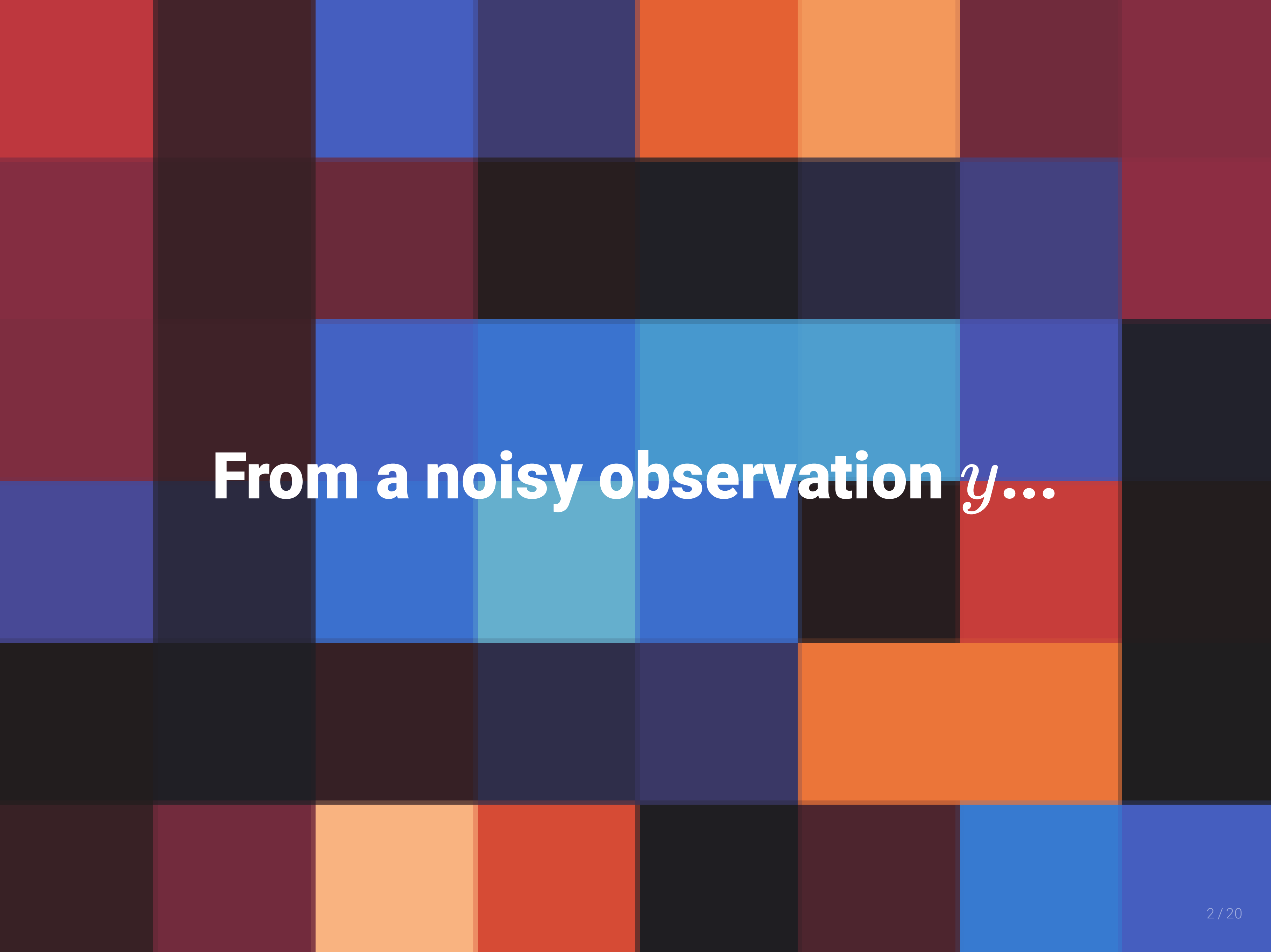
Inverting scientific images with score-based generative models

3rd Workshop on Uncertainty Quantification for Computer Vision, ECCV 2024
September 30, 2024

