DPS：扩散后验采样
数学推导与工程映射

Chung 等 · ICLR 2023 · arXiv:2209.14687v4
2026-11-25 · 原创中英双语深读 · 原文 CC BY-NC-ND 4.0

01 | 先区分测量模型、先验与许可

来源定位：原文 pp.1–3，§1–2；版本与许可见 arXiv 摘要页

DPS 要解决的是：已知物理前向算子 A，如何从有噪声的观测 y 恢复未知干净状态 x_0，并保留不适定性带来的多解。先验 $p(x_0)$ 来自扩散模型，似然 $p(y|x_0)$ 来自仪器和噪声统计。两者在后验中相乘，但作用不同：前者偏好训练分布中的合理状态，后者评价状态是否解释本次观测。视觉上逼真的样本与物理上可识别的参数不能混为一谈。

底本为 Hyungjin Chung、Jeongsol Kim、Michael T. Mccann、Marc L. Klasky、Jong Chul Ye 的 arXiv:2209.14687v4，2024-05-20 修订，英文 PDF 30 页，发表于 ICLR 2023。其许可实际为 CC BY-NC-ND 4.0：允许有条件分享未改动论文，但不授予传播改编稿的许可。本册因此采用原创中英双语数学推导与工程映射，而不是逐段译文；原版完整保留，推导与工程建议分别标注来源。

设状态维数 d、测量维数 m。白噪声模型 $y=A(x_0)+n$ 只是高斯观测的一种形式；光子计数应直接写成 $y_j \sim \text{Poisson}(\lambda_j(x_0))$，不能误当成与状态无关的加性噪声。实验前先写清 y 的单位、曝光时间、像素积分、归一化及有效掩码，才有可能正确选似然。若 y 是电压而 A 输出归一化亮度，梯度数值再漂亮也没有测量意义。

阅读本册时可沿三个层级核对：状态 x_0 是什么、测量 y 怎样生成、最终要估计哪个函数 $g(x_0)$。同一张重建图像可用于显示、定量定位或安全决策，三种用途要求不同误差证据。把目标函数及单位写在实验配置首行，可避免后面以视觉效果替代量测承诺。

$$p(x_0|y) \propto p(y|x_0)\, p(x_0)$$

$$\text{Gaussian: } y=A(x_0)+n,\ n\sim N(0,\sigma^2 I),\ x_0\in R^d,\ y\in R^m$$

$$\text{Photon counts: } y_j\sim \text{Poisson}(\lambda_j(x_0)),\ \lambda_j\geq 0$$

A posterior solver needs two separate contracts: the learned distribution of plausible clean states and the calibrated forward distribution of the current measurement. Neither visual plausibility nor a low measurement residual establishes uniqueness. The downloaded v4 paper is CC BY-NC-ND 4.0, so this companion develops original explanations and derivations instead of distributing a translated derivative. For a new experiment, define physical units, masks, exposure integration and the noise mechanism before implementing the likelihood. In particular, photon counts require a signal-dependent model; the additive Gaussian shorthand cannot represent every sensor.

02 | 从高斯扰动到时间依赖得分

来源定位：原文 pp.2–3，式(1)–(5)；本册统一时间记号

前向扩散逐步把干净样本变成噪声。VP 参数化可把任意时刻状态写成 $x_t=a_t x_0+b_t \varepsilon$，其中 $a_t=\sqrt{\bar{\alpha}_t}$、$b_t=\sqrt{1-\bar{\alpha}_t}$，$\varepsilon$ 是标准高斯。训练得到的 score $s_\theta(x_t,t)$ 近似该时刻边缘密度的梯度 $\nabla \log p_t(x_t)$，不是直接输出干净图像，也不是测量似然的梯度。若网络输出噪声 ε_θ，则转换关系为 $s_\theta=-\varepsilon_\theta/b_t$。

对 VP 连续过程，前向漂移为 $-\beta(t)x/2$，扩散系数为 $\sqrt{\beta(t)}$。反向 SDE 的漂移加上 $-\beta(t)$ score；这里时间从大到小积分，dt 为负，因此不要再按直觉自行翻一次符号。若将后验得分替换为先验得分与条件似然得分之和，就得到条件反向过程。真正困难在于第二项是在噪声状态 x_t 上定义，而测量仪器关联的是 x_0。

工程实现应选一种参数化并锁定它：网络可能预测噪声、得分、速度或干净状态，系数不能跨参数化直接挪用。本册只推导 VP/DDPM 情形；把公式用于 flow-matching 的速度场之前，必须重新建立路径条件期望关系。时间方向、噪声日程和归一化错误，会在简单线性问题中产生系统偏差，应该先通过可解析测试识别。

去噪得分匹配的监督目标在高斯路径下可直接写成 $-\varepsilon/b_t$。训练时平均的是扰动条件得分，最优网络才恢复边缘得分；不能把某一次采样的噪声除以 b_t 当作真实边缘得分。b_t 趋近零时该监督尺度变大，损失时间权重也因此属于训练契约。

