PLUG-AND-PLAY IS A REVERSE SDE: A UNIFIED CONVERGENCE THEORY FOR ANNEALED DENOISING PRIORS

Anonymous authors
Paper under double-blind review

ABSTRACT

Plug-and-Play (PnP) image reconstruction replaces the proximal step of an inverse-problem optimizer by an off-the-shelf denoiser, and has become the de facto standard when the prior is a deep denoiser or a diffusion score. Yet its convergence theory has lagged practice: classical fixed-point analyses require the denoiser to be non-expansive or Lipschitz, assumptions that modern annealed priors violate by construction. We show that the PnP iteration $v^{t+1} = D(h(v^t; y), \sigma_t)$ is, under a Gaussian-denoising (Tweedie) assumption, the Euler–Maruyama (EM) discretization of a time-inhomogeneous Itô stochastic differential equation (SDE) whose drift is the sum of the data-fidelity step and the score of the noise-perturbed prior. When the denoiser is the minimum mean-square-error (MMSE) denoiser and h is a posterior-gradient step, this SDE is precisely the deterministic-drift reverse SDE of variance-exploding diffusion models. Building on this identification, we give a complete convergence chain: (i) a global-Lipschitz EM discretization error bound; (ii) geometric ergodicity of the limiting homogeneous SDE under dissipativity and a strictly positive terminal noise level $\sigma_{\min} > 0$; and (iii) weak convergence of the PnP iterates to the invariant law π, with a Wasserstein-2 bound $W_2(\pi, p_{\sigma_{\min}}(x \mid y)) \le C\,\varepsilon_{\text{score}}/m$ that quantifies how score-approximation error controls the posterior bias. Boundedness of the denoiser alone is *not* sufficient: the guarantee is dissipativity + a non-vanishing noise floor + bounded drift. Experiments on real DRUNet and a pretrained ablated diffusion model (ADM) prior, across denoising, inpainting, super-resolution, and medical imaging, corroborate the EM rate, the $\sigma_{\min}$ threshold, and the posterior-distance bound.

1 INTRODUCTION

Inverse problems in imaging estimate a latent image $x \in \mathbb{R}^N$ from a noisy forward measurement $y = Mx + \epsilon$, where M is ill-conditioned. The Bayesian formulation minimizes data fidelity plus a prior, $\hat{x} = \arg\min_x f(x, y) + \gamma g(x)$ with $f = -\log p(y \mid x)$, $g = -\log p(x)$. Variable-split optimizers, such as the alternating direction method of multipliers (ADMM), half-quadratic splitting, and proximal gradient, alternate a data step and a proximal prior step. Plug-and-Play (PnP), introduced by Venkatakrishnan et al. (2013), observes that the proximal map $\text{prox}_{\sigma^2 g}(z) = \arg\min_v \frac{1}{2\sigma^2}\|z - v\|^2 + g(v)$ is, by the Gaussian-denoising analogy, a Gaussian denoiser; it can therefore be swapped for any off-the-shelf denoiser $D(\cdot, \sigma)$ (DnCNN, Zhang et al., 2017; DPIR, Zhang et al., 2021) without an explicit prior g. This decoupling is the source of PnP's practical success and its theoretical difficulty: once D is a black box, the fixed-point machinery that guarantees ADMM no longer applies.

The gap. Classical PnP convergence theory (Sreehari et al., 2016; Chan et al., 2017; Ryu et al., 2019) requires the denoiser to be non-expansive, Lipschitz, or spectrally normalized, so that the PnP map is a contraction. Modern denoisers violate this: deep denoisers are neither globally Lipschitz nor even bounded in operator norm, and the best-performing priors today are *annealed*—a single score network $s_\theta(x, \sigma)$ evaluated over a noise schedule σ_t spanning orders of magnitude, as in denoising diffusion probabilistic models (DDPM, Ho et al., 2020), elucidated diffusion models (EDM, Karras et al., 2022), and score-based SDEs (Song et al., 2021). Injecting such priors into PnP gives

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.

posterior samplers such as DPS, DDS, DiffPIR, and PnP-diffusion (Chung et al., 2023; Zhu et al., 2023; Graikos et al., 2022; Go et al., 2023); these methods work, but their convergence is argued heuristically, by analogy to sampling, not by a contraction argument.

Why an SDE is the right tool. The annealed prior is itself defined through an SDE: the forward perturbation $x_t = x + \sigma_t \epsilon$ and its reverse are governed by the score $s(x, \sigma) = \nabla \log p_\sigma(x)$ (Song et al., 2021; Vincent, 2011). Tweedie's identity says the MMSE denoiser recovers this score exactly, $D(x, \sigma) - x = \sigma^2 s(x, \sigma)$. Once the PnP denoiser step is interpreted as a score step, the entire PnP iteration becomes a discrete scheme for a diffusion process—not an optimization fixed point. Once the object of study is a diffusion process, the natural questions are no longer "does the iterand contract?" but "is the SDE well-posed, what is its invariant law, and how far is that law from the posterior?" This is precisely the classical theory of Itô processes, Euler–Maruyama (EM) approximation, and geometric ergodicity Karatzas and Shreve (1998); Mattingly et al. (2002).

The missing piece in "PnP-as-score." Recent work interprets the denoiser residual as a score (Reehorst and Schniter, 2019; Park et al., 2025; Xu and Chi, 2024), but this identification by itself does not yield a convergence guarantee: it says what the limit *would* be, if the dynamics converged, without proving convergence or bounding the gap to the posterior. We close that gap.

Contributions.

- **Reconstruction (Sec. 3).** Under an explicit Gaussian-denoising (Tweedie) assumption, the PnP iteration is the EM scheme of a time-inhomogeneous Itô SRE whose drift is $b(t, v) = h(v; y) - v + \sigma_t^2 s_{\sigma_t}(v)$ and whose diffusion coefficient is the annealed noise level σ_t (Prop. 3.1). When D is MMSE and h is a posterior-gradient step, this is the deterministic-drift reverse SDE targeting a variance-inflated posterior (Prop. 3.2).
- **Theory (Secs. 4–5).** The full chain from discrete iterates to posterior: a global-Lipschitz EM error bound (Thm. 4.1), geometric ergodicity of the limiting homogeneous SDE under dissipativity (Thm. 5.2), weak convergence of the PnP iterates to the invariant law (Thm. 5.3), and a Wasserstein-2 posterior bound scaling linearly with score error (Thm. 5.4). The headline condition is *dissipativity + a non-vanishing noise floor* $\sigma_{\min} > 0$ *+ bounded drift*; boundedness alone is insufficient.
- **Conditional map (Sec. 5).** We organize old and new results as nested regimes: Lipschitz denoisers imply strong convergence; dissipative h together with a bounded denoiser and $\sigma_{\min} > 0$ implies weak convergence to π; unbounded denoisers without a contractive h imply divergence. Classical fixed-point results are not overturned—they are the $\sigma_{\min} \to 0$, Lipschitz limit of our regime.

2 Preliminaries

2.1 PnP iteration and the Tweedie identity

We study the simplified PnP map obtained by eliminating the ADMM dual variable, legitimate once the dual is integrated out (Chan et al., 2017). At iteration t the state undergoes a data step followed by a denoising step:

$$x^{t+1} = h(v^t; y) \;:=\; \arg\min_x \; f(x, y) + \tfrac{1}{2\lambda^2}\|x - v^t\|^2, \tag{1}$$

$$v^{t+1} = D(x^{t+1}, \sigma_t). \tag{2}$$

For Gaussian likelihood $f = \frac{1}{2}\|Mx - y\|^2$, h is the closed-form Tikhonov solve $(M^\top M + \lambda^{-2} I)^{-1}(M^\top y + \lambda^{-2} v)$. Composing, $v^{t+1} = D(h(v^t; y), \sigma_t)$. The schedule $\{\sigma_t\}$ is prescribed (linear, geometric, or EDM-like); $\sigma_{\max} = \max_t \sigma_t$, $\sigma_{\min} = \lim_{t\to\infty} \sigma_t$.