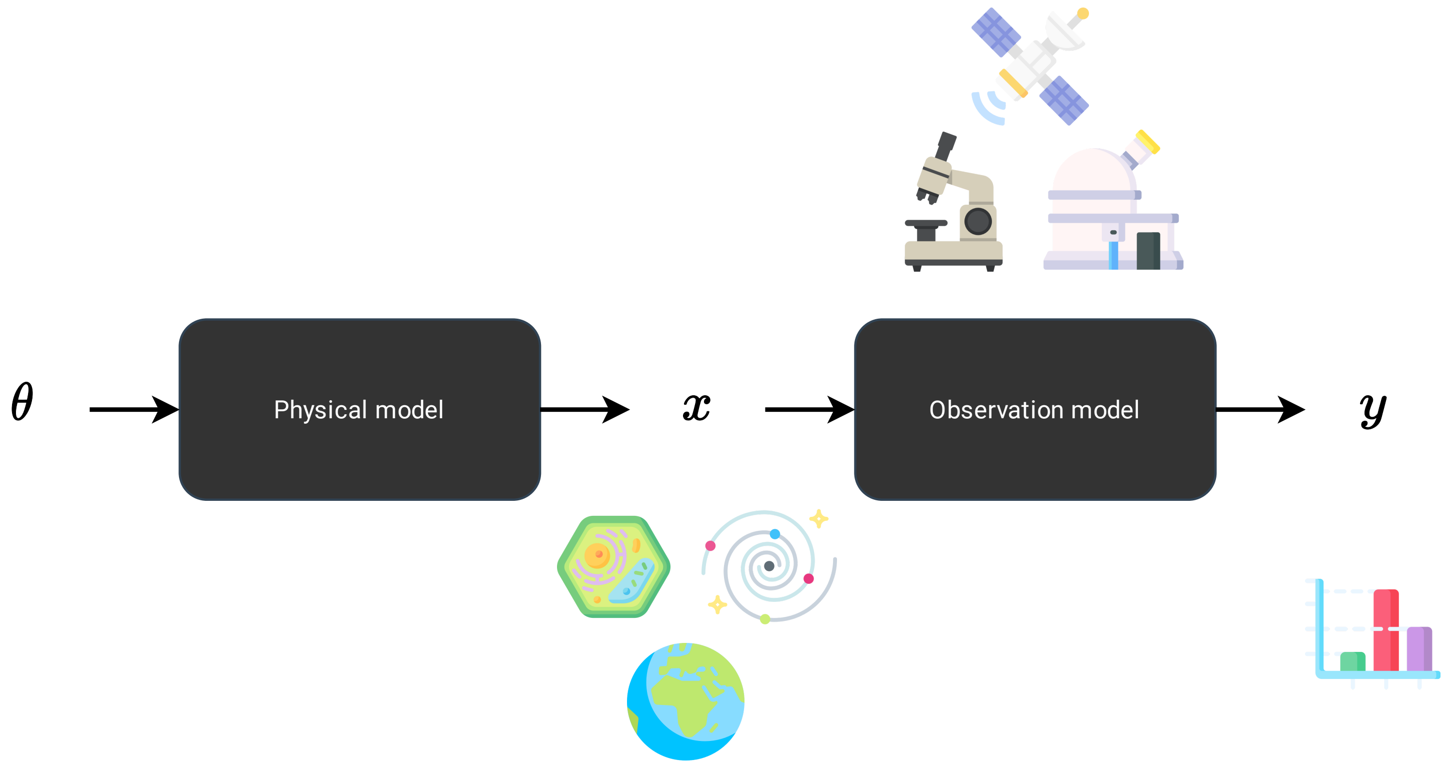
Gilles Louppe
g.louppe@uliege.be



From a noisy observation y...

… can we recover

all plausible images x?



Problem statement

Given a noisy observation y, estimate the posterior distribution $p(x|y) = \frac{p(y|x)p(x)}{p(y)}$ of plausible latent states x.

How do we estimate $p(x|y)$ when x is high-dimensional?

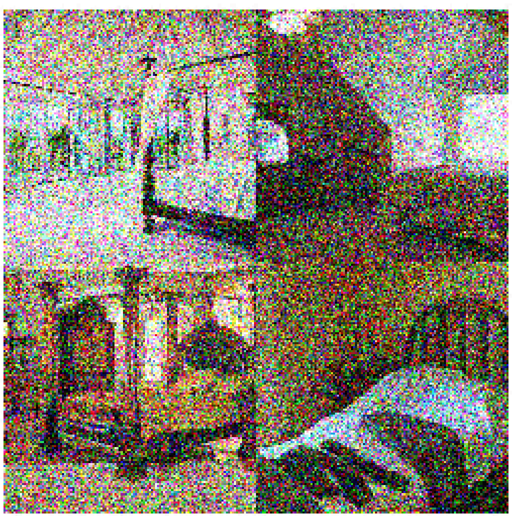
Score-based generative models 101

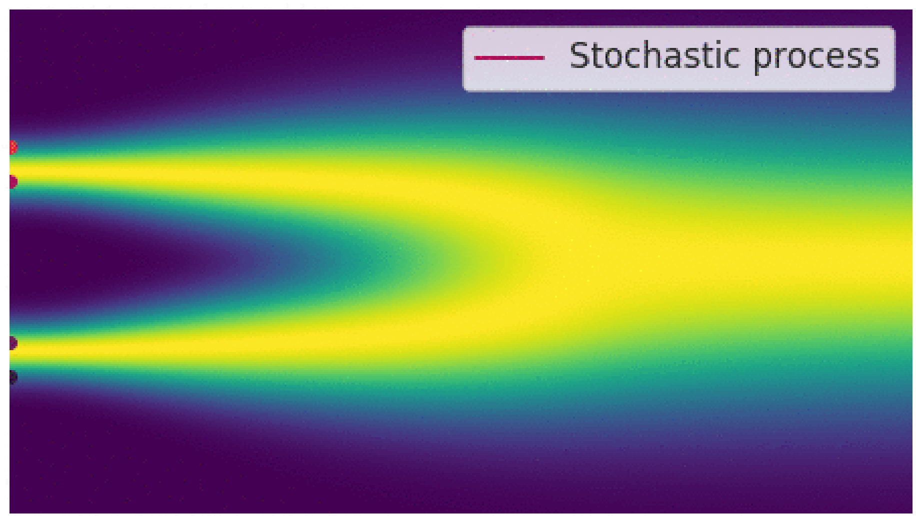
Samples $x \sim p(x)$ are progressively perturbed through a diffusion process described by the forward SDE

$$\mathrm{d}x_t = f_t x_t \mathrm{d}t + g_t \mathrm{d}w_t,$$

where x_t is the perturbed sample at time t, leading to a Gaussian diffusion kernel