> $$x_t=a_t x_0+b_t \varepsilon;\ a_t^2=\bar{\alpha}_t,\ b_t^2=1-\bar{\alpha}_t$$

> $$s_\theta(x_t,t)\approx\nabla \log p_t(x_t)=-\varepsilon_\theta(x_t,t)/b_t$$

> $$\text{reverse } dx=[-\beta(t)x/2-\beta(t)\{s_\theta+\nabla \log p_t(y|x_t)\}]dt+\sqrt{\beta(t)}\,d\bar{w}$$

The score is a gradient with respect to a noisy state, whereas the sensor likelihood is specified for a clean state. This mismatch is the central obstacle, not a missing Bayes identity. A VP noise-prediction model must be converted using $s=-\varepsilon/b$; a flow velocity is a different object. In the reverse SDE, the integration variable runs backward, so a negative time increment already carries the sign reversal. Choose one parameterization and validate its coefficients on a tractable model before adding a complicated forward operator.

03 | Tweedie 公式：把去噪均值推出来

来源定位：原文 p.4 命题1及附录 A；本册逐步推导

令 $q(x_t|x_0)=N(a_t\ x_0, b_t^2 I)$。对边缘密度 $p_t(x_t)=\int q(x_t|x_0)p(x_0)dx_0$ 求导，在可交换微分积分的正则条件下，密度梯度除以密度可写为条件期望。高斯核对 x_t 的对数梯度是 $-(x_t-a_t\ x_0)/b_t^2$，把它放进条件期望，就得到得分与干净状态后验均值的关系。这里的条件只有 x_t，还没有当前测量 y。

代数整理得到 $\hat{x}_0(x_t)=(x_t+b_t^2\ s\theta)/a_t$。若 score 精确，它给出 $E[x_0|x_t]$；若网络近似，则均值也有模型误差。它不是从后验里随机抽取的一个样本：多峰分布的均值可能位于两个模式之间，甚至不在高密度流形上。论文用它构造似然的可计算代理，不能据此宣称它总是物理可实现状态。

实际反传时 $\hat{x}_0$ 是 x_t 的函数。其雅可比包含 score 网络的雅可比，$J_{\hat{x}}=(I+b_t^2 J_s)/a_t$。若把 $\hat{x}_0$ detach，梯度无法通过网络传回 x_t；若只保留单位映射项，则已改成另一近似，必须在实验中明确。高噪声处 a_t 很小，也会放大误差，所以不应不加检查地在最极端时间点使用不稳定系数。

还可进一步求导：对于常数方差高斯扰动，$J_\mu=(a_t/b_t^2)\mathrm{Cov}(x_0|x_t)$。因此条件均值的响应与条件不确定性有直接联系。标量先验例中 $J_\mu=0.6$、$C=0.64$，而 $a_t\ C/b_t^2=0.6$，正好吻合。这个关系也提示忽略去噪器雅可比并非无关紧要的加速。

> $$\nabla \log p_t(x_t)=E[-(x_t-a_t\ x_0)/b_t^2 \mid x_t]$$

> $$a_t\ E[x_0|x_t]=x_t+b_t^2\nabla \log p_t(x_t)$$

> $$\hat{x}_0=(x_t+b_t^2 s\theta)/a_t;\ J_{\hat{x}}=(I+b_t^2 J_s)/a_t$$

Differentiate the Gaussian corruption kernel under the integral, then express the result as a conditional expectation. This yields Tweedie's posterior mean directly. The conditioning is on x_t, not on the experimental measurement y. A mean can fall between plausible modes, so it is a computational bridge rather than a guarantee of physical realizability. Backpropagation must include the dependence of the denoised estimate on the noisy input. Detaching that estimate removes the intended guidance path; dropping the network Jacobian changes the approximation and should be evaluated as an explicit ablation.

04 | DPS 的核心近似：先取均值，再过似然

来源定位：原文 pp.4–5，式(7)–(15)；本册局部误差分析

条件独立关系 $y \leftarrow x_0 \rightarrow x_t$ 给出精确等式 $p(y|x_t)=E[p(y|x_0)|x_t]$。这个期望通常难算，因为需要整个 $p(x_0|x_t)$，而网络容易给出的是其均值。DPS 把外层似然的期望近似为均值处的似然，即 $p(y|x_t)\approx p(y|\hat{x}_0)$。这是把分布压缩成一个点，而不是贝叶斯公式本身。

令 $f(x)=p(y|x)$，$\mu=E[x_0|x_t]$、$\Sigma=Cov(x_0|x_t)$。若 f 足够平滑，二阶泰勒展开后，一阶项的条件期望消失，领先误差约为二分之一 $\mathrm{tr}(H_f(\mu)\Sigma)$。于是局部曲率大、去噪不确定性大时，点代入更危险。即使前向算子 A 是线性的，高斯似然仍是非线性函数，均值代入通常也不精确。