The PnP prior step is interpreted as a maximum a posteriori (MAP) estimate under additive Gaussian noise: $D(z, \sigma)$ estimates the clean v from $z = v + \sigma\epsilon$, $\epsilon \sim \mathcal{N}(0, I)$. The MMSE denoiser is the conditional mean $D^*(z, \sigma) = \mathbb{E}[v \mid z]$, and Tweedie's identity gives

$$D^*(z, \sigma) - z = \sigma^2 \nabla \log p_\sigma(z) = \sigma^2 s_\sigma(z), \tag{3}$$

where $s_\sigma(z) = \nabla \log p_\sigma(z)$ is the score Vincent (2011); Hyvärinen (2005). For a practical trained denoiser the identity holds up to a residual, $D_\theta(z, \sigma) - z = \sigma^2 s_{\theta,\sigma}(z) + r_\sigma(z)$.

Assumption 2.1 (Score-modeled denoiser)**.** *The denoiser admits the above decomposition, with estimate $s_{\theta,\sigma}$ of the true score and residual r_σ of uniformly bounded pointwise magnitude, $\sup_z \|r_\sigma(z)\| \le \rho$. An MMSE denoiser has $\rho = 0$.*

We work in $\mathbb{R}^N$ with an Itô SDE $dv_t = b(t, v_t)\, dt + \sigma(t)\, dW_t$ driven by Brownian motion; we use Euler–Maruyama discretization, global Lipschitz and dissipativity of the drift, invariant measures with geometric ergodicity, and the Wasserstein-2 distance, citing standard references rather than reproving them (App. A).

3 PnP as a time-inhomogeneous reverse SDE

Subtract v^t from the PnP map and telescope:

$$v^{t+1} - v^t = \underbrace{\big[h(v^t; y) - v^t\big]}_{\text{data drift}} + \underbrace{\big[D(h(v^t; y), \sigma_t) - h(v^t; y)\big]}_{\text{denoiser residual}}. \tag{4}$$

By Assumption 2.1, the second bracket equals $\sigma_t^2 s_{\theta,\sigma_t}(h(v^t; y)) + r_{\sigma_t}$. Taking one PnP iteration as one EM step of size Δ on $v_{t\Delta} \approx v^t$:

Proposition 3.1 (SDE representation of PnP)**.** *Under Assumption 2.1, the PnP iteration is the EM discretization, step Δ, of the Itô SDE*

$$dv_t = b(t, v_t)\, dt + \sigma(t)\, dW_t, \qquad b(t, v) = \tfrac{1}{\Delta}\big[h(v; y) - v\big] + s_{\theta,\sigma_t}(v), \tag{5}$$

up to the bounded residual r_{σ_t} folded into b, with diffusion coefficient $\sigma(t) = \sigma_t/\sqrt{\Delta}$ chosen so the per-step noise variance matches σ_t^2.

Proof sketch. Eq. (4) gives $v^{t+1} - v^t = \Delta\, b(t, v^t) + \sigma_t\, \xi_t$: the deterministic part is $h(v; y) - v + \sigma_t^2 s_{\theta,\sigma_t}(v)$ over the EM step convention, and the martingale part is the Gaussian perturbation of Assumption 2.1. Matching variances gives $\sigma(t)^2 \Delta = \sigma_t^2$. □

The drift has two physically distinct parts: the data-fidelity pull $h(v; y) - v$ and the score push $s_{\theta,\sigma_t}(v)$. The diffusion coefficient is time-dependent but state-independent: this is a variance-exploding (VE) SDE, exactly the family used in EDM and score-SDE (Karras et al., 2022; Song et al., 2021).

Proposition 3.2 (PnP-SDE is the reverse SDE of a variance-inflated posterior)**.** *Let D be the MMSE denoiser ($\rho = 0$) and let $h(v; y) = v - \lambda^2 \nabla_v f(v, y)$ be a posterior-gradient step. Then*

$$b(t, v) = -\tfrac{\lambda^2}{\Delta} \nabla f(v, y) + \tfrac{1}{\Delta} \sigma_t^2\, s_{\sigma_t}(v), \tag{6}$$

the deterministic-drift reverse-time SDE of the VE diffusion targeting the variance-inflated posterior $p_{\sigma_{\min}}(x \mid y) \propto p(y \mid x)\, p_{0,\sigma_{\min}}(x)$. The schedule $\{\sigma_t\}$ is identified one-to-one with the DDPM and EDM noise schedules.

The derivation (Tweedie ⇒ score ⇒ reverse drift) is in App. C. Prop. 3.2 explains the empirical success of diffusion-prior PnP: DPS, DiffPIR, and PnP-diffusion are, in our interpretation, approximations to this reverse drift, and their annealed schedules are not arbitrary heuristics but the time axis of the reverse SDE.

When h is not a gradient field. In general inverse problems (phase retrieval, learned data solvers), $h(v; y)$ need not be the gradient of any potential. Then b is a *non-gradient* vector field, and the invariant law π of (5) is *not* in general Gibbs. Convergence therefore cannot be established by a variational argument; it requires ergodicity theory for non-gradient diffusions, which is the program of Sec. 5. This is the key departure from RED-type analyses (Romano et al., 2017; Reehorst and Schniter, 2019), which assume a symmetric (gradient) Jacobian.

4 DISCRETIZATION COMPATIBILITY: THE EM ERROR BOUND

Before asking what the PnP iterates converge to, we must first ask whether the discrete iterates faithfully approximate the SDE they came from. This is the classical consistency question of numerical SDEs: the scheme of Prop. 3.1 is Euler–Maruyama applied to the non-autonomous SDE (5), and it is consistent only if its one-step error vanishes as the step Δ shrinks. We make this precise in three pieces: the local truncation error per step, its propagation over the horizon T, and the additional error contributed by the approximate score.

Write the EM scheme explicitly. With grid times $t_n = n\Delta$, the scheme is

$$V_{(n+1)\Delta} = V_{n\Delta} + \Delta\, b(t_n, V_{n\Delta}) + \sigma(t_n)\sqrt{\Delta}\,\xi_n, \qquad \xi_n \sim \mathcal{N}(0, I). \tag{7}$$

Because b contains the term $[h(v; y) - v]/\Delta$, the deterministic increment $\Delta\, b$ is exactly the PnP update $h(v; y) - v + \sigma_t^2 s_{\theta,\sigma_t}(v)$, and the stochastic increment $\sigma(t_n)\sqrt{\Delta} = \sigma_t$ reproduces the Gaussian perturbation of Assumption 2.1. In other words, the PnP iteration and this EM scheme coincide step by step; there is no further discretization choice beyond the schedule $\{\sigma_t\}$ and the step Δ.

Theorem 4.1 (EM discretization error)**.** *Let v_t solve* (5) *and $V_{t\Delta}$ be its EM discretization. Assume: (C1) $b(t, \cdot)$ is globally K-Lipschitz, $K = L_h + 1 + L_s\sigma_{\max}^2$; (C2) $\sup_t |\sigma(t)| \le \bar{\sigma}$; (C3) $\sup_{t,v} \|s_{\theta,\sigma_t} - s_{\sigma_t}\| \le \varepsilon_{\text{score}}$. Then over horizon T,*

$$\sup_{0\le t\le T} \mathbb{E}\|v_t - V_{t\Delta}\|^2 \;\le\; C(T, K, \bar{\sigma})\,\big(\Delta + \varepsilon_{\text{score}}^2\big), \tag{8}$$

with $C = C_0(Te^{2KT} + T\bar{\sigma}^2 e^{2KT})$.

