

GSNR: Graph Smooth Null-Space Representation for Inverse Problems

Romario Gualdrón-Hurtado*, Roman Jacome*, Rafael S. Suárez, Henry Arguello
Universidad Industrial de Santander, Colombia, 680002

{yesid2238324, rajaccar, rafael2269082}@correo.uis.edu.co, henarfu@uis.edu.co

Abstract

Inverse problems in imaging are ill-posed, leading to infinitely many solutions consistent with the measurements due to the non-trivial null-space of the sensing matrix. Common image priors promote solutions on the general image manifold, such as sparsity, smoothness, or score function. However, as these priors do not constrain the null-space component, they can bias the reconstruction. Thus, we aim to incorporate meaningful null-space information in the reconstruction framework. Inspired by smooth image representation on graphs, we propose Graph-Smooth Null-Space Representation (GSNR), a mechanism that imposes structure only into the invisible component. Particularly, given a graph Laplacian, we construct a null-restricted Laplacian that encodes similarity between neighboring pixels in the null-space signal, and we design a low-dimensional projection matrix from the p -smoothest spectral graph modes (lowest graph frequencies). This approach has strong theoretical and practical implications: i) improved convergence via a null-only graph regularizer, ii) better coverage, how much null-space variance is captured by p modes, and iii) high predictability, how well these modes can be inferred from the measurements. GSNR is incorporated into well-known inverse problem solvers, e.g., PnP, DIP, and diffusion solvers, in four scenarios: image deblurring, compressed sensing, demosaicing, and image super-resolution, providing consistent improvement of up to 4.3 dB over baseline formulations and up to 1 dB compared with end-to-end learned models in terms of PSNR.

1. Introduction

Recovering an image from a limited number of noisy measurements requires estimating a solution to an undetermined inverse problem, which is of the form

$$\mathbf{y} = \mathbf{H}\mathbf{x}^* + \boldsymbol{\omega}, \quad \boldsymbol{\omega} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}). \quad (1)$$

where $\mathbf{y} \in \mathbb{R}^m$ denotes the measurement vector, $\mathbf{x}^* \in \mathbb{R}^n$ is the unknown target image, $\mathbf{H} \in \mathbb{R}^{m \times n}$ is the sensing matrix with $m \leq n$, and $\boldsymbol{\omega}$ models additive Gaussian noise. This formulation is used in a broad range of imaging tasks

by using a proper structure of $\mathbf{H}$. Compressed sensing (CS) tasks employ dense sensing matrices [5]. Image restoration problems, such as deblurring [53], inpainting [51], and super-resolution [46], are modeled by structured Toeplitz or block-diagonal matrices. In medical imaging, modalities like magnetic resonance imaging (MRI) [26] and computed tomography (CT) [2] use structured physics-based operators (e.g., undersampled Fourier or Radon transforms). The formulation in (1) becomes ill-posed due to matrix instability and/or the undersampling scheme. Therefore, incorporating prior knowledge or regularization is essential to promote reconstructions that are consistent with the expected structure of $\mathbf{x}^*$ [4]. The estimation of $\mathbf{x}^*$ is commonly formulated as a variational problem that combines a data-fidelity term and a regularization term:

$$\hat{\mathbf{x}} = \arg \min_{\tilde{\mathbf{x}}} \underbrace{\frac{1}{2} \|\mathbf{H}\tilde{\mathbf{x}} - \mathbf{y}\|_2^2}_{\text{data fidelity: } g(\tilde{\mathbf{x}})} + \eta \underbrace{f(\tilde{\mathbf{x}})}_{\text{prior}} \quad (2)$$

where $g(\tilde{\mathbf{x}})$ enforces consistency with the measurements and $f(\tilde{\mathbf{x}})$ incorporates prior knowledge about the structure of $\mathbf{x}^*$. Recent advances have introduced data-driven regularizers for $f(\tilde{\mathbf{x}})$, most notably Plug-and-Play (PnP) [6, 20, 48, 52] and Regularization by Denoising (RED) [9, 34, 35], which have achieved remarkable success across a wide range of imaging tasks [20] by implicitly promoting solutions within the training image dataset manifold i.e., well-trained denoiser learns the data score-function [27]. PnP and RED use learned denoisers, allowing the integration of powerful deep restoration networks within algorithms [27, 44]. Despite their versatility and success, these priors operate in the image domain and do not explicitly exploit the structure of $\mathbf{H}$, including its null-space (NS). Only a portion of $\mathbf{x}^*$ is observable through $\mathbf{H}$. The range-null-space decomposition (RNSD) states that any $\mathbf{x} \in \mathbb{R}^n$ vector is decomposed as $\mathbf{x} = \mathbf{x}_r + \mathbf{x}_n$, where $\mathbf{x}_r = \mathbf{P}_r \mathbf{x}$ denotes the component lying in the range space of $\mathbf{H}$, with $\mathbf{P}_r = \mathbf{H}^\dagger \mathbf{H}$, and $\mathbf{x}_n = \mathbf{P}_n \mathbf{x}$ denotes the component lying in its NS, with the NS projector $\mathbf{P}_n = \mathbf{I} - \mathbf{H}^\dagger \mathbf{H}$. $\mathbf{x}_n$ is invisible to $\mathbf{H}$ because it belongs to its NS, defined as

$$\begin{aligned} \text{Null}(\mathbf{H}) &= \{\mathbf{x} \in \mathbb{R}^n : \mathbf{H}\mathbf{x} = \mathbf{0}\} \\ &= \{\mathbf{x} : \mathbf{x} \perp \mathbf{h}_j, \forall j \in \{1, \dots, m\}\}, \end{aligned}$$

*Equal contribution.

where $\mathbf{h}_j \in \mathbb{R}^n$ is the j -th row of $\mathbf{H}$. This interpretation has inspired approaches that explicitly exploit $\text{Null}(\mathbf{H})$. Null-Space Networks (NSN) [41] and Deep Decomposition [7] operate across the entire $\text{Null}(\mathbf{H})$, treating all invisible directions as equally relevant, neglecting the fact that natural, perceptually plausible images occupy only a low-dimensional manifold within that space [17]. To address this gap, the recent Nonlinear Projections of the Null-Space (NPN) framework [17] introduces a learnable low-dimensional NS projection matrix $\mathbf{S} \in \mathbb{R}^{p \times n}$, with $p \leq (n - m)$. $\mathbf{S}$ is designed depending on the inverse problem, making its rows orthogonal to $\mathbf{H}$, then $\mathbf{S}$ is optimized jointly a NS component predictor $\mathbf{G}(\mathbf{y}) \equiv \mathbf{S}\mathbf{x}^*$ to adapt the $\text{span}(\mathbf{S})$ into a subspace of $\text{Null}(\mathbf{H})$. During reconstruction, NPN penalizes deviations $\|\mathbf{G}(\mathbf{y}) - \mathbf{S}\tilde{\mathbf{x}}\|_2^2$. However, natural images occupy a low-dimensional, structured subset inside $\text{Null}(\mathbf{H})$, so blindly learning an arbitrary $\text{span}(\mathbf{S}^\top) \subset \text{Null}(\mathbf{H})$ can waste capacity and induce bias. In this work, we present a principled framework that provides a new representation of the NS, enabling the optimal selection of informative NS directions. To address this issue, we took into account two main optimality criteria:

Coverage: *How much of the null-space of $\mathbf{H}$ is represented with the projection matrix $\mathbf{S}$.*

Predictability: *How easily can the projections $\mathbf{S}\mathbf{x}$ be estimated from the measurements $\mathbf{y}$ only.*

Inspired by the graph (smooth) representation of images [28, 29], we introduce Graph-Smooth Null-Space Representation (GSNR) to overcome the shortcomings of NS representation and prediction. Different from the well-studied graph-structure preserving approach in imaging inverse problems [19, 28] that promotes smoothness over the image $\mathbf{x}$, our approach is the first to endow the NS component $\mathbf{x}_n$ with a graph structure, defined through a sparse, positive semidefinite graph Laplacian $\mathbf{L} \succeq \mathbf{0}$, such as 4/8-nearest-neighbor (NN) image grid whose edge weights encode local pixel similarity. We form the *null-restricted Laplacian* $\mathbf{T} := \mathbf{P}_n \mathbf{L} \mathbf{P}_n$ on $\text{Null}(\mathbf{H})$. Via spectral graph theory, we construct the graph-smooth NS projection matrix $\mathbf{S}$ as the p -smoothest NS modes, which corresponds to the p first graph Fourier modes. The proposed GSNR changes how we solve inverse problems by providing exact structure information where the sensor is blind. By selecting the p -smoothest eigenmodes of the null-restricted Laplacian, our method (i) reduces hallucinations and bias by constraining only the invisible component, not the whole image; (ii) improves conditioning, leading to faster and more stable convergence across PnP, DIP, and diffusion solvers; (iii) increases data efficiency by capturing most null-space variation with small p (high coverage) and by targeting the modes that are easiest to infer from measurements (high predictability); and (iv) offers operational di-

agnostics coverage/predictability curves that allows selecting p and justify regularization objectively. Because it is plug-compatible with standard pipelines and agnostic to the forward operator, the approach scales across different imaging tasks, turning an ill-posed ambiguity into a structured, measurable, and learnable component. The proposed construction has the following benefits:

1. Selected graph-smooth null-space components produce high coverage at low dimensionality projection, i.e., small p (Theorem 1 and Theorem 2).
2. Predicting the p null-space components from $\mathbf{y}$ using a GSNR is easier than the ones obtained with a common null-space basis (Proposition 1).
3. Predicted graph null-space components can be incorporated via regularization into solvers such as PnP, Deep Image Prior, and Diffusion Models.

