FROM GRADIENT MAGNITUDE TO LEVEL-SET GEOMETRY: A 2ND-ORDER ALIGNED REGULARIZER FOR 1ST-ORDER IMAGE PRIORS

Anonymous authors
Paper under double-blind review

ABSTRACT

Natural images exhibit coherent spatial organization of local variations, particularly in gradient orientations along contours and smoothly varying structures. Image corruption can disrupt this organization, causing neighboring gradient orientations to fluctuate irregularly even when individual variations remain plausible. In image denoising, 1st-order regularizers control local variation but do not explicitly constrain how these variations are organized across space. This raises a question: can spatial organization be regularized without introducing a preference that conflicts with the original 1st-order prior? In this paper, we show that it can be done using the first variation of a 1st-order prior. This defines a spatial response on the level-set geometry whose squared norm, the 2nd-order Aligned Regularizer (SOAR), penalizes how the normal profile and the level-set curvature vary across the image. In a local tangential Fourier picture, SOAR adds a 4th-order spectral response on top of the 2nd-order spectral response of the parent prior, acting as an additional low-pass filter on geometric fluctuations while preserving the parent prior's stationary points. For denoising, we combine image-adaptive weighting of the 1st-order regularizer with selective SOAR activation. The resulting single-image zero-shot denoising method consistently improves over the 1st-order baseline and achieves the highest PSNR on real-world benchmarks.

1 INTRODUCTION

Image noise disrupts both local variations and their spatial organization. Neighboring gradients along clean contours tend to have coherent orientations, whereas corruption can cause irregular distortions in these orientations even at comparable gradient strength; see Fig. 1. Such neighborhood-level coherence has been exploited in image processing (Weickert, 1999; Lefkimmiatis et al., 2015). This motivates distinguishing the strength of local image variation from its spatial organization.

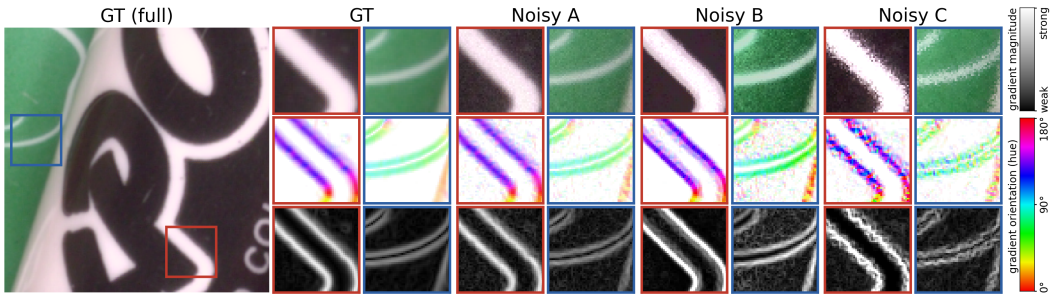


Figure 1: **Gradient magnitude and spatial organization capture different aspects of corruption. Noisy A** is a Gaussian-corrupted reference. Relative to A, **Noisy B** strengthens local gradients while largely preserving their orientation coherence; **Noisy C** has comparable gradient magnitude to B but disrupts the spatial organization of gradient orientations. Colored boxes mark the two enlarged regions. In the gradient maps, hue encodes orientation and brightness encodes magnitude.

A broad class of 1st-order image priors encodes local preferences by penalizing the gradient at each location and aggregating the resulting costs, e.g., Total Variation (TV) (Rudin et al., 1992)

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

and robust edge-preserving regularizers (Charbonnier et al., 1997). Different choices of the local penalty encode different preferences over image variations. However, these locally-evaluated-then-aggregated costs do not explicitly penalize how neighboring gradient states are organized across space. Spatial organization thus provides complementary information not represented by the local penalty itself.

Higher-order and neighborhood-aware regularizers exploit spatial information: total generalized variation (TGV) incorporates higher-order derivatives, Hessian-based models regularize 2nd-order differential structure, and structure-tensor methods aggregate 1st-order information over local neighborhoods (Bredies et al., 2010; Lefkimmiatis et al., 2012; 2015). These approaches motivate a more specific question:

> *Once a particular 1st-order prior on local variations is chosen, how should a compatible 2nd-order regularizer be designed?*

We address this by deriving the 2nd-order regularizer from the 1st-order regularizer. Our intuition is that a local gradient describes the state at a location, while spatial organization concerns how these states vary across locations. [1] We construct a *2nd-order Aligned Regularizer* (SOAR) from the first variation response of a 1st-order regularizer.

Our construction has two complementary properties: **1.** it introduces spatial sensitivity by responding to how local gradient states vary across the image. **2.** it remains variationally compatible with the underlying 1st-order regularizer. We study this principle through theoretical analysis and denoising experiments, examine when 2nd-order aligned regularization is beneficial, and investigate how this regime can be identified without access to the clean image. Our contributions are: **1.** We propose SOAR derived from a 1st-order image prior. We analyze its spatial effect through a proposed notion of gradient-orientation roughness and characterize a local condition under which additional 2nd-order regularization can improve reconstruction. **2.** We develop a single-image zero-shot denoising method that combines image-adaptive 1st-order weighting with selective SOAR activation, and evaluate it on real-world benchmarks against other zero-shot methods.

2 Background

2.1 Related Work

First-order regularization. Variational image restoration commonly combines data fidelity with 1st-order priors on image gradients. Classical choices are quadratic, Huber, and TV penalties, which differ in their treatment of gradient magnitudes (Pock et al., 2010). TV promotes piecewise smoothness while preserving sharp discontinuities (Rudin et al., 1992). Charbonnier et al. (1997) give a TV-related robust penalty that smoothly transitions between quadratic smoothing and edge-preserving behavior. These models share a common structure in which they evaluate gradient penalties locally and aggregate them over the image. Our work instead focuses on the spatial organization of local gradient states, which is not explicitly captured by these 1st-order constructions.

Higher-order and structure-aware regularization. Higher-order and neighborhood-aware methods incorporate spatial relationships beyond local gradient penalties. Popular approaches are curvature-based (Zhu & Chan, 2012; Zhu, 2020) and Euler-elastica regularizers (Mumford, 1994; Tai et al., 2011) that incorporate geometric curvature information, Hessian regularization (Lefkimmiatis et al., 2012), TGV (Bredies et al., 2010), and structure-tensor regularization (Lefkimmiatis et al., 2015). These methods demonstrate coherence of higher-order structure and orientation in image restoration. We differ in that we prescribe no independent spatial prior; instead, we derive a SOAR from the variational response of the chosen 1st-order regularizer itself.

Single-image and zero-shot denoising. Recent zero-shot methods target complex real-world noise: for example, MASH reduces local noise correlation through masked and shuffled blind-spot denoising (Chihaoui & Favaro, 2024); INS balances noise decorrelation with preservation of image correlations by implicit neural sampling (Quan et al., 2025); NLR uses cross-scale non-local pixel

[1] This intuition appears in hyperelasticity and anisotropic surface thermodynamics (Ogden, 1984; Cahn & Hoffman, 1974), where variational models convert a local energy law into a spatially coupled response.

refilling with distant similar observations (Guo et al., 2025); ZS-HPD constructs pseudo samples using complementary local-gradient and global self-similarity priors (Zhao et al., 2026). These methods focus on new sampling, masking, or pseudo-sample construction. Our work instead exploits the spatial organization of image gradients and encodes this property via explicit regularization of the single-image reconstruction.

2.2 PRELIMINARIES

Denoising problem. Let $\Omega \subset \mathbb{R}^2$ be the image domain and $u^\star : \Omega \to \mathbb{R}$ the unknown clean image, with noisy observation $y = u^\star + n$. We use $u : \Omega \to \mathbb{R}$ for scalar-valued images.[2] Assume boundary conditions under which integration-by-parts terms vanish. For an image prior J and strength $\lambda \geq 0$, we consider the variational denoising problem $u_0 \in \arg\min_u F(u) := \frac{1}{2}\|u - y\|^2_{L^2(\Omega)} + \lambda J(u)$.

Norm-based 1st-order regularizers. For $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ depending on the local gradient magnitude $|\nabla u|$, define $J_\phi(u) = |\phi(|\nabla u(x)|)|_1$, where $|\cdot|_1 = |\cdot|_{L^1(\Omega)}$. Examples are $J_{\mathrm{TV}_\epsilon}(u) = \big|\sqrt{|\nabla u|^2 + \epsilon^2_{\mathrm{TV}}}\big|_1$, $J_{\mathrm{H}}(u) = |h_\delta(|\nabla u|)|_1$, where h_δ is the Huber loss, and $J_{\mathrm{Q}}(u) = \frac{1}{2}|\nabla u|^2_2$, where $|\cdot|_2 = |\cdot|_{L^2(\Omega)}$.

