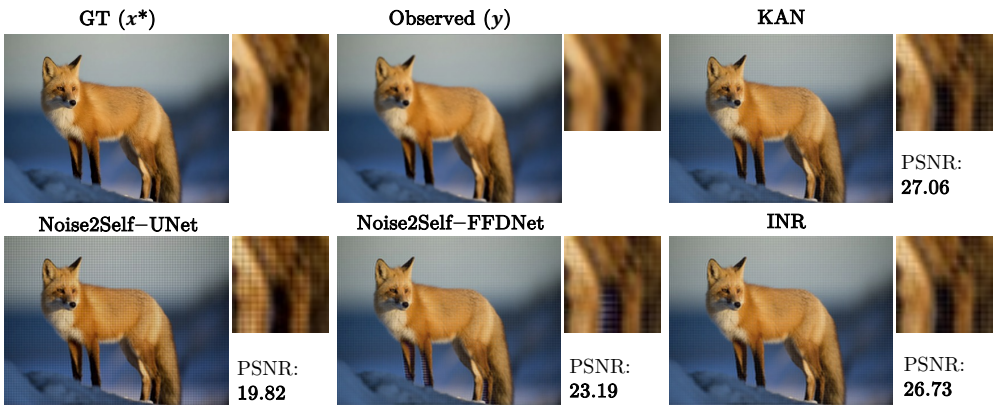

Fig. 2 The visualisation comparison of the 'Fox' example on 2× Super-Resolution task among the UNet, FFDNet, INR and KAN denoising prior in single-shot setting.

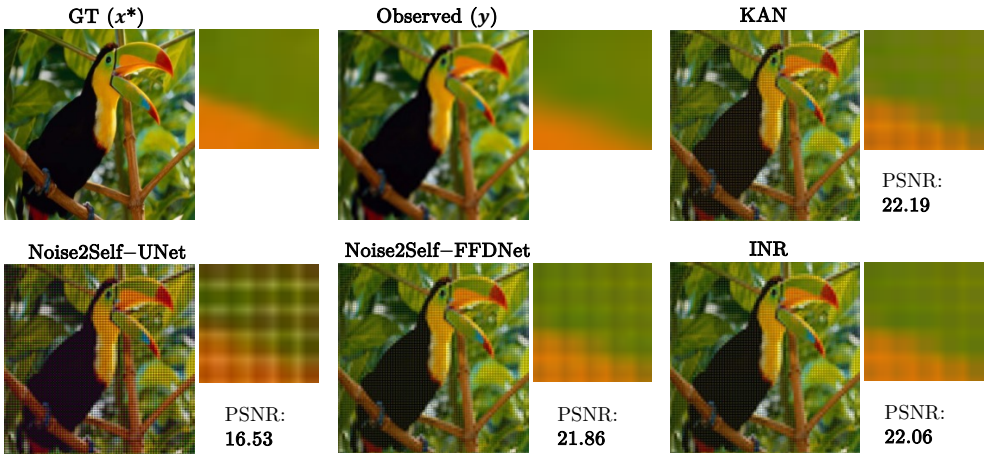

Fig. 3 Comparing the visual result of the 'Bird' example on 4× Super-Resolution task in single-shot setting with the UNet, FFDNet, INR and KAN denoising prior.

silhouette, as evidenced in the zoomed-in patches. In contrast, INR, despite a close PSNR, exhibits slight edge blurring, while FFDNet oversmooths the background, and UNet introduces grid-like distortions. In the more challenging 4x example (Figure 3, Bird), KAN-PnP maintains its edge, marginally surpassing INR and outperforming FFDNet and UNet. KAN-PnP excels in preserving edge sharpness and texture, as seen in the bird's feathers and beak, while retaining vibrant colour fidelity. INR, while competitive, shows blurring of edges, whereas FFDNet produces oversmoothed results, and UNet introduces significant artefacts and detail loss. These results reaffirm KAN-PnP's ability to deliver visually superior reconstructions with sharper details, fewer

Table 2 The performance (PSNR(dB), SSIM) comparison on joint deconvolution and demosaicing of single-shot deep denoising priors (KAN, INR, Noise2Self-UNET, and Noise2Self-FFDNet), pre-trained deep denoising priors following [35] (Pretrained-UNet and Pretrained-FFDNet), and classical denoising priors (TV) in Plug-and-Play framework.