2. Related Work

Variational and learned priors. A common scheme to solve (2) is proximal gradient descent (PGD) [30]. Recently, PnP [20] and RED [35] methods have replaced the proximal operator of $f(\tilde{\mathbf{x}})$ with learned denoisers, enabling the incorporation of data-driven priors without requiring an explicit analytic form [17]. Although these frameworks achieve state-of-the-art performance, they leave $\text{Null}(\mathbf{H})$ largely unconstrained; denoisers may freely modify components that are invisible to the sensing operator.

Nonlinear Projections of the Null-Space (NPN). To explicitly regularize image unobserved components, [17] introduced a task-specific low-dimensional projection within $\text{Null}(\mathbf{H})$, enforcing consistency between learned NS predictions and reconstructions. Building on this principle, we propose GSNR, which incorporates graph-smoothness analysis to guide the selection of a structured null-subspace. Our design ensures that the learned projection aligns with the intrinsic geometry of the data, yielding NS components that are both predictable from the measurements and consistent with the graph-induced smoothness prior.

Graph theory in imaging inverse problems. Variational regularization methods address inverse problems by imposing smoothness or sparsity priors that constrain the solution space. Total variation [38] enforces pixel-wise smoothness, whereas wavelet-based sparsity priors [13] promote sparse representations of images in the transform domain via soft-thresholding. Although these priors improve stability, they are typically limited to local structures and often oversmooth fine textures or complex geometries. To better capture local dependencies, graph-based and non-local regularization methods have been reintroduced. [32] proposed an adaptive nonlocal regularization framework that jointly estimates the image and the graph that encodes patch similarities. However, like classical priors, this framework operates solely on the reconstruction and doesn't constrain the image component lying in $\text{Null}(\mathbf{H})$.

3. Graph-Smooth Null-Space Representation

3.1. Preliminaries: Graphs

Graphs provide a geometric representation of data, where vertices correspond to data elements (e.g., pixels or image patches) and weighted edges encode pairwise similarities between them [42]. Signals defined on weighted undirected graphs are mathematically denoted by $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$, where $\mathcal{V}$ is the vertex set with $|\mathcal{V}| = n$, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the edge set, and $\mathbf{W} \in \mathbb{R}^{n \times n}$ is the symmetric weighted adjacency matrix with $\mathbf{W}_{ij} > 0$ if $(i, j) \in \mathcal{E}$ and $\mathbf{W}_{ij} = 0$ otherwise. The unnormalized graph Laplacian is

$$\mathbf{L} = \mathbf{D} - \mathbf{W}, \quad (\mathbf{L}\mathbf{x})_i = \sum_j \mathbf{W}_{ij}(\mathbf{x}_i - \mathbf{x}_j), \quad (3)$$

where the diagonal degree matrix $\mathbf{D} = \text{diag}(d_1, \dots, d_n)$ accomplishes $d_i = \sum_j \mathbf{W}_{ij}$. Its Dirichlet energy is

$$\mathbf{x}^\top \mathbf{L} \mathbf{x} = \frac{1}{2} \sum_{i,j} \mathbf{W}_{ij} (\mathbf{x}_i - \mathbf{x}_j)^2, \quad (4)$$

which penalizes discrepancies across edges with large weights. For symmetric $\mathbf{W}$, $\mathbf{L}$ is real symmetric positive semidefinite (SPSD) with eigenpairs $\{(\lambda_\ell, \mathbf{v}_\ell)\}_{\ell=0}^{n-1}$, $0 = \lambda_0 \leq \lambda_1 \leq \dots$, $\mathbf{v}_0 \propto \mathbf{1}$ on connected graphs. Eigenvectors with small λ_ℓ vary slowly across high-weight edges, motivating the interpretation of λ as a graph frequency. For the usage of $\mathbf{L}$, we focus on two graph topologies [29]:

1. **4NN**: This promotes piecewise smoothness aligned with the coordinate axes and attenuates horizontal/vertical fluctuations, yielding an anisotropic (axis-biased) smoothing effect. Here, $\mathcal{E}_4 = \{(i, j) : \|\mathbf{p}_i - \mathbf{p}_j\|_1 = 1\}$, $\mathbf{W}_{ij} = \mathbf{1}_{(i,j) \in \mathcal{E}_4}$, with $\mathbf{p}_i \in \mathbb{Z}^2$ pixel coordinates.
2. **8NN**: This reduces orientation bias and more closely approximates an isotropic Laplace operator, promoting rotation-invariant piecewise smoothness and suppressing fluctuations uniformly across directions. Here, $\mathcal{E}_8 = \{(i, j) : \|\mathbf{p}_i - \mathbf{p}_j\|_\infty = 1\}$, $\mathbf{W}_{ij} = w_o \mathbf{1}_{\|\cdot\|_\infty=1} + w_d \mathbf{1}_{\|\cdot\|_2=\sqrt{2}}$, where $w_d = w_o/\sqrt{2}$.

Nevertheless, we extend the analysis of the proposed approach to other topologies in Supp. A4.1.

3.2. Graph-smooth null modes

We study smoothness within the NS by restricting the Laplacian to

$$\mathbf{T} = \mathbf{P}_n \mathbf{L} \mathbf{P}_n, \quad (5)$$

the reason for using the null-graph operator $\mathbf{T} \in \mathbb{R}^{n \times n}$ stems from the need to pay attention to certain geometric properties of the image's NS, which are extracted by $\mathbf{L}$. Let $\mathbf{T}\mathbf{x} = \mathbf{P}_n \mathbf{L} \mathbf{P}_n \mathbf{x} = \mathbf{P}_n \mathbf{L} \mathbf{x}_n$, where $\mathbf{L}\mathbf{x}_n$ extracts, areas of greatest difference in the NS, and then re-project it to the NS with $\mathbf{P}_n$, thus ensuring that this operation results in a projection from the NS. In Fig. 1, from left to right, we show the ground-truth $\mathbf{x}^*$, back-projection $\mathbf{H}^\top \mathbf{y}$, true null component $\mathbf{P}_n \mathbf{x}^*$, and two graph NS projections $\mathbf{P}_n \mathbf{L}_{8nn} \mathbf{P}_n \mathbf{x}^*$ and $\mathbf{P}_n \mathbf{L}_{4nn} \mathbf{P}_n \mathbf{x}^*$, for super-resolution (SR).



Figure 1. For image SR task with $SRF = 4$, $n = 3 \cdot 128^2$, we show ground-truth, adjoint reconstruction, NS projection, projection onto graph-smooth NS with $\mathbf{L}_{4nn}$ and $\mathbf{L}_{8nn}$.

The map $\mathbf{P}_n \mathbf{x}^*$ isolates signal content invisible to the sensor, while $\mathbf{P}_n \mathbf{L} \mathbf{P}_n \mathbf{x}^*$ highlights where $\mathbf{L}\mathbf{x}_n$ falls into the NS. These graph projections highlight the smoothest *blind* signal components, e.g., textures near edges.

