Correcting misspecified score-based priors for inverse problems: An application to strong gravitational lensing

Gabriel Missael Barco[1,2,3] **Alexandre Adam**[1,2,3] **Connor Stone**[1,2,3]
Yashar Hezaveh[1,2,3,4,5,6] **Laurence Perreault-Levasseur**[1,2,3,4,5,6]
[1]Université de Montréal [2]Mila [3]Ciela Institute [4]CCA, Flatiron Institute
[5]Perimeter Institute for Theoretical Physics [6]Trottier Space Institute
{gabriel.missael.barco,alexandre.adam,connor.stone,yashar.hezaveh,
laurence.perreault.levasseur}@umontreal.ca

Abstract

Score-based generative models have gained popularity as expressive, data-driven priors for complex, high-dimensional inverse problems. However, in many scientific applications, it is often difficult or even impossible to acquire samples from the true distribution to train these models, in which case a surrogate, e.g. a simulator, is often used to produce training samples, meaning that the learned prior could be misspecified. This, in turn, can bias the inferred posteriors, which limits the potential applicability of these models in real-world scenarios. In this work, we propose addressing this issue by iteratively training new priors with posterior samples from different sets of observations. We showcase the potential of this method on the problem of background image reconstruction in strong gravitational lensing. We show that posterior sampling becomes less biased after several updates, and the learned distribution is closer to the true prior.

1 Introduction

In the era of precision science, Bayesian inference has become a cornerstone of modern statistical data analysis. It provides a mathematical framework for inferring the probability distribution of latent parameters of interest, the posterior, in the presence of noisy observations. In this paradigm, existing knowledge is encoded in a prior distribution, which can be updated with new, noisy information through the likelihood function, to produce updated belief over the parameters of a problem [e.g. 1]. In cases where there are no prior observational constraints for a given system, the only information regarding the system is the belief that it is a random sample from an underlying population. In those cases, a possible approach is to use expressive data-driven population-level priors [2], in which a sample of existing data representative of the population is used to learn their distribution (i.e., by using them to train a generative model). However, it is often the case that the true population-level distribution cannot be sampled directly to obtain training examples.

A possible alternative is to use simulators to obtain the needed training data. In astrophysics, high-quality simulators are often available and make this a promising avenue. For example, in cosmology, simulations of the Universe [e.g. 3, 4, 5] can provide samples of fields and objects of interest [e.g. 6]. Another alternative is to construct training datasets from existing data sources that exhibit some structural similarity to the physical phenomena of interest [e.g. 7, 8, 9]. However, all these options bear the risk of a distribution shift between the learned prior and the true population-level distribution.

In this work, we propose an iterative method to update an initially biased data-driven prior which approximates the process of hierarchical Bayesian inference of the population-level prior over multiple

Machine Learning and the Physical Sciences Workshop, NeurIPS 2024.

iterations of observations. We provide empirical results in high-dimensional settings, showing that this method can forget artifacts present in the initial prior but absent from the true data-generating distribution. We showcase our method on the problem of learning idealized brightness structures from noisy observations in the inverse problem setting of strong gravitational lensing source reconstruction. We use score-based models as priors, as they have emerged as state-of-the-art generative models for images in recent years [10, 11, 12, 13, 14, 15]. They have enjoyed a wide range of applicability in inverse problem settings due in part to the flexibility of the stochastic differential equation (SDE) formalism [16] that underpins the framework [17]. In astronomy, they have been successfully applied to, among others, interferometric imaging [e.g. 7, 9, 18, 19], deconvolution [e.g. 20, 8], mass modelling [21], cosmological fields or cosmological parameters recovery [e.g. 22, 23, 24, 25, 26, 27], and strong gravitational lensing source reconstruction [28, 29].

2 Methods

2.1 Strong gravitational lensing source reconstruction

Strong gravitational lensing occurs when a massive object curves space in such a way that the light emitted by a distant background source, for example, a galaxy, is deflected, causing the image we see from the source to be distorted and multiply imaged. Given a noisy lensed observation, reconstructing the source galaxy and the mass distribution in the lens is an important scientific problem, which can act as a probe to understand the nature of dark matter [e.g. 30, 31], the early formation of stars [e.g. 32], active galactic nuclei [e.g. 33], the expansion rate of the universe [e.g. 34], and other problems.

In the limit of a thin lens, and assuming that the mass distribution in the foreground lens is known, strong gravitational lensing of a background source into a distorted observation is a linear transformation. Under this regime, inferring the background source is a linear inverse problem, characterized by the equation $\mathbf{y} = A\mathbf{x} + \boldsymbol{\eta}$, where $\mathbf{x} \in \mathbb{R}^n$ are the parameters of interest, $\mathbf{y} \in \mathbb{R}^m$ is the observation and $\boldsymbol{\eta} \in \mathbb{R}^m$ is a vector of additive noise, which we consider to be Gaussian $\boldsymbol{\eta} \sim \mathcal{N}(0, \Sigma_{\boldsymbol{\eta}} = \sigma_{\boldsymbol{\eta}}\mathbb{I})$. The observation and parameters of interests are related by a constant matrix $A \in \mathbb{R}^{m\times n}$. In a Bayesian framework, the goal is to sample from the posterior distribution, $p(\mathbf{x} \mid \mathbf{y})$. In the next section, we explore this inference process with a data-driven expressive prior.

2.2 Score-based models as priors for inverse problems

Score-based models (SBM) are a class of generative models that aims to learn the score function of the data distribution convolved with noise, $\boldsymbol{\nabla}_{\mathbf{x}_t} \log p_t(\mathbf{x}_t) = \boldsymbol{\nabla}_{\mathbf{x_t}} \log \int p(\mathbf{x})p(\mathbf{x}_t \mid \mathbf{x})d\mathbf{x}$. The noising process is generally characterized by a Gaussian perturbation kernel, $p(\mathbf{x}_t \mid \mathbf{x})$ indexed by the time parameter, t, of an SDE [16]. A neural network, $\mathbf{s}_\theta(\mathbf{x}, t)$, typically a U-net [35], is trained to approximate $\boldsymbol{\nabla}_{\mathbf{x}} \log p_t(\mathbf{x})$ using samples $\mathbf{x} \in \mathbb{R}^n$ from data $\mathcal{D}{\sim}p(\mathbf{x})$ by minimising the denoising score-matching objective [36, 37]. Having access to an approximation of the score function allows one to create a generative model by solving the reverse-time SDE [38] associated with the noising process used during training

$$d\mathbf{x} = (f(\mathbf{x}, t) - g^2(t)\boldsymbol{\nabla}_{\mathbf{x}} \log p_t(\mathbf{x}))dt + g(t)d\bar{\mathbf{w}}\,, \tag{1}$$

where f is the drift, g is an homogeneous diffusion coefficient associated with the noising process and $\bar{\mathbf{w}}$ is a reverse-time Wiener process.

This generative process can also be used for posterior inference given new data by replacing the prior score function in the reverse-time SDE (1) with the posterior score function $\boldsymbol{\nabla}_{\mathbf{x}} \log p(\mathbf{x} \mid \mathbf{y})$, which is obtained using Bayes' theorem. However, the likelihood score is intractable, as it involves an expectation over backward trajectories of the reverse-time SDE [see e.g. 7]. For a Gaussian likelihood, we can construct an analytical estimate of its score using the convolved likelihood approximation [21, 28]

$$p_t(\mathbf{y} \mid \mathbf{x}) \approx \mathcal{N}(\mu(t)\mathbf{y} \mid A\mathbf{x}, \mu^2(t)\Sigma_{\boldsymbol{\eta}} + \sigma^2(t)AA^T)\,, \tag{2}$$

where $\Sigma_{\boldsymbol{\eta}}$ is the covariance associated with the additive noise distribution $\boldsymbol{\eta} \sim \mathcal{N}(0, \Sigma_{\boldsymbol{\eta}})$. With this machinery, any SBM trained on some dataset of parameters of interest can be used as an approximate posterior sampler without retraining or conditioning the neural network on the observations [39].

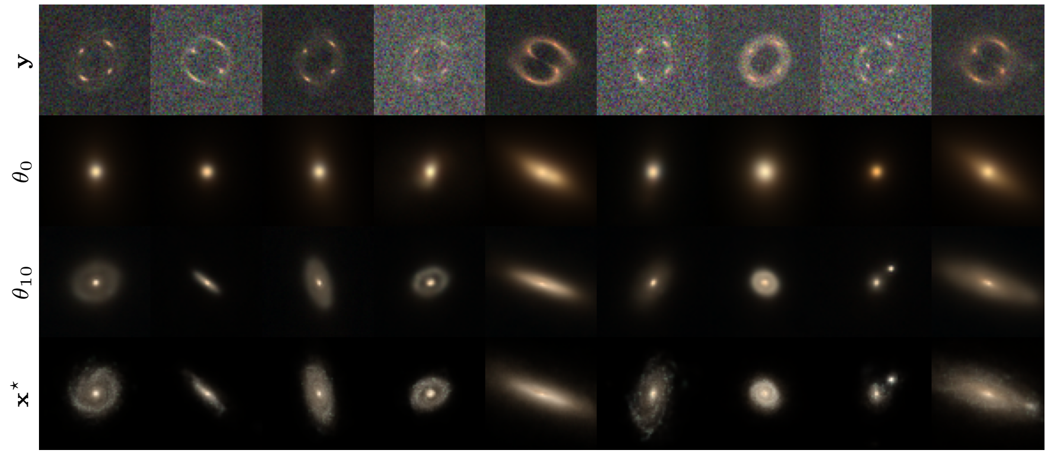

Figure 1: Improvement in strong gravitational lensing source reconstruction with galaxy sources under Algorithm 1 after 10 updates, highlighting the adaptation from a biased initial prior to better alignment with the target distribution. The top row shows noisy observations $\mathbf{y}$ with observational noise of $\sigma_{\boldsymbol{\eta}} = 3$. The second row displays posterior samples from the initial prior p_{θ_0}, characterized by significant bias. The third row presents samples after the final update $p_{\theta_{10}}$, demonstrating substantial improvements in matching the true sources $\mathbf{x}^\star$, which were sampled from $p_{\theta^\star}$.

2.3 Updating the prior with observations

We aim to study hierarchical inference in the context of moderately high-dimensional inference inverse problems using SBM as expressive priors. Assuming an initial SBM prior trained on a potentially corrupted dataset $\{\mathbf{x}_i^{(0)}\}$, our goal is to update the population-level parameters θ — the weights of the prior SBM network — given only a set of noisy and partial observations $\{\mathbf{y}_i\}_{i=1}^N$.

We introduce a method inspired by traditional generalized expectation maximization methods [e.g. 40, 41, 42, 43, 44], which consists of an iterative procedure that leverages the posterior sampling algorithm outlined in Section 2.2 to acquire increasingly plausible samples from the set of observations $\{\mathbf{y}_i^{(\alpha)}\}_{i=1}^N$. For each update, a set of posterior samples is aggregated from each observation to train a new prior distribution, encoded by the generative process of a SBM. The algorithm is summarized in Algorithm 1.

We can show that the updated prior will have larger log-evidence than the previous iteration (see Appendix F), and that, in the large data limit $N \to \infty$, there exists a stationary representation for the prior distribution, $p_{\hat{\theta}}(\mathbf{x})$, which has log-evidence equal to the true distribution $\theta^\star$ (see Appendix E). We show empirically in Section 3 that the procedure does converge to a prior close to the true underlying population distribution for the settings we explored.

Algorithm 1 Updating the prior with observations

Input: Initial prior $p_{\theta_0}(\mathbf{x})$, observations $\{\mathbf{y}_i\}_{i=1}^{M\times N}$, training procedure $\mathcal{A}$, number of posterior samples per observation K, number of updates M.
for $\alpha = 1$ **to** M **do**
  Select N observations $\{\mathbf{y}_i^{(\alpha)}\}_{i=1}^N = \{\mathbf{y}_i\}_{i=(\alpha-1)N}^{\alpha N}$
  for $i = 1$ **to** N **do**
    Get K posterior samples $\mathcal{D}_{\text{posterior}}^{i,\alpha} = \{\mathbf{x}_{i,j}^{(\alpha)}\}_{j=1}^K,\ \mathbf{x}_{i,j}^{(\alpha)} \sim p_{\theta_{\alpha-1}}(\mathbf{x} \mid \mathbf{y}_i^{(\alpha)})$
  end for
  Train new prior with posterior samples $p_{\theta_\alpha} = \mathcal{A}(\bigcup_{i=1}^N \mathcal{D}_{\text{posterior}}^{i,\alpha})$
end for

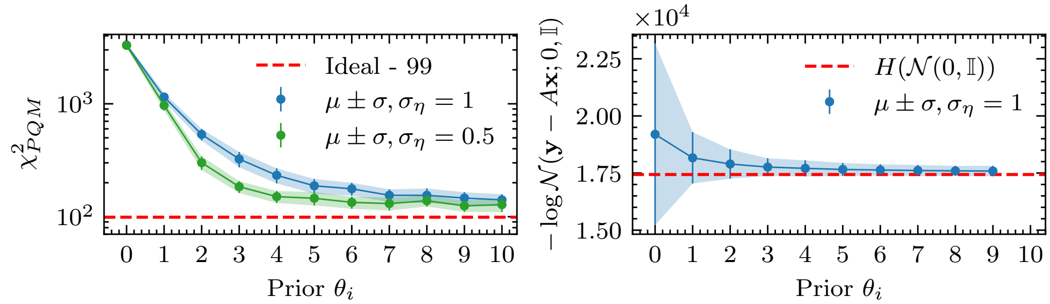

Figure 2: Evaluation of prior improvement for the galaxy experiment. **Left:** The χ^2_{PQM} statistic comparing the target distribution $p_{\theta^\star}$ with the proposal prior p_{θ_i} at each iteration. **Right:** Mean log-likelihood of the residuals $\mathbf{y} - A\mathbf{x}$ from $10\,240$ pairs $(\mathbf{y}, \mathbf{x} \sim p_{\theta_\alpha}(\mathbf{x} \mid \mathbf{y}))$.