		Raccoon		Fractals		Wolf		Butterfly	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Classic	TV	15.55	0.43	13.08	0.24	17.27	0.45	14.16	0.29
Pretrain	UNet	16.48	0.42	15.53	0.32	17.54	0.36	15.43	0.28
	FFDNet	23.27	0.75	20.26	0.57	25.66	0.81	20.26	0.50
Single Instance	Noise2Self-UNet	14.93	0.23	11.85	0.16	18.45	0.28	13.03	0.20
	Noise2Self-FFDNet	25.12	0.79	22.77	0.80	30.54	0.90	23.86	0.84
	INR	25.63	0.82	22.86	0.80	30.90	**0.91**	23.87	**0.86**
	KAN	**25.65**	**0.82**	**24.46**	**0.81**	**31.02**	**0.91**	**24.00**	**0.86**

		Giraffe		Mushroom		Baby		Monarch	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Classic	TV	13.82	0.21	17.83	0.53	13.13	0.33	16.35	0.41
Pretrain	UNet	16.71	0.26	18.09	0.49	17.77	0.50	17.10	0.41
	FFDNet	25.75	0.73	25.02	0.82	25.98	0.80	21.82	0.63
Single Instance	Noise2Self-UNet	12.62	0.11	18.78	0.34	12.95	0.13	15.39	0.26
	Noise2Self-FFDNet	29.15	0.83	26.86	0.88	28.32	0.83	26.31	**0.89**
	INR	30.64	**0.89**	26.99	**0.89**	29.09	0.86	26.48	**0.89**
	KAN	**30.80**	**0.89**	**27.00**	**0.89**	**29.27**	**0.87**	**26.52**	**0.89**

artefacts, and better preservation of textures and edges across varying scaling factors, setting it apart as a robust single-shot super-resolution method.

3.1.2 Joint Restoration Tasks

The results in Table 2 highlight the effectiveness of KAN-PnP when applied to the challenging joint task of demosaicing and deconvolution, demonstrating its superiority over classical, pre-trained, and other single-shot methods. These tasks compound the complexity of inverse problems, as they require the model to handle two distinct operations—recovering high-frequency details lost during the acquisition process and simultaneously reconstructing colour patterns. Despite these challenges, KAN-PnP consistently delivers the highest or comparable Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) scores across all datasets, showcasing its robustness and adaptability.

KAN-PnP outperforms its competitors, including INR and Noise2Self-FFDNet, in both PSNR and SSIM metrics. For example, on the Wolf dataset, KAN achieves a PSNR of 31.02 and an SSIM of 0.91, surpassing the already strong performance of INR (30.90, 0.91) and Noise2Self-FFDNet (30.54, 0.90). Similarly, for the Butterfly

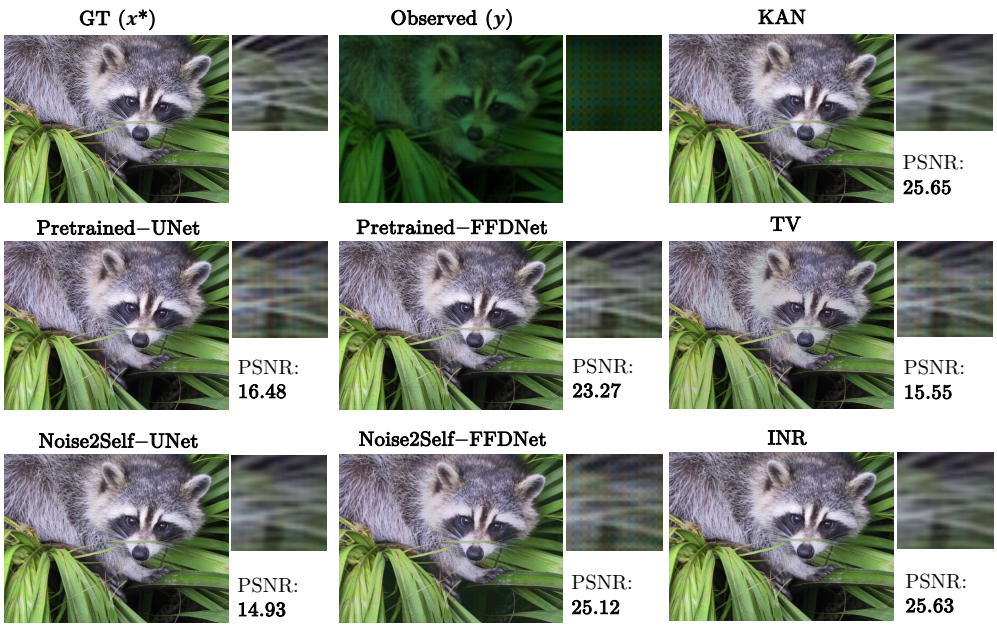

Fig. 4 Comparison of the visual output of the Plug-and-Play framework of the 'Raccoon' example on Joint Deconvoution and Demosaicing task with traditional prior (TV), pretrained prior (UNet and FFDNet) and Single-Shor prior (Noise2Self-UNet, Noise2Self-FFDNet, INR and KAN).