Then, by applying eigenvalue decomposition, we see that $\mathbf{T} = \mathbf{V} \text{diag}(\mu_1, \dots, \mu_n) \mathbf{V}^\top$ is SPSP, with $0 \leq \mu_1 \leq \dots \leq \mu_n$, where $\mathbf{V} \in \mathbb{R}^{n \times n}$ is orthogonal. Note that $\mathbf{V}$ diagonalizes $\mathbf{T}$, so its columns $\{\mathbf{v}_i\}$ are the *graph-Fourier modes within* $\text{Null}(\mathbf{H})$. If μ_i is smaller, then a smoother graph-mode is selected. $\mathbf{V}$ has the following properties: i) $\mathbf{V}$ has orthonormal columns $\mathbf{v}_i \in \text{Null}(\mathbf{H})$, μ_i -ordered. ii) These $\{\mathbf{v}_i\}$ are *graph-smooth null modes* probing $\mathbf{x}$. The graph-smooth NS projector $\mathbf{S}$ with p rows is obtained by taking the first p columns of $\mathbf{V}$:

$$\mathbf{S} = \overbrace{\mathbf{V}^\top}^{\mathbf{v}_p^\top}[:, 1:p]^\top = [\mathbf{v}_1, \dots, \mathbf{v}_p]^\top \in \mathbb{R}^{p \times n}. \quad (6)$$

Note that $\mathbf{S}$ keeps only the p -smoothest null directions and $\mathbf{S}\mathbf{x}$ expresses the NS part of $\mathbf{x}$ in an orthonormal basis sorted by increasing graph frequency. Particularly, based on RNSD: $\mathbf{S}\mathbf{x} = \mathbf{S}(\mathbf{P}_r + \mathbf{P}_n)\mathbf{x} = \mathbf{S}\mathbf{P}_r \mathbf{x} + \mathbf{S}\mathbf{P}_n \mathbf{x}$, since rows of $\mathbf{S}$ lie in $\text{Null}(\mathbf{H})$, hence $\mathbf{S}\mathbf{P}_r = \mathbf{0}$, leading to

$$\mathbf{S}\mathbf{x} = \mathbf{0} + \mathbf{S}\mathbf{P}_n \mathbf{x} = \mathbf{V}_p^\top \mathbf{P}_n \mathbf{x} = \begin{bmatrix} \langle \mathbf{v}_1, \mathbf{x}_n \rangle \\ \vdots \\ \langle \mathbf{v}_p, \mathbf{x}_n \rangle \end{bmatrix}. \text{ Here } \{\mathbf{v}_i\}_{i=1}^p$$

are the p Laplacian-smoothest orthonormal null modes; thus $\mathbf{S}\mathbf{x}$ collects the coefficients of $\mathbf{x}$ along these p directions.

3.3. Learning low-dimensional GSNR

We train the GSNR predictor $\mathbf{G}$ following:

$$\mathbf{G}^* = \min_{\mathbf{G}} \mathbb{E}[\|\mathbf{G}(\mathbf{y}) - \mathbf{S}\mathbf{x}^*\|_2^2]. \quad (7)$$

Here, we aim to predict the low-dimensional GSNR using the neural network $\mathbf{G}$. Subsequently, the theoretical foundations will be laid that greater predictability will guarantee better convergence in the task.

3.4. Reconstruction objective

Thus, we incorporate our method as

$$\min_{\tilde{\mathbf{x}}} g(\tilde{\mathbf{x}}) + \eta f(\tilde{\mathbf{x}}) + \gamma \|\mathbf{G}^*(\mathbf{y}) - \mathbf{S}\tilde{\mathbf{x}}\|_2^2 + \frac{\gamma_g}{2} \tilde{\mathbf{x}}^\top \overbrace{\mathbf{T}}^{\phi(\tilde{\mathbf{x}})} \tilde{\mathbf{x}}, \quad (8)$$

which includes the learned GSNR term, plus a graph regularizer acting only on the null component of the reconstruction, weighted by γ and γ_g , respectively. This retains compatibility with proximal/PnP/DM solvers.

4. Theory and Analysis

We develop a theoretical analysis on GSNR based on *null-restricted spectral* formulation that turns GSNR design/performance prediction into ordered-eigenvalue criteria on $\mathbf{T}$ for optimal subspace selection, obtaining a principled eigen-ordering and criteria used to choose $\mathbf{S}$, p , $\mathbf{L}$. The theoretical results are summarized as:

Coverage: the first p eigenmodes of $\mathbf{T}$ provide an optimal graph-smooth representation of $\text{Null}(\mathbf{H})$, by covering a high null-space spectrum energy with $p \ll n - m$.

Minimax optimality: the same GSNR basis is worst-case optimal over a graph-energy ellipsoid in $\text{Null}(\mathbf{H})$, giving a principled ordering of null directions

Predictability: via block-precision/Schur-complement analysis, we show that GSNR has a higher per-mode recoverability from measurements $\mathbf{y}$ than plain null-space projections.

Convergence: the effect of the graph regularizer provides better conditioning and convergence properties based on fixed-point analysis.

4.1. Null-space coverage

Before developing the formal analysis on the coverage properties of the NS projections, we validate the representation performance of three settings: i) $\mathbf{L} = \mathbf{I}$, ii) $\mathbf{L}_{4nn}$, and iii) $\mathbf{L}_{8nn}$. Note that the first setting is equivalent to just taking the NS basis. In Fig. 2 (a), each row shows projections ($\mathbf{x} = (\mathbf{H}^\top \mathbf{H} + \mathbf{S}^\top \mathbf{S})\mathbf{x}^*$) as the dimension of the null-space subspace increases, $p \in \{153, 1536, 4608, 9216, 15360\}$, for three graph choices in the design of $\mathbf{S}$: (top, orange) $\mathbf{L} = \mathbf{I}$; (middle, green) grid $\mathbf{L}_{8nn}$; (bottom, blue) grid $\mathbf{L}_{4nn}$. For small p , the grid Laplacians already yield visually plausible faces, while $\mathbf{L} = \mathbf{I}$ still produces noisy, poorly structured estimates. The number printed above each image is the MSE of the representation. The plot at the bottom shows the MSE for all values of $p = \{1, \dots, n - m\}$, showing that the graph-null-space projections provide better signal representation with smaller p than just the NS basis. In general, to measure the amount of energy from the NS that the low-dimensional operator $\mathbf{S}$ can capture, we use the spectral coverage defined as:

$$C(p) = \frac{\text{tr}(\mathbf{S}\text{Cov}(\mathbf{x}_n)\mathbf{S}^\top)}{\text{tr}(\text{Cov}(\mathbf{x}_n))} \in [0, 1], \quad (9)$$

where $\text{Cov}(\mathbf{x})$ is the covariance matrix of $\mathbf{x}$. View the Laplacian eigenvectors as a graph-Fourier basis: the first modes are smooth, low “graph-frequency” patterns, while later modes are increasingly oscillatory [29, 42]. The $C(p)$ indicates what fraction of the total NS variance lies in the span of the first p graph-Fourier modes retained by $\mathbf{S}$, similarly as explained variance in PCA [18]. Intuitively, if the signal is graph-smooth, most variability concentrates in low

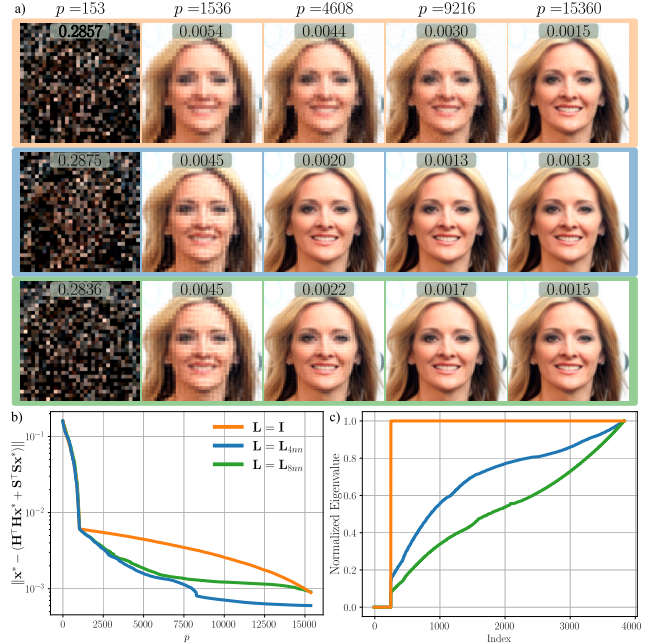


Figure 2. Coverage and spectral analysis for SR task. (a) $\mathbf{x}^* = \mathbf{H}^\top \mathbf{H}\mathbf{x}^* + \mathbf{S}^\top \mathbf{S}\mathbf{x}^*$ for $\mathbf{S}$ given by (6) using $\mathbf{L} = \mathbf{I}$ (orange), $\mathbf{L} = \mathbf{L}_{4nn}$ (green) and $\mathbf{L} = \mathbf{L}_{8nn}$ (blue) for different values of p . (b) RNSD representation error, varying p . (c) Variation of $\mathbf{T}$ normalized eigenvalues with respect to their index.

frequencies, so coverage climbs rapidly with small p ; if not, coverage grows more slowly, like in PCA. A4.2 shows coverage results and eigenvalue analysis that validate the usefulness of graph Laplacians, rather than a learnable matrix. Mathematical insights for the selection of the appropriate p value depending on the coverage are given in A5.