3 Experiments and results

For our experiments, we intentionally create a misspecified initial prior p_{θ_0} compared to the true distribution $p_{\theta^\star}$. We gather two datasets of galaxy images ($3 \times 64 \times 64$ resolution) of different classes: simulated observations of blue spiral galaxies, taken from a subset of the SKIRT TNG dataset [6], and real observations of red elliptical galaxies, obtained from the DESI Legacy Imaging Surveys [45] DR10. The datasets are described in Appendix D. We define the true distribution $p_{\theta^\star}$, used to simulate observations, to be what an SBM learned when trained on the spiral dataset, and the initial distribution p_{θ_0} is an SBM trained on the elliptical dataset. This choice is motivated by the fact that initially, only noisy telescope observations of local, evolved red galaxies would be available to train an initial prior, but we want to demonstrate that the proposed method can discover new features.

We perform two experiments with 10 updates, $10\,240$ observations per update, 1 posterior sample per observation, and with relatively low noise levels, $\sigma_{\boldsymbol{\eta}} = 0.5, 1$. To simulate the forward strong gravitational lensing model, we use the `Caustics` python package [46]. We also assume that the forward model is the same for all observations and is known. In Figure 1, we show a selection of 10 observations and their posterior samples for the first and last iterations. This figure demonstrates how the structure of generated posterior samples evolves as a function of the update, increasing in complexity from elliptical shapes to showcasing rings and satellites in the last update. The samples obtained using the initial prior are strongly biased toward elliptical galaxies and differ from the ground truths $\mathbf{x}^\star$ in color, flux, and morphology. Thus, important morphological information about the data is learned and encoded in the updated SBM which only had access to noisy observations.

To further test the improvement of the prior, in Figure 2 (right), we compute the log-likelihood of the residuals as a function of the update index and compute the mean and variance of the log-likelihood. The ideal value corresponds to the entropy of the noise model $H(\mathcal{N}(0, \Sigma_{\boldsymbol{\eta}}))$. This metric informs us about the information left in the data to be extracted by the posterior sampling algorithm. We observe that the mean converges to the ideal scenario while the variance reduces accordingly.

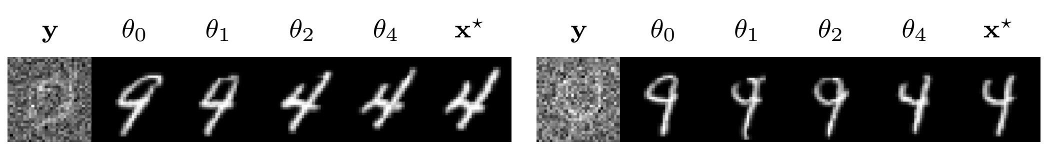

Figure 3: Sequences showcasing the model's ability to accurately reconstruct the digit 4 from noisy observations in posterior samples after updating the prior with Algorithm 1, despite 4 not being included in the initial prior. The evolution of posterior samples $\mathbf{x} \sim p_{\theta_\alpha}(\mathbf{x} \mid \mathbf{y})$ is shown for each update.

Finally, we also compute the PQMass metric between each updated prior $p_{\theta_\alpha}(\mathbf{x})$ and $p_{\theta^\star}(\mathbf{x})$. PQMass is a sample-based metric to assess the quality of generative models [47] which is based on partitions of the space by Voronoi cells to estimate the probability that both samples come from the same distribution using a χ^2_{PQM} using counts in those cells. We use $n_r = 100$ regions and estimate a mean and variance by using 5 independent sets of samples from both distributions, with $2\,048$ samples in each set. The results are reported in the left panel of Figure 2. We observed that the value of χ^2_{PQM} improves as a function of the update index, and depends on the noise level of the observations. The improvement in these metrics and posterior samples showcase the potential of this method.

We also performed experiments using MNIST digits [48] as pixelated background sources. Here, we explore mode mismatch, which we define as the situation where the initial prior and the true underlying population-level distribution do not share the same modes. As an example, we train the initial prior $p_{\theta_0}(\mathbf{x})$ on a subset of MNIST with the digits 1 and 4 removed, while the true population distribution, $p_{\theta^\star}(\mathbf{x})$ is constructed with the digits 1 and 6 missing. We perform $M = 4$ iterations using our algorithm.

As observed in Figure 3, when we use the initial prior $p_{\theta_0}(\mathbf{x})$ with an observation where the source is a number 4 (not seen during training), the digit 9 is obtained from posterior samples. We hypothesize this to be due to the similarity between the digits 4 and 9. In the first few columns, the posterior samples are biased towards reconstructing the digit 9 because 4 is missing from this prior. Crucially, our algorithm can recover the correct shape of the digit 4 after a few updates, even though this digit was never seen during the initial training of the prior SBM. We further analyze digit proportion in prior samples at each iteration in Appendix A.

4 Limitations

One of the main limitations of our current framework is that it requires a substantial amount of data and compute resources. The iterative retraining of SBM from scratch alone can be an important computational burden. Since the improvement per update slows down drastically after a few updates, an argument can be made to keep the number of iterations low generally. Another possible improvement is to use fine-tuning techniques, such as using LoRA weights [49], instead of training from scratch. Regarding data requirements, rather than performing a single update per observation set, it may be worth exploring multiple updates per set or even using a single dataset for all the iterations. This approach could approximate maximum-likelihood fitting for the given dataset, which introduces the risk of overfitting. Finally, it is important to mention that in this work, we assumed that the physical forward model A, determined by the lensing configuration, is known. However, this is typically not the case with real data. Jointly sampling the lensing parameters and the pixelated source in the presence of real (potentially non-Gaussian) noise would be necessary to apply this approach to real data.

5 Conclusion

In conclusion, we have introduced an algorithm that addresses the misspecification of a prior SBM by updating it using only partial and noisy observations, allowing us to learn a distribution over high-dimensional spaces accurately. We have demonstrated empirically that this method can learn new features in linear inverse problem settings. Such a method is of high value considering the volume of partially corrupted observations currently available and upcoming in large surveys of the sky like the *Euclid* space telescope [50] and the Vera Rubin Observatory [51]. Extracting information from these surveys and encoding it in SBM neural networks for future inference tasks is an important subject for the development of computational imaging techniques in astronomy.

Acknowledgments and Disclosure of Funding

This work is partially supported by Schmidt Futures, a philanthropic initiative founded by Eric and Wendy Schmidt as part of the Virtual Institute for Astrophysics (VIA). The work is in part supported by computational resources provided by Calcul Quebec and the Digital Research Alliance of Canada. Y.H. and L.P.-L. acknowledge support from the Canada Research Chairs Program, the National Sciences and Engineering Council of Canada through grants RGPIN-2020-05073 and 05102, and the Fonds de recherche du Québec through grants CFQCU-2024-348060, 2022-NC-301305 and 300397. C.S. acknowledges the support of a NSERC Postdoctoral Fellowship and a CITA National Fellowship.

References

[1] Rens van de Schoot, Sarah Depaoli, Ruth King, Bianca Kramer, Kaspar Märtens, Mahlet G. Tadesse, Marina Vannucci, Andrew Gelman, Duco Veen, Joukje Willemsen, and Christopher Yau. Bayesian statistics and modelling. *Nature Reviews Methods Primers*, 1(1):1, January 2021.

[2] Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Akti Vehtari, and Donald B. Rubin. *Bayesian data analysis*. Chapman & Hall/CRC Texts in Statistical Science Series. CRC, Boca Raton, Florida, third edition, 2013.

[3] Shy Genel, Mark Vogelsberger, Volker Springel, Debora Sijacki, Dylan Nelson, Greg Snyder, Vicente Rodriguez-Gomez, Paul Torrey, and Lars Hernquist. Introducing the Illustris project: the evolution of galaxy populations across cosmic time. *Monthly Notices of the Royal Astronomical Society*, 445(1):175–200, November 2014.

[4] Jonah C. Rose, Paul Torrey, Francisco Villaescusa-Navarro, Mariangela Lisanti, Tri Nguyen, Sandip Roy, Kassidy E. Kollmann, Mark Vogelsberger, Francis-Yan Cyr-Racine, Mikhail V. Medvedev, Shy Genel, Daniel Anglés-Alcázar, Nitya Kallivayalil, Bonny Y. Wang, Belén Costanza, Stephanie O'Neil, Cian Roche, Soumyodipta Karmakar, Alex M. Garcia, Ryan Low, Shurui Lin, Olivia Mostow, Akaxia Cruz, Andrea Caputo, Arya Farahi, Julian B. Muñoz, Lina Necib, Romain Teyssier, Julianne J. Dalcanton, and David Spergel. Introducing the DREAMS Project: DaRk mattEr and Astrophysics with Machine learning and Simulations. *arXiv e-prints*, page arXiv:2405.00766, May 2024.

[5] Francisco Villaescusa-Navarro, Shy Genel, Daniel Anglé s-Alcázar, Leander Thiele, Romeel Dave, Desika Narayanan, Andrina Nicola, Yin Li, Pablo Villanueva-Domingo, Benjamin Wandelt, David N. Spergel, Rachel S. Somerville, Jose Manuel Zorrilla Matilla, Faizan G. Mohammad, Sultan Hassan, Helen Shao, Digvijay Wadekar, Michael Eickenberg, Kaze W. K. Wong, Gabriella Contardo, Yongseok Jo, Emily Moser, Erwin T. Lau, Luis Fernando Machado Poletti Valle, Lucia A. Perez, Daisuke Nagai, Nicholas Battaglia, and Mark Vogelsberger. The CAMELS multifield data set: Learning the universe's fundamental parameters with artificial intelligence. *The Astrophysical Journal Supplement Series*, 259(2):61, 04 2022.

[6] Connor Bottrell, Hassen M. Yesuf, Gergö Popping, Kiyoaki Christopher Omori, Shenli Tang, Xuheng Ding, Annalisa Pillepich, Dylan Nelson, Lukas Eisert, Hua Gao, Andy D. Goulding, Boris S. Kalita, Wentao Luo, Jenny E. Greene, Jingjing Shi, and John D. Silverman. IllustrisTNG in the HSC-SSP: image data release and the major role of mini mergers as drivers of asymmetry and star formation. *MNRAS*, 527(3):6506–6539, January 2024.

[7] Berthy T. Feng, Jamie Smith, Michael Rubinstein, Huiwen Chang, Katherine L. Bouman, and William T. Freeman. Score-based diffusion models as principled priors for inverse imaging. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 10486–10497, 2023.

[8] Alexandre Adam, Connor Stone, Connor Bottrell, Ronan Legin, Yashar Hezaveh, and Laurence Perreault-Levasseur. Echoes in the Noise: Posterior Samples of Faint Galaxy Surface Brightness Profiles with Score-Based Likelihoods and Priors. In *Machine Learning and the Physical Sciences Workshop, NeurIPS 2023*, 11 2023.

[9] Noe Dia, M. J. Yantovski-Barth, Alexandre Adam, Micah Bowles, Pablo Lemos, Anna M. M. Scaife, Yashar Hezaveh, and Laurence Perreault-Levasseur. Bayesian Imaging for Radio Interferometry with Score-Based Priors. In *Machine Learning and the Physical Sciences Workshop, NeurIPS 2023*, 11 2023.

[10] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 6840–6851. Curran Associates, Inc., 2020.

[11] Yang Song and Stefano Ermon. Improved techniques for training score-based generative models. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 12438–12448. Curran Associates, Inc., 2020.

[12] Diederik Kingma, Tim Salimans, Ben Poole, and Jonathan Ho. Variational diffusion models. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan, editors, *Advances in Neural Information Processing Systems*, volume 34, pages 21696–21707. Curran Associates, Inc., 2021.