dataset, KAN achieves 24.00 PSNR and 0.86 SSIM, outperforming INR (23.87, 0.86) and significantly exceeding pre-trained methods like FFDNet (20.26, 0.50). These results underscore the ability of KAN-PnP to preserve fine details, edges, and textures while accurately reconstructing complex color patterns, making it particularly suited for high-dimensional inverse problems.

KAN-PnP's superior performance can be attributed to its mathematical foundation based on the Kolmogorov-Arnold representation, which ensures smooth, stable approximations while preserving fine details and structures. The Lipschitz continuity of the KAN denoiser ensures stability during optimisation, a critical factor for solving complex inverse problems like joint demosaicing and deconvolution. By effectively balancing local and global features, KAN-PnP surpasses both classical and deep learning-based methods, even in scenarios with limited data availability.

The visualisations in Figures 4 and 5 further support these numerical results, showcasing KAN-PnP's ability to deliver sharper reconstructions and preserve critical details compared to other methods. For instance, in the Raccoon example (Figure 4), KAN-PnP produces the most visually accurate result, effectively restoring the texture of the fur and the structure of the leaves, while methods like Noise2Self-FFDNet and TV fail to maintain texture fidelity and introduce artifacts. Similarly, in the Baby example (Figure 5), KAN-PnP achieves a PSNR of 29.27, recovering fine features such as the softness of the lips and smoothness of the skin, outperforming pre-trained methods like FFDNet, which oversmooths details, and classical methods like TV, which

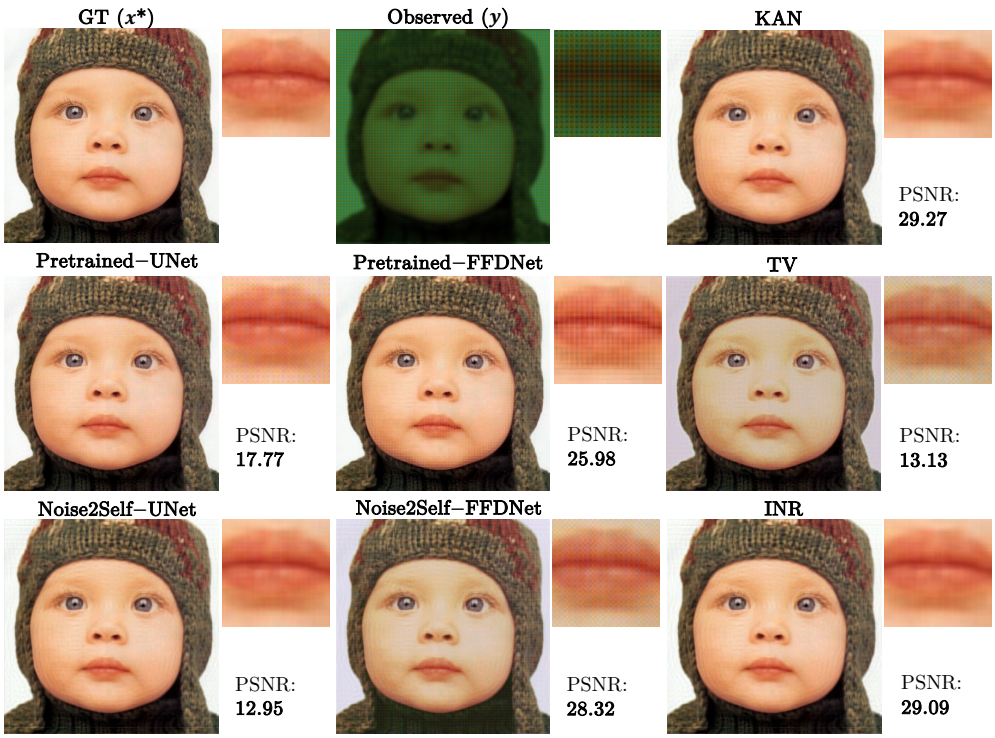

Fig. 5 The visualisation comparison of the 'Baby' example on Joint Deconvoution and Demosaicing task with traditional prior (TV), pretrained prior (UNet and FFDNet) and single-shot prior (Noise2Self-UNet, Noise2Self-FFDNet, INR and KAN) in the Plug-and-Play framework.

exhibit heavy artifacts. These visual comparisons highlight KAN-PnP's ability to consistently produce high-quality reconstructions that align with its strong numerical performance.