$$p(x_t|x) = \mathcal{N}(x_t|\alpha_t x, \Sigma_t).$$



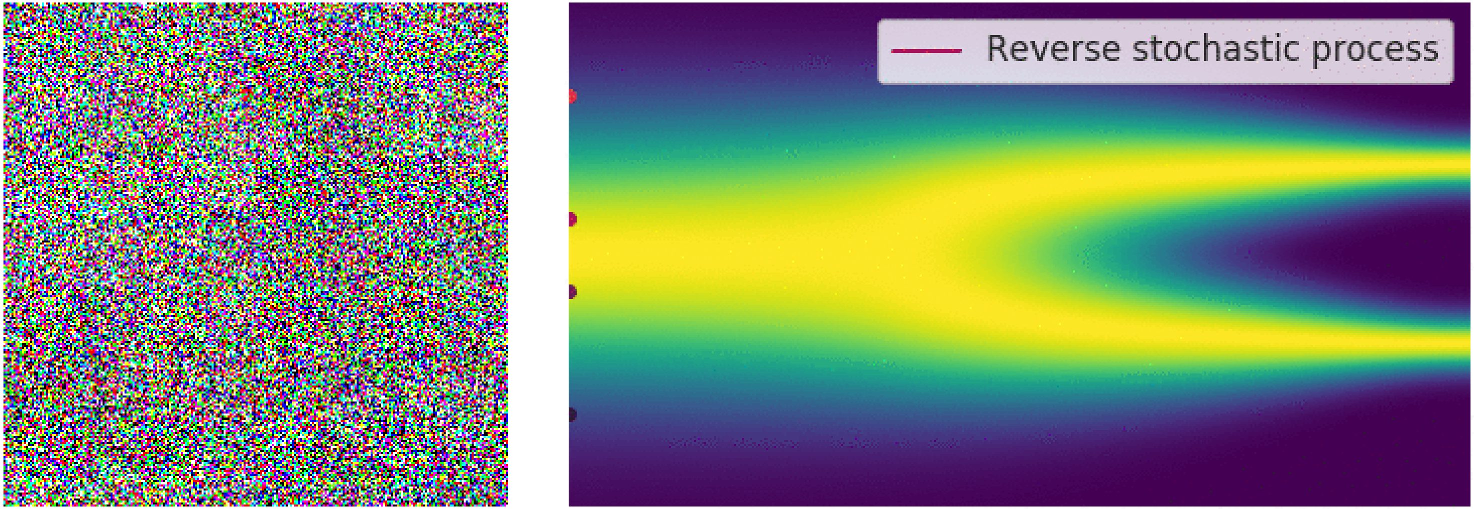


Forward diffusion process.

The reverse process satisfies a reverse-time SDE that can be derived analytically from the forward SDE as

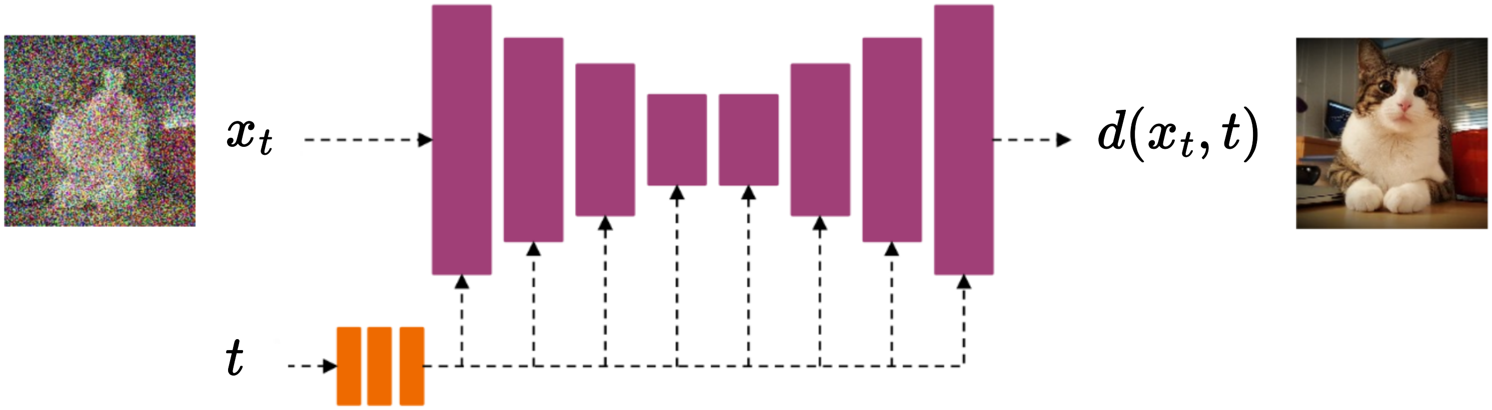
$$\mathrm{d}x_t = \left[f_t x_t - \frac{1+\eta^2}{2} g_t^2 \nabla_{x_t} \log p(x_t) \right] \mathrm{d}t + \eta g_t \mathrm{d}w_t.$$

Therefore, to generate data samples $x_0 \sim p(x_0) \approx p(x)$, we can draw noise samples $x_1 \sim p(x_1) \approx \mathcal{N}(0, \Sigma_1)$ and gradually remove the noise therein by simulating the reverse SDE from $t = 1$ to 0.



Reverse denoising process.

Credits: Song, 2021.