[13] Alexander Quinn Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 8162–8171. PMLR, 18–24 Jul 2021.

[14] Prafulla Dhariwal and Alexander Nichol. Diffusion models beat gans on image synthesis. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan, editors, *Advances in Neural Information Processing Systems*, volume 34, pages 8780–8794. Curran Associates, Inc., 2021.

[15] Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. In *Proc. NeurIPS*, 2022.

[16] Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net, 2021.

[17] Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In Francis Bach and David Blei, editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 2256–2265, Lille, France, 07–09 Jul 2015. PMLR.

[18] M. Drozdova, V. Kinakh, O. Bait, O. Taran, E. Lastufka, M. Dessauges-Zavadsky, T. Holotyak, D. Schaerer, and S. Voloshynovskiy. Radio-astronomical image reconstruction with a conditional denoising diffusion model. *A&A*, 683:A105, March 2024.

[19] Berthy T. Feng, Katherine L. Bouman, and William T. Freeman. Event-horizon-scale Imaging of M87* under Different Assumptions via Deep Generative Image Priors. *arXiv e-prints*, page arXiv:2406.02785, June 2024.

[20] Zhiwei Xue, Yuhang Li, Yash Patel, and Jeffrey Regier. Diffusion Models for Probabilistic Deconvolution of Galaxy Images. *arXiv e-prints*, page arXiv:2307.11122, July 2023.

[21] B. Remy, F. Lanusse, N. Jeffrey, J. Liu, J. L. Starck, K. Osato, and T. Schrabback. Probabilistic mass-mapping with neural score estimation. *A&A*, 672:A51, April 2023.

[22] Nayantara Mudur and Douglas P. Finkbeiner. Can denoising diffusion probabilistic models generate realistic astrophysical fields? In *36th Conference on Neural Information Processing Systems: Workshop on Machine Learning and the Physical Sciences*, 11 2022.

[23] David Heurtel-Depeiges, Blakesley Burkhart, Ruben Ohana, and Bruno Régaldo-Saint Blancard. Removing Dust from CMB Observations with Diffusion Models. *arXiv e-prints*, page arXiv:2310.16285, October 2023.

[24] Thomas Flöss, William R. Coulton, Adriaan J. Duivenvoorden, Francisco Villaescusa-Navarro, and Benjamin D. Wandelt. Denoising Diffusion Delensing Delight: Reconstructing the Non-Gaussian CMB Lensing Potential with Diffusion Models. *arXiv e-prints*, page arXiv:2405.05598, May 2024.

[25] Victoria Ono, Core Francisco Park, Nayantara Mudur, Yueying Ni, Carolina Cuesta-Lazaro, and Francisco Villaescusa-Navarro. Debiasing with Diffusion: Probabilistic reconstruction of Dark Matter fields from galaxies with CAMELS. *arXiv e-prints*, page arXiv:2403.10648, March 2024.

[26] Nayantara Mudur, Carolina Cuesta-Lazaro, and Douglas P. Finkbeiner. Diffusion-HMC: Parameter Inference with Diffusion Model driven Hamiltonian Monte Carlo. *arXiv e-prints*, page arXiv:2405.05255, May 2024.

[27] Ronan Legin, Matthew Ho, Pablo Lemos, Laurence Perreault-Levasseur, Shirley Ho, Yashar Hezaveh, and Benjamin Wandelt. Posterior sampling of the initial conditions of the universe from non-linear large scale structures using score-based generative models. *Monthly Notices of the Royal Astronomical Society: Letters*, 527(1):L173–L178, October 2023. _eprint: https://academic.oup.com/mnrasl/article-pdf/527/1/L173/54609829/slad152.pdf.

[28] Alexandre Adam, Adam Coogan, Nikolay Malkin, Ronan Legin, Laurence Perreault-Levasseur, Yashar Hezaveh, and Yoshua Bengio. Posterior samples of source galaxies in strong gravitational lenses with score-based priors. In *Machine Learning and the Physical Sciences Workshop*, page E1, January 2022.

[29] Konstantin Karchev, Noemi Anau Montel, Adam Coogan, and Christoph Weniger. Strong-Lensing Source Reconstruction with Denoising Diffusion Restoration Models. In *36th Conference on Neural Information Processing Systems: Workshop on Machine Learning and the Physical Sciences*, 11 2022.

[30] Simona Vegetti and Mark Vogelsberger. On the density profile of dark matter substructure in gravitational lens galaxies. *MNRAS*, 442(4):3598–3603, August 2014.

[31] Yashar Hezaveh, Neal Dalal, Gilbert Holder, Theodore Kisner, Michael Kuhlen, and Laurence Perreault Levasseur. Measuring the power spectrum of dark matter substructure using strong gravitational lensing. *JCAP*, 2016(11):048, November 2016.

[32] Brian Welch, Dan Coe, Erik Zackrisson, S. E. de Mink, Swara Ravindranath, Jay Anderson, Gabriel Brammer, Larry Bradley, Jinmi Yoon, Patrick Kelly, Jose M. Diego, Rogier Windhorst, Adi Zitrin, Paola Dimauro, Yolanda Jiménez-Teja, Abdurro'uf, Mario Nonino, Ana Acebron, Felipe Andrade-Santos, Roberto J. Avila, Matthew B. Bayliss, Alex Benítez, Tom Broadhurst, Rachana Bhatawdekar, Maruša Bradač, Gabriel B. Caminha, Wenlei Chen, Jan Eldridge, Ebraheem Farag, Michael Florian, Brenda Frye, Seiji Fujimoto, Sebastian Gomez, Alaina Henry, Tiger Y. Y. Hsiao, Taylor A. Hutchison, Bethan L. James, Meridith Joyce, Intae Jung, Gourav Khullar, Rebecca L. Larson, Guillaume Mahler, Nir Mandelker, Stephan McCandliss, Takahiro Morishita, Rosa Newshore, Colin Norman, Kyle O'Connor, Pascal A. Oesch, Masamune Oguri, Masami Ouchi, Marc Postman, Jane R. Rigby, Jr. Ryan, Russell E., Soniya Sharma, Keren Sharon, Victoria Strait, Louis-Gregory Strolger, F. X. Timmes, Sune Toft, Michele Trenti, Eros Vanzella, and Anton Vikaeus. JWST Imaging of Earendel, the Extremely Magnified Star at Redshift z = 6.2. *ApJL*, 940(1):L1, November 2022.

[33] Chien Y. Peng, Chris D. Impey, Hans-Walter Rix, Christopher S. Kochanek, Charles R. Keeton, Emilio E. Falco, Joseph Lehár, and Brian A. McLeod. Probing the Coevolution of Supermassive Black Holes and Galaxies Using Gravitationally Lensed Quasar Hosts. *ApJ*, 649(2):616–634, October 2006.

[34] Kenneth C. Wong, Sherry H. Suyu, Geoff C. F. Chen, Cristian E. Rusu, Martin Millon, Dominique Sluse, Vivien Bonvin, Christopher D. Fassnacht, Stefan Taubenberger, Matthew W. Auger, Simon Birrer, James H. H. Chan, Frederic Courbin, Stefan Hilbert, Olga Tihhonova, Tommaso Treu, Adriano Agnello, Xuheng Ding, In Jee, Eiichiro Komatsu, Anowar J. Shajib, Alessandro Sonnenfeld, Roger D. Blandford, Léon V. E. Koopmans, Philip J. Marshall, and Georges Meylan. H0LiCOW - XIII. A 2.4 per cent measurement of H_0 from lensed quasars: 5.3σ tension between early- and late-Universe probes. *MNRAS*, 498(1):1420–1439, October 2020.

[35] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015.

[36] Aapo Hyvärinen. Estimation of non-normalized statistical models by score matching. *Journal of Machine Learning Research*, 6(24):695–709, 2005.

[37] Pascal Vincent. A connection between score matching and denoising autoencoders. *Neural Comput.*, 23(7):1661–1674, 2011.

[38] Brian D.O. Anderson. Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12(3):313–326, 1982.

[39] Alexandros Graikos, Nikolay Malkin, Nebojsa Jojic, and Dimitris Samaras. Diffusion models as plug-and-play priors. In *Thirty-Sixth Conference on Neural Information Processing Systems*, 2022.

[40] H. O. Hartley. Maximum likelihood estimation from incomplete data. *Biometrics*, 14(2):174–194, 1958.

[41] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum Likelihood from Incomplete Data Via the EM Algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22, 12 1977.

[42] G.J. McLachlan and T. Krishnan. *The EM Algorithm and Extensions*. Wiley Series in Probability and Statistics. Wiley, 2007.

[43] William Ruth. A review of Monte Carlo-based versions of the EM algorithm. *arXiv e-prints*, page arXiv:2401.00945, January 2024.

[44] François Rozet, Gérôme Andry, François Lanusse, and Gilles Louppe. Learning Diffusion Priors from Observations by Expectation Maximization. *arXiv e-prints*, page arXiv:2405.13712, May 2024.

[45] Arjun Dey, David J. Schlegel, Dustin Lang, Robert Blum, Kaylan Burleigh, Xiaohui Fan, Joseph R. Findlay, Doug Finkbeiner, David Herrera, Stéphanie Juneau, Martin Landriau, Michael Levi, Ian McGreer, Aaron Meisner, Adam D. Myers, John Moustakas, Peter Nugent, Anna Patej, Edward F. Schlafly, Alistair R. Walker, Francisco Valdes, Benjamin A. Weaver, Christophe Yèche, Hu Zou, Xu Zhou, Behzad Abareshi, T. M. C. Abbott, Bela Abolfathi, C. Aguilera, Shadab Alam, Lori Allen, A. Alvarez, James Annis, Behzad Ansarinejad, Marie Aubert, Jacqueline Beechert, Eric F. Bell, Segev Y. BenZvi, Florian Beutler, Richard M. Bielby, Adam S. Bolton, César Briceño, Elizabeth J. Buckley-Geer, Karen Butler, Annalisa Calamida, Raymond G. Carlberg, Paul Carter, Ricard Casas, Francisco J. Castander, Yumi Choi, Johan Comparat, Elena Cukanovaite, Timothée Delubac, Kaitlin DeVries, Sharmila Dey, Govinda Dhungana, Mark Dickinson, Zhejie Ding, John B. Donaldson, Yutong Duan, Christopher J. Duckworth, Sarah Eftekharzadeh, Daniel J. Eisenstein, Thomas Etourneau, Parker A. Fagrelius, Jay Farihi, Mike Fitzpatrick, Andreu Font-Ribera, Leah Fulmer, Boris T. Gänsicke, Enrique Gaztanaga, Koshy George, David W. Gerdes, Satya Gontcho A Gontcho, Claudio Gorgoni, Gregory Green, Julien Guy, Diane Harmer, M. Hernandez, Klaus Honscheid, Lijuan (Wendy) Huang, David J. James, Buell T. Jannuzi, Linhua Jiang, Richard Joyce, Armin Karcher, Sonia Karkar, Robert Kehoe, Kneib Jean-Paul, Andrea Kueter-Young, Ting-Wen Lan, Tod R. Lauer, Laurent Le Guillou, Auguste Le Van Suu, Jae Hyeon Lee, Michael Lesser, Laurence Perreault Levasseur, Ting S. Li, Justin L. Mann, Robert Marshall, C. E. Martínez-Vázquez, Paul Martini, Hélion du Mas des Bourboux, Sean McManus, Tobias Gabriel Meier, Brice Ménard, Nigel Metcalfe, Andrea Muñoz-Gutiérrez, Joan Najita, Kevin Napier, Gautham Narayan, Jeffrey A. Newman, Jundan Nie, Brian Nord, Dara J. Norman, Knut A. G. Olsen, Anthony Paat, Nathalie Palanque-Delabrouille, Xiyan Peng, Claire L. Poppett, Megan R. Poremba, Abhishek Prakash, David Rabinowitz, Anand Raichoor, Mehdi Rezaie, A. N. Robertson, Natalie A. Roe, Ashley J. Ross, Nicholas P. Ross, Gregory Rudnick, Sasha Safonova, Abhijit Saha, F. Javier Sánchez, Elodie Savary, Heidi Schweiker, Adam Scott, Hee-Jong Seo, Huanyuan Shan, David R. Silva, Zachary Slepian, Christian Soto, David Sprayberry, Ryan Staten, Coley M. Stillman, Robert J. Stupak, David L. Summers, Suk Sien Tie, H. Tirado, Mariana Vargas-Magaña, A. Katherina Vivas, Risa H. Wechsler, Doug Williams, Jinyi Yang, Qian Yang, Tolga Yapici, Dennis Zaritsky, A. Zenteno, Kai Zhang, Tianmeng Zhang, Rongpu Zhou, and Zhimin Zhou. Overview of the desi legacy imaging surveys. *The Astronomical Journal*, 157(5):168, April 2019.