Definition 1 (Gaussian Markov Random Field [39]) A GMRF on the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a zero-mean Gaussian vector $\mathbf{x} \in \mathbb{R}^n$ with precision matrix $\mathbf{Q} = \Sigma^{-1} \in \mathbb{R}^{n \times n}$, with $\Sigma = \text{Cov}(\mathbf{x})$, is sparse and encodes the conditional independence of the graph

$$p(\mathbf{x}) \propto \exp\left(-\frac{1}{2}\mathbf{x}^\top \mathbf{Q}\mathbf{x}\right), \quad \mathbf{Q}_{i,j} = 0 \iff (i,j) \notin \mathcal{E},$$

where smoothness is required in neighboring pixels, we define the precision matrix $\mathbf{Q} = \alpha\mathbf{L} + \epsilon\mathbf{I}$ with $\alpha > 0, \epsilon > 0$.

Under this setting, we derive the following guarantee on improved NS coverage.

Theorem 1 (Coverage for graph-smooth null-space)

Consider the construction of $\mathbf{T}$ in (5). The covariance of the NS, $\text{Cov}(\mathbf{x}_n)$ is a spectral function of $\mathbf{T}$, i.e., $\text{Cov}(\mathbf{x}_n) = \mathbf{V}\text{diag}(\lambda_1, \dots, \lambda_n)\mathbf{V}^\top$, where $\lambda_i = \frac{1}{\alpha\mu_i + \epsilon}$, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. With the construction of $\mathbf{S}$ in (6), the coverage of the NS using GMRF with a Laplacian matrix $\mathbf{L}$ as $C_L(p)$ and the coverage when $\mathbf{L} = \mathbf{I}$ is denoted $C_I(p)$, satisfies for every $p = 1, \dots, n - m$.

$$C_L(p) \geq C_I(p).$$

The proof can be found in Supp. A1.

We illustrate this effect in Fig. 2(b)-(c). In 2(b), the representation error of the RNSD decreases more rapidly when using $\mathbf{L}_{8nn}$ and $\mathbf{L}_{4nn}$ compared to $\mathbf{I}$. In 2(c), the eigenvalue spectra of $\mathbf{T}$ reveal that, unlike the flat spectrum of the pure NS projector $\mathbf{P}_n$, graph Laplacians exhibit smoothly increasing eigenvalues. From Theorem 1, a flat spectrum ($\mu_i = 1$ for all i as with $\mathbf{L} = \mathbf{I}$) yields $C_{\mathbf{I}}(p) = p/q$, so coverage grows only linearly with p , whereas Laplacians with increasing $\{\mu_i\}$ achieve faster coverage growth.

Our second theoretical result is a minimax optimality on the NS coverage by the NS scheme. First, define the feasible set of NS components

$$\mathcal{M}_\tau := \{\mathbf{x}_n \in \text{Null}(\mathbf{H}) : \mathbf{x}_n^\top \mathbf{L} \mathbf{x}_n \leq \tau\}$$

Theorem 2 (Minimax optimality) Let $V_p = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ and P_{V_p} be its projector. Then, among all p -dimensional $V \subset \text{Null}(\mathbf{H})$, provided $\mu_{p+1} > 0$,

$$\min_{\dim(V)=p} \sup_{\mathbf{x}_n \in \mathcal{M}_\tau} \|(\mathbf{I} - P_V)\mathbf{x}_n\|_2^2 = \sup_{\mathbf{x}_n \in \mathcal{M}_\tau} \|(\mathbf{I} - P_{V_p})\mathbf{x}_n\|_2^2 = \frac{\tau}{\mu_{p+1}}.$$

The proof of this theorem is in Supp. A2.

Remark 1 Selecting $\mathbf{S}$ so that its rows are $\{\mathbf{v}_1^\top, \dots, \mathbf{v}_p^\top\}$ is provably optimal for covering the NS under the graph energy $\mathbf{x}^\top \mathbf{T} \mathbf{x}$. As p grows, the guaranteed worst-case error decays with $1/\mu_{p+1}$, which is fast for graph Laplacians (where many μ_j are small). With $\mathbf{L} = \mathbf{I}$ one gets $\mathbf{T} = \mathbf{P}_n$ (flat spectrum), hence $\mu_{p+1} = 1$ and no improvement in the bound, explaining why geometry-free designs cannot cover the NS efficiently.

4.2. Null-space projections predictability

It is also important to quantify the predictability level of the NS projections $\mathbf{S}\mathbf{x}^*$ from the acquired measurements $\mathbf{y}$. We develop a *statistical coupling* analysis via linear estimation.

Proposition 1 (Per-mode predictability bound) Let $\mathbf{x} \in \mathbb{R}^n$ be zero-mean Gaussian with covariance $\mathbf{C}$ and precision $\mathbf{Q} = \mathbf{C}^{-1}$. Let $\mathbf{H} \in \mathbb{R}^{m \times n}$ and denote the orthogonal projectors onto $\mathcal{R} := \text{Range}(\mathbf{H}^\top)$ and $\mathcal{N} := \text{Null}(\mathbf{H})$ by $\mathbf{P}_r$ and $\mathbf{P}_n$, respectively. Block $\mathbf{Q}$ and $\mathbf{C}$ with respect to the decomposition $\mathbb{R}^n = \mathcal{R} \oplus \mathcal{N}$:

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{rr} & \mathbf{Q}_{rn} \\ \mathbf{Q}_{nr} & \mathbf{Q}_{nn} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} \mathbf{C}_{rr} & \mathbf{C}_{rn} \\ \mathbf{C}_{nr} & \mathbf{C}_{nn} \end{bmatrix}.$$

Assume $\mathbf{Q}_{nn} \succ 0$ on $\mathcal{N}$ and let $\{\mathbf{v}_j\}$ be an orthonormal eigenbasis of $\mathbf{Q}_{nn}$ with $\mathbf{Q}_{nn}\mathbf{v}_j = \mu_j \mathbf{v}_j$ and $\mu_j > 0$. Define the j -th null coefficient $a_j := \mathbf{v}_j^\top \mathbf{x}_n$ where $\mathbf{x}_n := \mathbf{P}_n \mathbf{x}$. Consider measurements from (1) and denote $\mathbf{C}_y := \mathbf{H}\mathbf{C}_{rr}\mathbf{H}^\top + \sigma^2 \mathbf{I}_m$. Then the population R^2 of the optimal linear predictor of a_j from $\mathbf{y}$ satisfies

$$\rho_j^2 := \frac{\text{Cov}(a_j, \mathbf{y})^\top \mathbf{C}_y^{-1} \text{Cov}(\mathbf{y}, a_j)}{\text{Var}(a_j)} \leq \frac{c_j}{c_j + \mu_j},$$

$$c_j := \mathbf{v}_j^\top (\mathbf{Q}_{rn} \mathbf{C}_{rr} \mathbf{Q}_{nr}) \mathbf{v}_j.$$

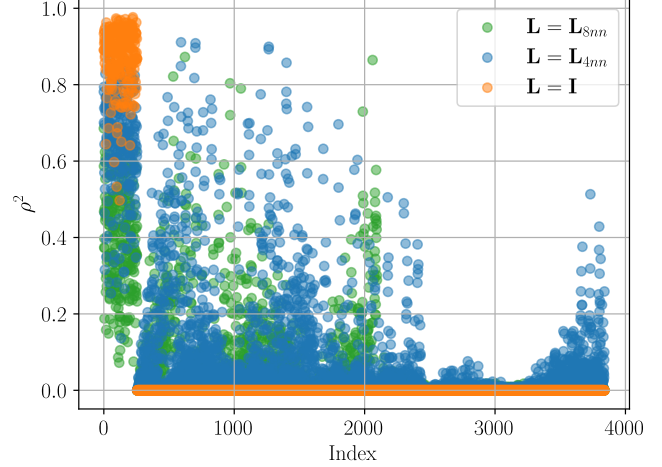


Figure 3. Per-mode predictability for each case when $\mathbf{L} \in \{\mathbf{I}, \mathbf{L}_{8nn}, \mathbf{L}_{4nn}\}$ for SR case with $\text{SRF} = 4$ with $n = 64^2$.