The score function $\nabla_{x_t} \log p(x_t)$ is unknown, but can be approximated by a neural network $d_\theta(x_t, t)$ by minimizing the denoising score matching objective

$$\mathbb{E}_{p(x)p(t)p(x_t|x)} \left[\lambda_t ||d_\theta(x_t, t) - x||_2^2\right].$$

The optimal denoiser d_θ is the mean $\mathbb{E}[x|x_t]$ which, via Tweedie's formula, allows to use $s_\theta(x_t, t) = \Sigma_t^{-1}(d_\theta(x_t, t) - x_t)$ as a score estimate in the reverse SDE.

Inverting single observations

Because of the Bayes' rule, the posterior score $\nabla_{x_t} \log p(x_t | y)$ to inject in the reverse SDE can be decomposed as

$$\nabla_{x_t} \log p(x_t | y) = \nabla_{x_t} \log p(x_t) + \nabla_{x_t} \log p(y | x_t) - \cancel{\nabla_{x_t} \log p(y)}.$$

This is particularly convenient as it enables **zero-shot posterior sampling** from a diffusion prior $p(x_0)$ without having to pre-wire the neural denoiser to the observation model $p(y | x)$.

Approximating $\nabla_{x_t} \log p(y|x_t)$

We want to estimate the score $\nabla_{x_t} \log p(y|x_t)$ of the noise-perturbed likelihood

$$p(y|x_t) = \int p(y|x)p(x|x_t)\mathrm{d}x.$$

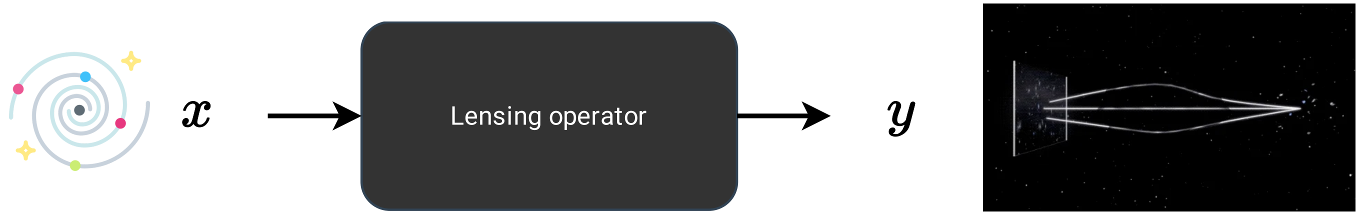
Our approach:

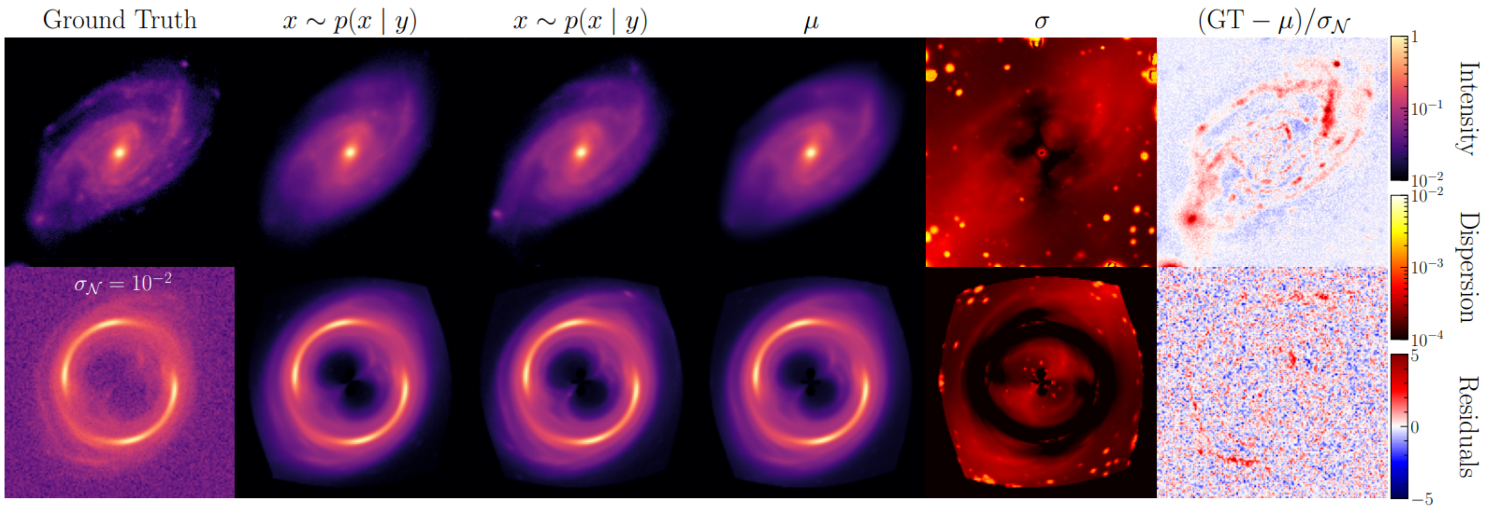
- Assume a linear Gaussian observation model $p(y|x) = \mathcal{N}(y|Ax, \Sigma_y)$.
- Assume the approximation $p(x|x_t) \approx \mathcal{N}(x|\mathbb{E}[x|x_t], \mathbb{V}[x|x_t])$, where $\mathbb{E}[x|x_t]$ is estimated by the denoiser and $\mathbb{V}[x|x_t]$ is estimated using Tweedie's covariance formula.
- Then $p(y|x_t) \approx \mathcal{N}(y|A\mathbb{E}[x|x_t], \Sigma_y + A\mathbb{V}[x|x_t]A^T)$.
- The score $\nabla_{x_t} \log p(y|x_t)$ then approximates to

$$\nabla_{x_t}\mathbb{E}[x|x_t]^T A^T (\Sigma_y + A\mathbb{V}[x|x_t]A^T)^{-1}(y - A\mathbb{E}[x|x_t]).$$

—

See also Daras et al (2024) for a survey on diffusion models for inverse problems.