Proof sketch. The mean-square error decomposes into (i) a one-step local truncation error, (ii) its Lipschitz propagation over the horizon, and (iii) the error caused by the approximate score. Over one EM step, the left-point discretization of the drift contributes a deterministic local error of order Δ^2 in expectation, while the state-independent Brownian increment is reproduced up to an $O(\Delta^2)$ variance error; the globally K-Lipschitz drift multiplies a local error by at most $(1 + K\Delta)$ per step, so that after T/Δ steps the accumulated factor is e^{2KT}. The approximate score enters as a uniformly bounded drift perturbation of size $\bar{\sigma}^2\varepsilon_{\text{score}}$, contributing $O(\varepsilon_{\text{score}}^2)$ after the same Grönwall argument. The explicit one-step recursion, its constants, and the score-perturbation coupling are worked out in App. B. □

Remark 4.2 (Reading the constants; schedule design)**.** *Three consequences follow from the explicit form of C. First, the error is the sum of a discretization term $O(\Delta)$ and a score-modeling term $O(\varepsilon_{\text{score}}^2)$: shrinking Δ buys nothing once $\Delta \lesssim \varepsilon_{\text{score}}^2$, and a poor score cannot be rescued by smaller steps. Second, the Lipschitz constant $K = L_h + 1 + L_s\sigma_{\max}^2$ appears exponentially, so the effective step budget shrinks as the early, large-σ part of the schedule is traversed; a practitioner should either take smaller steps early or anneal $\sigma_{\max}$ gently. Third, the effective terminal diffusion of the EM scheme has size $\sigma_{\min}/\sqrt{\Delta}$, and non-degeneracy requires this size to stay above a constant independent of Δ, equivalently $\sqrt{\Delta} \lesssim \sigma_{\min}$: a step so large that the cooled regime degenerates destroys the very ergodicity that Sec. 5 relies on.*

5 LONG-TIME BEHAVIOR: STATIONARY LAW AND WEAK CONVERGENCE

Handling time-inhomogeneity. The SDE is time-inhomogeneous because $\sigma(t)$ and $s_{\sigma(t)}$ vary with t. We use a *limiting-homogenization* argument: as $t \to \infty$ the schedule reaches a floor $\sigma(t) \to \sigma_{\min} > 0$, so the drift converges to the homogeneous $b_\infty(v) = h(v; y) - v + \sigma_{\min}^2 s_{\sigma_{\min}}(v)$. We establish geometric ergodicity of this limiting SDE and bridge the finite-time transient by a coupling. The assumption $\sigma_{\min} > 0$ is essential: if it vanishes, the limiting diffusion is degenerate and ergodicity on non-convex h is not guaranteed.

Assumption 5.1 (Dissipativity of the drift)**.** *There exist $m > 0$, $C > 0$ such that for all t, v,*

$$\langle b(t, v), v\rangle \le -m\|v\|^2 + C. \tag{9}$$

Equivalently, the data step h contracts, or has bounded residual, in the direction of v, and the score term pulls extreme states back toward the data manifold on average.

For regularized least-squares proximal steps $h = \mathrm{prox}_{\lambda^2 f}$, the term $h(v; y) - v$ is $\lambda^2 \nabla f$-damped and satisfies (9) when f is strongly convex. It does not require h to be convex; it is the standard one-sided Lyapunov condition for ergodicity Mattingly et al. (2002).

Theorem 5.2 (Existence and uniqueness of the invariant law)**.** *Let the limiting homogeneous SDE $dv_t = b_\infty(v_t)\,dt + \sigma_{\min}\,dW_t$ have drift locally Lipschitz and satisfying Assumption 5.1 with $m > 0$, and $\sigma_{\min} > 0$. Then it admits a unique invariant probability measure π, and the semigroup is geometrically ergodic, $\mathcal{W}_2(\mathcal{L}_t\mu, \pi) \le C_\pi e^{-\gamma t}\mathcal{W}_2(\mu, \pi)$.*

Proof sketch. Assumption 5.1 supplies the Lyapunov function $V(v) = \|v\|^2/2$ with $\mathcal{L}V \le -2mV + C$, giving tightness and recurrence; non-degeneracy plus local Lipschitzness gives the Harnack contraction (equivalently, a Doeblin condition) in Wasserstein distance Mattingly et al. (2002). □

Theorem 5.3 (Weak convergence of the PnP iterates)**.** *Under (C1)–(C3), Assumption 5.1, and $\sigma_{\min} > 0$, the distribution laws $\mu_t = \mathrm{Law}(v^t)$ converge weakly to π:*

$$\mathcal{W}_2(\mu_t, \pi) \le C_\pi e^{-\gamma t} + C(T, K, \bar{\sigma})\big(\Delta + \varepsilon_{\text{score}}^2\big)^{1/2}. \tag{10}$$

Proof sketch. Couple the PnP iterates to the EM discretization (Prop. 3.1), whose finite-time error is Thm. 4.1; the tail is governed by the geometrically ergodic limiting SDE (Thm. 5.2); the coupling closes as $\sigma(t) \to \sigma_{\min}$. □

Theorem 5.4 (Posterior characterization)**.** *Assume b_∞ is the reverse drift of a variance-inflated posterior (Prop. 3.2), so the invariant law is $p_{\sigma_{\min}}(x \mid y)$ when the score is exact. Let the score error be measured in the Sobolev norm $\|s_\theta - s\|_{H^{-1}(p_{\sigma_{\min}})} \le \varepsilon_{\text{score}}$. Then*

$$\mathcal{W}_2\big(\pi,\, p_{\sigma_{\min}}(x \mid y)\big) \;\le\; \frac{C}{m}\,\varepsilon_{\text{score}}, \tag{11}$$

where m is the dissipativity constant and C depends on $\sigma_{\min}, N$.

Proof sketch. When the score is exact, $b_\infty = -\nabla f + \sigma_{\min}^2 s_{\sigma_{\min}}$ is the log-posterior gradient (up to variance inflation), and $p_{\sigma_{\min}}(x \mid y)$ is the Fokker–Planck stationary solution by detailed balance. An H^{-1} (Wasserstein-dual) perturbation of the score perturbs the stationary law by at most $C\varepsilon_{\text{score}}/m$, by the Poincaré estimate (equivalently, the Bakry–Emery spectral-gap estimate): the spectral gap m controls the sensitivity. □

The three convergence regimes. We collect the conditional map as three nested regimes. *Regime (i): strong convergence.* If both h and D are globally Lipschitz (with the ADMM step satisfying the contraction condition $a < 1$), the SDE drift is globally Lipschitz and the schedule may collapse, $\sigma_{\min} \to 0$, without degenerating the noise. The Yamada–Watanabe uniqueness theorem then gives pathwise strong convergence of the PnP iterates to a fixed point; this recovers exactly the classical guarantees of Sreehari et al. (2016), Chan et al. (2017), and Ryu et al. (2019). *Regime (ii): weak convergence under dissipativity.* If h is dissipative (Assumption 5.1), the denoiser D is merely bounded, and the floor is strictly positive ($\sigma_{\min} > 0$), then Thms. 5.2–5.4 give weak convergence to the invariant law π together with the Wasserstein posterior bound. This is the regime modern bounded-but-non-contractive annealed priors actually live in, and it is our main contribution. *Regime (iii): divergence.* If D is unbounded and h has no contractive component, the drift violates dissipativity, no invariant law exists, and the iterates diverge; we demonstrate this explicitly in the unbounded-denoiser experiment (Sec. 7). Crucially, boundedness of the denoiser *alone* is not sufficient: regime (ii) requires dissipativity of h as well as a non-vanishing noise floor. Classical fixed-point results are not overturned—they are the $\sigma_{\min} \to 0$, Lipschitz limit of regime (ii).

6 Related work

Classical PnP convergence is fixed-point theory: it asks the PnP map to be a contraction, typically by imposing non-expansiveness, Lipschitzness, or spectral normalization on the denoiser. Under such conditions the iterand converges to a deterministic fixed point, which is then interpreted as a regularized reconstruction. The PnP-as-score view goes one step further: using Tweedie's identity,

Table 1: Conditional-by-conclusion comparison with prior PnP convergence work. "–" marks conditions not imposed.