论文用 Jensen gap 讨论似然值近似误差，并给出条件上界；这不自动成为对数似然梯度或最终采样分布的误差界。除以很小的似然会放大相对误差，求导还需要额外光滑性控制。因而较小似然值差异、较低 FID、正确后验方差是三个不同验证对象。工程上可以从标量高斯例子清楚看到：忽略条件协方差，往往会把测量指导做得过强。

当条件分布近似高斯且 A 为线性矩阵时，积分可以显式完成：$y|x_t$ 的协方差为 $R+A\Sigma A^T$。点代入把额外项 $A\Sigma A^T$设成零，因此高噪声时会对观测过度自信。利用这一项构造协方差感知指导是一个可检验候选，应先在低维问题确认，再讨论复杂图像近似。

> $$p(y|x_t)=\int p(y|x_0)p(x_0|x_t)dx_0$$

> $$\text{DPS plug-in: } E[f(x_0)|x_t]\approx f(E[x_0|x_t])$$

> $$E[f(x_0)]-f(\mu)\approx \tfrac{1}{2}\,\mathrm{tr}(H_f(\mu)\Sigma)$$

DPS replaces a likelihood averaged over the clean-state conditional distribution with the likelihood evaluated at its mean. This is a point approximation, even when the forward map is linear. A local Taylor expansion exposes the interaction between likelihood curvature and conditional uncertainty. A bound on likelihood values does not by itself bound log-likelihood gradients or establish exact posterior sampling. Small likelihood values and differentiation introduce additional sensitivities. A useful validation strategy therefore separates perceptual reconstruction quality, guidance accuracy and posterior uncertainty rather than treating any one of them as a substitute for the others.

05 | 高斯似然：链式法则与二倍系数

来源定位：原文 pp.5–6，式(16)、(21)；本册严格区分残差平方与半平方

对方差 σ² 的独立高斯噪声，负对数似然除常数为 L=||A(x₀)−y||²/(2σ²)。将 x₀ 换成 x̂₀ 后，对 x_t 求梯度必须经过两层映射：先过 A，再过去噪器。若 r=A(x̂₀)−y，则 ∇L=J_x̂ᵀJ_AᵀR/σ²。对于非线性可微 A，不必显式生成巨型雅可比，反向自动微分的向量雅可比积即可计算。

必须固定损失约定。若代码的 loss 是 residual.square().sum()，它的梯度自带 2；若是 0.5*sum，系数就没有 2。原文将若干常数吸收到指导步长，不能把其 ρ 与严格概率式中的 1/σ² 在不同损失定义下机械等同。本册在推导中始终保留半平方，使 σ 的物理意义清楚；实际有限步调参另用 η_i 表示。

对一般协方差 R，应把欧氏残差改成 rᵀR⁻¹r/2；白化后维度和单位才一致。一个像素一个样本还是对全部像素求平均，也会改变梯度规模。输入分辨率翻倍、测量掩码改变，若仍用同一未经归一化的步长，实验差异可能主要来自损失缩放，而非模型能力。

一个有限差分检查是定义 φ(x_t)=L(y,A(x̂₀(x_t)))，比较方向 v 上 [φ(x_t+hv)−φ(x_t−hv)]/(2h) 与 vᵀ∇φ。先用双精度、光滑小算子和一组逐减 h；过大 h 有截断误差、过小有消减误差。不能只用一次极小步长就宣称梯度验证通过。

$$L(x_t)=\tfrac{1}{2}[A(\hat{x}_0)-y]^{T}R^{-1}[A(\hat{x}_0)-y]$$

$$\nabla_x_t \log p(y|\hat{x}_0)=-J_\hat{x}^{T}J_A^{T}R^{-1}r$$

$$R=\sigma^2 I \Rightarrow \text{gradient}=-J_\hat{x}^{T}J_A^{T}r/\sigma^2$$

Apply the chain rule through both the forward operator and the denoiser. For the half-squared Gaussian residual, the coefficient is exactly $1/\sigma^2$; for an unhalved squared residual, the gradient introduces an additional factor of two. Algorithmic step sizes can absorb constants, but a physically calibrated noise parameter cannot silently change meaning. Correlated measurements require their covariance or a whitening operator. Sum versus mean reduction, image resolution and masking also rescale guidance, so they must remain part of the reproducible numerical contract.