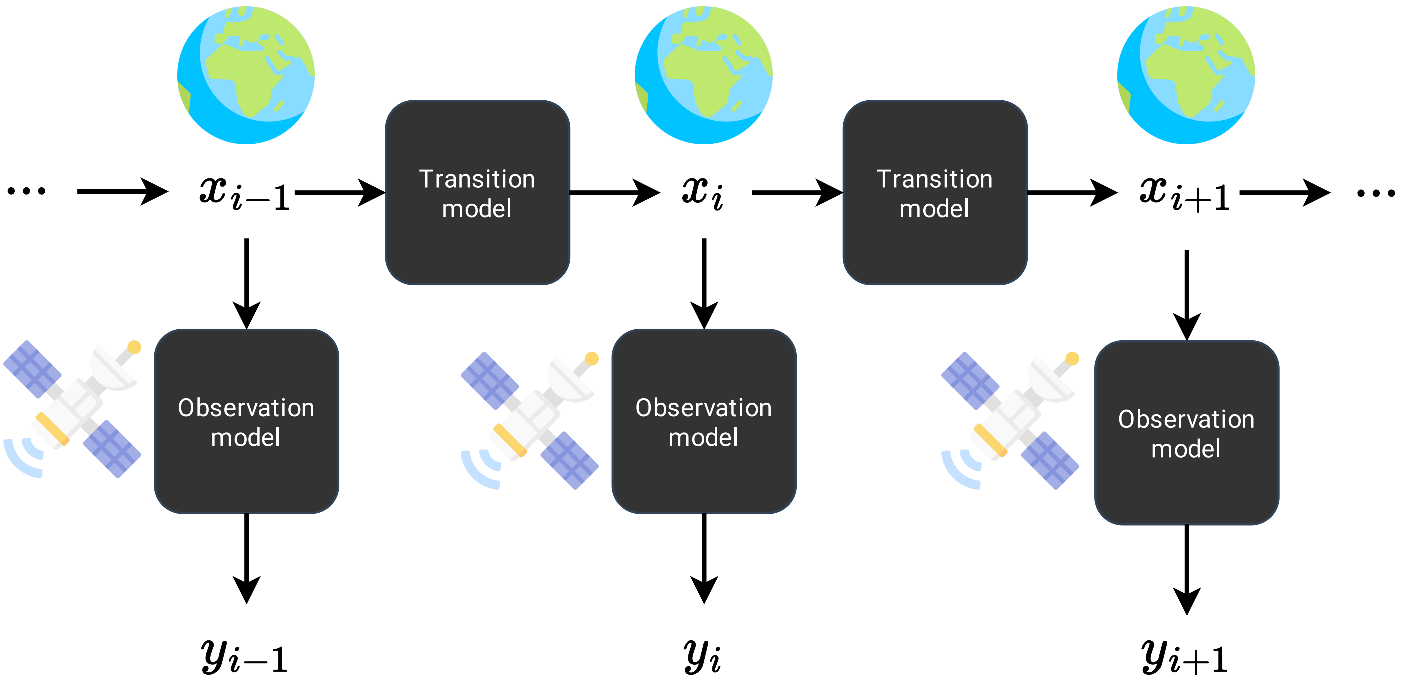




Plausible galaxy images x can be recovered from lensed observations y by zero-shot posterior sampling from a diffusion prior $p(x)$.

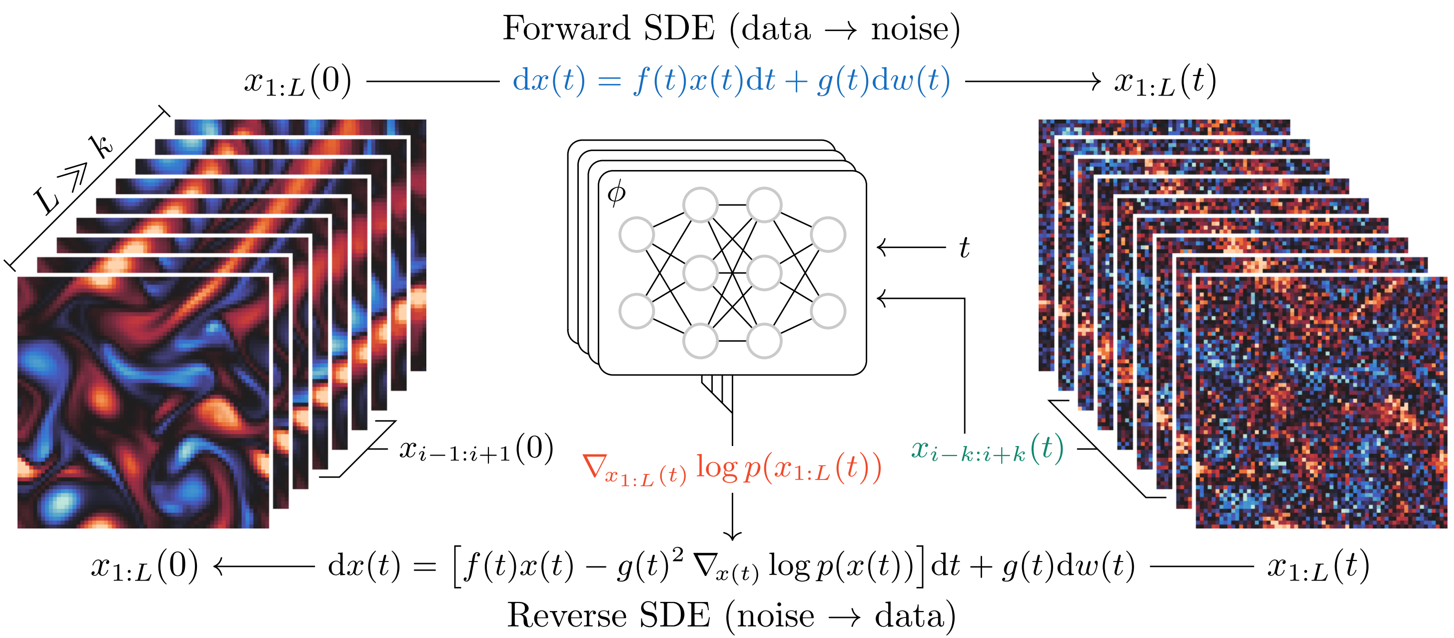
—
Credits: Adam et al., 2022 (arXiv:2211.03812).