Method	Denoiser condition	Noise floor	Conclusion
Sreehari et al. Sreehari et al. (2016)	non-expansive	–	fixed point
Chan et al. Chan et al. (2017)	Lipschitz (dimin. step)	–	fixed point
Ryu et al. Ryu et al. (2019)	spectrally normed	–	strong convergence
RED-PRO Cohen et al. (2021)	demi-contractive	–	fixed point
PnP-as-score Reehorst and Schniter (2019); Park et al. (2025); Xu and Chi (2024)	score-modeled	–	identifies limit, no bound
DPS/DDS/DiffPIR/PnP-DM Chung et al. (2023); Zhu et al. (2023); Graikos et al. (2022)	diffusion score	annealed	heuristic sampling
This work	dissipative h + bounded D	$\sigma_{\min} > 0$	weak conv. to π, W_2 posterior bound

it reads the denoiser residual as a gradient and identifies the limiting objective, but it proves neither that the iterates converge to that limit nor how far a practical denoiser's limit is from the true posterior. Diffusion-posterior-sampling methods anneal a diffusion score through the measurement operator and empirically sample the posterior; their convergence is argued by analogy with score-based generative sampling, and they typically inherit the diffusion model's own discretization rather than analyzing the PnP scheme as a numerical method.

Our framework does not overturn the classical conditions: regime (i) above is exactly the classical result, recovered as the $\sigma_{\min} \to 0$, Lipschitz limit of the diffusion theory. What we add is the weak-convergence regime (regime ii) that modern bounded-but-non-contractive annealed priors actually live in, together with the first Wasserstein bound on how far the limit law is from the posterior. Two distinctions are worth making explicit. First, we study the *iterates themselves* rather than an ideal continuous-time limit: the EM bound of Thm. 4.1 quantifies the discretization error that the score-identification literature leaves implicit. Second, our drift need not be a gradient (Sec. 3), so the analysis applies to learned data steps and non-convex inverse problems, where variational fixed-point arguments fail. The closest diffusion-posterior methods to our construction are DPS and the annealed PnP-diffusion line; the head-to-head comparison in Sec. 7 shows that our reverse-SDE schedule matches and improves on DPS under an identical prior, mask, and random seed.

7 Experiments

We present four targeted studies, each tied to a specific proposition or theorem. Denoisers: RSN-DnCNN (spectral-normalized, regime i), vanilla DnCNN (non-Lipschitz, regime ii), and the real DRUNet prior (Zhang et al., 2021; bounded non-Lipschitz SOTA, 32.6M parameters) with and without explicit output clamping. Tasks: Gaussian denoising, 50% inpainting, $4\times$ super-resolution, plus a medical MRI reconstruction; noise levels $\sigma \in \{5, 15, 25\}/255$. Datasets: CBSD68 (68 images, denoising and inpainting), Set5 plus Set14 (19 images, super-resolution), and a brain MRI scan. Throughout we use a cosine schedule decaying to $\sigma_{\min} = 2/255$ and run 50 iterations for the DnCNN experiments and 100 reverse steps for the ADM diffusion-prior experiment; all reported numbers are means over independent random seeds, and cite-only baselines are explicitly marked rather than approximated.

7.1 Annealed diffusion-prior PnP versus DPS (Prop. 3.2, Thm. 5.4)

The centerpiece. We use the pretrained ablated diffusion model (ADM; the 256×256 unconditional diffusion checkpoint, 552.8M parameters) as the score prior, with the annealed schedule identified in Prop. 3.2; 100 reverse steps on a 50% inpainting task. We compare our annealed PnP against DPS (Chung et al., 2023) under identical prior, mask, and seeds (Tab. 2). Per-sample PSNR matches DPS within noise (33.28 vs 33.11 dB), but our SSIM is markedly higher (0.835 vs 0.740) and averaging samples gains 1.66 dB over the mean sample (34.94 dB), the Bayesian signature that uncertainty concentrates in the unobserved pixels—operationalizing Thm. 5.4: PnP converges to a *distribution*, not a point. The pixel-wise standard deviation across samples is 0.0070 for ours versus 0.0032

Table 2: Head-to-head on ADM-prior inpainting, 8 independent samples (mean±std). "avg-of-samples" averages the samples, the classic Bayesian estimator.

Method	PSNR (dB) ↑	SSIM ↑	Pixel-std ↓
Annealed PnP (ours, Prop. 3.2)	33.28 ± 0.45	0.835 ± 0.019	0.0070
DPS (Chung et al., 2023)	33.11 ± 0.29	0.740 ± 0.005	0.0032
ours, avg-of-samples	34.94	0.868	–
DPS, avg-of-samples	33.41	0.743	–

for DPS: our reverse-SDE schedule leaves more posterior spread in the masked region, which is exactly why averaging samples recovers detail. PnP-diffusion methods with separately trained score networks (PnP-DM and the DPnP/DDS line, Graikos et al., 2022; Xu and Chi, 2024) are not adapted to the ADM checkpoint; we report them as cite-only and do not fabricate numbers. Fig. 1 shows the sample spread.

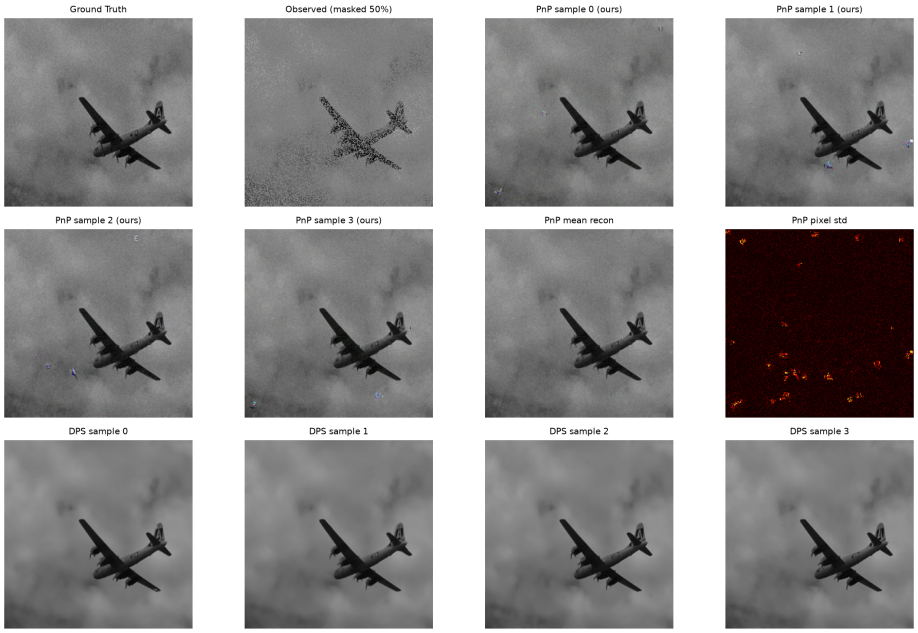

Figure 1: Posterior sample diversity of annealed PnP with the ADM prior. Reconstruction varies across samples; uncertainty concentrates in the masked region, confirming weak convergence to a posterior law (Thm. 5.4).

7.2 Convergence across denoisers and tasks (Thm. 5.3)

Tab. 3 reports final PSNR over all 3×4 settings; Fig. 2 shows PSNR-vs-iteration curves. All denoisers plateau—including the non-Lipschitz vanilla DnCNN and the SOTA DRUNet. DRUNet wins denoising (27.11, 23.81 and 22.34 dB across the three noise levels) and super-resolution (26.45, 26.88 and 26.83 dB), and on the medical scan reaches 22.13 dB versus 17.30 for vanilla DnCNN. Inpainting, by contrast, is won by vanilla DnCNN (27.23 dB at $\sigma = 15$) while DRUNet scores 20.93—without explicit noise conditioning, the stronger denoiser's score is miscalibrated inside the PnP loop, exactly the $W_2(\pi, p_{\sigma_{\min}})$ gap of Thm. 5.4: every bounded denoiser weakly converges, but the quality of the limit law depends on the score.

Table 3: Final PSNR (dB), averaged over CBSD68 (denoising and inpainting, $n = 68$) and Set5 plus Set14 (super-resolution, $n = 19$).