Derivative and Euler–Lagrange stationarity. For differentiable F, its v-directional derivative at u is $\partial F(u; v) = \frac{\mathrm{d}}{\mathrm{d}t} F(u + tv)|_{t=0} = \langle \frac{\partial F}{\partial u}(u), v\rangle_{L^2(\Omega)}$, when the derivative admits an L^2 representation. An interior minimizer satisfies the Euler–Lagrange condition $\frac{\partial F}{\partial u}(u) = 0$ (Gelfand & Fomin, 1963).

Deep Image Prior. Ulyanov et al. (2018) parameterized u as $\mathtt{Net}(\vartheta, z)$ using an untrained neural network with input z. Substituting this parameterization into $F(u)$ yields an optimization problem over network parameters ϑ.

3 PROPOSED VARIATIONAL METHOD, IDEA AND ANALYSIS

We introduce the 2nd-order Aligned Regularizer (SOAR) and its local mechanism for image denoising. Starting from a chosen 1st-order prior, SOAR derives a 2nd-order regularizer from the prior's own variational response. We first establish variational compatibility, then characterize the aligned regularizer's geometric response, and finally derive a local orientation-gap condition motivating a data-driven activation rule.

SOAR. Starting from J_ϕ, we apply the Euler–Lagrange equation[3] to define SOAR $K_\phi(u)$:

$$\frac{\partial J_\phi}{\partial u}(u) = -\operatorname{div}\big(\phi'(|\nabla u|)\tfrac{\nabla u}{|\nabla u|}\big), \quad |\nabla u| \neq 0, \qquad K_\phi(u) := \tfrac{1}{2}\|\tfrac{\partial J_\phi}{\partial u}(u)\|^2_2. \tag{SOAR}$$

SOAR introduces no independently chosen 2nd-order prior; its higher-order response is generated directly from the 1st-order regularizer. With weight μ, the resulting reconstruction problem is

$$u_\mu \in \arg\min_u F_\mu(u) \;:=\; \{\tfrac{1}{2}\|u - y\|^2_{L^2(\Omega)} + \lambda J_\phi(u) + \mu K_\phi(u)\}, \qquad \mu \geq 0. \tag{1}$$

Examples. For the priors in Sect. 2, $K_{\mathrm{TV}_\epsilon}(u) = \frac{1}{2}\|\operatorname{div}(\nabla u/\sqrt{|\nabla u|^2 + \epsilon^2_{\mathrm{TV}}})\|^2_2$, $K_{\mathrm{H}}(u) = \frac{1}{2}\,\|\mathrm{div}(q_\delta(|\nabla u|)\nabla u)\|^2_2$, $q_\delta(x) = 1$ if $x \leq \delta$, $q_\delta(x) = \delta/x$ if $x > \delta$, and $K_{\mathrm{Q}}(u) = \frac{1}{2}\|\Delta u\|^2_2$.

SOAR generated by the variational response of J_ϕ preserves the prior's stationary minimizers.

Proposition 1 (Preservation of prior minimizers) *If $\partial J_\phi/\partial u$ is differentiable in a neighbourhood of $\bar{u}$ and $\frac{\partial J_\phi}{\partial u}(\bar{u}) = 0$, then $K_\phi(\bar{u}) = 0$ and $\frac{\partial K_\phi}{\partial u}(\bar{u}) = 0$. Consequently, in the unconstrained differentiable case, every minimizer of J_ϕ is also a minimizer of $J_\phi + \mu K_\phi$ for $\mu \geq 0$.*

3.1 GEOMETRIC ANALYSIS: MAGNITUDE–DIRECTION DECOMPOSITION

Our goal in this subsection is to show why SOAR suppresses corruption-induced geometric fluctuations. Viewing γ_y as a perturbation of the clean contour $\gamma^\star$, and letting k denote a Fourier mode of the displacement ξ, we will see that the 1st-order prior contributes a k^2 spectral response, while

[2] For RGB images, the same spatial operators are applied channelwise.

[3] At $|\nabla u| = 0$, we define the vector field inside the divergence by continuity.

SOAR adds k^4. The SOAR term is thus a strictly stronger low-pass filter on ξ in the tangential direction. The rest of the subsection derives this from the prior's Euler–Lagrange response.

At each point, the gradient ∇u contains a magnitude $r := |\nabla u|$, describing transition strength, and a direction $\mathbf{n} := \nabla u / r$ describing transition orientation. The associated curvature is $\kappa := \operatorname{div}(\mathbf{n})$, and we define $\partial_{\mathbf{n}} r := \nabla r \cdot \mathbf{n}$. Using these quantities, the variational response decomposes as

$$\begin{aligned} \frac{\partial J_\phi}{\partial u}(u) \overset{SOAR}{=} -\operatorname{div}(\phi'(r)\mathbf{n}) &\overset{\text{product rule}}{=} -[\nabla(\phi'(r)) \cdot \mathbf{n} + \phi'(r)\operatorname{div}(\mathbf{n})] \\ &\overset{\text{chain rule}}{=} -[\phi''(r)\nabla r \cdot \mathbf{n} + \phi'(r)\kappa] \\ &= -\phi''(r)\,\partial_{\mathbf{n}} r - \phi'(r)\,\kappa, \end{aligned} \tag{2}$$

Here, the term $-\phi''(r)\partial_{\mathbf{n}} r$ measures changes in transition strength along the normal, while $-\phi'(r)\kappa$ measures changes in transition orientation.

Thus, although J_ϕ is evaluated pointwise in gradient magnitude, its variational response captures the spatial organization of neighboring gradients. SOAR makes this response explicit by penalizing its magnitude.[4] We next study its geometric effect through a local contour perturbation model.

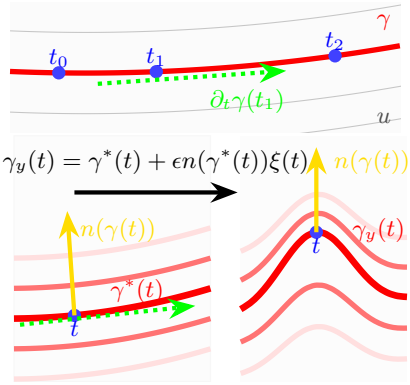
Local model and notation. We illustrate the mechanism locally along an edge $\gamma(t)$ of the image u, parameterized by the tangential coordinate t. Consider a perturbation

$$\gamma_y(t) = \gamma^\star(t) + \epsilon n(\gamma^\star(t))\xi(t),$$

where $\gamma^\star$ is the clean curve, $n(\gamma^\star(t))$ is the unit normal direction along $\gamma^\star$, $\epsilon > 0$ is the corruption strength, and $\xi(t)$ is the smooth Gaussian bump function that describes the perturbation profile. Here

$$\partial_t \gamma_y(t) = (1 + \epsilon\kappa\xi(t))\partial_t\gamma^\star(t) + \epsilon n(\gamma^\star(t))\partial_t\xi(t)$$

describes the local edge tilt, while $\partial_{tt}\gamma_y(t)$ describes how rapidly that tilt changes.



From edge length to bending. Perturbing an edge increases its length, which can be suppressed by the chosen 1st-order prior through the penalty $\|\partial_t\gamma\|_2^2$. The corresponding bending response is captured by a 2nd-order term proportional to $\|\partial_{tt}\gamma\|_2^2$ (Appendix A.2). A local Taylor approximation model gives

$$\mathcal{Q}_\mu(\gamma;\gamma_y) = \underbrace{\tfrac{\alpha}{2}\|\gamma-\gamma_y\|_2^2}_{\text{follow the observed edge}} + \underbrace{\tfrac{\beta}{2}\|\partial_t\gamma\|_2^2}_{\text{resist extra length}} + \underbrace{\tfrac{\mu c_K}{2}\|\partial_{tt}\gamma\|_2^2}_{\text{resist bending}}. \tag{3}$$

Here $\alpha > 0$ comes from the intensity profile, $\beta \geq 0$, $c_K > 0$, $\mu \geq 0$ are specified below.