Score-based data assimilation in dynamical systems



The goal of **data assimilation** is to estimate plausible trajectories $x_{1:L}$ given one or more noisy observations y (or $y_{1:L}$) as the posterior

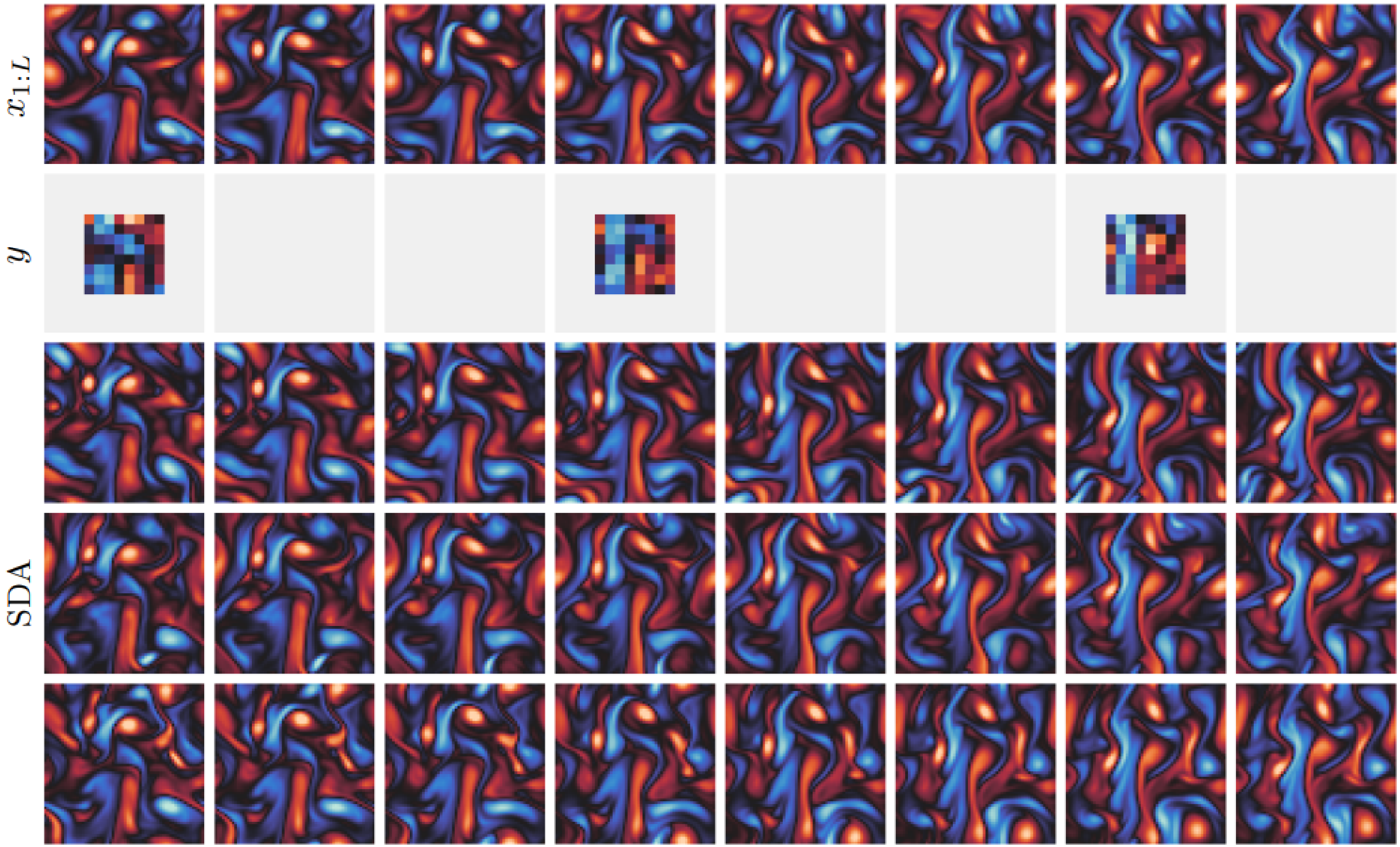
$$p(x_{1:L}|y) = \frac{p(y|x_{1:L})}{p(y)} p(x_0) \prod_{i=1}^{L-1} p(x_{i+1}|x_i).$$