The proof of this proposition is found in Supp. A3.

Remark 2 If the prior is a GMRF with precision $\mathbf{Q} = \alpha \mathbf{L} + \varepsilon \mathbf{I}$ on $\mathcal{L}$, then $\mathbf{Q}_{nn} = \alpha \mathbf{T} + \varepsilon \mathbf{P}_n$ with $\mathbf{T} := \mathbf{P}_n \mathbf{L} \mathbf{P}_n$, and one can take $\{\mathbf{v}_j\}$ as the $\mathbf{T}$ -eigenmodes (graph-smooth null modes). The bound becomes

$$\rho_j^2 \leq \frac{c_j}{c_j + \alpha \lambda_j(\mathbf{T}) + \varepsilon}, \quad c_j = \mathbf{v}_j^\top (\mathbf{Q}_{rn} \mathbf{C}_{rr} \mathbf{Q}_{nr}) \mathbf{v}_j,$$

with $\mathbf{Q}_{rn} = \alpha \mathbf{L}_{rn}$. Thus smoother null modes (small $\lambda_j(\mathbf{T})$) are more predictable. Now, if $\mathbf{L} = \mathbf{I}$ (isotropic prior), then $\mathbf{Q} = \gamma \mathbf{I}$ and $\mathbf{Q}_{rn} = \mathbf{P}_r \mathbf{Q} \mathbf{P}_n = \mathbf{0}$, so $c_j = 0$ and $\rho_j^2 \leq 0$ for all j : there is no statistical coupling between $\mathcal{R}$ and $\mathcal{N}$.

To evaluate the developed theory, Fig. 3, the empirical per-mode predictability ρ_j^2 (one dot per eigenmode), for three choices of $\mathbf{L} \in \{\mathbf{L}_{8nn}, \mathbf{L}_{4nn}, \mathbf{I}\}$, where $\mathbf{L} = \mathbf{I}$ has same index ordering across panels. From indices $[0, m-1]$, the predictability is high for all settings, since they correspond to the eigenpairs related to the Range space, thus, the predictability is high. The observations are: (i) for $\mathbf{L} = \mathbf{I}$, the spectrum is constant over the NS and then drops to 0 outside of it, while both grid Laplacians exhibit a smooth, monotone growth of μ_j ; (ii) correspondingly, with grid Laplacians, a long tail of modes has non-zero predictability, (ρ_j^2 spread over a wide index range), whereas with $\mathbf{L} = \mathbf{I}$ only a small initial block of modes shows any predictability, and the rest are essentially 0. Note also that $\mathbf{L}_{4nn}$ has higher predictability than $\mathbf{L}_{8nn}$ in more eigenmodes, following the fast increase eigenvalue distribution Fig 2 (c).

Note that this analysis was using a linear estimation framework, harnessing the NS GMRF prior. In practice, the *statistical coupling* can also be achieved with the supervised learning of the predictor G in Eq. (7), but the trend of GS prior holds on this setting. To evaluate the predictability of the learned predictor, we used

$$R^2(p) \triangleq 1 - \frac{\mathbb{E} [\|\mathbf{G}^*(\mathbf{y}) - \mathbf{S}\mathbf{x}^*\|_2^2]}{\mathbb{E} [\|\mathbf{S}\mathbf{x}^*\|_2^2]}. \quad (10)$$

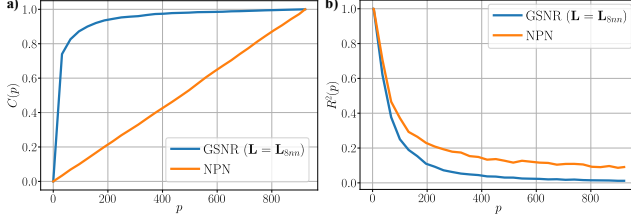


Figure 4. a) Coverage and b) Predictability vs. p for CS with CIFAR-10 dataset. In this case, $m/n = 0.1$ and $p = 1 \cdots n - m$.

4.3. Graph-based regularizer

Consider the PnP-PGD iteration solving the optimization problem (8) without the term $\|G^*(\mathbf{y}) - \mathbf{S}\mathbf{x}\|$ (for more detailed analysis of this term, see [17, Theorem 1]),

$$\mathbf{x}_{k+1} = D_\sigma(\mathbf{x}_k - \alpha(\nabla g(\mathbf{x}_k) + \gamma_g \nabla \phi(\mathbf{x}_k))) \quad (11)$$

$$= D_\sigma(\mathbf{x}_k - \alpha(\mathbf{H}^\top(\mathbf{H}\mathbf{x}_k - \mathbf{y}) + \gamma_g \mathbf{T}\mathbf{x}_k)), \quad (12)$$

where D_σ is a bounded denoiser with $\|D_\sigma(\mathbf{u}) - D_\sigma(\mathbf{v})\| \leq \sqrt{1 + \delta} \|\mathbf{u} - \mathbf{v}\|$ (as in [17, Assumption 2]), since $\nabla \phi(\mathbf{x}_k) = \mathbf{T}^\top \mathbf{x}_k = \mathbf{T}\mathbf{x}_k$. Let $\lambda_{\min} > 0$ and $\lambda_{\max}$ be the extreme eigenvalues of $\mathbf{A}_{\gamma_g} = \mathbf{H}^\top \mathbf{H} + \gamma_g \mathbf{T}$ on $\text{span}(\text{Range}(\mathbf{H}^\top) \cup \text{Null}(\mathbf{H})) = \mathbb{R}^n$. Then for any step size $\alpha \in (0, 2/\lambda_{\max})$,

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\| \leq (1 + \delta) \|\mathbf{I} - \alpha \mathbf{A}_{\gamma_g}\|_2 \|\mathbf{x}_k - \mathbf{x}^*\| \quad (13)$$

$$\leq \rho \|\mathbf{x}_k - \mathbf{x}^*\|, \quad (14)$$

where $\rho = (1 + \delta) \max\{1 - \alpha \lambda_{\min}, \alpha \lambda_{\max} - 1\}$. In particular, if $\rho < 1$, The iteration is a contraction and converges linearly to a fixed point. The optimal α (for this linear bound) is $\alpha^* = \frac{2}{\lambda_{\min} + \lambda_{\max}}$ with rate $\rho^* = (1 + \delta) \frac{\kappa - 1}{\kappa + 1}$, $\kappa = \frac{\lambda_{\max}}{\lambda_{\min}}$. Compared with the baseline ($\gamma_g = 0$), adding $\gamma_g \mathbf{T}$: (i) makes $\mathbf{A}_{\gamma_g} \succ 0$ even on $\text{Null}(\mathbf{H})$, guaranteeing uniqueness of the quadratic subproblem, and (ii) reduces the condition number κ because $\mathbf{T}$ acts precisely on directions where $\mathbf{H}^\top \mathbf{H}$ is small, improving convergence and ρ^* .

5. Experiments

To validate the proposed approach, we implemented it in PyTorch [31] and used the DeepInverse library [43] to benchmark different solvers, as detailed below. Since our work does not propose a specific network G , we validate the approach using a U-Net model [37] with a four-level 2D encoder-decoder architecture: the encoder applies double convolution blocks with $C, 2C, 4C$, and $8C$ channels (with $C = 64$) and 2×2 max-pooling, while the decoder uses nearest-neighbor upsampling with skip connections, refining features back to C channels before a final 1×1 convolution produces the output. The network is optimized using the Adam optimizer [21] with a learning rate of 10^{-3} for 50 epochs. For all experiments, we used a noise variance $\sigma^2 = 0.05$. Code at ¹. We consider the inverse problems:

Compressed Sensing (CS): $\mathbf{H} \in \{-1, 1\}^{m \times n}$ are the first 10% rows of a Hadamard matrix $\mathbf{A} \in \{-1, 1\}^{n \times n}$ [5]. We validate with CIFAR-10 [23] and CelebA [25] datasets.