[46] Connor Stone, Alexandre Adam, Laurence Perreault-Levasseur, Yashar Hezaveh, Cordero Core, Landung Setiawan, Adam Coogan, Andreas Filipp, Charles Wilson, Misha Yantovski-Barth, and Ronan Legin. caustics, April 2024.

[47] Pablo Lemos, Sammy Sharief, Nikolay Malkin, Laurence Perreault-Levasseur, and Yashar Hezaveh. PQMass: Probabilistic Assessment of the Quality of Generative Models using Probability Mass Estimation, February 2024. arXiv:2402.04355 [cs, stat].

[48] Yann LeCun, Corinna Cortes, and Christopher J. C. Burges. The mnist database of handwritten digits, 1998.

[49] Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-Rank Adaptation of Large Language Models. *arXiv e-prints*, page arXiv:2106.09685, June 2021.

[50] Euclid Collaboration, Y. Mellier, Abdurro'uf, J. A. Acevedo Barroso, A. Achúcarro, J. Adamek, R. Adam, G. E. Addison, N. Aghanim, M. Aguena, V. Ajani, Y. Akrami, A. Al-Bahlawan, A. Alavi, I. S. Albuquerque, G. Alestas, G. Alguero, A. Allaoui, S. W. Allen, V. Allevato, A. V. Alonso-Tetilla, B. Altieri, A. Alvarez-Candal, A. Amara, L. Amendola, J. Amiaux, I. T. Andika, S. Andreon, A. Andrews, G. Angora, R. E. Angulo, F. Annibali, A. Anselmi, S. Anselmi, S. Arcari, M. Archidiacono, G. Aricò, M. Arnaud, S. Arnouts, M. Asgari, J. Asorey, L. Atayde, H. Atek, F. Atrio-Barandela, M. Aubert, E. Aubourg, T. Auphan, N. Auricchio, B. Aussel, H. Aussel, P. P. Avelino, A. Avgoustidis, S. Avila, S. Awan, R. Azzollini, C. Baccigalupi, E. Bachelet, D. Bacon, M. Baes, M. B. Bagley, B. Bahr-Kalus, A. Balaguera-Antolinez, E. Balbinot, M. Balcells, M. Baldi, I. Baldry, A. Balestra, M. Ballardini, O. Ballester, M. Balogh, E. Bañados, R. Barbier, S. Bardelli, T. Barreiro, J. C. Barriere, B. J. Barros, A. Barthelemy, N. Bartolo, A. Basset, P. Battaglia, A. J. Battisti, C. M. Baugh, L. Baumont, L. Bazzanini, J. P. Beaulieu, V. Beckmann, A. N. Belikov, J. Bel, F. Bellagamba, M. Bella, E. Bellini, K. Benabed, R. Bender, G. Benevento, C. L. Bennett, K. Benson, P. Bergamini, J. R. Bermejo-Climent, F. Bernardeau, D. Bertacca, M. Berthe, J. Berthier, M. Bethermin, F. Beutler, C. Bevillon, S. Bhargava, R. Bhatawdekar, L. Bisigello, A. Biviano, R. P. Blake, A. Blanchard, J. Blazek, L. Blot, A. Bosco, C. Bodendorf, T. Boenke, H. Böhringer, M. Bolzonella, A. Bonchi, M. Bonici, D. Bonino, L. Bonino, C. Bonvin, W. Bon, J. T. Booth, S. Borgani, A. S. Borlaff, E. Borsato, A. Bosco, B. Bose, M. T. Botticella, A. Boucaud, F. Bouche, J. S. Boucher, D. Boutigny, T. Bouvard, H. Bouy, R. A. A. Bowler, V. Bozza, E. Bozzo, E. Branchini, S. Brau-Nogue, P. Brekke, M. N. Bremer, M. Brescia, M. A. Breton, J. Brinchmann, T. Brinckmann, C. Brockley-Blatt, M. Brodwin, L. Brouard, M. L. Brown, S. Bruton, J. Bucko, H. Buddelmeijer, G. Buenadicha, F. Buitrago, P. Burger, C. Burigana, V. Busillo, D. Busonero, R. Cabanac, L. Cabayol-Garcia, M. S. Cagliari, A. Caillat, L. Caillat, M. Calabrese, A. Calabro, G. Calderone, F. Calura, B. Camacho Quevedo, S. Camera, L. Campos, G. Canas-Herrera, G. P. Candini, M. Cantiello, V. Capobianco, E. Cappellaro, N. Cappelluti, A. Cappi, K. I. Caputi, C. Cara, C. Carbone, V. F. Cardone, E. Carella, R. G. Carlberg, M. Carle, L. Carminati, F. Caro, J. M. Carrasco, J. Carretero, P. Carrilho, J. Carron Duque, B. Carry, A. Carvalho, C. S. Carvalho, R. Casas, S. Casas, P. Casenove, C. M. Casey, P. Cassata, F. J. Castander, D. Castelao, M. Castellano, L. Castiblanco, G. Castignani, T. Castro, C. Cavet, S. Cavuoti, P. Y. Chabaud, K. C. Chambers, Y. Charles, S. Charlot, N. Chartab, R. Chary, F. Chaumeil, H. Cho, G. Chon, E. Ciancetta, P. Ciliegi, A. Cimatti, M. Cimino, M. R. L. Cioni, R. Claydon, C. Cleland, B. Clément, D. L. Clements, N. Clerc, S. Clesse, S. Codis, F. Cogato, J. Colbert, R. E. Cole, P. Coles, T. E. Collett, R. S. Collins, C. Colodro-Conde, C. Colombo, F. Combes, V. Conforti, G. Congedo, S. Conseil, C. J. Conselice, S. Contarini, T. Contini, L. Conversi, A. R. Cooray, Y. Copin, P. S. Corasaniti, P. Corcho-Caballero, L. Corcione, O. Cordes, O. Corpace, M. Correnti, M. Costanzi, A. Costille, F. Courbin, L. Courcoult Mifsud, H. M. Courtois, M. C. Cousinou, G. Covone, T. Cowell, C. Cragg, G. Cresci, S. Cristiani, M. Crocce, M. Cropper, P. E Crouzet, B. Csizi, J. G. Cuby, E. Cucchetti, O. Cucciati, J. C. Cuillandre, P. A. C. Cunha, V. Cuozzo, E. Daddi, M. D'Addona, C. Dafonte, N. Dagoneau, E. Dalessandro, G. B. Dalton, G. D'Amico, H. Dannerbauer, P. Danto, I. Das, A. Da Silva, R. da Silva, G. Daste, J. E. Davies, S. Davini, T. de Boer, R. Decarli, B. De Caro, H. Degaudenzi, G. Degni, J. T. A. de Jong, L. F. de la Bella, S. de la Torre, F. Delhaise, D. Delley, G. Delucchi, G. De Lucia, J. Denniston, F. De Paolis, M. De Petris, A. Derosa, S. Desai, V. Desjacques, G. Despali, G. Desprez, J. De Vicente-Albendea, Y. Deville, J. D. F. Dias, A. Díaz-Sánchez, J. J. Diaz, S. Di Domizio, J. M. Diego, D. Di Ferdinando, A. M. Di Giorgio, P. Dimauro, J. Dinis, K. Dolag, C. Dolding, H. Dole, H. Domínguez Sánchez, O. Doré, F. Dournac, M. Douspis, H. Dreihahn, B. Droge, B. Dryer, F. Dubath, P. A. Duc, F. Ducret, C. Duffy, F. Dufresne, C. A. J. Duncan, X. Dupac, V. Duret, R. Durrer, F. Durret, S. Dusini, A. Ealet, A. Eggemeier, P. R. M. Eisenhardt, D. Elbaz, M. Y. Elkhashab, A. Ellien, J. Endicott, A. Enia, T. Erben, J. A. Escartin Vigo, S. Escoffier, I. Escudero Sanz, J. Essert, S. Ettori, M. Ezziati, G. Fabbian, M. Fabricius, Y. Fang, A. Farina, M. Farina, R. Farinelli, S. Farrens, F. Faustini, A. Feltre, A. M. N. Ferguson, P. Ferrando, A. G. Ferrari, A. Ferré-Mateu, P. G. Ferreira, I. Ferreras, I. Ferrero, S. Ferriol, P. Ferruit, D. Filleul, F. Finelli, S. L. Finkelstein,

A. Finoguenov, B. Fiorini, F. Flentge, P. Focardi, J. Fonseca, A. Fontana, F. Fontanot, F. Fornari, P. Fosalba, M. Fossati, S. Fotopoulou, D. Fouchez, N. Fourmanoit, M. Frailis, D. Fraix-Burnet, E. Franceschi, A. Franco, P. Franzetti, J. Freihoefer, G. Frittoli, P. A. Frugier, N. Frusciante, A. Fumagalli, M. Fumagalli, M. Fumana, Y. Fu, L. Gabarra, S. Galeotta, L. Galluccio, K. Ganga, H. Gao, J. García-Bellido, K. Garcia, J. P. Gardner, B. Garilli, L. M. Gaspar-Venancio, T. Gasparetto, V. Gautard, R. Gavazzi, E. Gaztanaga, L. Genolet, R. Genova Santos, F. Gentile, K. George, Z. Ghaffari, F. Giacomini, F. Gianotti, G. P. S. Gibb, W. Gillard, B. Gillis, M. Ginolfi, C. Giocoli, M. Girardi, S. K. Giri, L. W. K. Goh, P. Gómez-Alvarez, A. H. Gonzalez, E. J. Gonzalez, J. C. Gonzalez, S. Gouyou Beauchamps, G. Gozaliasl, J. Gracia-Carpio, S. Grandis, B. R. Granett, M. Granvik, A. Grazian, A. Gregorio, C. Grenet, C. Grillo, F. Grupp, C. Gruppioni, A. Gruppuso, C. Guerbuez, S. Guerrini, M. Guidi, P. Guillard, C. M. Gutierrez, P. Guttridge, L. Guzzo, S. Gwyn, J. Haapala, J. Haase, C. R. Haddow, M. Hailey, A. Hall, D. Hall, N. Hamaus, B. S. Haridasu, J. Harnois-Déraps, C. Harper, W. G. Hartley, G. Hasinger, F. Hassani, N. A. Hatch, S. V. H. Haugan, B. Häußler, A. Heavens, L. Heisenberg, A. Helmi, G. Helou, S. Hemmati, K. Henares, O. Herent, C. Hernández-Monteagudo, T. Heuberger, P. C. Hewett, S. Heydenreich, H. Hildebrandt, M. Hirschmann, J. Hjorth, J. Hoar, H. Hoekstra, A. D. Holland, M. S. Holliman, W. Holmes, I. Hook, B. Horeau, F. Hormuth, A. Hornstrup, S. Hosseini, D. Hu, P. Hudelot, M. J. Hudson, M. Huertas-Company, E. M. Huff, A. C. N. Hughes, A. Humphrey, L. K. Hunt, D. D. Huynh, R. Ibata, K. Ichikawa, S. Iglesias-Groth, O. Ilbert, S. Ilić, L. Ingoglia, E. Iodice, H. Israel, U. E. Israelsson, L. Izzo, P. Jablonka, N. Jackson, J. Jacobson, M. Jafariyazani, K. Jahnke, H. Jansen, M. J. Jarvis, J. Jasche, M. Jauzac, N. Jeffrey, M. Jhabvala, Y. Jimenez-Teja, A. Jimenez Muñoz, B. Joachimi, P. H. Johansson, S. Joudaki, E. Jullo, J. J. E. Kajava, Y. Kang, A. Kannawadi, V. Kansal, D. Karagiannis, M. Kärcher, A. Kashlinsky, M. V. Kazandjian, F. Keck, E. Keihänen, E. Kerins, S. Kermiche, A. Khalil, A. Kiessling, K. Kiiveri, M. Kilbinger, J. Kim, R. King, C. C. Kirkpatrick, T. Kitching, M. Kluge, M. Knabenhans, J. H. Knapen, A. Knebe, J. P. Kneib, R. Kohley, L. V. E. Koopmans, H. Koskinen, E. Koulouridis, R. Kou, A. Kovács, I. Kova{č}ić, A. Kowalczyk, K. Koyama, K. Kraljic, O. Krause, S. Kruk, B. Kubik, U. Kuchner, K. Kuijken, M. Kümmel, M. Kunz, H. Kurki-Suonio, F. Lacasa, C. G. Lacey, F. La Franca, N. Lagarde, O. Lahav, C. Laigle, A. La Marca, O. La Marle, B. Lamine, M. C. Lam, A. Lançon, H. Landt, M. Langer, A. Lapi, C. Larcheveque, S. S. Larsen, M. Lattanzi, F. Laudisio, D. Laugier, R. Laureijs, G. Lavaux, A. Lawrenson, A. Lazanu, T. Lazeyras, Q. Le Boulc'h, A. M. C. Le Brun, V. Le Brun, F. Leclercq, S. Lee, J. Le Graet, L. Legrand, K. N. Leirvik, M. Le Jeune, M. Lembo, D. Le Mignant, M. D. Lepinzan, F. Lepori, G. F. Lesci, J. Lesgourgues, L. Leuzzi, M. E. Levi, T. I. Liaudat, G. Libet, P. Liebing, S. Ligori, P. B. Lilje, C. C. Lin, D. Linde, E. Linder, V. Lindholm, L. Linke, S. S. Li, S. J. Liu, I. Lloro, F. S. N. Lobo, N. Lodieu, M. Lombardi, L. Lombriser, P. Lonare, G. Longo, M. López-Caniego, X. Lopez Lopez, J. Lorenzo Alvarez, A. Loureiro, J. Loveday, E. Lusso, J. Macias-Perez, T. Maciaszek, M. Magliocchetti, F. Magnard, E. A. Magnier, A. Magro, G. Mahler, G. Mainetti, D. Maino, E. Maiorano, E. Maiorano, N. Malavasi, G. A. Mamon, C. Mancini, R. Mandelbaum, M. Manera, A. Manjón-García, F. Mannucci, O. Mansutti, M. Manteiga Outeiro, R. Maoli, C. Maraston, S. Marcin, P. Marcos-Arenal, B. Margalef-Bentabol, O. Marggraf, D. Marinucci, M. Marinucci, K. Markovic, F. R. Marleau, J. Marpaud, J. Martignac, J. Martín-Fleitas, P. Martin-Moruno, E. L. Martin, M. Martinelli, N. Martinet, H. Martin, C. J. A. P. Martins, F. Marulli, D. Massari, R. Massey, D. C. Masters, S. Matarrese, Y. Matsuoka, S. Matthew, B. J. Maughan, N. Mauri, L. Maurin, S. Maurogordato, K. McCarthy, A. W. McConnachie, H. J. McCracken, I. McDonald, J. D. McEwen, C. J. R. McPartland, E. Medinaceli, V. Mehta, S. Mei, M. Melchior, J. B. Melin, B. Ménard, J. Mendes, J. Mendez-Abreu, M. Meneghetti, A. Mercurio, E. Merlin, R. B. Metcalf, G. Meylan, M. Migliaccio, M. Mignoli, L. Miller, M. Miluzio, B. Milvang-Jensen, J. P. Mimoso, R. Miquel, H. Miyatake, B. Mobasher, J. J. Mohr, P. Monaco, M. Monguió, A. Montoro, A. Mora, A. Moradinezhad Dizgah, M. Moresco, C. Moretti, G. Morgante, N. Morisset, T. J. Moriya, P. W. Morris, D. J. Mortlock, L. Moscardini, D. F. Mota, L. A. Moustakas, T. Moutard, T. Müller, E. Munari, G. Murphree, C. Murray, N. Murray, P. Musi, S. Nadathur, B. C. Nagam, T. Nagao, K. Naidoo, R. Nakajima, C. Nally, P. Natoli, A. Navarro-Alsina, D. Navarro Girones, C. Neissner, A. Nersesian, S. Nesseris, H. N. Nguyen-Kim, L. Nicastro, R. C. Nichol, M. Nielbock, S. M. Niemi, S. Nieto, K. Nilsson, J. Noller, P. Norberg, A. Nourizonoz, P. Ntelis, A. A. Nucita, P. Nugent, N. J. Nunes, T. Nutma, I. Ocampo, J. Odier, P. A. Oesch, M. Oguri, D. Magalhaes Oliveira, M. Onoue, T. Oosterbroek, F. Oppizzi, C. Ordenovic, K. Osato, F. Pacaud, F. Pace, C. Padilla, K. Paech, L. Pagano, M. J. Page, E. Palazzi, S. Paltani, S. Pamuk, S. Pandolfi, D. Paoletti, M. Paolillo, P. Papaderos, K. Pard-