Straight-line simplification. To isolate the corruption-induced component, assume locally that $\gamma^\star(t)$ is straight (without loss of generality), so that $\kappa^\star = 0$, $\partial_t n(\gamma^\star(t)) = 0$, and $\partial_{tt}\gamma^\star(t) = 0$. Hence

$$\partial_t\gamma_y(t) = \partial_t\gamma^\star(t) + \epsilon\, n(\gamma^\star(t))\,\partial_t\xi(t), \qquad \partial_{tt}\gamma_y = \epsilon\, n(\gamma^\star(t))\,\partial_{tt}\xi(t).$$

Relative to the clean straight reference $\gamma^\star$, consider the change in the two regularization terms evaluated along the perturbed path γ_y. This increase is

$$\begin{aligned} \Delta\mathcal{Q}_\mu^{\text{reg}} &:= \tfrac{\beta}{2}\left(\|\partial_t\gamma_y\|_2^2 - \|\partial_t\gamma^\star\|_2^2\right) + \tfrac{\mu c_K}{2}\left(\|\partial_{tt}\gamma_y\|_2^2 - \|\partial_{tt}\gamma^\star\|_2^2\right) \\ &= \tfrac{\beta\epsilon^2}{2}\|\partial_t\xi\|_2^2 + \tfrac{\mu c_K\epsilon^2}{2}\|\partial_{tt}\xi\|_2^2. \end{aligned} \tag{4}$$

Fourier interpretation. The distinction between the two penalties becomes particularly clear in the Fourier domain. Let $\widehat{\xi}(\omega)$ denote the Fourier transform of the perturbation profile $\xi(t)$. Then $\widehat{\partial_t\xi}(\omega) = i\omega\hat{\xi}(\omega)$ and $\widehat{\partial_{tt}\xi}(\omega) = -\omega^2\hat{\xi}(\omega)$ (Stein & Shakarchi, 2003). Hence, by the Parseval–Plancherel theorem,

$$\|\partial_t\xi\|_2^2 = \int_{\mathbb{R}} \omega^2|\widehat{\xi}(\omega)|^2\,d\omega, \qquad \|\partial_{tt}\xi\|_2^2 = \int_{\mathbb{R}} \omega^4|\widehat{\xi}(\omega)|^2\,d\omega,$$

[4] SOAR is not an independently selected curvature penalty: the curvature response arises from the variational derivative of J_ϕ, so changing the parent prior ϕ changes both the 1st-order preference and its induced aligned regularizer K_ϕ.

up to the normalization convention of the Fourier transform. Hence, restricting attention to the normal perturbation component, the regularization part of Eq.4 can be written as

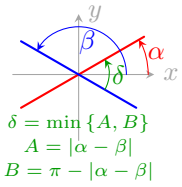
$$\frac{\epsilon^2}{2}\int_{\mathbb{R}} \left(\beta\omega^2 + \mu c_K \omega^4\right) |\widehat{\xi}(\omega)|^2 \, d\omega.$$

Thus the first-order prior penalizes Fourier modes quadratically in frequency, whereas the bending term penalizes them quartically. Equivalently, on a periodic interval of length L, with Fourier modes $\omega_k = \frac{2\pi k}{L}$, the two penalties scale respectively as $\beta\omega_k^2 = O(k^2)$, $\mu c_K \omega_k^4 = O(k^4)$. Consequently, modes associated with rapidly varying edge tilt are increasingly suppressed as the spatial frequency grows. Geometrically, the first-order term penalizes the edge tilt $\partial_t \gamma$, while the second-order term penalizes how rapidly that tilt changes along the edge, $\partial_{tt}\gamma$. The latter therefore provides an additional mechanism for suppressing localized oscillatory or bump-like deformations that may incur only a moderate first-order cost but a substantially larger bending cost. In the image domain, a spatially varying edge tilt causes the local edge normal, and hence the gradient orientation θ, to vary along the edge. The rate of this orientation change, $\partial_t \theta$, and the local curvature κ are, to leading order, both determined by the second derivative $\partial_{tt}\gamma$, up to sign (Appendix A.2). Thus, the SOAR term responds directly to spatial changes in gradient orientation: the faster the local orientation changes along an edge, the stronger the SOAR response.

From spectral response to an activation criterion. However, such changes can come from either corruption or genuine image structure, such as curved edges, fine details, or complex contours. To determine when SOAR is likely to be beneficial, we therefore need a way to quantify the variation in local gradient orientation within an image. If we can measure this variation, we can compare the noisy and clean images directly: a larger value for the noisy image would indicate additional orientation variation introduced by corruption.

For a discrete image, we measure this variation through the disagreement between neighboring orientations. Let $\theta_i^u \in [0, \pi)$ be the local axial isophote orientation at pixel i (Mardia & Jupp, 2000), estimated from a structure tensor (Bigün et al., 1991), and let $\mathcal{E}(M)$ denote valid neighboring pixel pairs within a structural mask M. We define

$$R_\theta(u; M) := \left[\frac{1}{|\mathcal{E}(M)|}\sum_{(i,j)\in\mathcal{E}(M)} \min\left\{|\theta_i^u - \theta_j^u|,\ \pi - |\theta_i^u - \theta_j^u|\right\}^2\right]^{1/2}, \quad (5)$$



where R_θ measures the RMS disagreement of neighboring isophote orientations in structural regions. It is an unsigned, mask-averaged discrete analogue of spatial variation in the normal field and avoids explicit level-line tracing.

Ideally, if the clean image were available, we could compare $R_\theta(y; M_\star)$ with $R_\theta(u^\star; M_\star)$ directly. Their gap measures the additional orientation variation present in the noisy observation. Within the local model, Appendix A.3 shows that when this gap is sufficiently large relative to the clean orientation variation, adding SOAR gives a positive local gain. This motivates an ideal diagnostic for when SOAR is expected to be beneficial.

Selective activation in practice. We now instantiate the framework with smoothed TV as the parent 1st-order prior and its corresponding SOAR term. The ideal diagnostic above suggests comparing the noisy observation with a cleaner reference. We obtain this reference by first solving Eq. 1 with $\mu = 0$, giving the TV reconstruction u_{TV}. This reconstruction is both the result that SOAR may refine and the reference used to estimate how much orientation variation is introduced by corruption.

The quality of this reference therefore matters. If λ in Eq.1 is too small, corruption may remain; if it is too large, genuine structure may be removed. We address this by using an image-adaptive TV weight, selecting λ separately for each input. With the corresponding structural mask M_{TV} (Appendix B.1), we define

$$\widehat{G} := \frac{R_\theta(y; M_{\mathrm{TV}}) - R_\theta(u_{\mathrm{TV}}; M_{\mathrm{TV}})}{R_\theta(u_{\mathrm{TV}}; M_{\mathrm{TV}})}, \qquad R_\theta(u_{\mathrm{TV}}; M_{\mathrm{TV}}) > 0. \quad (6)$$

Here, $\widehat{G}$ measures the additional orientation variation in the noisy observation relative to the TV reconstruction. If $\widehat{G} > \widehat{\tau}$, we activate SOAR with $\mu = \lambda/r$, where r controls the relative SOAR strength; otherwise, we retain u_{TV}.

4 EXPERIMENTS

4.1 EXPERIMENTAL SETUP

Datasets and metrics. We use 100 images from PolyU (Xu et al., 2018), 160 images from SIDD-Small (Abdelhamed et al., 2018), and 48 microscopy images from three FMD (Zhang et al., 2019) subcategories.[5] We report dataset-average PSNR and SSIM (Wang et al., 2004) for all methods.

Implementation. Within the DIP framework, we use smoothed TV ($\epsilon_{\mathrm{TV}} = 10^{-3}$) and its induced SOAR regularizer, with weights λ and μ. We set λ adaptively from the noisy input using the fixed rule in Appendix B. We first obtain the TV reconstruction u_{TV} and compute $\widehat{G}$ in Eq. 6. If $\widehat{G} > \widehat{\tau}$, we activate SOAR with $\mu = \lambda/r$; otherwise, we retain u_{TV}. We use a single global configuration, $\widehat{\tau} = 0.9$ and $r = 12$ for all datasets, and we examine their sensitivity in Sect. 4.3. We optimize for 10^4 iterations using Adam with learning rate 10^{-2} and $\epsilon_{\mathrm{Adam}} = 10^{-3}$ (Kingma & Ba, 2015). All reported results use the EMA output at the final iteration.

Baselines. DIP and DIP+TV follow the same optimization settings as our method, with DIP+TV using the same smoothed TV term and adaptive weight λ. We also compare with MASH (Chihaoui & Favaro, 2024), INS (Quan et al., 2025), and NLR (Guo et al., 2025) using their released implementations and settings. For ZS-HPD† (Zhao et al., 2026), we report the published results.

4.2 REAL-WORLD DENOISING PERFORMANCE

We evaluate our method on three real-world denoising benchmarks.

Quantitative results. Table 1 shows a consistent advantage across all three real-noise benchmarks. Our method achieves the highest PSNR on PolyU, SIDD, and FMD, while the strongest competing method differs across datasets. Compared with the controlled DIP+TV baseline, it further improves PSNR by 0.04, 0.45, and 0.24 dB[6], respectively, with corresponding SSIM improvements on all three datasets. For SSIM, our method ranks first on PolyU and second on SIDD and FMD.