06｜离散采样：先走先验，再做测量指导

来源定位：原文 p.5 算法1、2，附录 D；本册无歧义下标写法

采用 DDPM 记号 α_i=1−β_i，ᾱ_i 为累积乘积。令 i 从 N 递减到 1：由 x_i 计算 score 和 $\hat{x}_0$，按祖先采样构造一个先验候选 x′_(i−1)，再沿基于 x_i 的测量损失梯度修正它。这个顺序体现先验生成与测量约束交替作用。最后输出所选终态估计；实现应明确返回去噪均值还是带残余噪声状态。

先验候选的均值是 x_i 与 $\hat{x}_0$ 的时间系数线性组合，噪声方差可以沿用已训练扩散模型。η_i 不仅取决于 σ，也受离散化、损失约定和指导归一化影响。论文附录讨论按残差范数缩放的步长启发式；它有稳定性经验价值，却不是从精确后验直接推导的通用常数。残差为零时分母需要明确停止规则或数值正则，不能依赖偶然浮点噪声。

显存和计算代价主要来自每步网络调用及反传。当前制作只做 CPU 文献处理与小型解析校验，不运行扩散网络，也不占用 GPU。若未来部署实时机器人，应把昂贵重建放在异步感知层，给控制层输出时标、有效区域、误差估计和过期判据，不能将论文离线图像采样循环直接放进高频伺服。

$$x'_{(i-1)} = [\sqrt{\alpha_i}(1-\bar{\alpha}_{(i-1)})/(1-\bar{\alpha}_i)]x_i + [\sqrt{\bar{\alpha}_{(i-1)}}\beta_i/(1-\bar{\alpha}_i)]\hat{x}_0 + \tilde{\sigma}_i z$$

$$x_{(i-1)} = x'_{(i-1)} - \eta_i \nabla_{x_i} L(y, A(\hat{x}_0(x_i)))$$

$$i = N,\ldots,1;\ z \sim N(0,I);\ \bar{\alpha}_0 = 1$$

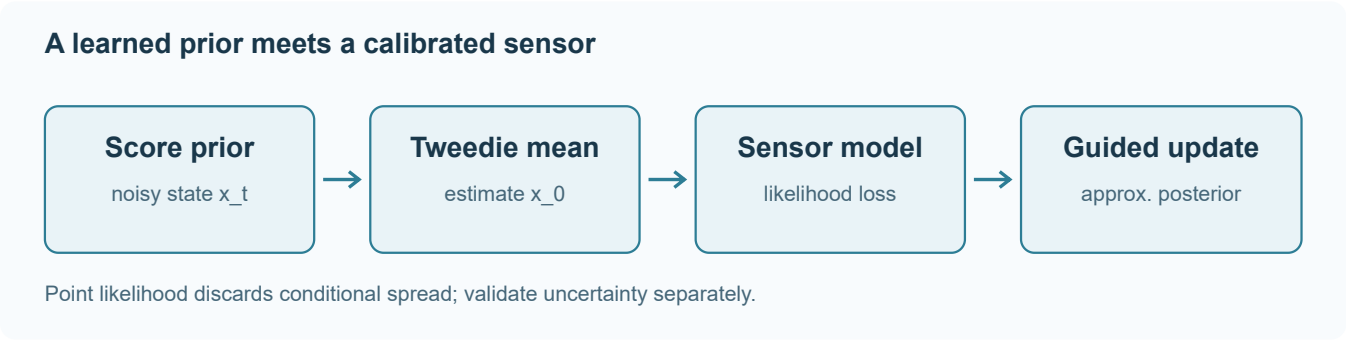

图 D1：本册原创；先验、去噪均值、测量模型与更新各有独立契约。

Use an explicit index convention: at level i, compute the denoised estimate and its measurement loss, draw the ancestral proposal for i−1, then apply guidance differentiated with respect to x_i. State whether the final return value is the denoised estimate or the last stochastic state. Residual-normalized step sizes are practical heuristics, not universal posterior coefficients. Their zero-residual behavior also needs a defined rule. A full network forward/backward pass per step makes this an offline or asynchronous perception method, not an automatic fit for a high-frequency servo loop.

07 | Poisson：低计数不能套高斯权重

来源定位：原文 p.6，式(17)–(22)；本册精确计数似然及零计数处理

对光子计数 y_j，设前向预测 λ_j(x)>0。精确 Poisson 负对数似然是 Σ_j[λ_j−y_j log λ_j+log(y_j!)]。对状态求导，测量域残差变成 1−y/λ，再由前向模型与去噪器反传。λ=曝光增益×预测辐照度+暗计数时，增益与暗计数应在算子内，不应先把 y 任意归一化到 [0,1] 再沿用计数公式。

论文在计数足够大时采用高斯近似，并以测量 y_j 替代方差中的预测 λ_j，得到按 1/(2y_j) 加权的平方残差。这个近似在低计数和零计数处不可靠，y_j=0 还会产生无穷权重。把分母简单改成 y_j+ε 是一个额外估计器选择，ε 的单位必须是计数，并应报告其偏差；更清楚的替代是使用带正值参数化的精确 Poisson 损失。

精确似然也不等于精确 DPS：$\hat{x}_0$ 点代入仍忽略 p(x₀|x_t) 的宽度。热像素、读出噪声、阈值失配和事件相机漏触发未必属于理想 Poisson，需要鲁棒混合、额外标定或剔除模型。把噪声族误配造成的误差归给网络先验，会掩盖真正可修复的传感器问题。