ede, G. Parimbelli, A. Parmar, C. Partmann, F. Pasian, F. Passalacqua, K. Paterson, L. Patrizii, C. Pattison, A. Paulino-Afonso, R. Paviot, J. A. Peacock, F. R. Pearce, K. Pedersen, A. Peel, R. F. Peletier, M. Pellejero Ibanez, R. Pello, M. T. Penny, W. J. Percival, A. Perez-Garrido, L. Perotto, V. Pettorino, A. Pezzotta, S. Pezzuto, A. Philippon, O. Piersanti, M. Pietroni, L. Piga, L. Pilo, S. Pires, A. Pisani, A. Pizzella, L. Pizzuti, C. Plana, G. Polenta, J. E. Pollack, M. Poncet, M. Pöntinen, P. Pool, L. A. Popa, V. Popa, J. Popp, C. Porciani, L. Porth, D. Potter, M. Poulain, A. Pourtsidou, L. Pozzetti, I. Prandoni, G. W. Pratt, S. Prezelus, E. Prieto, A. Pugno, S. Quai, L. Quilley, G. D. Racca, A. Raccanelli, G. Rácz, S. Radinović, M. Radovich, A. Ragagnin, U. Ragnit, F. Raison, N. Ramos-Chernenko, C. Ranc, N. Raylet, R. Rebolo, A. Refregier, P. Reimberg, T. H. Reiprich, F. Renk, A. Renzi, J. Retre, Y. Revaz, C. Reylé, L. Reynolds, J. Rhodes, F. Ricci, M. Ricci, G. Riccio, S. O. Ricken, S. Rissanen, I. Risso, H. W. Rix, A. C. Robin, B. Rocca-Volmerange, P. F. Rocci, M. Rodenhuis, G. Rodighiero, M. Rodriguez Monroy, R. P. Rollins, M. Romanello, J. Roman, E. Romelli, M. Romero-Gomez, M. Roncarelli, P. Rosati, C. Rosset, E. Rossetti, W. Roster, H. J. A. Rottgering, A. Rozas-Fernández, K. Ruane, J. A. Rubino-Martin, A. Rudolph, F. Ruppin, B. Rusholme, S. Sacquegna, I. Sáez-Casares, S. Saga, R. Saglia, M. Sahlén, T. Saifollahi, Z. Sakr, J. Salvalaggio, R. Salvaterra, L. Salvati, M. Salvato, J. C. Salvignol, A. G. Sánchez, E. Sanchez, D. B. Sanders, D. Sapone, M. Saponara, E. Sarpa, F. Sarron, S. Sartori, B. Sassolas, L. Sauniere, M. Sauvage, M. Sawicki, R. Scaramella, C. Scarlata, L. Scharré, J. Schaye, J. A. Schewtschenko, J. T. Schindler, E. Schinnerer, M. Schirmer, F. Schmidt, F. Schmidt, M. Schmidt, A. Schneider, M. Schneider, P. Schneider, N. Schöneberg, T. Schrabback, M. Schultheis, S. Schulz, J. Schwartz, D. Sciotti, M. Scodeggio, D. Scognamiglio, D. Scott, V. Scottez, A. Secroun, E. Sefusatti, G. Seidel, M. Seiffert, E. Sellentin, M. Selwood, E. Semboloni, M. Sereno, S. Serjeant, S. Serrano, F. Shankar, R. M. Sharples, A. Short, A. Shulevski, M. Shuntov, M. Sias, G. Sikkema, A. Silvestri, P. Simon, C. Sirignano, G. Sirri, J. Skottfelt, E. Slezak, D. Sluse, G. P. Smith, L. C. Smith, R. E. Smith, S. J. A. Smit, F. Soldano, B. G. B. Solheim, J. G. Sorce, F. Sorrenti, E. Soubrie, L. Spinoglio, A. Spurio Mancini, J. Stadel, L. Stagnaro, L. Stanco, S. A. Stanford, J. L. Starck, P. Stassi, J. Steinwagner, D. Stern, C. Stone, P. Strada, F. Strafella, D. Stramaccioni, C. Surace, F. Sureau, S. H. Suyu, I. Swindells, M. Szafraniec, I. Szapudi, S. Taamoli, M. Talia, P. Tallada-Crespí, K. Tanidis, C. Tao, P. Tarrío, D. Tavagnacco, A. N. Taylor, J. E. Taylor, P. L. Taylor, E. M. Teixeira, M. Tenti, P. Teodoro Idiago, H. I. Teplitz, I. Tereno, N. Tessore, V. Testa, G. Testera, M. Tewes, R. Teyssier, N. Theret, C. Thizy, P. D. Thomas, Y. Toba, S. Toft, R. Toledo-Moreo, E. Tolstoy, E. Tommasi, O. Torbaniuk, F. Torradeflot, C. Tortora, S. Tosi, S. Tosti, M. Trifoglio, A. Troja, T. Trombetti, A. Tronconi, M. Tsedrik, A. Tsyganov, M. Tucci, I. Tutusaus, C. Uhlemann, L. Ulivi, M. Urbano, L. Vacher, L. Vaillon, I. Valdes, E. A. Valentijn, L. Valenziano, C. Valieri, J. Valiviita, M. Van den Broeck, T. Vassallo, R. Vavrek, B. Venemans, A. Venhola, S. Ventura, G. Verdoes Kleijn, D. Vergani, A. Verma, F. Vernizzi, A. Veropalumbo, G. Verza, C. Vescovi, D. Vibert, M. Viel, P. Vielzeuf, C. Viglione, A. Viitanen, F. Villaescusa-Navarro, S. Vinciguerra, F. Visticot, K. Voggel, M. von Wietersheim-Kramsta, W. J. Vriend, S. Wachter, M. Walmsley, G. Walth, D. M. Walton, N. A. Walton, M. Wander, L. Wang, Y. Wang, J. R. Weaver, J. Weller, D. J. Whalen, M. Wiesmann, J. Wilde, O. R. Williams, H. A. Winther, A. Wittje, J. H. W. Wong, A. H. Wright, V. Yankelevich, H. W. Yeung, S. Youles, L. Y. A. Yung, A. Zacchei, L. Zalesky, G. Zamorani, A. Zamorano Vitorelli, M. Zanoni Marc, M. Zennaro, F. M. Zerbi, I. A. Zinchenko, J. Zoubian, E. Zucca, and M. Zumalacarregui. Euclid. I. Overview of the Euclid mission. *arXiv e-prints*, page arXiv:2405.13491, May 2024.

[51] Bob Blum, Seth W. Digel, Alex Drlica-Wagner, Salman Habib, Katrin Heitmann, Mustapha Ishak, Saurabh W. Jha, Steven M. Kahn, Rachel Mandelbaum, Phil Marshall, Jeffrey A. Newman, Aaron Roodman, and Christopher W. Stubbs. Snowmass2021 Cosmic Frontier White Paper: Rubin Observatory after LSST. *arXiv e-prints*, page arXiv:2203.07220, March 2022.

[52] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In Yoshua Bengio and Yann LeCun, editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015.

[53] P. Camps and M. Baes. Skirt 9: Redesigning an advanced dust radiative transfer code to allow kinematics, line transfer and polarization by aligned dust grains. *Astronomy and Computing*, 31:100381, 2020.

[54] Dylan Nelson, Volker Springel, Annalisa Pillepich, Vicente Rodriguez-Gomez, Paul Torrey, Shy Genel, Mark Vogelsberger, Ruediger Pakmor, Federico Marinacci, Rainer Weinberger,

Luke Kelley, Mark Lovell, Benedikt Diemer, and Lars Hernquist. The IllustrisTNG simulations: public data release. *Computational Astrophysics and Cosmology*, 6(1):2, May 2019.