3.1.3 Unplugging KAN: Ablation Study

• **KAN Denoiser Basis.** The results in Table 3 provide valuable insights into the impact of different basis functions used in the KAN denoiser on image reconstruction performance. Three basis functions—B-Spline, Fourier, and Symbolic—are evaluated across multiple datasets, with metrics presented in terms of Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). These findings allow us to better understand the contributions of the selected basis to the overall reconstruction quality and robustness. This consistent superiority highlights the strength of B-Spline's smooth and adaptive nature, which is particularly effective at capturing local features and preserving fine textures, such as those in the Wolf and Butterfly datasets. The smoothness of B-Spline functions enables it to balance noise suppression and detail recovery, making it an ideal choice for complex reconstruction tasks. The Fourier basis shows a notable decline in performance compared to B-Spline and Symbolic bases,

Table 3 The performance (PSNR(dB) and SSIM) comparison across different KAN layers (B-Spline, Fourier, and Symbolic) for image reconstruction.

Method	Metric	Raccoon	Fractals	Cartoon	Wolf	Butterfly	Dog	Forest
B-Spline	PSNR	**25.65**	**24.46**	**23.73**	**31.02**	**24.00**	**28.80**	**20.97**
	SSIM	**0.82**	**0.81**	**0.79**	**0.91**	**0.86**	**0.89**	**0.64**
Fourier	PSNR	23.29	20.82	19.78	23.37	19.70	23.77	18.90
	SSIM	0.71	0.67	0.54	0.69	0.64	0.69	0.51
Symbolic	PSNR	25.48	24.40	23.51	30.58	23.86	28.29	20.91
	SSIM	0.81	0.80	0.77	0.90	0.84	0.87	**0.64**

Table 4 The comparison with other methods that using few data (BM3D-PnP and DIP-PnP) with the KAN prior in the Plug-and-Play framework on different number of iterations.

		Koala		Squirrel		Bird		Fractals	
		2×	4×	2×	4×	2×	4×	2×	4×
100 iter	BM3D-PnP	9.06	8.36	16.04	15.05	11.42	10.87	16.07	11.56
	DIP-PnP	14.81	14.55	18.57	18.01	16.31	15.81	16.14	16.38
200 iter	BM3D-PnP	9.13	8.54	16.91	15.70	11.66	11.35	17.07	12.21
	DIP-PnP	21.57	18.39	23.84	20.70	22.18	20.06	21.58	18.45
5 iter	KAN-PnP	**25.00**	**20.56**	**27.07**	**23.05**	**26.43**	**22.19**	**25.47**	**21.18**

particularly on texture-rich datasets such as Raccoon and Fractals. The limitations of the Fourier basis may stem from its sensitivity to discontinuities and noise, which can hinder its ability to reconstruct images with complex structures or textures. While Symbolic performs well overall, it tends to lag behind B-Spline on datasets requiring higher levels of smoothness and adaptability, such as the Wolf and Butterfly samples.

• **KAN-PnP against other methods with fewer data.** The results in Table 4 compare KAN-PnP with other schemes such as BM3D-PnP, which relies on classical denoisers with minimal data, and DIP-PnP, which uses Deep Image Prior to learn priors with small datasets (typically around 15 images). This comparison highlights KAN-PnP's ability to converge faster and achieve superior performance, even in settings where data is extremely limited. While DIP-PnP follows a philosophy closer to KAN-PnP than BM3D-PnP, its reliance on small datasets for training the denoiser sets it apart, emphasising KAN-PnP's unique ability to perform robustly with a single-instance prior. KAN-PnP consistently achieves the highest PSNR values across all datasets and scaling factors, outperforming both BM3D-PnP and DIP-PnP. These results highlight the strength of KAN-PnP in reconstructing high-quality images with a single-instance prior, where BM3D-PnP struggles with limited flexibility, and DIP-PnP is constrained by its dependence on small training datasets.

KAN-PnP converges in just 5 iterations, demonstrating remarkable efficiency compared to the 100 or 200 iterations required by BM3D-PnP and DIP-PnP. This rapid

convergence is a significant advantage, especially in scenarios where computational resources or time are limited. The ability to achieve superior performance in fewer iterations can be attributed to the Kolmogorov-Arnold Network (KAN), which provides a mathematically grounded, adaptive framework that captures intricate features and textures efficiently.