精确计数似然要求预测 λ 非负，可由正强度物理模型或 softplus 参数化保证；若把负强度简单裁成常数，裁切区的梯度可能为零，优化器无法恢复。读出噪声与暗计数同时存在时，Poisson–Gaussian 复合模型比单纯泊松更贴近相机，但其计算代价和参数标定须另行验证。

> $L_P=\Sigma_j[\lambda_j-y_j \log \lambda_j+\log \Gamma(y_j+1)]$

> $\nabla_{x_t} L_P=J_{\hat{x}}^{T} J_{\lambda}^{T}(1-y/\lambda)$

> large-count surrogate: $L\approx\Sigma_j (y_j-\lambda_j)^2/(2y_j)$, only when justified

---

The exact Poisson objective remains well-defined at an observed count of zero when the predicted intensity is positive. The high-count weighted least-squares surrogate does not: replacing its variance by y creates a singular weight at zero. An epsilon fix introduces an additional estimator and requires a unit-consistent justification. Exposure gain and dark counts belong in the forward model. Even an exact count likelihood leaves the DPS plug-in approximation intact. Hot pixels, threshold errors and readout noise should be modelled or calibrated rather than silently forced into an inappropriate noise family.

08｜一个能算到底的高斯反例

来源定位：本册原创解析例；对照原文 p.4–5 近似机制

取 $x_0 \sim N(0,\tau^2)$，测量 $y=h\,x_0+n$、$n\sim N(0,\sigma^2)$。真实后验仍是高斯：方差 $V=(\tau^{-2}+h^2/\sigma^2)^{-1}$，均值 $m=V\,h\,y/\sigma^2$。这个例子无需训练网络，就能检验观测噪声、先验权重、单位换算和采样不确定性。取 $\tau^2=1$、$h=2$、$\sigma=0.5$、$y=1$，得到 $V=1/17\approx0.058824$、$m=8/17\approx0.470588$。

再看噪声时刻：若 $a=0.6$、$b=0.8$，则 $\mathrm{Var}(x_t)=1$，去噪条件均值 $\mu=0.6x_t$，条件方差 $C=0.64$。精确 $y|x_t$ 为 $N(h\mu,\sigma^2+h^2C)$，方差为 2.81；而 DPS 点代入只保留 $\sigma^2=0.25$。因此在本例中，两者的指导方向相同，但点代入指导强度是真实值的 11.24 倍。

在 $x_t=0.3$ 时，$\mu=0.18$、残差 $y-h\mu=0.64$。真实对数似然梯度为 $(h\times0.6)\times0.64/2.81\approx0.273310$；点代入为同一分子除 0.25，得到 3.072000。这不是模型训练失败，而是忽略条件不确定性产生的结构误差。工程上可先用此测试验证代码，再测有限步采样的均值和方差是否落在真实后验附近。

本例还有一个便利性质：先验方差为 1 且 $a^2+b^2=1$，所以 x_t 的边缘分布仍为标准高斯。由此 $\mathrm{score}=-x_t$，Tweedie 公式直接给出 $\hat{x}_0=(x_t-b^2x_t)/a=ax_t$。它把网络、去噪均值和条件似然连接起来，是检查三个模块是否采用相同记号的闭环测试。

> $$V=(1/\tau^2+h^2/\sigma^2)^{-1};\ m=Vhy/\sigma^2$$

> $$C=\tau^2b^2/(a^2\tau^2+b^2);\ \mu=a\tau^2x_t/(a^2\tau^2+b^2)$$

> $$\text{exact guidance / plug-in guidance}=\sigma^2/(\sigma^2+h^2C)=1/11.24$$

A scalar Gaussian model gives an exact unit test for the entire conditional-sampling argument. With prior variance one, gain two and noise standard deviation 0.5, the clean posterior has mean $8/17$ and variance $1/17$ for $y=1$. At corruption coefficients $a=0.6$ and $b=0.8$, conditional clean-state variance adds 2.56 to measurement variance. The exact noisy-state likelihood therefore has variance 2.81, whereas the point approximation uses 0.25. At $x_t=0.3$ the respective gradients are approximately 0.273310 and 3.072000. This is approximation bias, not a training defect.