Table 1: Quantitative comparison on three real-world datasets. Dataset-average PSNR (dB) and SSIM are reported. Best and second-best results are shown in bold and underlined, respectively.

Method	PolyU		SIDD		FMD	
	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
MASH	34.78	0.9265	35.15	0.8961	<u>33.15</u>	**0.8690**
INS	33.53	0.9131	34.53	0.8996	31.94	0.8392
NLR	35.65	0.9437	<u>36.50</u>	**0.9210**	31.42	0.8158
ZS-HPD†	36.24	0.9504	35.57	0.8705	32.54	0.8301
DIP	38.52	0.9630	33.66	0.8056	30.65	0.7585
DIP+TV	<u>38.56</u>	<u>0.9633</u>	36.30	0.9132	33.02	0.8572
Ours	**38.60**	**0.9636**	**36.76**	<u>0.9153</u>	**33.25**	<u>0.8613</u>

Qualitative results. Fig. 2 provides complementary visual comparisons. In the first example, which contains large smooth regions, several methods reduce the noise effectively, but MASH, DIP and DIP+TV retain visible background fluctuations; our method produces a cleaner reconstruction that more closely matches the ground truth. Several zero-shot baselines also denoise this region effectively. The second example presents a different challenge: INS and NLR noticeably oversmooth the fine cellular structures, while MASH and DIP-based reconstructions preserve them more faithfully. Our method achieves the highest PSNR in both examples, showing a more consistent balance between noise suppression and structural preservation across the two cases.

4.3 ANALYSIS AND ABLATION

We analyze how SOAR affects gradient organization beyond adaptive TV, and isolate the contributions of adaptive TV weighting, 2nd-order regularization, and selective activation. We then evaluate

[5] We follow ZS-HPD (Zhao et al., 2026) to centrally crop all test images to 256×256 pixels.

[6] Differences are computed from unrounded values.

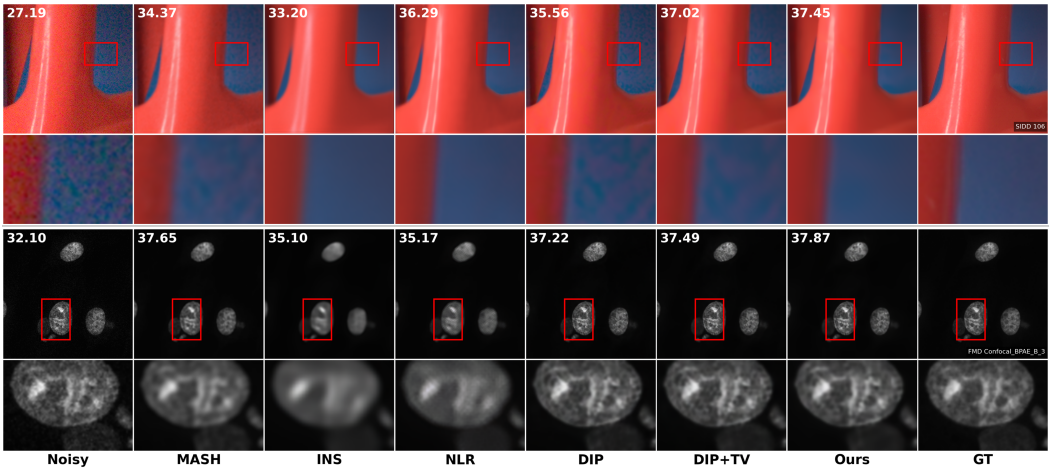


Figure 2: **Qualitative denoising comparison.** Columns show the noisy input, method reconstructions, and ground truth. Red boxes mark the regions enlarged directly below, and upper-left values report PSNR; Our method achieves the highest PSNR in both examples.

whether $\widehat{G}$ identifies images that benefit from SOAR and examine sensitivity to the SOAR strength and activation threshold.

Effect on gradient organization. Fig. 3 shows two activated examples with two local regions per image, where hue represents gradient orientation and opacity represents gradient magnitude. DIP+TV already removes much of the visible noise, but its orientation maps still show irregular changes along the reconstructed contours. With SOAR, the orientations become more continuous along these contours and closer to the clean reference, together with higher PSNR in both examples. Table 2 quantifies this effect: on all datasets, our method reduces R_θ relative to DIP+TV and brings it closer to the clean reference. The table also suggests why the gain on PolyU is small: its noisy images are only $1.75°$ rougher than the clean ones (vs. $5.89°$ on SIDD and $5.50°$ on FMD). These results support our motivation that SOAR complements TV by encouraging smoother spatial organization of neighboring gradient orientations.

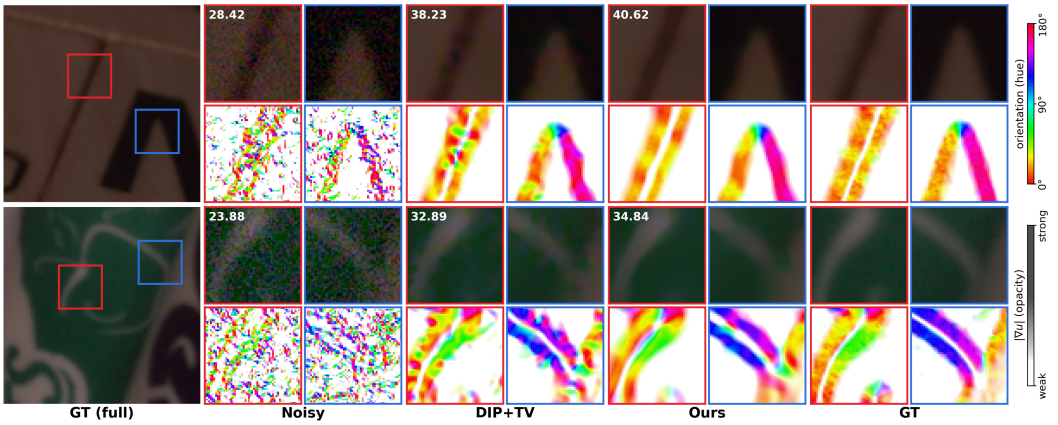


Figure 3: **Gradient-orientation visualization.** For two activated examples, we show the full image context and two cropped regions for Noisy, DIP+TV, Ours, and GT. Hue denotes gradient orientation, and opacity denotes gradient magnitude.

Component contributions. Table 3 isolates the three components of our method: image-adaptive TV weighting, the SOAR regularizer, and its selective activation. We first compare our adaptive TV weighting with a strong globally fixed baseline, whose shared weight is selected by a cross-dataset sweep (Appendix B.2). The adaptive weight improves PSNR by 0.49, 0.57, and 0.44 dB on PolyU,

Table 2: **Gradient-orientation roughness.** Mean R_θ (degrees) on the clean structural mask $M_\star$, used only for evaluation. Bold marks the reconstruction closest to GT.

Dataset	Noisy	DIP+TV	Ours	GT
PolyU	4.98	3.85	**3.69**	3.23
SIDD	9.44	5.44	**4.12**	3.55
FMD	10.38	8.45	**7.17**	4.88

SIDD, and FMD, respectively. It is also competitive with fixed weights tuned separately for each dataset, reaching comparable performance on PolyU and FMD while improving SIDD by 0.46 dB.

We next examine the contribution of the second-order regularizer while keeping the adaptive TV weights fixed, comparing SOAR with Laplacian and Euler–Elastica regularization. Their formulations and second-order strength sweeps are detailed in Appendix B. To isolate the effect of SOAR from selective activation, we apply SOAR to every image with $r = 12$, which we refer to as always-on SOAR. Always-on SOAR gives the largest PSNR gain on SIDD ($+0.31$ dB, compared with $+0.04$ and $+0.18$ dB for Laplacian and Euler–Elastica) and matches the strongest alternative on PolyU. However, it lowers SSIM on all datasets and reduces PSNR on FMD, indicating that SOAR is effective but image-dependent.

This motivates selective activation: we apply SOAR only when $\widehat{G} > \widehat{\tau}$, keeping all other settings unchanged. The resulting method achieves the best SSIM on all three datasets, with the highest PSNR on SIDD and FMD and tied-for-best PSNR on PolyU. Selective activation is particularly important on FMD, where it turns the -0.11 dB loss of always-on SOAR into a $+0.24$ dB gain. This dataset-wise pattern is consistent with the orientation-roughness analysis above in Sect. 4.3: SOAR brings larger gains on SIDD and FMD, where noise induces substantially greater orientation roughness, whereas its effect on PolyU is small, consistent with the much smaller noisy-to-clean roughness gap.