4 Conclusion

In this work, we introduced KAN-PnP, an optimisation framework that integrated Kolmogorov-Arnold Networks (KANs) as denoisers within the Plug-and-Play (PnP) paradigm. KAN-PnP was specifically designed to tackle inverse problems with single-instance priors, where only a single noisy observation was available, thus eliminating the need for large datasets typically required by traditional denoising methods. We demonstrated that KANs, based on the Kolmogorov-Arnold representation theorem, effectively served as priors in these settings, offering a robust solution for denoising. We provided theoretical guarantees for KAN-PnP, proving its convergence under key assumptions. Extensive experimental results showed that KAN-PnP outperformed existing denoising methods, providing superior performance in single-shot learning scenarios with minimal data. The method exhibited reliable convergence, achieving high accuracy with fewer iterations. Comparisons with state-of-the-art approaches further highlighted KAN-PnP's advantages in both convergence speed and solution quality, demonstrating its potential for solving real-world inverse problems with limited data. An interesting direction for future work lies in explicitly enforcing strict non-expansiveness of the KAN denoiser (i.e., ensuring $L_H \leq 1$) through architectural constraints. This could lead to tighter theoretical convergence guarantees and further improve the empirical stability of Plug-and-Play methods in the single-instance setting.

Acknowledgements

YC is funded by an AstraZeneca studentship and a Google studentship. CBS acknowledges support from the Philip Leverhulme Prize, the Royal Society Wolfson Fellowship, the EPSRC advanced career fellowship EP/V029428/1, EPSRC grants EP/S026045/1 and EP/T003553/1, EP/N014588/1, EP/T017961/1, the Wellcome Innovator Awards 215733/Z/19/Z and 221633/Z/20/Z, the European Union Horizon 2020 research and innovation programme under the Marie Skodowska-Curie grant agreement No. 777826 NoMADS, the Cantab Capital Institute for the Mathematics of Information and the Alan Turing Institute. AIAR gratefully acknowledges the support from Yau Mathematical Sciences Center, Tsinghua University.

References

[1] Fessler, J.A.: Model-based image reconstruction for mri. IEEE signal processing magazine **27**(4), 81–89 (2010)

[2] Elbakri, I.A., Fessler, J.A.: Statistical image reconstruction for polyenergetic x-ray computed tomography. IEEE transactions on medical imaging **21**(2), 89–99 (2002)

[3] Tikhonov, A.N., Goncharsky, A., Bloch, M.: Ill-posed problems in the natural sciences. Mir (1987)

[4] Engl, H.: Regularization of inverse problems. Mathematics and its Applications **375** (1996)

[5] Chambolle, A., Pock, T.: An introduction to continuous optimization for imaging. Acta Numerica **25**, 161–319 (2016)

[6] Gribonval, R., Nikolova, M.: A characterization of proximity operators. Journal of Mathematical Imaging and Vision **62**(6), 773–789 (2020)

[7] Gabay, D., Mercier, B.: A dual algorithm for the solution of nonlinear variational problems via finite element approximation. Computers & mathematics with applications **2**(1), 17–40 (1976)

[8] Venkatakrishnan, S.V., Bouman, C.A., Wohlberg, B.: Plug-and-play priors for model based reconstruction. In: 2013 IEEE Global Conference on Signal and Information Processing, pp. 945–948 (2013). IEEE

[9] Meinhardt, T., Moller, M., Hazirbas, C., Cremers, D.: Learning proximal operators: Using denoising networks for regularizing inverse imaging problems. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1781–1790 (2017)

[10] Ryu, E., Liu, J., Wang, S., Chen, X., Wang, Z., Yin, W.: Plug-and-play methods provably converge with properly trained denoisers. In: International Conference on Machine Learning, pp. 5546–5557 (2019). PMLR

[11] Sun, Y., Wohlberg, B., Kamilov, U.S.: An online plug-and-play algorithm for regularized image reconstruction. IEEE Transactions on Computational Imaging **5**(3), 395–408 (2019)

[12] Hurault, S., Leclaire, A., Papadakis, N.: Proximal denoiser for convergent plug-and-play optimization with nonconvex regularization. In: International Conference on Machine Learning, pp. 9483–9505 (2022). PMLR

[13] Hurault, S., Leclaire, A., Papadakis, N.: Gradient step denoiser for convergent plug-and-play. arXiv preprint arXiv:2110.03220 (2021)