09 | 非线性算子：可微不等于可识别

来源定位：原文 pp.6–9 相位恢复与非均匀去模糊；本册对称性分析

DPS 的一项实用价值是不必显式求前向算子的奇异值分解。只要能从标量损失对输入做向量雅可比积，就可处理可微非线性模型。相位恢复以傅里叶幅值为测量，非均匀去模糊可用神经网络近似前向过程。这里“能反传”只是计算条件，不是问题可逆的充分条件；若前向模型错了，优化器会精确解决错误模型。

对复信号 x，幅值测量 A(x)=|Fx| 对全局相位 x→e^(iφ)x 不变；在特定边界约定下，还可能有平移或共轭反演等歧义。观测无法区分的解需要在商空间比较，或明确额外参考。把原始坐标逐像素差异直接当作重建失败，会惩罚合法等价解；反过来，仅选择最像真值的后验样本，也会夸大普通部署表现。

在机械状态、粒子像与事件融合中也要问同一问题：哪个参数真正进入测量？相机只看投影时，遮挡深度、未采样区域速度或某些尺度可能处在不可观测方向。先验可选出一个合理解，却不能把先验补全写成直接测量。报告应区分观测残差、场景视觉质量、目标量误差和不可识别等价类。

前向模型来自神经网络时，应另留独立物理标定数据：在多种状态下比较真实仪器与 A 的输出，检查残差是否与状态、亮度或运动速度相关。若残差有系统结构，增加逆解迭代只能更努力拟合这个结构性偏差，不能靠更强先验补救算子身份错误。

> A(x)=|Fx|; A(e^(iφ)x)=A(x)

> J_A(x)δx=0 ⇒ δx is locally unobserved to first order

> differentiate a valid forward model; do not infer injectivity from differentiability

---

Automatic differentiation extends guidance to nonlinear operators without an explicit SVD, but it does not remove physical ambiguities. Fourier magnitude is invariant to a global phase, and other symmetries can survive under specific boundary conventions. Evaluate either after an agreed alignment or on the corresponding equivalence class. The same principle applies to hidden robot states and unobserved flow regions: a prior selects plausible solutions but cannot turn an unobserved degree of freedom into a direct measurement. Keep operator validity, computational differentiability and identifiability as separate questions.

10 | 为何硬投影在有噪声时会伤害解

来源定位：原文 pp.6–7，图3与 MCG 比较；本册线性几何解释

无噪声线性观测 Ax=y 定义仿射可行集。当它与真实信号分布相容时，投影到该集合可能有效。但含噪观测 y=Ax_true+n 的等式集合偏离真实状态；若 A 的小奇异值方向存在，强迫精确一致就把噪声除以很小的奇异值，产生大幅状态偏移。若有测量维度不在 A 的值域，甚至可能根本不存在精确解。

令 A=UΣV^{T}，最小范数反演把 u_k 方向的测量误差乘 1/s_k。小 s_k 不代表前向梯度大，而代表逆问题敏感。DPS 用柔性的似然梯度而非每步严格投影，允许按噪声尺度保留残差，因此避免把全部噪声当作必须拟合的信号。论文的流形图可帮助直觉理解，但复杂神经网络去噪器并非处处严格正交投影。

选择软约束仍需验证：指导过弱会退回先验幻觉，指导过强会过拟合噪声，时间相关误差又可能累积。可用残差白化后的统计量检查拟合尺度，但其自由度受正则化、参数估计和模型偏差影响，不能无条件把每像素残差等于 σ 当成通用停止规则。

从正则化观点，在线性高斯先验下，后验均值等同于带二次先验项的最小化解；在 A 的每个奇异方向，其滤波增益为 $\tau^2 s_k/(\tau^2 s_k^2+\sigma^2)$。当 s_k 很小时增益趋近零，而不再像伪逆的 1/s_k 那样发散。这给出了软约束抑制噪声放大的可计算例证。

> $\delta x = A^{\dagger} n = \Sigma_k v_k (u_k^{T} n)/s_k$

> hard consistency: $Ax=y$; noisy likelihood: penalize $||Ax-y||^2/(2\sigma^2)$

> small forward singular value ⇒ large inverse noise amplification

---

Exact data consistency can be the wrong target when the measurement is noisy. In a linear SVD, a small forward singular value amplifies the corresponding inverse noise component. Soft likelihood guidance permits a noise-sized discrepancy instead of fitting every fluctuation. This explains the practical distinction between a strict projection and a probabilistic data term without asserting that every neural denoiser is an exact manifold projector. Guidance strength remains a modelling and numerical choice: too little discards the observation, while too much can reproduce its noise.

11 | 读原文实验：感知指标与测量指标分开

来源定位：原文 pp.7–10，表1–4与附录 E；本册指标解释

原文在 FFHQ 与 ImageNet 的 256×256 图像上各取 1000 验证样本，考察四倍超分、方框 / 随机修补、高斯 / 运动去模糊等线性问题，以及相位恢复、非均匀去模糊等非线性问题。主要高斯测量噪声 σ=0.05，图像归一化到 [0,1]；这些条件必须随数值一起引用。FFHQ 扩散模型用排除验证集后的 49,000 训练图像、100 万步训练；比较中的扩散方法共享 score，以控制先验差异。