Super-Resolution (SR): We set $n = 3 \cdot 128^2$ and use an SR factor (SRF) of 4, i.e., $m = 3 \cdot 32^2$. We validate on the CelebA dataset [25], using 10000 training images resized to 128×128 . We also show results for the Synthetic Aperture Radar (SAR) image dataset [11].

Demosaicing: We used the CelebA dataset, resized to 64×64 , and the Bayer filter [3] acquisition pattern.

Deblurring: We set $n = 64^2 \cdot 3$ and used a 2-D Gaussian kernel with a bandwidth $\sigma_k = 1$. We validated our experiments on the CelebA [25] dataset and the Places365 [54] dataset, both resized to 64×64 , using 10000 images for training and a batch size of 32.

For all of these imaging tasks, we construct the null-restricted Laplacian $\mathbf{T}$ and the matrix $\mathbf{S}$, leveraging properties of $\mathbf{H}$. Details on the computation of these matrices in A5. We provide comprehensive ablations of GSNR in A6.

5.1. Predictability and Coverage

The coverage definition in (9), has a general expression for the covariance matrix computation. In practice, we computed as follows. Given centered images $\mathbf{x}_i$ and their null projections $\mathbf{x}_{n,i} = \mathbf{P}_n(\mathbf{x}_i - \bar{\mathbf{x}})$, where $\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i$ and the NS covariance $\text{Cov}(\mathbf{x}_n) = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_{n,i} \mathbf{x}_{n,i}^\top$. Fig. 4 shows the results in terms of (a) Coverage and (b) Predictability in CS for different operators $\mathbf{S}$, under the same acquisition matrix $\mathbf{H}$. We used the $N = 10000$ image of the CIFAR-10 dataset. In this case, NPN ($\mathbf{L} = \mathbf{I}$) is obtained following [17, Eq. 3], and GSNR uses $\mathbf{L} = \mathbf{L}_{8nm}$. Note that here we take the first p rows of each $\mathbf{S}$. Supp. A9 shows additional results for CS. A7 shows the coverage curves for the SR task.

5.2. Diffusion Solvers

We used two diffusion model (DM) solvers, DPS [8] and DiffPIR [55]. See Supp. A8 for more details on the adaptation of these models with the proposed approach. Fig. 5 compares DPS and DiffPIR with and without the proposed GSNR term. In all cases, the baseline DM already produces plausible faces, but it often leaves residual blur and aliasing around high-frequency structures such as hair, eyebrows, and facial contours. When we replace the generic NS basis by the GSNR, the reconstructions become visibly sharper and more faithful to the ground truth: edges around the jawline and nose are better defined, textures are less washed out, and blocky artifacts from the low-resolution observation $\mathbf{y}$ are further suppressed. These qualitative improvements are consistent with the quantitative SSIM/P-SNR numbers printed in each panel: the gains over the baseline DiffPIR/DPS are modest in dB but systematic (e.g., $30.19 \rightarrow 30.31$; $30.19 \rightarrow 30.48$), indicating that the GSNR term helps the diffusion prior resolve NS ambiguities rather than hallucinating arbitrary details. In other words, the DM

¹github.com/yromariogh/GSNR

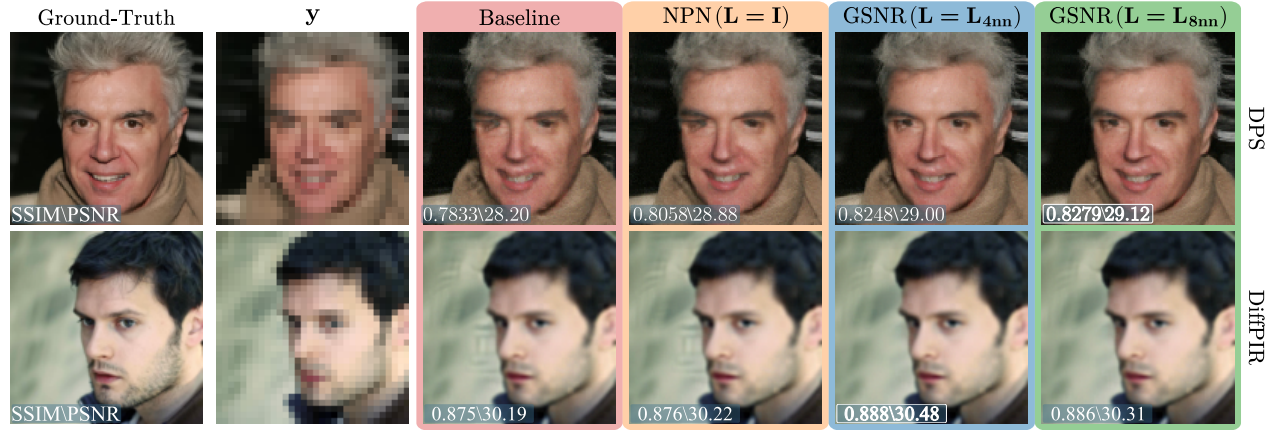


Figure 5. Results of DM-based solvers (DPS [8] & DiffPIR [55]) for Baseline, NPN [17], and GSNR with L_{4nn} and L_{8nn} . Here, $p = 0.1n$.

Table 1. Quantitative comparison on the super-resolution and demosaicing tasks (higher is better) using the CelebA test set. For each task, the **best** result and **second-best** results.

Method	Task	
	Super-Resolution	Demosaicing
PnP [20]	27.37	39.35
NPN [17]	29.21	39.77
DDN-Cascade [7]	28.92	37.83
DDN-Independent [7]	28.81	38.92
DNSN [41]	26.52	39.33
Unrolling [50]	28.97	38.82
GSNR-PnP w. L_{8nn}	29.38	39.88
GSNR-PnP w. L_{4nn}	29.42	39.89

still provides the powerful image prior, but GSNR steers it along graph-smooth NS directions that are predictable from y and compatible with the forward model, yielding more accurate and stable reconstructions. Supp. A8 shows the integration of GSNR in latent-space DMs [15].

5.3. Comparison with End-to-End Methods

Table 1 reports PSNR on SR and demosaicing, comparing classical PnP, the proposed GSNR-PnP variants, NPN, and several strong deep-learning baselines. Across both tasks, GSNR-PnP with graph Laplacians L_{4nn} and L_{8nn} achieves the best performance. For SR, GSNR-PnP yields roughly a 2 dB gain over vanilla PGD-PnP and a clear margin over both NPN (with $L = I$) and task-specific unrolled/cascade architectures. For demosaicing, the problem is less ill-posed, so the absolute gains are smaller, but GSNR-PnP still provides a consistent improvement over PGD-PnP, NPN, and the deep baselines. Overall, these results show that injecting graph-smooth NS structure into a standard PnP solver consistently improves upon (i) geometry-free PnP/NPN baselines and (ii) specialized deep unrolled or cascade architectures, with particularly strong benefits on the most underdetermined SR task. In A10 and A11, we show additional results for Demosaicing and SR, respectively.

Table 2. Final PSNR (dB) for deblurring ($\sigma_k = 1$) with $n = 64^2 \cdot 3$ and $p = 0.8 \cdot n$. We report results on Places365 [11] and CelebA [25] test sets using Lip-DnCNN [40] and a wavelet denoiser [12]. For each column, **best** and **second-best** results.

Method	γ_g	Places365		CelebA	
		Lip-DnCNN	Wavelet	Lip-DnCNN	Wavelet
Baseline (PGD-PnP)	–	31.58	30.78	35.26	33.64
NPN [17]	–	33.22	33.17	37.86	37.68
GSNR w. L_{8nn}	✗	33.60	33.59	38.18	37.99
GSNR w. L_{4nn}	✓	33.60	33.59	38.18	37.99

5.4. Plug-and-Play

We used the PGD-PnP [16, 20], but GSNR can be easily adapted to other formulations, such as ADMM [6] or HQS [33]. We used three denoisers: DnCNN with Lipchitz training (Lip-DnCNN) [40], DRUNet [52], Wavelet denoiser (Daubechies 8 with 4 levels) using soft-thresholding proximal [12]. For improving stability and performance of the PGD-PnP, we include the equivariant denoising technique [45] with two random 90 rotations and two reflections.