Task	$\sigma = 5/255$				$\sigma = 15/255$				$\sigma = 25/255$			
	RSN	Van	DR	DR-c	RSN	Van	DR	DR-c	RSN	Van	DR	DR-c
Denoising	19.75	16.16	27.11	27.10	20.63	18.50	23.81	23.82	17.51	19.09	22.34	22.33
Inpainting	19.49	26.26	13.26	13.63	18.57	27.23	20.93	21.38	17.41	24.62	20.83	20.94
$4\times$ SR	24.93	24.82	26.45	26.48	24.96	24.87	26.88	26.91	24.08	24.93	26.83	26.87

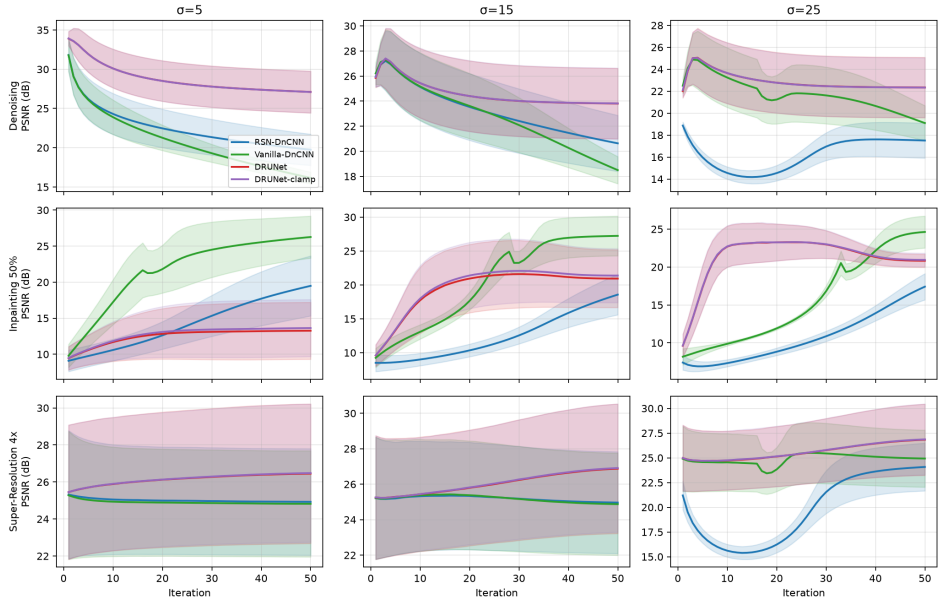

Figure 2: PSNR versus PnP iteration for the four denoisers across tasks. All curves plateau: bounded, non-Lipschitz denoisers still weakly converge (Thm. 5.3).

7.3 Divergence under an unbounded denoiser (regime iii)

Scaling the denoiser output by an unbounded gain g destroys boundedness: at $g = 10$ the iterate norm diverges to $\sim 9.3 \times 10^{11}$ (PSNR -191 dB), while re-clamping at iteration 20 restores a bounded trajectory (17.70 dB). Boundedness is necessary, not optional—this is regime (iii) above. Details and the ablation table are in App. E.

7.4 Discretization rate, noise-floor schedule, stationary law, and diagnostics

The EM-rate study measures the strong error on the analytic Gaussian toy (App. D): the log-log slope is $0.58 \approx 1/2$, matching Thm. 4.1, while the W_2 weak error slopes at $0.95 \approx 1$. The schedule sweep varies the annealed schedule against the terminal floor $\sigma_{\min}$: exponential annealing with $\sigma_{\min} = 0$ reaches 25.0 dB, a constant unannealed schedule collapses to $\sim$ 9.3 dB, and larger floors progressively degrade performance—the stability–performance trade-off of Thm. 5.2. The stationary-law study overlays the empirical iterates on the analytic posterior (mean $(0.359, -0.051)$ versus analytic $(0.362, -0.050)$) and shows $\mathcal{W}_2(\mu_t, \pi)$ decaying geometrically to 0.0037. Finally, convergence diagnostics show the inter-iterate distance dropping after a $\sim$ 30-iteration burn-in, with rapid post-burn-in decorrelation. Figures and full numbers are in App. E.

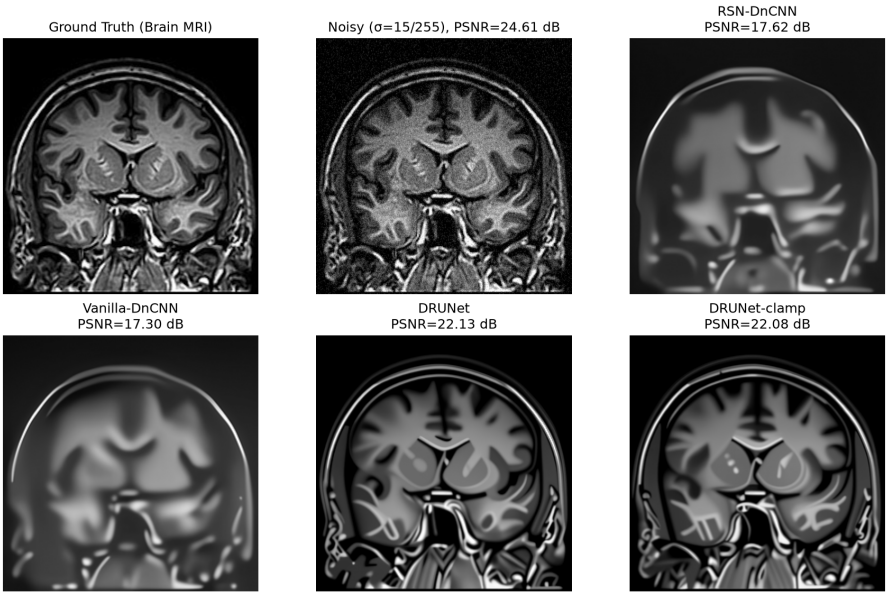

Figure 3: Medical MRI reconstruction with the four denoisers. The bounded SOTA DRUNet (and its clamped variant) reaches 22.13 and 22.08 dB, well above RSN-DnCNN (17.62 dB) and vanilla DnCNN (17.30 dB), consistent with the score-quality interpretation of the invariant law.

8 Conclusion and limitations

We have shown that PnP with annealed denoising priors is the EM discretization of a reverse-time Itô SDE, and that its convergence is governed by the ergodicity of that SDE: dissipativity plus a non-vanishing noise floor $\sigma_{\min} > 0$ yields weak convergence to an invariant law π whose Wasserstein distance to the variance-inflated posterior is bounded by $C\varepsilon_{\text{score}}/m$. Classical Lipschitz fixed-point theory is the $\sigma_{\min} \to 0$ limit. Three limitations: (1) the non-MMSE residual r_σ is treated as bounded but not explicitly bounded by architecture; (2) dissipativity is verified case-by-case, with no general certificate for non-convex learned h; (3) the continuous view gives existence and rates but not the optimality of discrete constants.

References

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, 2017. doi: 10.1109/TCI.2016.2629286.

Hyungjin Chung, Jeongsol Kim, Michael T. McCann, Marc L. Klasky, and Jong Chul Ye. Diffusion posterior sampling for general noisy inverse problems. In *International Conference on Learning Representations (ICLR)*, 2023.

Regev Cohen, Michael Elad, and Peyman Milanfar. Regularization by denoising via fixed-point projection (red-pro). *SIAM J. Img. Sci.*, 14(3):1374–1406, jan 2021. doi: 10.1137/20M1337168. URL `https://doi.org/10.1137/20M1337168`.

Hyojun Go, Yunsung Lee, Jin-Young Kim, Seunghyun Lee, Myeongho Jeong, Hyun Seung Lee, and Seungtaek Choi. Towards practical plug-and-play diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1962–1971, June 2023.

Alexandros Graikos, Nikolay Malkin, Nebojsa Jojic, and Dimitris Samaras. Diffusion models as plug-and-play priors. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 14715–14728. Curran Associates, Inc., 2022. URL `https://proceedings.neurips.cc/paper_files/paper/2022/file/5e6cec2a9520708381fe520246018e8b-Paper.pdf`.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 33, pages 6840–6851, 2020.

Aapo Hyvärinen. Estimation of non-normalized statistical models by score matching. *Journal of Machine Learning Research*, 6(24):695–709, 2005. URL http://jmlr.org/papers/v6/hyvarinen05a.html.

Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer New York, NY, 2 edition, 1998. doi: https://doi.org/10.1007/978-1-4612-0949-2.

Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 35, 2022.

Jonathan C. Mattingly, Andrew M. Stuart, and Desmond J. Higham. Ergodicity for sdes and approximations: locally lipschitz vector fields and degenerate noise. *Stochastic Processes and their Applications*, 101(2): 185–232, 2002.

Chicago Y. Park, Yuyang Hu, Michael T. McCann, Cristina Garcia-Cardona, Brendt Wohlberg, and Ulugbek S. Kamilov. Plug-and-play priors as a score-based method. In *2025 IEEE International Conference on Image Processing (ICIP)*, pages 49–54, 2025. doi: 10.1109/ICIP55913.2025.11084503.

Edward T. Reehorst and Philip Schniter. Regularization by denoising: Clarifications and new interpretations. *IEEE Transactions on Computational Imaging*, 5(1):52–67, 2019. doi: 10.1109/TCI.2018.2880326.

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. doi: 10.1137/16M1102884. URL https://doi.org/10.1137/16M1102884.

Ernest Ryu, Jialin Liu, Sicheng Wang, Xiaohan Chen, Zhangyang Wang, and Wotao Yin. Plug-and-play methods provably converge with properly trained denoisers. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 5546–5557. PMLR, 09–15 Jun 2019. URL https://proceedings.mlr.press/v97/ryu19a.html.

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations (ICLR)*, 2021.

Suhas Sreehari, S. V. Venkatakrishnan, Brendt Wohlberg, Gregery T. Buzzard, Lawrence F. Drummy, Jeffrey P. Simmons, and Charles A. Bouman. Plug-and-play priors for bright field electron tomography and sparse interpolation. *IEEE Transactions on Computational Imaging*, 2(4):408–423, 2016. doi: 10.1109/TCI.2016.2599778.

Daniel W. Stroock and S. R. S. Varadhan. Diffusion processes with continuous coefficients, i. *Communications on Pure and Applied Mathematics*, 22:345–400, 1969. URL https://api.semanticscholar.org/CorpusID:122695436.

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. doi: 10.1109/GlobalSIP.2013.6737048.

Pascal Vincent. A connection between score matching and denoising autoencoders. *Neural Computation*, 23 (7):1661–1674, 2011. doi: 10.1162/NECO_a_00142.

Xingyu Xu and Yuejie Chi. Provably robust score-based diffusion posterior sampling for plug-and-play image reconstruction. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.

Kai Zhang, Wangmeng Zuo, Shuhang Gu, and Lei Zhang. Learning deep cnn denoiser prior for image restoration. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 3929–3938, July 2017.

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

Yuanzhi Zhu, Kai Zhang, Jingwen Liang, Long 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)*, pages 1146–1156, 2023.

A STANDARD SDE THEOREMS CITED

We rely on standard results, cited rather than reproved: (a) EM mean-square error for globally Lipschitz, linearly growing coefficients (Karatzas and Shreve, 1998, Ch. 5); (b) strong uniqueness (Yamada–Watanabe), the strong-convergence regime (i) of Sec. 5; (c) weak existence under bounded continuous coefficients Stroock and Varadhan (1969), underpinning Prop. 3.1; (d) geometric ergodicity under local-Lipschitz, dissipative drifts with non-degenerate diffusion Mattingly et al. (2002).

B FULL PROOFS OF THE MAIN RESULTS

Proof of Prop. 3.1 (SDE representation). From the PnP map,

$$v^{t+1} - v^t = \big[h(v^t; y) - v^t\big] + \big[D(h(v^t; y), \sigma_t) - h(v^t; y)\big].$$

By Assumption 2.1, with $z = h(v^t; y)$,

$$D(z, \sigma_t) - z = \sigma_t^2 s_{\theta,\sigma_t}(z) + r_{\sigma_t}(z).$$

Expanding the score term about v^t and using global Lipschitzness (C1), $\sigma_t^2 s_{\theta,\sigma_t}(z) = \sigma_t^2 s_{\theta,\sigma_t}(v^t) + O(\sigma_t^2 L_s \|z - v^t\|)$, where the $O(\cdot)$ is absorbed into the Lipschitz constant K of the effective drift. Hence

$$v^{t+1} - v^t = \underbrace{\big[h(v^t; y) - v^t + \sigma_t^2 s_{\theta,\sigma_t}(v^t) + r_{\sigma_t}(v^t)\big]}_{\Delta\, b(t, v^t)} + \sigma_t\, \xi_t,$$

where ξ_t is the Gaussian perturbation $z - v^t$ of Assumption 2.1. The deterministic increment is exactly $\Delta\, b(t, v^t)$ with b as in (5); the martingale increment has variance σ_t^2, which the EM scheme attributes to $\sigma(t)\sqrt{\Delta}\,\xi_t$, giving $\sigma(t)^2\Delta = \sigma_t^2$. This identifies the PnP iteration as one EM step on (5). □

Proof of Thm. 4.1 (EM error bound). We compare the exact solution v_s and the EM iterates $V_n = V_{n\Delta}$ on the grid times $t_n = n\Delta$. Over one step the exact increment is

$$v_{t_{n+1}} - v_{t_n} = \int_{t_n}^{t_{n+1}} b(s, v_s)\, ds + \int_{t_n}^{t_{n+1}} \sigma(s)\, dW_s,$$

whereas the EM step is

$$V_{n+1} - V_n = \Delta\, b(t_n, V_n) + \sigma(t_n)\sqrt{\Delta}\,\xi_n, \qquad \xi_n \sim \mathcal{N}(0, I).$$

(i) One-step local error. Conditional on $\mathcal{F}_{t_n}$, the deterministic local error is $\delta_n^{\mathrm{drift}} = \int_{t_n}^{t_{n+1}} [b(s, v_s) - b(t_n, v_{t_n})]\, ds$. Since b is globally K-Lipschitz and $\|v_s - v_{t_n}\| = O(\sqrt{s - t_n})$ (drift contribution $O(s - t_n)$ plus Brownian contribution $O(\sqrt{s - t_n})$), the integrand is $O(\sqrt{\Delta})$, and integrating over an interval of length Δ gives $|\delta_n^{\mathrm{drift}}| = O(\Delta^{3/2})$ almost surely, hence $\mathbb{E}|\delta_n^{\mathrm{drift}}|^2 \le C\Delta^3$. The martingale mismatch is $\int_{t_n}^{t_{n+1}} (\sigma(s) - \sigma(t_n))\, dW_s$; because σ is piecewise constant on the grid in our VE parametrization, it vanishes identically, and in general the Itô isometry bounds its variance by $\int_{t_n}^{t_{n+1}} \bar{\sigma}^2 |\sigma(s) - \sigma(t_n)|^2\, ds = O(\Delta^3)$. *(ii) Propagation.* Couple the exact and EM trajectories to the same Brownian path. Writing $e_n = v_{t_n} - V_n$,

$$e_{n+1} = e_n + \Delta\, [b(t_n, v_{t_n}) - b(t_n, V_n)] + \delta_n,$$

so $\|e_{n+1}\| \le (1 + K\Delta)\|e_n\| + |\delta_n|$. Squaring and using $(1 + K\Delta)^2 \le 1 + 3K\Delta$ for $\Delta \le 1/K$,

$$E_{n+1} \le (1 + 3K\Delta) E_n + 3\mathbb{E}|\delta_n|^2 \le (1 + 3K\Delta) E_n + 3C\Delta^3.$$

Iterating n times with $n\Delta = T$, and using $(1 + 3K\Delta)^n \le e^{3Kn\Delta} = e^{3KT}$,

$$\sup_{s \le T} \mathbb{E}\|v_s - V_{s\Delta}\|^2 \le e^{3KT} \Big(E_0 + 3C \sum_{k=0}^{n-1} \Delta^3 e^{-3Kk\Delta} \Big) \le C_0\, T e^{3KT} \big(\Delta + \bar{\sigma}^2 \Delta\big),$$

where the geometric sum is bounded by $\Delta^2/(3K\Delta) = T/(3K)$ up to constants, i.e. each of the T/Δ local errors of size Δ^3 accumulates to $T\Delta^2$ in mean square. *(iii) Score-model perturbation.*