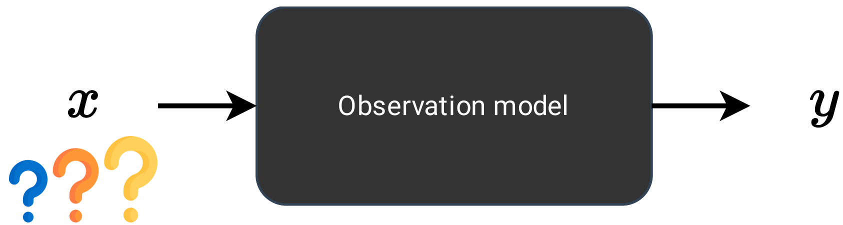
Our approach:

- Build a score-based generative model $p(x_{1:L})$ of arbitrary-length trajectories*.
- Use zero-shot posterior sampling to generate plausible trajectories from noisy observations y.

*:The score of a (noise perturbed) trajectory can be approximated by a sum of scores. See paper for details.



Sampling trajectories $x_{1:L}$ from
noisy, incomplete and coarse-grained observations y.



Learning priors from noisy observations

Assume only observations $y \sim p(y)$ and a known observation model $p(y|x)$.

The objective of **Empirical Bayes** is find a prior model $q_\theta(x)$ such that

$$q_\theta(y) = \int p(y|x) q_\theta(x) \mathrm{d}x$$

is closest to $p(y)$.

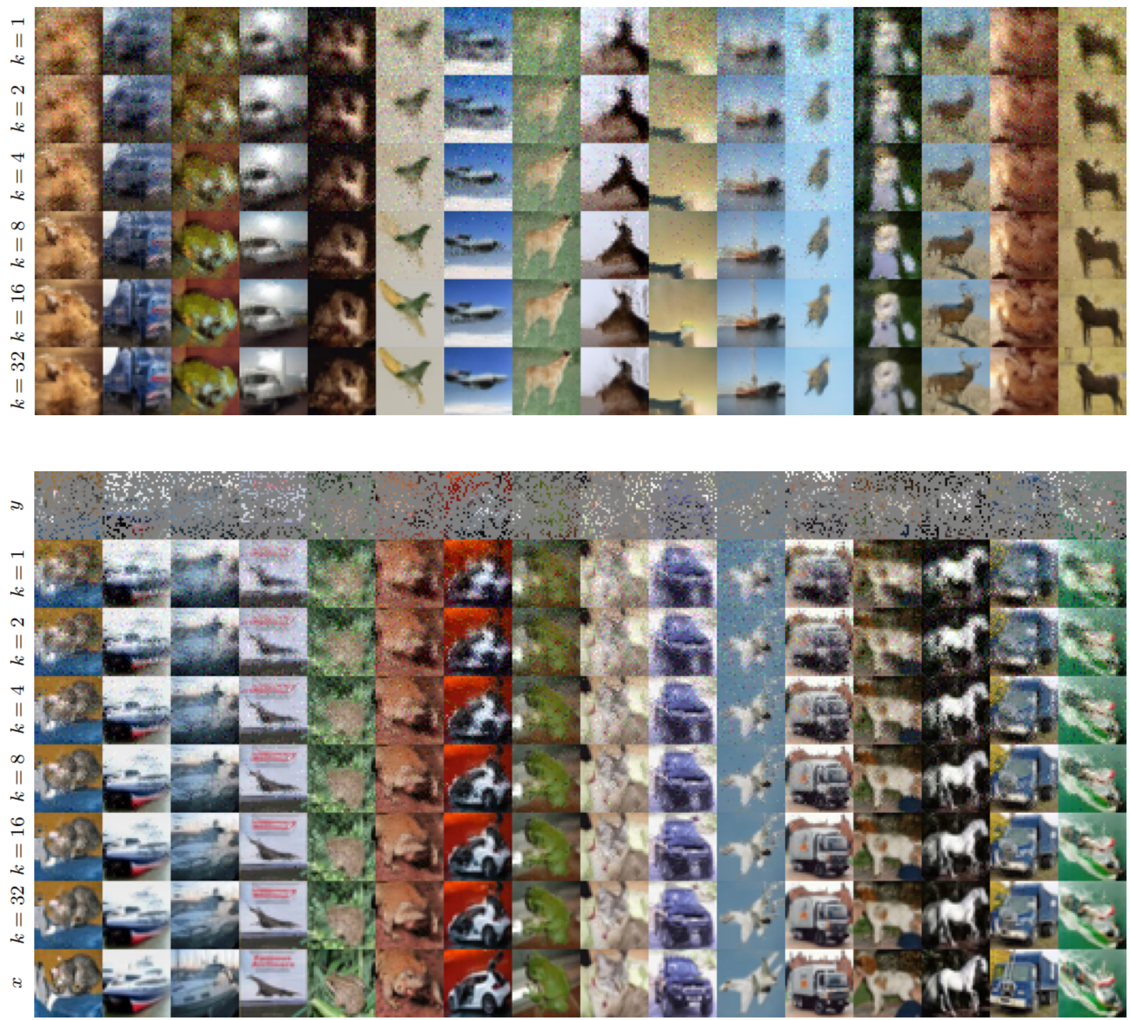
Our approach:

- If we parameterize the latent state x with a diffusion prior $q_\theta(x)$, then Expectation-Maximization can be used to maximize $q_\theta(y)$.