Deblurring: Table 2 reports the final PSNR for deblurring on Places365 and CelebA under different graph variants and denoisers. In all settings, both NPN and GSNR substantially improve over the baseline PGD-PnP solver: on CelebA, the gain is on the order of +2 dB for both Lip-DnCNN and the wavelet denoiser, and on Places365, the gain is around +1.5–2 dB. The GSNR variant with L_{8nn} provides a small but consistent improvement over NPN on both datasets and for both denoisers, indicating that injecting NS structure into the NS helps the solver recover sharper details without overfitting. Interestingly, turning the additional graph penalty on or off (γ_g ✓ vs. ✗) has a negligible effect on the final PSNR in this experiment, suggesting that the learned GSNR already captures most of the useful graph structure for deblurring. Figure 6 shows the deblurring convergence behaviour with the DnCNN denoiser. The baseline PnP-PGD curve increases steadily but saturates at a lower PSNR plateau. In contrast, both GSNR variants (with $\gamma_g = 0$ and $\gamma_g = 0.1$) converge to a noticeably higher final PSNR, confirming that the graph-smooth null-space repre-

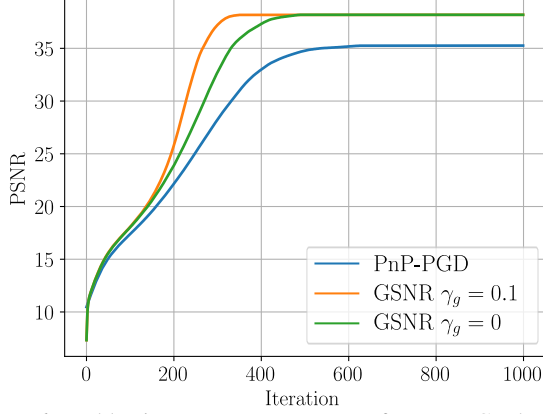


Figure 6. Deblurring convergence curves for PnP-PGD baseline using Lip-DnCNN and GSNR variants with L_{8nn} with and without the null-restricted Laplacian regularization.

sensation improves the fixed point of the algorithm. The explicit null-only graph penalty ($\gamma_g = 0.1$) mainly affects the transient regime: it accelerates convergence in the first few hundred iterations, reaching near-peak PSNR significantly earlier than GSNR with $\gamma_g = 0$ (aligned with the theoretical analysis conducted in Sec. 4.3), while all three methods eventually stabilize. Additional visual results in A12.

SR: Fig. 7 shows an SR example for different graph designs and denoisers. Columns correspond to the choice of Laplacian L or to the baseline PnP without GSNR, while rows correspond to different denoisers. In all three rows, moving from the baseline and $L = I$ columns to the graph-based L_{4nn} and L_{8nn} columns produces visibly sharper details: facial contours are crisper, hair strands are better resolved, and blocky artifacts inherited from the low-resolution backprojection $H^T y$ are reduced. The improvements are most pronounced for the DnCNN and DRUNet denoisers. Still, they are also clearly present for the simpler Wavelet denoiser, indicating that the benefit comes from the GSNR itself rather than from the denoiser. Among the graph designs, L_{8nn} yields the visually clearest reconstruction, suggesting that richer graph connectivity enables the solver to exploit structure in the null space rather than hallucinating high-frequency content. Results with Deep Image Prior [49] and SAR images in SR can be found in A11.

6. Discussion

We demonstrate that GSNR is a versatile approach for a broad range of inverse problems, data modalities, and neural networks (A14). Nevertheless, the method requires precise knowledge of the sensing matrix, which is unavailable in some applications, e.g., real-world image deblurring. In A15, we provide analysis on the effect of uncertainty in the sensing operator, where a performance drop is evident, but the improvement compared with the baseline algorithm is maintained. Follow-up ideas will harness the null-space construction under sensing matrix uncertainty for calibrated null-space regularization functions. Another aspect



Figure 7. Results of PnP, NPN, and GSNR varying L and denoisers. Best results for each denoiser are in bold. Here, $p = 0.1n$.

of GSNR is the requirement to perform EVD on the null-restricted Laplacian T , which, at large image scales, necessitates substantial computational resources. However, this computation is performed *offline* once per (H, L, n, p) , and reused for learning the network G and the reconstruction step. In A13, we provide a detailed runtime analysis of the approach. Additionally, our approach is designed for linear inverse problems; the adaptation for nonlinear inverse problems, e.g., phase retrieval, requires careful modeling of the invertibility properties of the sensing operator and its input-dependent ambiguities. In recent work [14], a generalization of the null-space property was developed for neural networks (non-linear functions), opening new frontiers for learning optimal null-space components.

7. Conclusion and Future Outlook

We introduced *GSNR: Graph-Smooth Null-Space Representation*, which equips an inverse problem solver not only with *what* the measurements constrain, but also with *how* to use the degrees of freedom that the sensor never sees (null-space). Instead of regularizing the full image, GSNR builds a low-dimensional basis for $\text{Null}(H)$ aligned with graph smoothness via the null-restricted Laplacian $T = P_n L P_n$, and can optionally add a null-only graph penalty at inference. Theoretically, the smoothest T -modes are minimax-optimal for null-space components, and under a GMRF prior, they provably capture more null-space variance than geometry-free choices while enjoying stronger statistical coupling to the measurements. In practice, this yields better coverage of plausible solutions, higher predictability of null-space signal components, and improved conditioning, translating into faster and more stable convergence of PnP and diffusion-based solvers. Looking ahead, rather than fixing L to a hand-crafted grid Laplacian, learning the graph structure or employing interpretable graph neural operators promises data-adaptive GSNR bases that better reflect semantic structure. Also, a more rigorous study of GSNR within generative solvers could provide a principled way to inject graph-smooth null constraints into these frameworks without breaking their probabilistic interpretation.

Acknowledgements

This work was supported by the OFFICE OF NAVAL RESEARCH under Grant W911NF2510165, Project titled “ADVANCED ALGORITHMS FOR DISTRIBUTED BEAMFORMING”. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the OFFICE OF NAVAL RESEARCH or the U.S. Government.

References

- [1] Eirikur Agustsson and Radu Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2017. 11
- [2] Daniel Otero Bague, Johannes Leuschner, and Maximilian Schmidt. Computed tomography reconstruction using deep image prior and learned reconstruction methods. *Inverse Problems*, 36(9):094004, 2020. 1
- [3] Bryce E Bayer. Color imaging array, 1976. US Patent 3,971,065. 6
- [4] Martin Benning and Martin Burger. Modern regularization methods for inverse problems. *Acta numerica*, 27:1–111, 2018. 1
- [5] E. J. Candes and M. B. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25(2):21–30, 2008. 1, 6
- [6] Stanley H Chan, Xiran Wang, and Omar A Elgendy. Plug-and-play admm for image restoration: Fixed-point convergence and applications. *IEEE Transactions on Computational Imaging*, 3(1):84–98, 2016. 1, 7
- [7] Dongdong Chen and Mike E Davies. Deep decomposition learning for inverse imaging problems. In *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXVIII 16*, pages 510–526. Springer, 2020. 2, 7
- [8] Hyungjin Chung, Jeongsol Kim, Michael Thompson McCann, Marc Louis Klasky, and Jong Chul Ye. Diffusion posterior sampling for general noisy inverse problems. In *The Eleventh International Conference on Learning Representations*, 2023. 6, 7, 1
- [9] Regev Cohen, Michael Elad, and Peyman Milanfar. Regularization by denoising via fixed-point projection (red-pro). *SIAM Journal on Imaging Sciences*, 14(3):1374–1406, 2021. 1
- [10] Yuning Cui, Syed Waqas Zamir, Salman Khan, Alois Knoll, Mubarak Shah, and Fahad Shahbaz Khan. AdaIR: Adaptive all-in-one image restoration via frequency mining and modulation. In *The Thirteenth International Conference on Learning Representations*, 2025. 11
- [11] Emanuele Dalsasso, Xiangli Yang, Loïc Denis, Florence Tupin, and Wen Yang. Sar image despeckling by deep neural networks: from a pre-trained model to an end-to-end training strategy. *Remote Sensing*, 12(16), 2020. 6, 7
- [12] D.L. Donoho. De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3):613–627, 1995. 7
- [13] David L Donoho and Iain M Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994. 2
- [14] Yamit Ehrlich, Nimrod Berman, and Assaf Shocher. Pseudo-invertible neural networks, 2026. 8
- [15] Yutong He, Naoki Murata, Chieh-Hsin Lai, Yuhta Takida, Toshimitsu Uesaka, Dongjun Kim, Wei-Hsiang Liao, Yuki Mitsufuji, J Zico Kolter, Ruslan Salakhutdinov, and Stefano Ermon. Manifold preserving guided diffusion. In *The Twelfth International Conference on Learning Representations*, 2024. 7, 1, 8
- [16] Samuel Hurault, Antonin Chambolle, Arthur Leclaire, and Nicolas Papadakis. A relaxed proximal gradient descent algorithm for convergent plug-and-play with proximal denoiser. In *International Conference on Scale Space and Variational Methods in Computer Vision*, pages 379–392. Springer, 2023. 7
- [17] Roman Jacome, Romario Gualdrón-Hurtado, León Suárez-Rodríguez, and Henry Arguello. NPN: Non-Linear Projections of the Null-Space for Imaging Inverse Problems. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. 2, 6, 7, 8, 9, 11
- [18] Ian T Jolliffe and Jorge Cadima. Principal component analysis: a review and recent developments. *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, 374(2016):20150202, 2016. 4
- [19] Vassilis Kalofolias. How to learn a graph from smooth signals. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, pages 920–929, Cadiz, Spain, 2016. PMLR. 2
- [20] Ulugbek S. Kamilov, Charles A. Bouman, Gregory T. Buzard, and Brendt Wohlberg. Plug-and-play methods for integrating physical and learned models in computational imaging: Theory, algorithms, and applications. *IEEE Signal Processing Magazine*, 40(1):85–97, 2023. 1, 2, 7
- [21] DP Kingma and J Ba. Adam: a method for stochastic optimization, 2014. 6
- [22] Lingshun Kong, Jiangxin Dong, Jinhui Tang, Ming-Hsuan Yang, and Jinshan Pan. Efficient visual state space model for image deblurring. In *Proceedings of the computer vision and pattern recognition conference*, pages 12710–12719, 2025. 11
- [23] A. Krizhevsky. Learning multiple layers of features from tiny images. *Master’s thesis, University of Tront*, page 1, 2009. 6
- [24] Richard B Lehoucq, Danny C Sorensen, and Chao Yang. *ARPACK users’ guide: solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods*. SIAM, 1998. 4
- [25] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Deep learning face attributes in the wild. In *Proceedings of International Conference on Computer Vision (ICCV)*, 2015. 6, 7
- [26] Michael Lustig, David L Donoho, Juan M Santos, and John M Pauly. Compressed sensing mri. *IEEE signal processing magazine*, 25(2):72–82, 2008. 1
- [27] Peyman Milanfar and Mauricio Delbracio. Denoising: a powerful building block for imaging, inverse problems and