[14] Teodoro, A.M., Bioucas-Dias, J.M., Figueiredo, M.A.: A convergent image fusion algorithm using scene-adapted gaussian-mixture-based denoising. IEEE Transactions on Image Processing **28**(1), 451–463 (2018)

[15] Yuan, X., Liu, Y., Suo, J., Dai, Q.: Plug-and-play algorithms for large-scale snapshot compressive imaging. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1447–1457 (2020)

[16] Zhang, K., Zuo, W., Gu, S., Zhang, L.: Learning deep cnn denoiser prior for image restoration. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3929–3938 (2017)

[17] Ono, S.: Primal-dual plug-and-play image restoration. IEEE Signal Processing Letters **24**(8), 1108–1112 (2017)

[18] Wei, K., Aviles-Rivero, A., Liang, J., Fu, Y., Schönlieb, C.-B., Huang, H.: Tuning-free plug-and-play proximal algorithm for inverse imaging problems. In: International Conference on Machine Learning, pp. 10158–10169 (2020). PMLR

[19] Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3-d transform-domain collaborative filtering. IEEE Transactions on image processing **16**(8), 2080–2095 (2007)

[20] Jordan, C.: Sur la series de fourier. CR Acad. Sci., Paris **92**, 228–230 (1881)

[21] Laumont, R., Bortoli, V.D., Almansa, A., Delon, J., Durmus, A., Pereyra, M.: Bayesian imaging using plug & play priors: when langevin meets tweedie. SIAM Journal on Imaging Sciences **15**(2), 701–737 (2022)

[22] Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L.: Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. IEEE transactions on image processing **26**(7), 3142–3155 (2017)

[23] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, pp. 234–241 (2015). Springer

[24] Goujon, A., Neumayer, S., Unser, M.: Learning weakly convex regularizers for convergent image-reconstruction algorithms. SIAM Journal on Imaging Sciences **17**(1), 91–115 (2024)

[25] Ulyanov, D., Vedaldi, A., Lempitsky, V.: Deep image prior. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 9446–9454 (2018)

[26] Sun, Z., Latorre, F., Sanchez, T., Cevher, V.: A plug-and-play deep image prior. In: ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 8103–8107 (2021). IEEE

[27] Cheng, Y., Zhang, L., Shen, Z., Wang, S., Yu, L., Chan, R.H., Schönlieb, C.-B., Aviles-Rivero, A.I.: Single-shot plug-and-play methods for inverse problems. Transactions on Machine Learning Research (2024)

[28] Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., Halverson, J., Soljačić, M., Hou, T.Y., Tegmark, M.: Kan: Kolmogorov-arnold networks. arXiv preprint arXiv:2404.19756 (2024)

[29] Buades, A., Coll, B., Morel, J.-M.: A non-local algorithm for image denoising. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 2, pp. 60–65 (2005). Ieee

[30] Kolmogorov, A.N.: On the representation of continuous functions of several variables by superpositions of continuous functions of a smaller number of variables. American Mathematical Society (1961)

[31] Kolmogorov, A.N.: On the representations of continuous functions of many variables by superposition of continuous functions of one variable and addition. In: Dokl. Akad. Nauk USSR, vol. 114, pp. 953–956 (1957)

[32] Braun, J., Griebel, M.: On a constructive proof of kolmogorov's superposition theorem. Constructive approximation **30**, 653–675 (2009)

[33] Bevilacqua, M., Roumy, A., Guillemot, C., Alberi-Morel, M.L.: Low-complexity single-image super-resolution based on nonnegative neighbor embedding (2012)

[34] Zeyde, R., Elad, M., Protter, M.: On single image scale-up using sparse-representations. In: Curves and Surfaces: 7th International Conference, Avignon, France, June 24-30, 2010, Revised Selected Papers 7, pp. 711–730 (2012). Springer

[35] Lai, Z., Wei, K., Fu, Y., Härtel, P., Heide, F.: δ-prox: Differentiable proximal algorithm modeling for large-scale optimization. ACM Transactions on Graphics (TOG) **42**(4), 1–19 (2023)

[36] Krull, A., Buchholz, T.-O., Jug, F.: Noise2void-learning denoising from single noisy images. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2129–2137 (2019)

[37] Zhang, K., Zuo, W., Zhang, L.: Ffdnet: Toward a fast and flexible solution for cnn-based image denoising. IEEE Transactions on Image Processing **27**(9), 4608–4622 (2018)