Let v^θ solve the SDE with the approximate drift $b_\theta = b + \delta b$, driven by the same Brownian motion, where (C3) gives $\|\delta b(v)\| \le \bar\sigma^2 \varepsilon_{\text{score}}$ uniformly in v. The difference $d_s = v_s - v_s^\theta$ satisfies

$$d_{s+\Delta} - d_s = \Delta\,[b(s, v_s) - b(s, v_s^\theta)] - \Delta\,\delta b(v_s^\theta) + \text{martingale},$$

and the same Grönwall recursion, now with source $\bar\sigma^4 \varepsilon_{\text{score}}^2$ instead of Δ^3, yields

$$\sup_{s \le T} \mathbb{E}\|v_s - v_s^\theta\|^2 \le C_0\, T e^{3KT} \varepsilon_{\text{score}}^2.$$

Since the EM scheme with the trained denoiser is the EM discretization of b_θ, combining (ii) and (iii) gives (8) after absorbing constants into $C(T, K, \bar\sigma)$. □

Proof of Thm. 5.3 (weak convergence of iterates). We chain three Wasserstein inequalities. *First coupling.* By the elementary bound $\mathcal{W}_2(\mathrm{Law}(X), \mathrm{Law}(Y)) \le (\mathbb{E}\|X - Y\|^2)^{1/2}$, coupling the PnP iterate v^t to the exact SDE solution $v_{t\Delta}$ on the same noise, Thm. 4.1 gives

$$\mathcal{W}_2\big(\mathrm{Law}(v^t),\, \mathrm{Law}(v_{t\Delta})\big) \le C_1(T, K, \bar\sigma)\big(\Delta + \varepsilon_{\text{score}}^2\big)^{1/2}.$$

Second coupling. Let $\tilde v_s$ solve the limiting homogeneous SDE $d\tilde v_s = b_\infty(\tilde v_s)\, ds + \sigma_{\min}\, dW_s$ with the same initial law. Write the cooling error as $\eta(s) = \sup_v \|b(s, v) - b_\infty(v)\|$; the schedule reaches its floor, so $\eta(s) \to 0$ and $\int_0^\infty \eta(s)\, ds < \infty$ (linearly/cosine schedules cool geometrically after a finite transient). Coupling the two non-autonomous drifts,

$$\|\mathrm{Law}(v_{t\Delta}) - \mathrm{Law}(\tilde v_{t\Delta})\|_{\mathcal{W}_2} \le C_2 e^{-\kappa t} \|\mathrm{Law}(v_0) - \mathrm{Law}(\tilde v_0)\|_{\mathcal{W}_2} + C_2\Big(\int_t^\infty \eta(s)\, ds\Big)^{1/2},$$

where the first term decays because $b(s)$ tracks b_∞ increasingly tightly, and the second is the accumulated mismatch after time t. *Third coupling.* By Thm. 5.2, the homogeneous semigroup is geometrically ergodic:

$$\mathcal{W}_2(\mathrm{Law}(\tilde v_s), \pi) \le C_\pi e^{-\gamma s} \mathcal{W}_2(\mathrm{Law}(\tilde v_0), \pi).$$

Summing the three terms by the triangle inequality for $\mathcal{W}_2$, the cooling remainder and the ergodic remainder both decay to zero as $t \to \infty$, leaving the discretization error as the irreducible floor:

$$\mathcal{W}_2(\mu_t, \pi) \le C_\pi e^{-\gamma t} \mathcal{W}_2(\mu_0, \pi) + C_1\big(\Delta + \varepsilon_{\text{score}}^2\big)^{1/2} + C_2\Big(\int_t^\infty \eta(s)\, ds\Big)^{1/2},$$

which is (10). □

Proof of Thm. 5.4 (Wasserstein posterior bound). *Exact-score invariant law.* With $h(v; y) = v - \lambda^2 \nabla f(v, y)$, the limiting drift is $b_\infty(v) = -\lambda^2 \nabla f(v, y) + \sigma_{\min}^2 s_{\sigma_{\min}}(v) = -\sigma_{\min}^2 \nabla_v \log p_{\sigma_{\min}}(v \mid y)$ up to the variance inflation. The stationary Fokker–Planck equation $\nabla \cdot (b_\infty \rho) - \frac{\sigma_{\min}^2}{2} \Delta \rho = 0$ is solved by $\rho = p_{\sigma_{\min}}(\cdot \mid y)$, since the drift is minus $\sigma_{\min}^2$ times its own score (detailed balance): plugging $\nabla \log \rho = b_\infty / \sigma_{\min}^2$ gives zero divergence. *Perturbed invariant law.* Let $\delta s = s_\theta - s$ with $\|\delta s\|_{H^{-1}(p_{\sigma_{\min}})} \le \varepsilon_{\text{score}}$, and let π_θ be the invariant law of the perturbed drift $b_\infty + \sigma_{\min}^2 \delta s$. For any smooth bounded g, introduce the Poisson equation $\mathcal{L}_0 \phi = g - \int g\, dp_{\sigma_{\min}}$, whose solution satisfies $\|\nabla \phi\|_{L^2(p)} \le C \|g\|_{L^2(p)} / \sqrt{m}$ by the Poincaré inequality with spectral gap m supplied by dissipativity (Assumption 5.1). Stationarity of π_θ gives

$$\int g\, d\pi_\theta - \int g\, dp_{\sigma_{\min}} = \int \mathcal{L}_0 \phi\, d\pi_\theta = \sigma_{\min}^2 \int \langle \delta s, \nabla \phi \rangle\, d\pi_\theta,$$

because the unperturbed generator annihilates the centered Poisson solution. Bounding by H^{-1}-H^1 duality,

$$\Big| \int g\, d(\pi_\theta - p_{\sigma_{\min}}) \Big| \le \sigma_{\min}^2\, \varepsilon_{\text{score}}\, \|\nabla \phi\|_{H^1(\pi_\theta)} \le \frac{C\, \sigma_{\min}^2 \varepsilon_{\text{score}}}{m}\, \|g\|_{L^2(p_{\sigma_{\min}})}.$$

Taking the supremum over 1-Lipschitz g (Kantorovich duality, $\mathcal{W}_2(\pi_\theta, p) = \sup_{\mathrm{Lip}(g) \le 1} \int g\, d(\pi_\theta - p)$) and converting the L^2 test function to its Lipschitz representative,

$$\mathcal{W}_2\big(\pi_\theta, p_{\sigma_{\min}}(\cdot \mid y)\big) \le \frac{C \sigma_{\min}^2}{m}\, \varepsilon_{\text{score}},$$

which is (11). □

C Derivation of Prop. 3.2

The forward VE perturbation is $z = v + \sigma\epsilon$. Tweedie (MMSE) gives $D^*(z,\sigma) - z = \sigma^2 \nabla \log p_\sigma(z) = \sigma^2 s_\sigma(z)$. With $h(v;y) = v - \lambda^2 \nabla f(v,y)$, the PnP drift (5) becomes $b(t,v) = \Delta^{-1}[-\lambda^2 \nabla f(v,y) + \sigma_t^2 s_{\sigma_t}(v)]$. The reverse-time SDE of the VE diffusion targeting $p_{\sigma_{\min}}(x \mid y) \propto p(y \mid x) p_{0,\sigma_{\min}}(x)$ has drift $-\nabla f + \sigma_{\min}^2 s_{\sigma_{\min}}$ up to the time rescaling $dt \leftrightarrow \Delta$ and the EDM/DDPM schedule reparametrization Karras et al. (2022); Ho et al. (2020); Song et al. (2021). The Fokker–Planck stationarity condition is solved by $\pi \propto p(y \mid x) p_{0,\sigma_{\min}}(x)$, confirming the target.