- machine learning. *Philosophical Transactions A*, 383(2299): 20240326, 2025. [1](#)
- [28] Sunil K Narang, Yung Hsuan Chao, and Antonio Ortega. Graph-wavelet filterbanks for edge-aware image processing. In *2012 IEEE Statistical Signal Processing Workshop (SSP)*, pages 141–144, 2012. [2](#)
- [29] Antonio Ortega, Pascal Frossard, Jelena Kovačević, José MF Moura, and Pierre Vandergheynst. Graph signal processing: Overview, challenges, and applications. *Proceedings of the IEEE*, 106(5):808–828, 2018. [2](#), [3](#), [4](#)
- [30] Neal Parikh and Stephen Boyd. Proximal algorithms. *Foundations and Trends in Optimization*, 1(3):127–239, 2014. [2](#)
- [31] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2019. [6](#)
- [32] Gabriel Peyré, Sébastien Bogleux, and Laurent Cohen. Non-local regularization of inverse problems. In *Computer Vision – ECCV 2008*, pages 57–68, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg. [2](#)
- [33] Arash Rasti-Meymandi, Aboozar Ghaffari, and Emad Fatemizadeh. Plug and play augmented hqs: Convergence analysis and its application in mri reconstruction. *Neuro-computing*, 518:1–14, 2023. [7](#)
- [34] Edward T Reehorst and Philip Schniter. Regularization by denoising: Clarifications and new interpretations. *IEEE transactions on computational imaging*, 5(1):52–67, 2018. [1](#)
- [35] Yaniv Romano, Michael Elad, and Peyman Milanfar. The little engine that could: Regularization by denoising (red). *SIAM journal on imaging sciences*, 10(4):1804–1844, 2017. [1](#), [2](#)
- [36] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10684–10695, 2022. [8](#)
- [37] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015. [6](#)
- [38] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992. [2](#)
- [39] Havard Rue and Leonhard Held. *Gaussian Markov random fields: theory and applications*. Chapman and Hall/CRC, 2005. [4](#)
- [40] Ernest Ryu, Jialin Liu, Sicheng Wang, Xiaohan Chen, Zhangyang Wang, and Wotao Yin. Plug-and-play methods provably converge with properly trained denoisers. In *International Conference on Machine Learning*, pages 5546–5557. PMLR, 2019. [7](#), [9](#)
- [41] Johannes Schwab, Stephan Antholzer, and Markus Haltmeier. Deep null space learning for inverse problems: convergence analysis and rates. *Inverse Problems*, 35(2): 025008, 2019. [2](#), [7](#)
- [42] David I Shuman, Sunil K Narang, Pascal Frossard, Antonio Ortega, and Pierre Vandergheynst. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE signal processing magazine*, 30(3):83–98, 2013. [3](#), [4](#)
- [43] Julián Tachella, Matthieu Terris, Samuel Hurault, Andrew Wang, Leo Davy, Jérémy Scanvic, Victor Sechaud, Romain Vo, Thomas Moreau, Thomas Davies, Dongdong Chen, Nils Laurent, Brayan Monroy, Jonathan Dong, Zhiyuan Hu, Minh-Hai Nguyen, Florian Sarron, Pierre Weiss, Paul Escande, Mathurin Massias, Thibaut Modrzyk, Brett Levac, Tobías I. Liaudat, Maxime Song, Johannes Hertrich, Sebastian Neumayer, and Georg Schramm. Deepinverse: A python package for solving imaging inverse problems with deep learning. *Journal of Open Source Software*, 10(115): 8923, 2025. [6](#)
- [44] Hong Ye Tan, Subhadip Mukherjee, and Junqi Tang. Solving imaging inverse problems using plug-and-play denoisers: Regularization and optimization perspectives, 2025. [1](#)
- [45] Matthieu Terris, Thomas Moreau, Nelly Pustelnik, and Julian Tachella. Equivariant plug-and-play image reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 25255–25264, 2024. [7](#)
- [46] Jing Tian and Kai-Kuang Ma. A survey on super-resolution imaging. *Signal, Image and Video Processing*, 5:329–342, 2011. [1](#)
- [47] Xiangpeng Tian, Xiangyu Liao, Xiao Liu, Meng Li, and Chao Ren. Degradation-aware feature perturbation for all-in-one image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 28165–28175, 2025. [11](#)
- [48] Singanallur V. Venkatakrishnan, Charles A. Bouman, and Brendt Wohlberg. Plug-and-play priors for model based reconstruction. In *2013 IEEE Global Conference on Signal and Information Processing*, pages 945–948, 2013. [1](#)
- [49] Lizhi Wang, Chen Sun, Ying Fu, Min H Kim, and Hua Huang. Hyperspectral image reconstruction using a deep spatial-spectral prior. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8032–8041, 2019. [8](#)
- [50] Jinxi Xiang, Yonggui Dong, and Yunjie Yang. Fista-net: Learning a fast iterative shrinkage thresholding network for inverse problems in imaging. *IEEE Transactions on Medical Imaging*, 40(5):1329–1339, 2021. [7](#)
- [51] Jiahui Yu, Zhe Lin, Jimei Yang, Xiaohui Shen, Xin Lu, and Thomas S Huang. Generative image inpainting with contextual attention. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5505–5514, 2018. [1](#)
- [52] 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 Pat-*

- tern Analysis and Machine Intelligence*, 44(10):6360–6376, 2021. [1](#), [7](#), [9](#)
- [53] Kaihao Zhang, Wenqi Ren, Wenhan Luo, Wei-Sheng Lai, Björn Stenger, Ming-Hsuan Yang, and Hongdong Li. Deep image deblurring: A survey. *International Journal of Computer Vision*, 130(9):2103–2130, 2022. [1](#)
- [54] Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba. Places: A 10 million image database for scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017. [6](#)
- [55] Yuanzhi Zhu, Kai Zhang, Jingyun Liang, Jiezhong Cao, Bihan Wen, Radu Timofte, and Luc Van Gool. Denoising diffusion models for plug-and-play image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 1219–1229, 2023. [6](#), [7](#), [1](#), [8](#)