[55] Hiroaki Aihara, Nobuo Arimoto, Robert Armstrong, Stéphane Arnouts, Neta A. Bahcall, Steven Bickerton, James Bosch, Kevin Bundy, Peter L. Capak, James H. H. Chan, Masashi Chiba, Jean Coupon, Eiichi Egami, Motohiro Enoki, Francois Finet, Hiroki Fujimori, Seiji Fujimoto, Hisanori Furusawa, Junko Furusawa, Tomotsugu Goto, Andy Goulding, Johnny P. Greco, Jenny E. Greene, James E. Gunn, Takashi Hamana, Yuichi Harikane, Yasuhiro Hashimoto, Takashi Hattori, Masao Hayashi, Yusuke Hayashi, Krzysztof G. Hełminiak, Ryo Higuchi, Chiaki Hikage, Paul T. P. Ho, Bau-Ching Hsieh, Kuiyun Huang, Song Huang, Hiroyuki Ikeda, Masatoshi Imanishi, Akio K. Inoue, Kazushi Iwasawa, Ikuru Iwata, Anton T. Jaelani, Hung-Yu Jian, Yukiko Kamata, Hiroshi Karoji, Nobunari Kashikawa, Nobuhiko Katayama, Satoshi Kawanomoto, Issha Kayo, Jin Koda, Michitaro Koike, Takashi Kojima, Yutaka Komiyama, Akira Konno, Shintaro Koshida, Yusei Koyama, Haruka Kusakabe, Alexie Leauthaud, Chien-Hsiu Lee, Lihwai Lin, Yen-Ting Lin, Robert H. Lupton, Rachel Mandelbaum, Yoshiki Matsuoka, Elinor Medezinski, Sogo Mineo, Shoken Miyama, Hironao Miyatake, Satoshi Miyazaki, Rieko Momose, Anupreeta More, Surhud More, Yuki Moritani, Takashi J. Moriya, Tomoki Morokuma, Shiro Mukae, Ryoma Murata, Hitoshi Murayama, Tohru Nagao, Fumiaki Nakata, Mana Niida, Hiroko Niikura, Atsushi J. Nishizawa, Yoshiyuki Obuchi, Masamune Oguri, Yukie Oishi, Nobuhiro Okabe, Sakurako Okamoto, Yuki Okura, Yoshiaki Ono, Masato Onodera, Masafusa Onoue, Ken Osato, Masami Ouchi, Paul A. Price, Tae-Soo Pyo, Masao Sako, Marcin Sawicki, Takatoshi Shibuya, Kazuhiro Shimasaku, Atsushi Shimono, Masato Shirasaki, John D. Silverman, Melanie Simet, Joshua Speagle, David N. Spergel, Michael A. Strauss, Yuma Sugahara, Naoshi Sugiyama, Yasushi Suto, Sherry H. Suyu, Nao Suzuki, Philip J. Tait, Masahiro Takada, Tadafumi Takata, Naoyuki Tamura, Manobu M. Tanaka, Masaomi Tanaka, Masayuki Tanaka, Yoko Tanaka, Tsuyoshi Terai, Yuichi Terashima, Yoshiki Toba, Nozomu Tominaga, Jun Toshikawa, Edwin L. Turner, Tomohisa Uchida, Hisakazu Uchiyama, Keiichi Umetsu, Fumihiro Uraguchi, Yuji Urata, Tomonori Usuda, Yousuke Utsumi, Shiang-Yu Wang, Wei-Hao Wang, Kenneth C. Wong, Kiyoto Yabe, Yoshihiko Yamada, Hitomi Yamanoi, Naoki Yasuda, Sherry Yeh, Atsunori Yonehara, and Suraphong Yuma. The Hyper Suprime-Cam SSP Survey: Overview and survey design. *PASJ*, 70:S4, January 2018.

[56] Michael R. Blanton, Matthew A. Bershady, Bela Abolfathi, Franco D. Albareti, Carlos Allende Prieto, Andres Almeida, Javier Alonso-García, Friedrich Anders, Scott F. Anderson, Brett Andrews, Erik Aquino-Ortíz, Alfonso Aragón-Salamanca, Maria Argudo-Fernández, Eric Armengaud, Eric Aubourg, Vladimir Avila-Reese, Carles Badenes, Stephen Bailey, Kathleen A. Barger, Jorge Barrera-Ballesteros, Curtis Bartosz, Dominic Bates, Falk Baumgarten, Julian Bautista, Rachael Beaton, Timothy C. Beers, Francesco Belfiore, Chad F. Bender, Andreas A. Berlind, Mariangela Bernardi, Florian Beutler, Jonathan C. Bird, Dmitry Bizyaev, Guillermo A. Blanc, Michael Blomqvist, Adam S. Bolton, Médéric Boquien, Jura Borissova, Remco van den Bosch, Jo Bovy, William N. Brandt, Jonathan Brinkmann, Joel R. Brownstein, Kevin Bundy, Adam J. Burgasser, Etienne Burtin, Nicolás G. Busca, Michele Cappellari, Maria Leticia Delgado Carigi, Joleen K. Carlberg, Aurelio Carnero Rosell, Ricardo Carrera, Nancy J. Chanover, Brian Cherinka, Edmond Cheung, Yilen Gómez Maqueo Chew, Cristina Chiappini, Peter Doohyun Choi, Drew Chojnowski, Chia-Hsun Chuang, Haeun Chung, Rafael Fernando Cirolini, Nicolas Clerc, Roger E. Cohen, Johan Comparat, Luiz da Costa, Marie-Claude Cousinou, Kevin Covey, Jeffrey D. Crane, Rupert A. C. Croft, Irene Cruz-Gonzalez, Daniel Garrido Cuadra, Katia Cunha, Guillermo J. Damke, Jeremy Darling, Roger Davies, Kyle Dawson, Axel de la Macorra, Flavia Dell'Agli, Nathan De Lee, Timothée Delubac, Francesco Di Mille, Aleks Diamond-Stanic, Mariana Cano-Díaz, John Donor, Juan José Downes, Niv Drory, Hélion du Mas des Bourboux, Christopher J. Duckworth, Tom Dwelly, Jamie Dyer, Garrett Ebelke, Arthur D. Eigenbrot, Daniel J. Eisenstein, Eric Emsellem, Mike Eracleous, Stephanie Escoffier, Michael L. Evans, Xiaohui Fan, Emma Fernández-Alvar, J. G. Fernandez-Trincado, Diane K. Feuillet, Alexis Finoguenov, Scott W. Fleming, Andreu Font-Ribera, Alexander Fredrickson, Gordon Freischlad, Peter M. Frinchaboy, Carla E. Fuentes, Lluís Galbany, R. Garcia-Dias, D. A. García-Hernández, Patrick Gaulme, Doug Geisler, Joseph D. Gelfand, Héctor Gil-Marín, Bruce A. Gillespie, Daniel Goddard, Violeta Gonzalez-Perez, Kathleen Grabowski, Paul J. Green, Catherine J. Grier, James E. Gunn, Hong Guo, Julien Guy, Alex Hagen, ChangHoon Hahn, Matthew Hall, Paul Harding, Sten Hasselquist, Suzanne L. Hawley, Fred Hearty, Jonay I. Gonzalez Hernández, Shirley Ho, David W. Hogg, Kelly Holley-Bockelmann, Jon A. Holtzman,

Parker H. Holzer, Joseph Huehnerhoff, Timothy A. Hutchinson, Ho Seong Hwang, Héctor J. Ibarra-Medel, Gabriele da Silva Ilha, Inese I. Ivans, KeShawn Ivory, Kelly Jackson, Trey W. Jensen, Jennifer A. Johnson, Amy Jones, Henrik Jönsson, Eric Jullo, Vikrant Kamble, Karen Kinemuchi, David Kirkby, Francisco-Shu Kitaura, Mark Klaene, Gillian R. Knapp, Jean-Paul Kneib, Juna A. Kollmeier, Ivan Lacerna, Richard R. Lane, Dustin Lang, David R. Law, Daniel Lazarz, Youngbae Lee, Jean-Marc Le Goff, Fu-Heng Liang, Cheng Li, Hongyu Li, Jianhui Lian, Marcos Lima, Lihwai Lin, Yen-Ting Lin, Sara Bertran de Lis, Chao Liu, Miguel Angel C. de Icaza Lizaola, Dan Long, Sara Lucatello, Britt Lundgren, Nicholas K. MacDonald, Alice Deconto Machado, Chelsea L. MacLeod, Suvrath Mahadevan, Marcio Antonio Geimba Maia, Roberto Maiolino, Steven R. Majewski, Elena Malanushenko, Viktor Malanushenko, Arturo Manchado, Shude Mao, Claudia Maraston, Rui Marques-Chaves, Thomas Masseron, Karen L. Masters, Cameron K. McBride, Richard M. McDermid, Brianne McGrath, Ian D. McGreer, Nicolás Medina Peña, Matthew Melendez, Andrea Merloni, Michael R. Merrifield, Szabolcs Meszaros, Andres Meza, Ivan Minchev, Dante Minniti, Takamitsu Miyaji, Surhud More, John Mulchaey, Francisco Müller-Sánchez, Demitri Muna, Ricardo R. Munoz, Adam D. Myers, Preethi Nair, Kirpal Nandra, Janaina Correa do Nascimento, Alenka Negrete, Melissa Ness, Jeffrey A. Newman, Robert C. Nichol, David L. Nidever, Christian Nitschelm, Pierros Ntelis, Julia E. O'Connell, Ryan J. Oelkers, Audrey Oravetz, Daniel Oravetz, Zach Pace, Nelson Padilla, Nathalie Palanque-Delabrouille, Pedro Alonso Palicio, Kaike Pan, John K. Parejko, Taniya Parikh, Isabelle Pâris, Changbom Park, Alim Y. Patten, Sebastien Peirani, Marcos Pellejero-Ibanez, Samantha Penny, Will J. Percival, Ismael Perez-Fournon, Patrick Petitjean, Matthew M. Pieri, Marc Pinsonneault, Alice Pisani, Radosław Poleski, Francisco Prada, Abhishek Prakash, Anna Bárbara de Andrade Queiroz, M. Jordan Raddick, Anand Raichoor, Sandro Barboza Rembold, Hannah Richstein, Rogemar A. Riffel, Rogério Riffel, Hans-Walter Rix, Annie C. Robin, Constance M. Rockosi, Sergio Rodríguez-Torres, A. Roman-Lopes, Carlos Román-Zúñiga, Margarita Rosado, Ashley J. Ross, Graziano Rossi, John Ruan, Rossana Ruggeri, Eli S. Rykoff, Salvador Salazar-Albornoz, Mara Salvato, Ariel G. Sánchez, D. S. Aguado, José R. Sánchez-Gallego, Felipe A. Santana, Basílio Xavier Santiago, Conor Sayres, Ricardo P. Schiavon, Jaderson da Silva Schimoia, Edward F. Schlafly, David J. Schlegel, Donald P. Schneider, Mathias Schultheis, William J. Schuster, Axel Schwope, Hee-Jong Seo, Zhengyi Shao, Shiyin Shen, Matthew Shetrone, Michael Shull, Joshua D. Simon, Danielle Skinner, M. F. Skrutskie, Anže Slosar, Verne V. Smith, Jennifer S. Sobeck, Flavia Sobreira, Garrett Somers, Diogo Souto, David V. Stark, Keivan Stassun, Fritz Stauffer, Matthias Steinmetz, Thaisa Storchi-Bergmann, Alina Streblyanska, Guy S. Stringfellow, Genaro Suárez, Jing Sun, Nao Suzuki, Laszlo Szigeti, Manuchehr Taghizadeh-Popp, Baitian Tang, Charling Tao, Jamie Tayar, Mita Tembe, Johanna Teske, Aniruddha R. Thakar, Daniel Thomas, Benjamin A. Thompson, Jeremy L. Tinker, Patricia Tissera, Rita Tojeiro, Hector Hernandez Toledo, Sylvain de la Torre, Christy Tremonti, Nicholas W. Troup, Octavio Valenzuela, Inma Martinez Valpuesta, Jaime Vargas-González, Mariana Vargas-Magaña, Jose Alberto Vazquez, Sandro Villanova, M. Vivek, Nicole Vogt, David Wake, Rene Walterbos, Yuting Wang, Benjamin Alan Weaver, Anne-Marie Weijmans, David H. Weinberg, Kyle B. Westfall, David G. Whelan, Vivienne Wild, John Wilson, W. M. Wood-Vasey, Dominika Wylezalek, Ting Xiao, Renbin Yan, Meng Yang, Jason E. Ybarra, Christophe Yèche, Nadia Zakamska, Olga Zamora, Pauline Zarrouk, Gail Zasowski, Kai Zhang, Gong-Bo Zhao, Zheng Zheng, Zheng Zheng, Xu Zhou, Zhi-Min Zhou, Guangtun B. Zhu, Manuela Zoccali, and Hu Zou. Sloan Digital Sky Survey IV: Mapping the Milky Way, Nearby Galaxies, and the Distant Universe. *The Astronomical Journal*, 154(1):28, July 2017.