Table 3: Ablation of TV weighting and SOAR selection. Best results are bold.

Configuration	PolyU		SIDD		FMD	
	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
DIP+TV (global fixed)	38.07	0.9545	35.73	0.8759	32.58	0.8479
DIP+TV (dataset fixed)	**38.60**	0.9630	35.84	0.8899	33.05	0.8606
DIP+TV (adaptive)	38.56	0.9633	36.30	0.9132	33.02	0.8572
+ Laplacian	38.58	0.9633	36.35	0.9136	33.00	0.8568
+ Euler–Elastica	**38.60**	0.9635	36.48	0.9125	33.02	0.8544
+ SOAR (always-on)	**38.60**	0.9627	36.62	0.9044	32.90	0.8418
+ SOAR (selective)	**38.60**	**0.9636**	**36.76**	**0.9153**	**33.25**	**0.8613**

Validity of selective SOAR activation. The purpose of $\widehat{G}$ is to determine whether adding SOAR to the adaptive-TV reconstruction is likely to improve reconstruction quality. We therefore evaluate the SOAR candidate for every image, including those for which SOAR is not activated by the final method, and measure $\Delta\mathrm{PSNR} = \mathrm{PSNR}(u_{\mathrm{SOAR}}) - \mathrm{PSNR}(u_{\mathrm{TV}})$. We then compare this gain between images selected and rejected by $\widehat{G} > 0.9$. As shown in Table 4, the selected groups exhibit consistently larger mean gains than the rejected groups in both PSNR and SSIM. For PSNR, the selected–rejected gaps are 0.17, 1.11, and 1.31 dB, respectively; selected images achieve positive mean gains on all three datasets, whereas rejected images show negligible or negative gains. To test whether the PSNR separation could arise by chance, we compare against random selection with the same activation count; a one-sided permutation test (10^5 permutations) gives $p \le 10^{-3}$ on all three datasets. These results indicate that $\widehat{G}$ preferentially selects images that benefit more from SOAR while rejecting those for which additional regularization is less effective or detrimental, supporting its use as the activation criterion.

Sensitivity to SOAR strength and activation threshold. We finally study the sensitivity to the SOAR strength r and activation threshold $\widehat{\tau}$. For clarity, we report PSNR in the main text; SSIM

Table 4: **SOAR gains under the proposed activation rule.** Values are mean PSNR/SSIM changes after adding SOAR. K/N is the number selected by $\widehat{G} > 0.9$ out of the total number of images. Gap is Selected minus Rejected.

Dataset	K/N	**Selected**	**Rejected**	**Gap**
PolyU	$26/100$	$+0.167/+0.0010$	$-0.003/-0.0012$	$+0.169/+0.0022$
SIDD	$109/160$	$+0.665/+0.0030$	$-0.448/-0.0342$	$+1.114/+0.0372$
FMD	$30/48$	$+0.378/+0.0065$	$-0.933/-0.0519$	$+1.311/+0.0584$

shows the same overall trends, with full results given in Appendix B. We first fix $\widehat{\tau} = 0.9$ and vary r, where a smaller r gives stronger SOAR through $\mu = \lambda/r$. In Table 5(a), selective activation gives positive mean PSNR gains over adaptive TV for all tested r values on all three datasets. It also outperforms always-on SOAR on SIDD and FMD for every r. On FMD, selection turns the losses at $r = 12$ and $r = 6$ into gains of 0.236 and 0.185 dB.

We then fix $r = 12$ and vary $\widehat{\tau}$. Table 5(b) shows positive mean PSNR gains for all tested thresholds. FMD is more sensitive to the threshold than PolyU and SIDD. For $\widehat{\tau}$ between 0.7 and 0.9, both PSNR and SSIM improve on all three datasets. Overall, the gains remain robust to the choices of r and $\widehat{\tau}$.

Table 5: Sensitivity to SOAR strength and activation threshold. Values report mean ΔPSNR (dB) over adaptive TV. (a) Always-on and selective SOAR across r at $\widehat{\tau} = 0.9$. (b) Selective SOAR across $\widehat{\tau}$ at $r = 12$. Bold labels mark the settings used in the main comparison.

(a) SOAR strength, $\widehat{\tau} = 0.9$

r	**PolyU**		**SIDD**		**FMD**	
	Always-on	Selective	Always-on	Selective	Always-on	Selective
50	$+0.035$	$+0.024$	$+0.321$	$+0.338$	$+0.056$	$+0.159$
25	$+0.041$	$+0.030$	$+0.364$	$+0.417$	$+0.025$	$+0.217$
12	$+0.041$	$+0.043$	$+0.310$	$+0.453$	-0.113	$+0.236$
6	$+0.016$	$+0.053$	$+0.123$	$+0.396$	-0.353	$+0.185$

(b) Activation threshold, $r = 12$

$\widehat{\tau}$	**PolyU**	**SIDD**	**FMD**
0.6	$+0.046$	$+0.462$	$+0.059$
0.7	$+0.043$	$+0.465$	$+0.152$
0.8	$+0.043$	$+0.453$	$+0.152$
0.9	$+0.043$	$+0.453$	$+0.236$
1.0	$+0.040$	$+0.479$	$+0.027$

5 Conclusion

We introduced SOAR, a 2nd-order aligned regularizer derived directly from the first variation of a chosen 1st-order prior. Rather than introducing an independent higher-order prior, SOAR remains tied to the same underlying energy and preserves its stationary points. Its variational response links changes in gradient strength with level-set curvature. This provides a direct way to regularize spatial organization while remaining consistent with the original local prior.

We instantiate this idea with smoothed TV in single-image denoising. The TV weight is adapted from the noisy input. The SOAR weight is tied to it through a fixed ratio. SOAR is then activated only when the orientation-roughness score indicates that extra regularization is useful. Experiments show that adaptive TV weighting strengthens the DIP+TV baseline. They also show that the effect of SOAR is image-dependent. Selective activation turns this variable response into consistent gains over adaptive TV in mean PSNR and SSIM on all three real-noise benchmarks. It also brings gradient orientations closer to the clean reference.

Future work will study SOAR with other parent priors and inverse problems, and explore more flexible strength selection and more efficient implementations.

AI USE STATEMENT

In this work, we used generative AI tools for language enhancement, including grammar and spelling correction. We have not used generative AI tools for idea generation. We have reviewed all AI-generated text in the paper. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REFERENCES

Abdelrahman Abdelhamed, Stephen Lin, and Michael S. Brown. A high-quality denoising dataset for smartphone cameras. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1692–1700, 2018.

Josef Bigün, Gösta H. Granlund, and Johan Wiklund. Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(8):775–790, 1991.

Kristian Bredies, Karl Kunisch, and Thomas Pock. Total generalized variation. *SIAM Journal on Imaging Sciences*, 3(3):492–526, 2010. doi: 10.1137/090769521.

John W. Cahn and David W. Hoffman. A vector thermodynamics for anisotropic surfaces—ii. curved and faceted surfaces. *Acta Metallurgica*, 22(10):1205–1214, 1974. doi: 10.1016/0001-6160(74)90134-5.

Pierre Charbonnier, Laure Blanc-Féraud, Gilles Aubert, and Michel Barlaud. Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, 6(2):298–311, 1997. doi: 10.1109/83.551699.

Hamadi Chihaoui and Paolo Favaro. Masked and shuffled blind spot denoising for real-world images. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3025–3034. IEEE, 2024.

David L. Donoho and Iain M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994. doi: 10.1093/biomet/81.3.425.

Izrail Moiseevich Gelfand and Sergei Vasilevich Fomin. *Calculus of Variations*. Prentice-Hall, Englewood Cliffs, NJ, 1963. Revised English edition, translated and edited by Richard A. Silverman.

Qilong Guo, Tianjing Zhang, Zhiyuan Ma, and Hui Ji. Zero-shot blind-spot image denoising via cross-scale non-local pixel refilling. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.

Stamatios Lefkimmiatis, Aurélien Bourquard, and Michael Unser. Hessian-based norm regularization for image restoration with biomedical applications. *IEEE Transactions on Image Processing*, 21(3):983–995, 2012. doi: 10.1109/TIP.2011.2168232.

Stamatios Lefkimmiatis, Anastasios Roussos, Petros Maragos, and Michael Unser. Structure tensor total variation. *SIAM Journal on Imaging Sciences*, 8(2):1090–1122, 2015. doi: 10.1137/14098154X.

Kanti V. Mardia and Peter E. Jupp. *Directional Statistics*. John Wiley & Sons, Chichester, 2000. ISBN 978-0-471-95333-3.