以原文表 1 的 FFHQ 为例，DPS 四倍超分 FID=39.35、LPIPS=0.214，DDRM 为 62.15、0.294；高斯去模糊 DPS 为 44.05、0.257，DDRM 为 74.92、0.332。这些是特定图像分布与算子下的作者实测结果，本册未重跑网络。FID 比较样本分布的特征统计，LPIPS 衡量感知差异；它们都不是某颗粒子的质心误差，也不是速度梯度或流量守恒误差。

论文还在附录给出 PSNR、SSIM。更强感知先验可能使结果更自然，同时改变细弱粒子、狭缝或边界位置。若目标是实验流体力学，应同时评估像素、粒子定位、匹配、速度与导数量，而不是从图像指标推断量测精度。重复随机种子用于区分采样波动，场景外数据用于揭示先验失配；两者解决不同问题。

复用作者数字时要保留任务和数据分组，不能把 FFHQ 的 FID 与 ImageNet 的 LPIPS 拼成一个模型配置；不能把表格中未提供的结果填成零。附录 PSNR 指标也应与其成像算子、噪声和对齐方式一起读取，不可转写成未实际测过的亚像素误差。

> author-reported FFHQ ×4 SR: DPS (FID 39.35, LPIPS 0.214), DDRM (62.15, 0.294)

> author-reported Gaussian deblur: DPS (44.05, 0.257), DDRM (74.92, 0.332)

> perceptual image quality ≠ subpixel localization accuracy ≠ velocity-field accuracy

---

The paper reports experiments on 1,000 validation images per dataset at 256×256 resolution, typically with Gaussian noise standard deviation 0.05 after normalization. Its perceptual gains are measured under those operators and distributions, using a shared diffusion score across diffusion-based baselines. They are not particle-localization or fluid-velocity benchmarks. This companion cites the reported results without running the networks. A fluid-imaging evaluation needs a chain of task-specific measurements: image reconstruction, particle centers, correspondence, velocity and spatial derivatives. Random seeds quantify sampler variation; out-of-domain scenes test a different failure mode.

12 | 事件视觉迁移：写出曝光与触发的同一物理链

来源定位：本册原创工程映射；联系库内06事件综述、19 EDI及27 E2VID

对曝光帧 B，应把潜在时变强度沿曝光时间积分，而不是把某一时刻清晰图像直接与 B 相减。事件相机观测则来自对数强度变化越过阈值的时空点列，包含极性、时间和位置；阈值、漏事件、噪声及时间同步误差都进入观测模型。DPS 风格的先验可放在潜在图像序列或运动状态上，但必须通过同一前向链生成帧与事件代理，再合并各自数据项。

一个可执行的低维验证路线是：用已知粒子轨迹生成时序图像、积分曝光、生成阈值事件，先不训练任何网络，验证前向算子和自动微分；然后对比仅帧、仅事件、融合三种反演。保持同一运动、曝光和噪声种子，分别测 PSNR、质心偏差及轨迹误差；这样可以区分更多观测的信息收益与更强先验的收益。

若实际追踪段从未穿过某片反流区域，图像先验和不可压缩约束可能补出反流，但那是物理场推断。以紧支撑流函数扰动 δψ 构造 δu=(∂_yδψ,−∂_xδψ)，其支撑避开已观测轨迹，则速度在观测位置不变且散度仍为零。这样存在一族同观测场，说明无散约束本身不提供缺失测量。

若事件阈值 C 随像素变化，把它与亮度或运动一同自由优化可能产生新的尺度歧义。可先独立标定 C 与曝光时间，再只反演状态；随后用有约束的小范围阈值扰动做敏感性分析。把标定参数与状态参数分离，是减少“模型能解释一切却不能识别任何量”的直接办法。

$$B(u)=T^{-1}\int_{[t_0,t_0+T]} I(u,t)\,dt;\ \ell(u,t)=\log I(u,t)$$

$$\text{event trigger: } \ell(u,t_k)-\ell(u,t_{(k-1)})\approx p_k\, C(u)$$

$$\delta u=(\partial_y\delta\psi,-\partial_x\delta\psi);\ \operatorname{div}\delta u=0;\ \operatorname{supp}(\delta\psi)\cap\text{observations}=\emptyset$$

A blurred frame is an exposure integral; an event stream records thresholded log-intensity changes. A fused inverse problem must generate both through a common latent scene and motion model, with separate noise and calibration terms. Validate this forward chain on CPU using synthetic trajectories before introducing a learned prior. If no observed trajectories enter a region, an incompressibility prior may infer a plausible flow there but does not create direct particle evidence. A compactly supported streamfunction perturbation away from observations preserves both measured velocities and zero divergence, giving an explicit nonidentifiability construction.