[57] Abdurro'uf, Katherine Accetta, Conny Aerts, Víctor Silva Aguirre, Romina Ahumada, Nikhil Ajgaonkar, N. Filiz Ak, Shadab Alam, Carlos Allende Prieto, Andrés Almeida, Friedrich Anders, Scott F. Anderson, Brett H. Andrews, Borja Anguiano, Erik Aquino-Ortíz, Alfonso Aragón-Salamanca, Maria Argudo-Fernández, Metin Ata, Marie Aubert, Vladimir Avila-Reese, Carles Badenes, Rodolfo H. Barbá, Kat Barger, Jorge K. Barrera-Ballesteros, Rachael L. Beaton, Timothy C. Beers, Francesco Belfiore, Chad F. Bender, Mariangela Bernardi, Matthew A. Bershady, Florian Beutler, Christian Moni Bidin, Jonathan C. Bird, Dmitry Bizyaev, Guillermo A. Blanc, Michael R. Blanton, Nicholas Fraser Boardman, Adam S. Bolton, Médéric Boquien, Jura Borissova, Jo Bovy, W. N. Brandt, Jordan Brown, Joel R. Brownstein, Marcella Brusa, Johannes Buchner, Kevin Bundy, Joseph N. Burchett, Martin Bureau, Adam Burgasser, Tuesday K. Cabang, Stephanie Campbell, Michele Cappellari, Joleen K. Carlberg, Fábio Carneiro Wanderley, Ricardo Carrera, Jennifer Cash, Yan-Ping Chen, Wei-Huai Chen, Brian Cherinka,

Cristina Chiappini, Peter Doohyun Choi, S. Drew Chojnowski, Haeun Chung, Nicolas Clerc, Roger E. Cohen, Julia M. Comerford, Johan Comparat, Luiz da Costa, Kevin Covey, Jeffrey D. Crane, Irene Cruz-Gonzalez, Connor Culhane, Katia Cunha, Y. Sophia Dai, Guillermo Damke, Jeremy Darling, Jr. Davidson, James W., Roger Davies, Kyle Dawson, Nathan De Lee, Aleksandar M. Diamond-Stanic, Mariana Cano-Díaz, Helena Domínguez Sánchez, John Donor, Chris Duckworth, Tom Dwelly, Daniel J. Eisenstein, Yvonne P. Elsworth, Eric Emsellem, Mike Eracleous, Stephanie Escoffier, Xiaohui Fan, Emily Farr, Shuai Feng, José G. Fernández-Trincado, Diane Feuillet, Andreas Filipp, Sean P. Fillingham, Peter M. Frinchaboy, Sebastien Fromenteau, Lluís Galbany, Rafael A. García, D. A. García-Hernández, Junqiang Ge, Doug Geisler, Joseph Gelfand, Tobias Géron, Benjamin J. Gibson, Julian Goddy, Diego Godoy-Rivera, Kathleen Grabowski, Paul J. Green, Michael Greener, Catherine J. Grier, Emily Griffith, Hong Guo, Julien Guy, Massinissa Hadjara, Paul Harding, Sten Hasselquist, Christian R. Hayes, Fred Hearty, Jesús Hernández, Lewis Hill, David W. Hogg, Jon A. Holtzman, Danny Horta, Bau-Ching Hsieh, Chin-Hao Hsu, Yun-Hsin Hsu, Daniel Huber, Marc Huertas-Company, Brian Hutchinson, Ho Seong Hwang, Héctor J. Ibarra-Medel, Jacob Ider Chitham, Gabriele S. Ilha, Julie Imig, Will Jaekle, Tharindu Jayasinghe, Xihan Ji, Jennifer A. Johnson, Amy Jones, Henrik Jönsson, Ivan Katkov, Dr. Khalatyan, Arman, Karen Kinemuchi, Shobhit Kisku, Johan H. Knapen, Jean-Paul Kneib, Juna A. Kollmeier, Miranda Kong, Marina Kounkel, Kathryn Kreckel, Dhanesh Krishnarao, Ivan Lacerna, Richard R. Lane, Rachel Langgin, Ramon Lavender, David R. Law, Daniel Lazarz, Henry W. Leung, Ho-Hin Leung, Hannah M. Lewis, Cheng Li, Ran Li, Jianhui Lian, Fu-Heng Liang, Lihwai Lin, Yen-Ting Lin, Sicheng Lin, Chris Lintott, Dan Long, Penélope Longa-Peña, Carlos López-Cobá, Shengdong Lu, Britt F. Lundgren, Yuanze Luo, J. Ted Mackereth, Axel de la Macorra, Suvrath Mahadevan, Steven R. Majewski, Arturo Manchado, Travis Mandeville, Claudia Maraston, Berta Margalef-Bentabol, Thomas Masseron, Karen L. Masters, Savita Mathur, Richard M. McDermid, Myles Mckay, Andrea Merloni, Michael Merrifield, Szabolcs Meszaros, Andrea Miglio, Francesco Di Mille, Dante Minniti, Rebecca Minsley, Antonela Monachesi, Jeongin Moon, Benoit Mosser, John Mulchaey, Demitri Muna, Ricardo R. Muñoz, Adam D. Myers, Natalie Myers, Seshadri Nadathur, Preethi Nair, Kirpal Nandra, Justus Neumann, Jeffrey A. Newman, David L. Nidever, Farnik Nikakhtar, Christian Nitschelm, Julia E. O'Connell, Luis Garma-Oehmichen, Gabriel Luan Souza de Oliveira, Richard Olney, Daniel Oravetz, Mario Ortigoza-Urdaneta, Yeisson Osorio, Justin Otter, Zachary J. Pace, Nelson Padilla, Kaike Pan, Hsi-An Pan, Taniya Parikh, James Parker, Sebastien Peirani, Karla Peña Ramírez, Samantha Penny, Will J. Percival, Ismael Perez-Fournon, Marc Pinsonneault, Frédérick Poidevin, Vijith Jacob Poovelil, Adrian M. Price-Whelan, Anna Bárbara de Andrade Queiroz, M. Jordan Raddick, Amy Ray, Sandro Barboza Rembold, Nicole Riddle, Rogemar A. Riffel, Rogério Riffel, Hans-Walter Rix, Annie C. Robin, Aldo Rodríguez-Puebla, Alexandre Roman-Lopes, Carlos Román-Zúñiga, Benjamin Rose, Ashley J. Ross, Graziano Rossi, Kate H. R. Rubin, Mara Salvato, Sebástian F. Sánchez, José R. Sánchez-Gallego, Robyn Sanderson, Felipe Antonio Santana Rojas, Edgar Sarceno, Regina Sarmiento, Conor Sayres, Elizaveta Sazonova, Adam L. Schaefer, Ricardo Schiavon, David J. Schlegel, Donald P. Schneider, Mathias Schultheis, Axel Schwope, Aldo Serenelli, Javier Serna, Zhengyi Shao, Griffin Shapiro, Anubhav Sharma, Yue Shen, Matthew Shetrone, Yiping Shu, Joshua D. Simon, M. F. Skrutskie, Rebecca Smethurst, Verne Smith, Jennifer Sobeck, Taylor Spoo, Dani Sprague, David V. Stark, Keivan G. Stassun, Matthias Steinmetz, Dennis Stello, Alexander Stone-Martinez, Thaisa Storchi-Bergmann, Guy S. Stringfellow, Amelia Stutz, Yung-Chau Su, Manuchehr Taghizadeh-Popp, Michael S. Talbot, Jamie Tayar, Eduardo Telles, Johanna Teske, Ani Thakar, Christopher Theissen, Andrew Tkachenko, Daniel Thomas, Rita Tojeiro, Hector Hernandez Toledo, Nicholas W. Troup, Jonathan R. Trump, James Trussler, Jacqueline Turner, Sarah Tuttle, Eduardo Unda-Sanzana, José Antonio Vázquez-Mata, Marica Valentini, Octavio Valenzuela, Jaime Vargas-González, Mariana Vargas-Magaña, Pablo Vera Alfaro, Sandro Villanova, Fiorenzo Vincenzo, David Wake, Jack T. Warfield, Jessica Diane Washington, Benjamin Alan Weaver, Anne-Marie Weijmans, David H. Weinberg, Achim Weiss, Kyle B. Westfall, Vivienne Wild, Matthew C. Wilde, John C. Wilson, Robert F. Wilson, Mikayla Wilson, Julien Wolf, W. M. Wood-Vasey, Renbin Yan, Olga Zamora, Gail Zasowski, Kai Zhang, Cheng Zhao, Zheng Zheng, Zheng Zheng, and Kai Zhu. The Seventeenth Data Release of the Sloan Digital Sky Surveys: Complete Release of MaNGA, MaStar, and APOGEE-2 Data. *The Astrophysical Journal Supplement Series*, 259(2):35, April 2022.

[58] Adam Ginsburg, Brigitta M. Sipőcz, C. E. Brasseur, Philip S. Cowperthwaite, Matthew W. Craig, Christoph Deil, James Guillochon, Giannina Guzman, Simon Liedtke, Pey Lian Lim, Kelly E. Lockhart, Michael Mommert, Brett M. Morris, Henrik Norman, Madhura Parikh, Magnus V. Persson, Thomas P. Robitaille, Juan-Carlos Segovia, Leo P. Singer, Erik J. Tollerud, Miguel de Val-Borro, Ivan Valtchanov, Julien Woillez, Astroquery Collaboration, and a subset of astropy Collaboration. astroquery: An Astronomical Web-querying Package in Python. *The Astronomical Journal*, 157(3):98, March 2019.

[59] Chris Lintott, Kevin Schawinski, Steven Bamford, Anå¾e Slosar, Kate Land, Daniel Thomas, Edd Edmondson, Karen Masters, Robert C. Nichol, M. Jordan Raddick, Alex Szalay, Dan Andreescu, Phil Murray, and Jan Vandenberg. Galaxy Zoo 1: data release of morphological classifications for nearly 900 000 galaxies. *Monthly Notices of the Royal Astronomical Society*, 410(1):166–178, January 2011.

[60] S. Kullback and R. A. Leibler. On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1):79 – 86, 1951.

[61] C. F. Jeff Wu. On the Convergence Properties of the EM Algorithm. *The Annals of Statistics*, 11(1):95 – 103, 1983.

[62] Pablo Lemos, Adam Coogan, Yashar Hezaveh, and Laurence Perreault-Levasseur. Sampling-based accuracy testing of posterior estimators for general inference. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett, editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 19256–19273. PMLR, 23–29 Jul 2023.

A Details of MNIST experiment

For the experiments with MNIST, in each update, $N = 60\,000$ observations are generated and $K = 1$ posterior samples are generated for each of them to train the next SBM prior. To systematically identify the digits, we train a CNN classifier on MNIST. This allows us to track the proportion of each mode in the prior distributions as a function of the prior update. These results are shown in Figure 4. We observe that the digit 6 is dropped after the first iteration. This is expected since our algorithm only uses posterior samples for its update, and no observations were consistent with the number 6 in the first round given the chosen inverse problem and noise level. On the other hand, learning to infer the digit 4 requires multiple iterations. Interestingly, the proportion of 9's correspondingly increases after the first iteration (as the digits 4 and 9 are morphologically very similar, and therefore true 4s get reconstructed as 9s by the initially misspecified prior) and gradually decreases as the proportion stabilizes closer to the true data-generating distribution. All other numbers keep almost the same proportion.

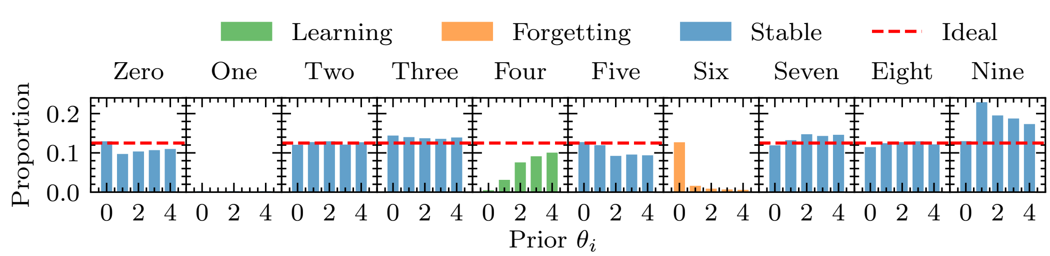

Figure 4: Learning and forgetting dynamics across updates using Algorithm 1 in the MNIST experiment. The plot shows the classification of $2\,048$ prior samples $\mathbf{x} \sim p_{\theta_\alpha}(\mathbf{x})$ at each update, with each panel corresponding to a digit category. $p_{\theta_0}(\mathbf{x})$ was trained excluding digits 4 and 1, while $p_{\theta^\star}(\mathbf{x})$ excluded 6 and 1. The red dashed line represents the proportion from the target distribution.

B Lensing forward model

In our experiments, we employ a Singular Isothermal Ellipsoid (SIE) lens model within a Flat-ΛCDM cosmology. The background source is represented by a grid of pixels. For simplicity, we fix the ellipticity of the lens model to $q = 0.5$, and the position angle to $\phi = \pi/5$. For MNIST we set $\theta_E = 0.5''$ and the source pixel scale to $0.03''$, while for galaxies, these constants are set to $0.8''$ and $0.04''$, respectively. The other parameters are set to the `Caustics` default values.

With these parameters defined, we can obtain the transformation matrix A by computing the Jacobian matrix of the simulator. Finally, we add Gaussian noise to the simulations.

C Models architecture and training

All the models used for the galaxy experiments share the same architecture and training hyperparameters. The same applies to the MNIST models independently. In this appendix, we specify the necessary details to reproduce the experiments. For the MNIST experiments, we use the Variance Exploding SDE [16] and train the SBM model accordingly. We use $\sigma_{min} = 10^{-5}$, and $\sigma_{max} = 100$. For the experiments with galaxies, we use the Variance Preserving SDE, with $\beta_{min} = 10^{-2}$, and $\beta_{max} = 20$. In both cases, we use the NCSN++ architecture [16] via the `score-models`[1] package.