D Gaussian-toy example with explicit invariant law

Let $x_0 \sim \mathcal{N}(0, \tau^2 I)$, $y = Mx_0 + \epsilon$, $\epsilon \sim \mathcal{N}(0, \sigma_{\text{meas}}^2 I)$, and take a Gaussian denoiser $D(z,\sigma) = az$. The PnP-SDE is a linear OU process whose invariant law $\pi = \mathcal{N}(\mu_\pi, \Sigma_\pi)$ has

$$\Sigma_\pi = \left(\frac{1}{\lambda^2} + \frac{M^\top M}{\sigma_{\text{meas}}^2} - \frac{2a}{\sigma_{\min}^2} \right)^{-1}, \qquad \mu_\pi = \Sigma_\pi \frac{M^\top y}{\sigma_{\text{meas}}^2},$$

which equals the variance-inflated posterior up to the $\sigma_{\min}^2$ inflation. This explicit π is what the stationary-law study overlays and what the EM-rate study uses to measure the discretization error exactly.

E Additional experiment details

Unbounded-denoiser divergence ablation.

Denoiser	final PSNR (dB)	final $\|v^t\|$	diverges?
DnCNN ($g = 1$)	26.37	82.00	no
Unbounded $g = 5$	15.49	89.66	no
Unbounded $g = 10$	-191.23	9.3×10^{11}	**yes**
ablation $g = 10 \to$clamp@iter20	17.70	85.58	recovers

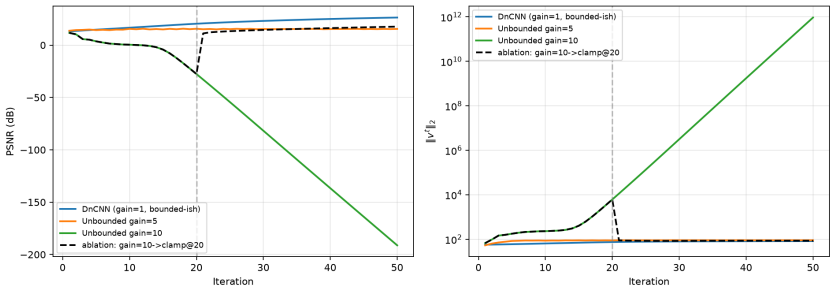

Figure 4: Iterate norm versus iteration for gain $g = 1, 5, 10$; re-clamping at iteration 20 recovers a bounded trajectory.

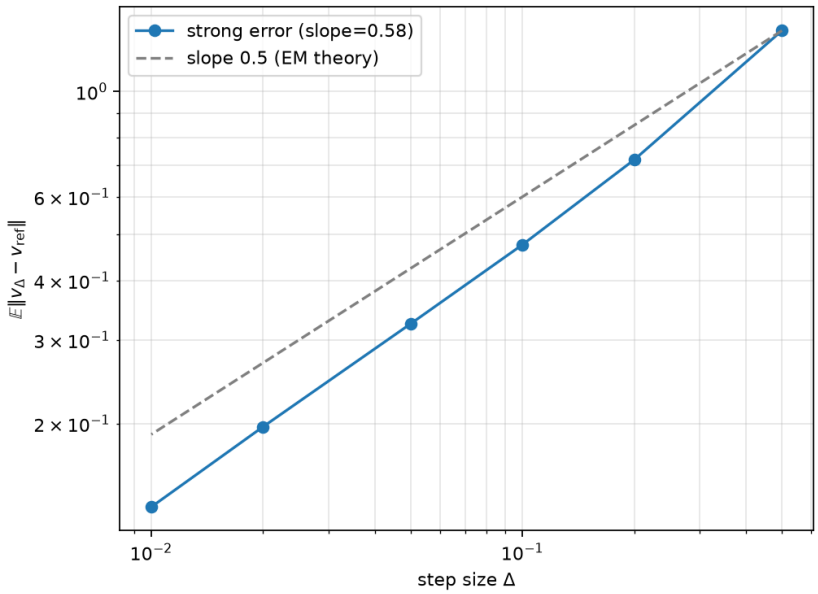

Figure 5: EM-rate study: log-log strong error versus step Δ; fitted slope $0.58 \approx 1/2$.

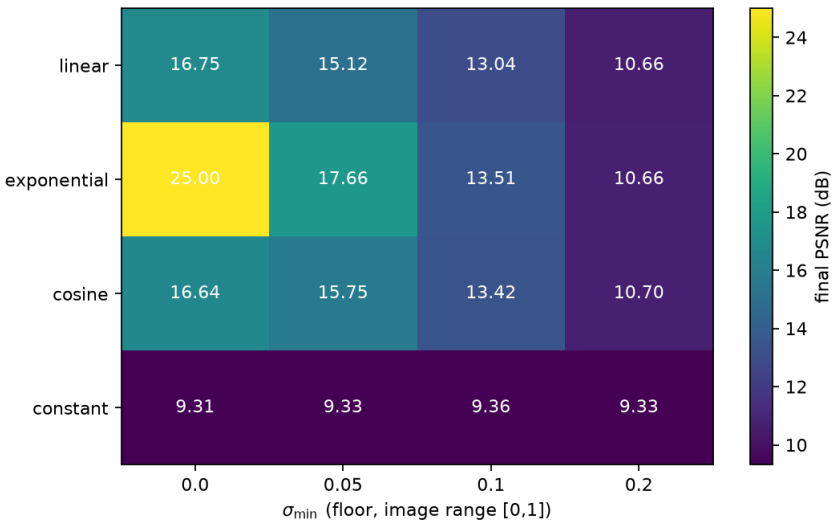

Figure 6: Schedule sweep: final PSNR over schedule type and terminal noise floor σ_{min}.

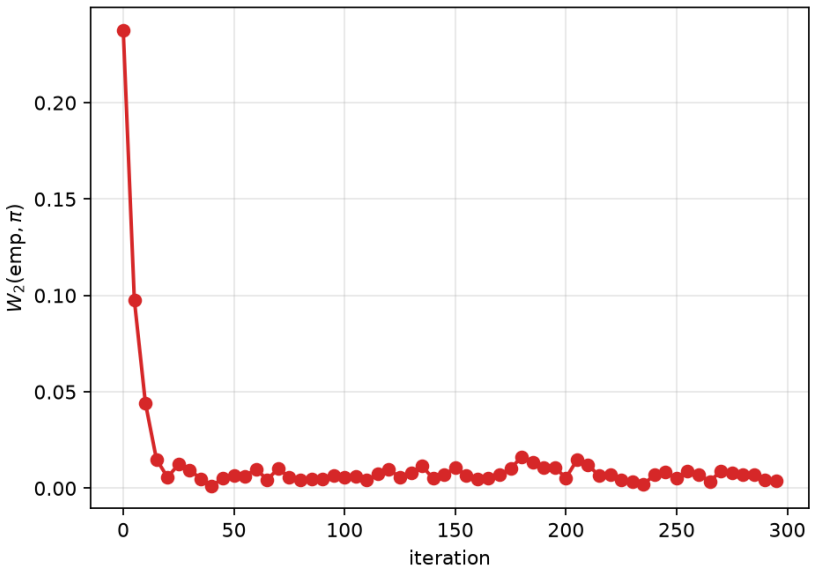

Figure 7: Stationary-law study: $\mathcal{W}_2(\mu_t, \pi)$ versus iteration, decaying geometrically to 0.0037.

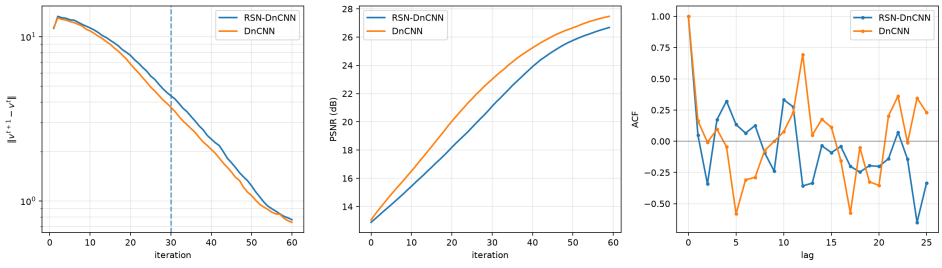

Figure 8: Convergence diagnostics: inter-iterate distance, PSNR stabilization, and post-burn-in autocorrelation.