David Mumford. Elastica and computer vision. In *Algebraic Geometry and its Applications: Collections of Papers from Shreeram S. Abhyankar's 60th Birthday Conference*, pp. 491–506. Springer, 1994.

Raymond W. Ogden. *Non-Linear Elastic Deformations*. Ellis Horwood, Chichester, 1984.

Thomas Pock, Daniel Cremers, Horst Bischof, and Antonin Chambolle. Global solutions of variational models with convex regularization. *SIAM Journal on Imaging Sciences*, 3(4):1122–1145, 2010.

Yuhui Quan, Tianxiang Zheng, Zhiyuan Ma, and Hui Ji. Zero-shot blind-spot image denoising via implicit neural sampling. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 7502–7512, 2025.

Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1–4):259–268, 1992. doi: 10.1016/0167-2789(92)90242-F.

Elias M Stein and Rami Shakarchi. *Fourier Analysis: An Introduction*, volume 1. Princeton University Press, 2003.

Xue-Cheng Tai, Jooyoung Hahn, and Ginmo Jason Chung. A fast algorithm for euler's elastica model using augmented lagrangian method. *SIAM Journal on Imaging Sciences*, 4(1):313–344, 2011.

Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 9446–9454, 2018.

Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

Joachim Weickert. Coherence-enhancing diffusion of colour images. *Image and Vision Computing*, 17(3–4):201–212, 1999. doi: 10.1016/S0262-8856(98)00102-4.

Jun Xu, Hui Li, Zhetong Liang, David Zhang, and Lei Zhang. Real-world noisy image denoising: A new benchmark. *arXiv preprint arXiv:1804.02603*, 2018.

Yide Zhang, Yinhao Zhu, Evan Nichols, Qingfei Wang, Siyuan Zhang, Cody Smith, and Scott Howard. A poisson-gaussian denoising dataset with real fluorescence microscopy images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019.

Xiaole Zhao, Qingsong Pang, Xiaobo Zhang, Xun Xu, Xun Gong, Yan Yang, and Tianrui Li. Zero-shot image denoising via hybrid prior-guided pseudo sample generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 22648–22657, 2026.

Wei Zhu. Image denoising using l p-norm of mean curvature of image surface. *Journal of Scientific Computing*, 83(2):32, 2020.

Wei Zhu and Tony Chan. Image denoising using mean curvature of image surface. *SIAM Journal on Imaging Sciences*, 5(1):1–32, 2012.

A Proofs

A.1 Prior-minimizer preservation

Write $g_\phi(u) := \partial J_\phi/\partial u(u)$. By the chain rule, $DK_\phi(u)[h] = \langle g_\phi(u), Dg_\phi(u)[h]\rangle_{L^2}$. Thus $g_\phi(\bar{u}) = 0$ implies both $K_\phi(\bar{u}) = 0$ and $DK_\phi(\bar{u}) = 0$. If $\bar{u}$ is an unconstrained global minimizer of J_ϕ, then for every admissible u and $\mu \geq 0$,

$$J_\phi(u) + \mu K_\phi(u) \geq J_\phi(u) \geq J_\phi(\bar{u}) = J_\phi(\bar{u}) + \mu K_\phi(\bar{u}).$$

This proves Proposition 1.

It concerns minimizers of the prior; the data-fidelity term generally changes the stationarity condition of the reconstruction problem.

The same construction also explains the word *aligned*. At a stationary point $\bar{u}$ of J_ϕ, let $H_\phi := Dg_\phi(\bar{u})$. Then $g_\phi(\bar{u} + h) = H_\phi h + o(\|h\|)$, so

$$K_\phi(\bar{u} + h) = \tfrac{1}{2}\|H_\phi h\|^2 + o(\|h\|^2).$$

Hence the local quadratic operator of $J_\phi + \mu K_\phi$ is $H_\phi + \mu H_\phi^* H_\phi$. For $J_{\mathrm{Q}}(u) = \frac{1}{2}\|\nabla u\|_2^2$, this gives $-\Delta + \mu\Delta^2$, with Fourier response $|k|^2 + \mu|k|^4$.

A.2 Local geometry and the curvature-dominated approximation

Work in local Cartesian coordinates (ρ_0, t), normal and tangential to a straight reference transition. For a smooth fixed profile q, set

$$u_{\mathrm{ref}}(\rho_0, t) = q(\rho_0), \qquad u_f(\rho_0, t) = q(\rho_0 - \gamma(t)), \qquad \rho \coloneqq \rho_0 - \gamma(t).$$

Assume $\|\partial_t\gamma\|_\infty \ll 1$, a nondegenerate monotone profile on the structural strip, and periodic tangential boundary conditions. Normal integrals are over a fixed profile interval I, with boundary effects neglected.

1st-order response and fidelity. Direct differentiation gives

$$\nabla u_f = q'(\rho)(1, -\partial_t\gamma), \qquad |\nabla u_f| = |q'(\rho)|\sqrt{1 + |\partial_t\gamma|^2}.$$

Taylor expansion gives

$$J_\phi(u_f) - J_\phi(u_{\mathrm{ref}}) = \frac{c_J}{2}\|\partial_t\gamma\|_2^2 + o(\|\partial_t\gamma\|_2^2), \qquad c_J \coloneqq \int_I \phi'(|q'(\rho)|)|q'(\rho)|\,\mathrm{d}\rho.$$

For the nondecreasing priors considered here, $c_J > 0$, giving $\beta = \lambda c_J$ before coordinate rescaling. Similarly, linearizing two nearby translates of q gives the leading fidelity $\frac{\alpha}{2}\|\gamma - \gamma_y\|_2^2$, where

$$\alpha = \int_I |q'(\rho)|^2\,\mathrm{d}\rho > 0.$$

Orientation and curvature. With the coordinate and sign convention above, $\mathbf{n}_f = (-1, \partial_t\gamma)/\sqrt{1 + |\partial_t\gamma|^2}$. Using a continuous local representation of the axial angle, we take $\theta_{\mathrm{ref}} = \pi/2$, so that

$$\theta_f - \theta_{\mathrm{ref}} = -\arctan(\partial_t\gamma) = -\partial_t\gamma + \mathcal{O}(|\partial_t\gamma|^3), \tag{7}$$

$$\partial_t\theta_f = -\frac{\partial_{tt}\gamma}{1 + |\partial_t\gamma|^2} = -\partial_{tt}\gamma + \mathcal{O}(|\partial_t\gamma|^2|\partial_{tt}\gamma|), \tag{8}$$

$$\kappa(u_f) = \frac{\partial_{tt}\gamma}{(1 + |\partial_t\gamma|^2)^{3/2}} = \partial_{tt}\gamma + \mathcal{O}(|\partial_t\gamma|^2|\partial_{tt}\gamma|). \tag{9}$$

Thus both orientation variation and curvature are represented, to leading order, by $\partial_{tt}\gamma$.

Which part of SOAR supplies the curvature response? Eq. 2 gives $g_\phi = g_{\phi,\mathrm{prof}} + g_{\phi,\kappa}$, with $g_{\phi,\mathrm{prof}} \coloneqq -\phi''(r)\partial_{\mathbf{n}} r$ and $g_{\phi,\kappa} \coloneqq -\phi'(r)\kappa$. Consequently,

$$\frac{1}{2}\|g_{\phi,\kappa}(u_f)\|_2^2 = \frac{c_K}{2}\|\partial_{tt}\gamma\|_2^2 + o(\|\partial_{tt}\gamma\|_2^2), \qquad c_K \coloneqq \int_I [\phi'(|q'(\rho)|)]^2\,\mathrm{d}\rho > 0.$$

However, the full SOAR regularizer is

$$K_\phi = \tfrac{1}{2}\|g_{\phi,\mathrm{prof}}\|_2^2 + \langle g_{\phi,\mathrm{prof}}, g_{\phi,\kappa}\rangle + \tfrac{1}{2}\|g_{\phi,\kappa}\|_2^2.$$

The curvature-dominated surrogate assumes that the geometry-dependent contributions of the first two terms are negligible on the patch and scales under consideration.

For unsmoothed TV on a smooth strip where $r > 0$, $\phi''(r) = 0$. For smoothed TV,

$$\phi_\epsilon'(r) = \frac{r}{\sqrt{r^2 + \epsilon^2}}, \qquad \phi_\epsilon''(r) = \frac{\epsilon^2}{(r^2 + \epsilon^2)^{3/2}}.$$

Thus $r \gg \epsilon$ makes ϕ_ϵ'' small. Curvature dominance additionally requires control of the normal-profile and cross terms.