13 | 无散与最优传输：约束应接在正确变量上

来源定位：本册原创跨论文推导；联系07流匹配、23随机插值与28 Sinkhorn

若未知变量是二维速度场 u，离散无散条件写为 Du=0。设 G 是与散度 D 相容的离散梯度，边界条件固定后，可用压力投影 $P=I-G(DG)^{-1}D$。它投影的是速度，不是任意图像像素；把去噪图像梯度当成流体压力更新没有物理意义。对无滑移、开放边界或不规则掩码，边界离散必须与投影一起定义。

一种研究候选是先在可行子空间参数化 u，再对粒子渲染残差构造 DPS 风格指导；另一种是交替执行概率更新与物理投影。后者一般改变目标分布，不能直接宣称保持原后验。若物理规律只是近似成立，例如测量为薄片平面、存在明显出平面流，则强制二维散度为零可能把真实三维效应误当噪声。

最优传输可连接两时刻粒子密度：成本矩阵描述移动代价，耦合 Π 描述质量分配。熵正则 Sinkhorn 提供高效软匹配，但概率归一化不等于真实粒子数守恒；漏检、遮挡和出平面运动会破坏等质量假设。应比较平衡、非平衡或带出生死亡模型的匹配，并将相机观测与物理先验分层，而非把一个 OT 损失当作所有守恒条件。

数值投影的零模态也要明确：纯 Neumann 压力边界通常使压力只确定到加法常数，可固定均值或一个参考自由度。这样处理的是压力规范，不是增加新的速度观测。若离散散度与梯度不相容，投影后残余散度可能由离散算子而非物理违约造成。

> $Du=0;\ P=I-G(DG)^{-1}D$ (with compatible discretization and boundary conditions)

> $OT\varepsilon(a,b)=\min_{\Pi 1=a,\ \Pi^{T}1=b}\langle C,\Pi\rangle+\varepsilon\ \Sigma_{ij}\ \Pi_{ij}(\log\Pi_{ij}-1)$

> a density-transport prior is not a certificate of particle identity

---

Physical constraints must act on the variables they describe. A compatible pressure projection enforces discrete velocity divergence, not image brightness. Alternating a probabilistic update with projection generally changes its target distribution unless justified. Planar observations of a three-dimensional flow may also violate a strict two-dimensional incompressibility assumption. Entropic optimal transport offers a soft coupling between particle densities, but equal normalized mass is not evidence of conserved detected particle count or persistent identity. Missing particles and out-of-plane motion require an appropriate transport model rather than an automatic balanced constraint.

14 | 从论文到可检验研究：最小决胜实验

来源定位：本册原创实施与验收方案；本册附带 CPU 解析校验

建议优先研究“物理先验能否在不牺牲粒子定位的前提下提升缺观测区速度推断”，而不是泛泛比较重建是否更好看。基线应至少包括解析观测反演、无先验数据项、相同先验但无物理约束、物理约束加传统平滑，以及所提概率方法。保持训练数据、前向算子、迭代预算和测量信息一致，才能把收益归因到新增机制。

把变量分三组：已观测区域、仅受物理约束影响区域、完全不可识别方向。分别报告定位误差（px）、速度误差（m/s）、散度（s^{-1}）、反流召回与 IoU、后验区间覆盖率。δψ 反例用于检验方法能否诚实表达歧义；前向失配、阈值偏差、曝光错时和域外纹理则检验鲁棒性。高 PSNR 与小残差只能通过各自指标成立，不能代替这些目标量。

本册附 math_checks.py 与 math_checks.json 在 CPU 上验证标量高斯后验、Tweedie 条件均值、真实与点代入指导以及非线性链式梯度；这些是数学和实现一致性检查，不是原论文网络复现。接到具身系统时，DPS 属于估计层：输出带时间戳的状态与不确定性；OpenVLA 或其他策略属于决策层；碰撞、安全和执行平滑应由明确控制层处理。把三层分开，才有机会定位失败究竟来自看错、选错还是执行错。

建议把第一轮资源投入到一个公开可复算的二维合成基准：同一流场生成曝光帧、事件和稀疏粒子轨迹，公开算子、噪声种子及不可观测扰动。先判断新增先验是否提高目标量且保留校准，再决定是否值得扩大到三维数据或训练大型模型；负结果也能准确指出缺的是测量还是模型。

> $\text{acceptance chain: sensor likelihood} \rightarrow \text{state inference} \rightarrow \text{action policy} \rightarrow \text{safety controller}$

> $\text{report px, m/s, } s^{-1} \text{ and interval coverage separately}$

> $\text{counterexample} + \text{calibrated synthetic test} + \text{held-out real measurements}$

A decisive research question is whether physical priors improve flow inference in unobserved regions without degrading measured particle localization. Compare against analytical inversion, data-only fitting, the same learned prior without physics, and conventional physical regularization under matched information and budgets. Evaluate observed, physically constrained and genuinely nonidentifiable regions separately. The included CPU checks establish mathematical consistency, not a reproduction of trained diffusion results. In an embodied system, keep perception, policy and safety control as separate layers with explicit timestamps, uncertainty and failure handling; otherwise an attractive reconstruction can obscure the actual source of a control failure.