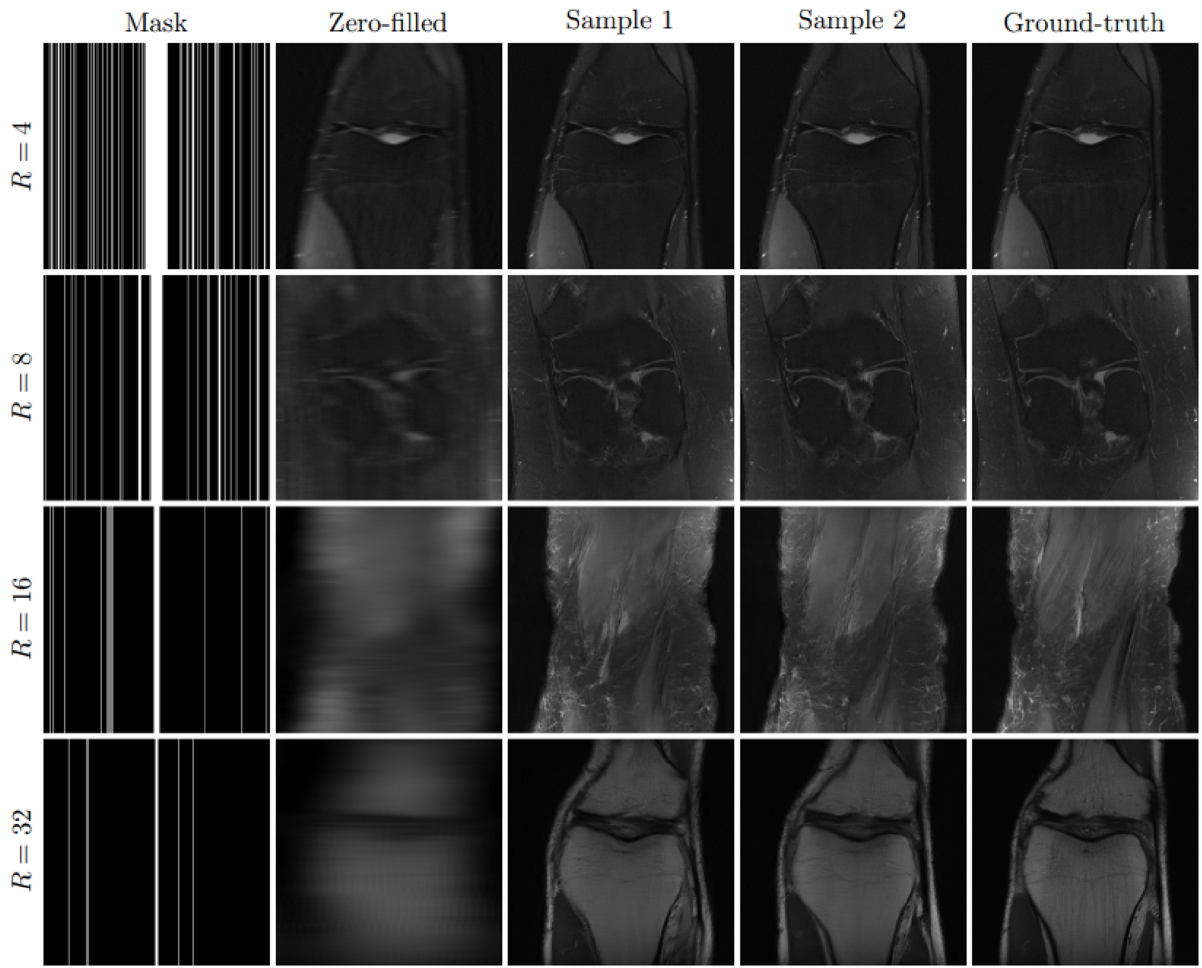
- It can be shown that the EM update

$$\theta_{k+1} = \arg\max_\theta \mathbb{E}_{p(y)} \mathbb{E}_{q_{\theta_k}(x|y)} \left[\log q_\theta(x) \right],$$

  where $q_{\theta_k}(x|y)$ is obtained by posterior sampling from $q_{\theta_k}(x)$, leads to a sequence of parameters θ_k such that $\mathbb{E}_{p(y)} \left[\log q_{\theta_k}(y) \right]$ is monotonically increasing and converges to a local optimum.



Samples from the prior $q_{\theta_k}(x)$ (top) and the posterior $q_{\theta_k}(x|y)$ (bottom) along the EM iterations when training from corrupted CIFAR-10 images.



Posterior samples for accelerated MRI using a diffusion prior trained only from observations with subsampled frequencies.

Conclusions

Score-based generative models...

- can be used for high-dimensional inverse problems;
- enable zero-shot posterior sampling, without pre-wiring the network to observations;
- do not require paired data.

Next challenges:

- Rigorous diagnostics for the quality of the approximation;
- Scalability to even larger dimensions (Earth-scale weather models, videos);




References:

- **Score-based data assimilation**
  François Rozet, Gilles Louppe.
  NeurIPS 2023, arXiv:2306.10574.
- **Score-based Data Assimilation for a Two-Layer Quasi-Geostrophic Model**
  François Rozet, Gilles Louppe.
  ML4PS workshop NeurIPS 2023, arXiv:2310.01853.
- **Learning Diffusion Priors from Observations by Expectation Maximization**
  François Rozet, Gérôme Andry, François Lanusse, Gilles Louppe.
  NeurIPS 2024, arXiv:2405.13712.

The end.