Discrete orientation consistency. The continuum calculation motivates the approximation of neighboring orientation differences by tangential orientation variation. We assume that, on the structural mask $M_\star$,

$$|R_\theta(u; M_\star) - \sqrt{c_\theta}\|\partial_t\theta_u\|_2| \le \delta_\theta\sqrt{c_\theta}\|\partial_t\theta_u\|_2, \qquad 0 \le \delta_\theta < 1,$$

with the same c_θ and δ_θ for the clean and observed transitions. Assume additionally that $\|\partial_t \gamma_u\|_\infty \le s$ for both the clean and observed transitions. By Eq. 8,

$$\frac{1}{1+s^2}\|\partial_{tt}\gamma_u\|_2 \le \|\partial_t \theta_u\|_2 \le \|\partial_{tt}\gamma_u\|_2.$$

Hence, defining

$$\widetilde{\delta}_\theta := \frac{\delta_\theta + s^2}{1+s^2},$$

we obtain

$$|R_\theta(u; M_\star) - \sqrt{c_\theta}\|\partial_{tt}\gamma_u\|_2| \le \widetilde{\delta}_\theta \sqrt{c_\theta}\|\partial_{tt}\gamma_u\|_2.$$

A.3 Proof of the local orientation-gap bound

All calculations below are within Eq. 3. For periodic tangential boundary conditions, Fourier expansion diagonalizes the quadratic objective. Representing the normal perturbation by a scalar displacement and absorbing its amplitude into ξ, we write

$$\gamma_y = \gamma^\star + \xi,$$

the minimizer satisfies

$$\widehat{\gamma}_{\mu,k} = b_k(\mu)\widehat{\gamma}_{y,k}, \qquad b_k(\mu) := \frac{\alpha}{\alpha + \beta k^2 + \mu c_K k^4}.$$

For $\mu = 0$, write

$$a_k := b_k(0) = \frac{\alpha}{\alpha + \beta k^2}.$$

Define

$$F_k := |\widehat{\gamma^\star}_k|^2, \qquad S_k := |\widehat{\xi}_k|^2, \qquad C_k := \mathrm{Re}\left(\widehat{\gamma^\star}_k \overline{\widehat{\xi}_k}\right).$$

The local gain relative to the 1st-order reconstruction is

$$\mathcal{G}_{\mathrm{loc}}(\mu) := \|\gamma_0 - \gamma^\star\|_2^2 - \|\gamma_\mu - \gamma^\star\|_2^2.$$

Assume that a finite nonzero frequency band $\mathcal{K}$ contains a fraction $\eta > 0$ of the corruption bending energy:

$$\sum_{k\in\mathcal{K}} k^4 S_k \ge \eta \sum_k k^4 S_k. \tag{10}$$

Corruption benefit, clean bias, and interaction. For fixed $\mu > 0$, abbreviate $b_k = b_k(\mu)$ and define

$$q_k(\mu) := a_k^2 - b_k^2, \qquad \mathcal{N} := \sum_k q_k(\mu) S_k,$$

$$\mathcal{B} := \sum_k \left[(1-b_k)^2 - (1-a_k)^2\right] F_k,$$

$$\mathcal{I} := 2\sum_k \left[a_k(a_k-1) - b_k(b_k-1)\right] C_k.$$

Then

$$\mathcal{G}_{\mathrm{loc}} = \mathcal{N} - \mathcal{B} + \mathcal{I}.$$

Set

$$X := \sum_k k^4 S_k = \|\partial_{tt}\xi\|_2^2, \qquad Y := \sum_k k^4 F_k = \|\partial_{tt}\gamma^\star\|_2^2, \qquad m := \eta \min_{k\in\mathcal{K}} \frac{q_k(\mu)}{k^4} > 0.$$

By Eq. 10, $\mathcal{N} \ge mX$.

Moreover,

$$a_k - b_k = \frac{\alpha\mu c_K k^4}{(\alpha+\beta k^2)(\alpha+\beta k^2+\mu c_K k^4)} \le \frac{\mu c_K}{\alpha} k^4.$$

Hence, with $\ell \coloneqq 2\mu c_K/\alpha$,

$$\mathcal{B} \le \ell Y.$$

Also,

$$a_k(a_k-1) - b_k(b_k-1) = (a_k-b_k)(a_k+b_k-1).$$

Since $|a_k+b_k-1| \le 1$ and $|C_k| \le \sqrt{F_k S_k}$, Cauchy–Schwarz gives

$$|\mathcal{I}| \le \ell\sqrt{XY}.$$

Therefore, by Young's inequality,

$$\mathcal{G}_{\mathrm{loc}}(\mu) \ge mX - \ell Y - \ell\sqrt{XY} \ge \frac{m}{2}X - LY, \qquad L \coloneqq \ell + \frac{\ell^2}{2m}. \tag{11}$$

Converting to the measured orientation gap. Write

$$R_y \coloneqq R_\theta(y; M_\star), \qquad R_c \coloneqq R_\theta(u^\star; M_\star), \qquad \Delta R_\theta \coloneqq R_y - R_c.$$

Using the combined consistency bound above and

$$\partial_{tt}\gamma_y = \partial_{tt}\gamma^\star + \partial_{tt}\xi,$$

we obtain

$$|\Delta R_\theta| \le \sqrt{c_\theta}\left((1+\widetilde{\delta}_\theta)\sqrt{X} + 2\widetilde{\delta}_\theta\sqrt{Y}\right), \qquad Y \le \frac{R_c^2}{c_\theta(1-\widetilde{\delta}_\theta)^2}.$$

Weighted Cauchy–Schwarz gives

$$\left((1+\widetilde{\delta}_\theta)\sqrt{X} + 2\widetilde{\delta}_\theta\sqrt{Y}\right)^2 \le (1+3\widetilde{\delta}_\theta)\left((1+\widetilde{\delta}_\theta)X + 2\widetilde{\delta}_\theta Y\right).$$

Consequently,

$$X \ge \frac{(\Delta R_\theta)^2}{c_\theta(1+\widetilde{\delta}_\theta)(1+3\widetilde{\delta}_\theta)} - \frac{2\widetilde{\delta}_\theta}{1+\widetilde{\delta}_\theta}Y.$$

Substitution into Eq. 11 yields

$$\mathcal{G}_{\mathrm{loc}}(\mu) \ge A_\mu(\Delta R_\theta)^2 - B_\mu R_c^2,$$

where

$$A_\mu \coloneqq \frac{m}{2c_\theta(1+\widetilde{\delta}_\theta)(1+3\widetilde{\delta}_\theta)}, \qquad B_\mu \coloneqq \frac{L + m\widetilde{\delta}_\theta/(1+\widetilde{\delta}_\theta)}{c_\theta(1-\widetilde{\delta}_\theta)^2}.$$

Thus a sufficiently large corruption-induced orientation gap relative to the intrinsic clean roughness is sufficient for positive local gain. The result applies to the local surrogate; it does not provide a guarantee for the practical score $\widehat{G}$, where the unavailable clean reference is replaced by u_{TV}.

B Experiments

B.1 Implementation Details

Full reconstruction pipeline. We use the smoothed TV instantiation of Eq. 1, with $\epsilon_{\mathrm{TV}} = 10^{-3}$ and $\mu = \lambda/r$ when SOAR is activated. The complete procedure is summarized in Algorithm 1.

DIP configuration. All DIP-based methods use the same skip-network architecture with five scales, 32 input channels, bilinear upsampling, and reflection padding. We use 128 channels for the down/up branches and 4 channels for the skip connections, add Gaussian perturbations with standard deviation $1/30$ to the fixed network input during optimization, and report the exponential moving-average output with decay 0.99. The optimization settings are given in Sect. 4.

Adaptive TV weight. For each image, we set the TV weight from the noisy observation as $\lambda = \frac{10\hat{\sigma}^2}{\rho_{\mathrm{edge}}+0.1}$, where $\hat{\sigma}$ is a MAD-based estimate computed from the high-pass residual of the noisy image following Donoho & Johnstone (1994), and ρ_{edge} is the fraction of pixels with sufficiently large Sobel gradient magnitude relative to the estimated noise level and image intensity range. Intuitively, noisier images receive stronger TV regularization, while images with richer gradient structure receive less regularization to limit over-smoothing. The same rule and constants are used for all images and datasets; full implementation details are provided in the supplementary code.

Algorithm 1 Selective SOAR reconstruction.

Require: Noisy image y
1: Estimate the image-adaptive TV weight $\lambda = \Lambda(y)$.
2: Optimize the 1st-order TV objective to obtain u_{TV}.
3: Construct M_{TV} and compute $\widehat{G}$ using Eq. 6.
4: **if** $\widehat{G} > \widehat{\tau}$ **then**
5: Set $\mu = \lambda/r$ and solve Eq. 1 to obtain u_μ.
6: $u_{\mathrm{final}} \leftarrow u_\mu$.
7: **else**
8: $u_{\mathrm{final}} \leftarrow u_{\mathrm{TV}}$.
9: **end if**
10: **return** u_{final}

Table 6: Representative fixed-TV weight sweep. Each dataset entry reports mean PSNR/SSIM. The last column is the equal average of the three dataset-mean PSNR values and is used to select the globally fixed weight. Only representative weights from the full sweep are shown.

λ_{TV}	PolyU	SIDD	FMD	Mean PSNR
0.01	38.60/0.9630	34.42/0.8310	31.33/0.7913	34.78
0.02	38.57/0.9616	34.93/0.8476	31.79/0.8135	35.10
0.05	38.07/0.9545	35.73/0.8759	32.58/0.8479	**35.46**
0.10	37.01/0.9422	35.84/0.8899	33.05/0.8606	35.30

Structural mask for R_θ. We estimate local orientations using a joint-channel structure tensor, with Gaussian pre-smoothing $\sigma = 0.8$ and tensor smoothing $\sigma = 2.0$. Reliable orientations are selected by the mask $M = \{i : e_i \geq P_{75}(e),\ c_i \geq 0.3\}$, where e_i is the local orientation energy and c_i is the corresponding structure-tensor coherence. For the practical score $\widehat{G}$, M_{TV} is computed once from u_{TV} and shared when evaluating $R_\theta(y; M_{\mathrm{TV}})$ and $R_\theta(u_{\mathrm{TV}}; M_{\mathrm{TV}})$. The neighbor set $\mathcal{E}(M)$ contains horizontal and vertical adjacent pairs whose endpoints both lie in M.

B.2 Supplementary Experiments

Selection of fixed TV weights. For the fixed-TV baselines in Table 3, we select the regularization weight from a common sweep shared across all three datasets. For the globally fixed baseline, each candidate λ_{TV} is evaluated separately on PolyU, SIDD, and FMD, and the three dataset-mean PSNR values are then averaged with equal weight:

$$S(\lambda_{\mathrm{TV}}) = \tfrac{1}{3} \sum\nolimits_{D \in \{\mathrm{PolyU,SIDD,FMD}\}} \overline{\mathrm{PSNR}}_D(\lambda_{\mathrm{TV}}).$$

This dataset-level averaging prevents the selection from being dominated by differences in dataset size. Table 6 reports representative points from the sweep. The full sweep is used for selection, with $\lambda_{\mathrm{TV}} = 0.05$ giving the highest cross-dataset score and therefore used as the globally fixed TV baseline. For the dataset-wise fixed baseline, the weight is instead selected independently for each dataset according to its mean PSNR.

Laplacian and Euler–Elastica baselines. For the controlled comparison in Table 3, we retain the same adaptive TV term and replace only the second-order regularizer. For the Laplacian baseline, we optimize

$$\mathcal{L}_{\mathrm{Lap}} = \mathcal{L}_{\mathrm{MSE}} + \lambda\, \mathrm{TV}_\epsilon(u) + \frac{\mu}{2} \|\Delta_h u\|_2^2,$$

where Δ_h is the centered five-point discrete Laplacian evaluated on the valid image interior. For Euler–Elastica, we use

$$\mathcal{L}_{\mathrm{EE}} = \mathcal{L}_{\mathrm{MSE}} + \lambda\, \mathrm{TV}_\epsilon(u) + \mu \left\langle \kappa_\epsilon(u)^2 |\nabla u|_\epsilon \right\rangle,$$

where κ_ϵ is the smoothed level-set curvature and $\langle \cdot \rangle$ denotes averaging over spatial locations and channels. Both baselines use the same adaptive TV weight λ and $\epsilon = 10^{-3}$ as SOAR, with the second-order strength parameterized by $\mu = \lambda/r$.

In Table 7, we sweep $r \in \{0.2, 0.5, 1, 6, 12, 25, 50\}$ for Laplacian and $r \in \{0.1, 0.2, 0.5, 1, 2, 12\}$ for Euler–Elastica. For each r, we first compute the mean PSNR on each dataset and then average the three dataset means with equal weight. We select the r with the highest resulting average PSNR and use this single value across all three datasets. This gives $r = 0.5$ for Laplacian and $r = 1$ for Euler–Elastica. SSIM is reported but is not used for parameter selection.

Table 7: Second-order strength sweeps for the Laplacian and Euler–Elastica baselines. Dataset entries report mean PSNR/SSIM. Avg. PSNR is the equally weighted mean of the three dataset-level PSNR values and is used for parameter selection. Bold indicates the selected setting.

Regularizer	r	**PolyU**	**SIDD**	**FMD**	**Avg. PSNR**
Laplacian	0.2	38.58/0.9633	36.36/0.9134	32.96/0.8557	35.968
	0.5	**38.58/0.9633**	**36.35/0.9136**	**33.00/0.8568**	**35.973**
	1	38.56/0.9633	36.33/0.9135	33.02/0.8572	35.970
	6	38.56/0.9633	36.31/0.9134	32.99/0.8570	35.954
	12	38.56/0.9634	36.31/0.9135	33.01/0.8574	35.961
	25	38.56/0.9632	36.31/0.9134	33.00/0.8572	35.954
	50	38.55/0.9632	36.31/0.9135	33.00/0.8576	35.954
Euler–Elastica	0.1	38.50/0.9613	36.54/0.9016	32.39/0.8328	35.814
	0.2	38.56/0.9626	36.58/0.9066	32.69/0.8425	35.944
	0.5	38.59/0.9633	36.54/0.9107	32.93/0.8514	36.021
	1	**38.60/0.9635**	**36.48/0.9125**	**33.02/0.8544**	**36.031**
	2	38.59/0.9635	36.43/0.9133	33.06/0.8567	36.027
	12	38.57/0.9633	36.34/0.9136	33.05/0.8574	35.987

Sensitivity to the activation threshold. We vary the activation threshold $\widehat{\tau}$ from 0.6 to 1.0 at the default $r = 12$. Table 8 shows that performance remains stable over a broad range, with $\widehat{\tau} \in [0.7, 0.9]$ improving both PSNR and SSIM over the 1st-order reconstruction on all three datasets. We use $\widehat{\tau} = 0.9$ throughout the main experiments and across all datasets.

Table 8: Sensitivity to the SOAR activation threshold $\widehat{\tau}$. Each entry reports ΔPSNR (dB) / ΔSSIM relative to the 1st-order reconstruction u_{TV}, with $r = 12$.

	$\widehat{\tau}$				
Dataset	**0.6**	**0.7**	**0.8**	**0.9**	**1.0**
PolyU	$+0.05/+0.0003$	$+0.04/+0.0003$	$+0.04/+0.0003$	$+0.04/+0.0003$	$+0.04/+0.0002$
SIDD	$+0.46/+0.0020$	$+0.47/+0.0021$	$+0.45/+0.0020$	$+0.45/+0.0021$	$+0.48/+0.0039$
FMD	$+0.06/-0.0035$	$+0.15/+0.0014$	$+0.15/+0.0014$	$+0.24/+0.0041$	$+0.03/-0.0020$

Full results for selective `SOAR` activation. Table 9 shows that selective activation becomes increasingly beneficial as the second-order regularization is strengthened, particularly on SIDD and FMD. At the default $r = 12$, it recovers the degradation caused by always-on SOAR on FMD while retaining the gain on SIDD and remaining stable on PolyU.

Table 9: Absolute performance of always-on and selectively activated `SOAR`. Each entry reports PSNR (dB) / SSIM. Bold indicates that selective activation improves both metrics over the corresponding always-on candidate (before rounding).

Dataset	**Activation**	$r = 50$	$r = 25$	$r = 12$	$r = 6$
PolyU	Always-on	38.59 / 0.9635	38.60 / 0.9632	38.60 / 0.9627	38.57 / 0.9614
	Selective	38.58 / 0.9635	38.59 / 0.9636	**38.60 / 0.9636**	**38.61 / 0.9636**
SIDD	Always-on	36.63 / 0.9128	36.67 / 0.9101	36.62 / 0.9044	36.43 / 0.8976
	Selective	**36.64 / 0.9164**	**36.72 / 0.9163**	**36.76 / 0.9153**	**36.70 / 0.9136**
FMD	Always-on	33.07 / 0.8545	33.04 / 0.8500	32.90 / 0.8418	32.66 / 0.8322
	Selective	**33.18 / 0.8628**	**33.23 / 0.8631**	**33.25 / 0.8613**	**33.20 / 0.8580**