For the galaxy experiment, the architecture parameters within the `score-models` package are:

```
"channels": 3,
"nf": 64,
"ch_mult": [1, 2, 2, 2],
"num_res_blocks": 2
```

And for MNIST:

```
"channels": 1,
"nf": 64,
"ch_mult": [2, 2, 2],
"num_res_blocks": 3,
```

We also use the `score-models` package to train the models. We use the Adam optimizer [52]. For the galaxy experiments we have $lr = 1e^{-4}$, batch size of 256, and `ema_decay` $= 0.999$, and for the MNIST experiments we have $lr = 5e^{-5}$, batch size of 256, and `ema_decay` $= 0.99$. For all experiments, we train for approximately 2.5×10^5 optimization steps. All hyperparameters not specified are left to the `score-models` default values.

We found these configurations for both sets of experiments by trying out 5 different parameter sets. In terms of compute resources, we perform training and inference (both prior and posterior sampling) in A100 GPUs. Each SBM model training and sampling routine was carried out on a single A100 GPU. The MNIST models required 14 hours of training (wall-time), with 16Gb of VRAM allocated. The galaxy models required 20 hours of training (wall-time), with 32Gb of VRAM allocated, while posterior/prior sampling of a set of 1 024 samples require 2 hours (wall-time) and 32Gb of VRAM. The posterior sampling procedure is explained in Appendix G.

In total, for all the experiments, we trained 12 SBM for MNIST, 30 SBM for galaxies, and performed 700 rounds of prior (to simulate observations) and posterior sampling (to create the training dataset for the next SBM prior), with 1 024 samples per set.

D Galaxy datasets

The **spiral dataset** is a synthetic dataset used as the true distribution for our experiments. It is taken from the SKIRT TNG dataset [6], made by a large public collection of images covering bands from 0.3-5 microns made by applying dust radiative transfer post-processing [53] to galaxies from the TNG cosmological magneto-hydrodynamical simulations[2] [54]. This synthetic data is simulated for the *grz* filters of the Hyper Suprime-Cam Subaru Strategic Program [55] and assigned to the (B, G, R) color channels, respectively, and serves as our ground truth sample since it contains no observational noise and can be taken at high resolution. We take 10 000 data points from this dataset, convert to flux in $\mu\mathrm{Jy}\,\mathrm{sr}^{-1}$ units, and downsample to 64×64 pixel images to train an SBM.

[1] github.com/AlexandreAdam/score_models
[2] `www.tng-project.org`

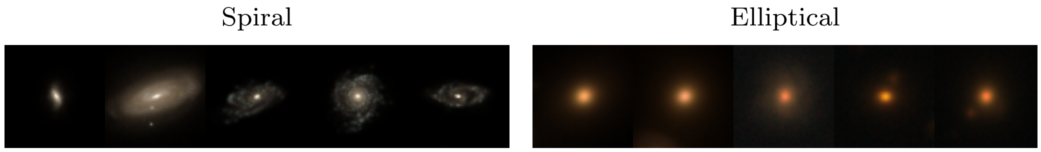

Figure 5: Random samples from the two galaxy datasets used in this work, highlighting the distinction between both. The *spiral galaxy dataset*, comes from a subset of the SKIRT TNG dataset [6]. The *elliptical galaxy dataset* is sourced from the DESI Legacy Imaging Surveys [45]

The **elliptical dataset** is used as the initial prior for our experiments. It is strongly out of distribution compared to the spiral dataset as it includes some corruption effects from real observation (e.g. observational noise and psf blurring) and is overall void of high-frequency features, unlike spiral galaxies. Moreover, the color channels are markedly different between the two sets (see Figure 5). We collected $10\,459$ galaxy images from the DESI Legacy Imaging Surveys [45] DR10, selected using the SDSS-IV [56] DR17 [57] database via Astroquery [58] to construct this dataset. We selected these galaxies based on the elliptical class from GalaxyZoo [59], using a threshold of at least 10 votes and a probability of at least 70%. We also filter postage stamps with thresholds for total magnitude ($5 \leq \texttt{modelMag\_r} \leq 22$), radius ($2' \leq r \leq 20'$), and flux criteria. Here we also select the *grz* bands for this dataset and assign them to the (B, G, R) color channels. The images are sampled at 64×64 pixel resolution, and the galaxy sample has been chosen to fit well in this size at the native resolution for the DESI observations.

Random samples from both datasets are shown in Figure 5.

E Stationary distribution

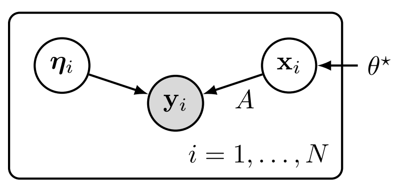

Figure 6: Graphical model of the inference problem. The *true* prior distribution is parametrized by the population-level parameters $\theta^\star$. Our goal is to learn an estimate $\hat{\theta} \approx \theta^\star$. In this work, we have access to the noise distribution that generates $\boldsymbol{\eta}_i$, the forward model A and a set of N observations $\{\mathbf{y}_i\}_{i=1}^N$.

We illustrate the data-generation process for the inference problem in Figure 6, where $p_{\theta^\star}$ is the true population-level distribution describing the underlying prior distribution we aim to approximate. At each iteration of Algorithm 1, we wish to train a SBM with parameters $\theta_{\alpha+1}$ that maximizes the log-likelihood of posterior samples obtained for a set of observations using the previous prior $p_{\theta_\alpha}(\mathbf{x})$. That is, we want to find $\theta_{\alpha+1}$ such that the updated prior given by $\mathbf{s}_{\theta_{\alpha+1}}(\mathbf{x}, t)$ approximates $\mathbb{E}_{\mathbf{y}\sim p(\mathbf{y})} p_{\theta_\alpha}(\mathbf{x} \mid \mathbf{y})$ in the large data limit. This is equivalent to finding the set of $\theta_{\alpha+1}$ minimizing the KL divergence [60]:

$$
\begin{aligned}
\theta_{\alpha+1} &= \arg\min_{\theta\in\Theta} \mathrm{KL}\left(\int d\mathbf{y}\, p(\mathbf{y}) p_{\theta_\alpha}(\mathbf{x} \mid \mathbf{y}) \,\middle\|\, p_\theta(\mathbf{x})\right) \\
&= \arg\max_{\theta\in\Theta} \int d\mathbf{y} d\mathbf{x}\, p(\mathbf{y}) p_{\theta_\alpha}(\mathbf{x} \mid \mathbf{y}) \log p_\theta(\mathbf{x}) = \arg\max_{\theta\in\Theta} \underset{\substack{\mathbf{y}\sim p(\mathbf{y}) \\ \mathbf{x}\sim p_{\theta_\alpha}(\mathbf{x}|\mathbf{y})}}{\mathbb{E}} [\log p_\theta(\mathbf{x})] \quad (3)
\end{aligned}
$$

Definition E.1. *Given a prior $p_{\theta_\alpha}(\mathbf{x})$ and a set of observations $\mathcal{S} = \{\mathbf{y}_i^{(\alpha)}\}_{i=1}^N$, we define the next prior distribution $p_{\theta_{\alpha+1}}(\mathbf{x})$ as the distribution encoded by the generative process of the SBM $\mathbf{s}_{\theta_{\alpha+1}}(\mathbf{x}, t)$ trained by minimizing the denoising objective with training set $\mathcal{D}$ given by:*

$$
\mathcal{D} = \{\mathbf{x}_{i,j} | \mathbf{x}_{i,j} \sim p_{\theta_\alpha}(\mathbf{x} \mid \mathbf{y}_j),\, \forall i \in [1, K],\, \forall \mathbf{y}_j \in \mathcal{S}\} \quad (4)
$$

Under Definition E.1, in the large data limit, the next prior $p_{\theta_{\alpha+1}}(\mathbf{x})$ is the expected posterior under prior $p_{\theta_\alpha}(\mathbf{x})$ with observations from $\mathbf{y} \sim p(\mathbf{y})$:

$$p_{\theta_{\alpha+1}}(\mathbf{x}) = \mathbb{E}_{\mathbf{y}\sim p(\mathbf{y})}[p_{\theta_\alpha}(\mathbf{x} \mid \mathbf{y})] \tag{5}$$

Samples from the observation distribution can be obtained by sampling the underlying population distribution $\mathbf{x} \sim p_{\theta^\star}(\mathbf{x})$ and using the forward process to calculate $\mathbf{y} = A\mathbf{x} + \boldsymbol{\eta}$. The distribution of these samples is given by:

$$p(\mathbf{y}) = \int p(\mathbf{y} \mid \mathbf{x}) p_{\theta^\star}(\mathbf{x})\, d\mathbf{x} \tag{6}$$

We observe that there is no change in the update if $p_{\theta_\alpha}(\mathbf{x})$ has already converged to the stationary distribution, $p_{\hat{\theta}}(\mathbf{x})$, which has marginal likelihood, or evidence, equal to the underlying population distribution, $p_{\hat{\theta}}(\mathbf{y}) = p(\mathbf{y})$:

$$p_{\theta_{\alpha+1}}(\mathbf{x}) = \mathbb{E}_{\mathbf{y}\sim p(\mathbf{y})}\left[\frac{p(\mathbf{y} \mid \mathbf{x}) p_{\theta_\alpha}(\mathbf{x})}{p_{\theta_\alpha}(\mathbf{y})}\right] = p_{\theta_\alpha}(\mathbf{x}) \int p(\mathbf{y} \mid \mathbf{x}) \frac{p(\mathbf{y})}{p_{\theta_\alpha}(\mathbf{y})}\, d\mathbf{y} = p_{\theta_\alpha}(\mathbf{x}) \tag{7}$$

Because if $p_{\theta_\alpha}(\mathbf{x})$ has converged to the stationary distribution $p_{\hat{\theta}}(\mathbf{x})$, then

$$p_{\theta_\alpha}(\mathbf{y}) = \int d\mathbf{x}\, p(\mathbf{y} \mid \mathbf{x}) p_{\theta_\alpha}(\mathbf{x}) = \int d\mathbf{x}\, p(\mathbf{y} \mid \mathbf{x}) p_{\hat{\theta}}(\mathbf{x}) = p_{\hat{\theta}}(\mathbf{y})\,. \tag{8}$$

However, in practice, we have finite samples, and we only approximate $\mathbb{E}_{y\sim p_{\theta^\star}(\mathbf{y})}[p_{\theta_\alpha}(\mathbf{x} \mid \mathbf{y})]$ at each iteration after training. For convergence to a distribution, we would need to have $p_{\theta_\alpha}(\mathbf{y}) \approx p_{\theta^\star}(\mathbf{y})$. We also note that the prior distribution found after iterating does not necessarily converge to the underlying prior $p_{\theta^*}(\mathbf{x})$, but it is such that it has equal marginal likelihood, and therefore equally explains the observed data, making it a plausible hyper distribution.

F Detailed proof of ascent property

There exist a long literature on the convergence properties of the generalized expectation maximization algorithm [41, 61, 43]. We wish to show that the procedure outlined in Algorithm 1 incrementally increases the log-likelihood of observations, that is, leads to successive priors models that allow us to incrementally increase the expected log-evidence of data for our fixed physical and noise models. More specifically,

$$\int d\mathbf{y} p(\mathbf{y}) \log p_{\theta_{\alpha+1}}(\mathbf{y}) \geq \int d\mathbf{y} p(\mathbf{y}) \log p_{\theta_\alpha}(\mathbf{y})\,, \tag{9}$$

where $p_{\theta_\alpha}(\mathbf{y}) = \int d\mathbf{x} p(\mathbf{y} \mid \mathbf{x}) p_{\theta_\alpha}(\mathbf{x})$ for every α, and $p(\mathbf{y})$ is the true observation distribution under the true prior. We have:

$$\int d\mathbf{y} p(\mathbf{y}) \left[\log p_{\theta_{\alpha+1}}(\mathbf{y}) - \log p_{\theta_\alpha}(\mathbf{y})\right] \tag{10}$$

$$= \int d\mathbf{y} p(\mathbf{y}) \log \left[\frac{p_{\theta_{\alpha+1}}(\mathbf{y})}{p_{\theta_\alpha}(\mathbf{y})}\right] \tag{11}$$

$$= \int d\mathbf{y} p(\mathbf{y}) \log \left[\frac{\int d\mathbf{x} p_{\theta_{\alpha+1}}(\mathbf{x}) p(\mathbf{y}|\mathbf{x})}{p_{\theta_\alpha}(\mathbf{y})}\right] \tag{12}$$

$$= \int d\mathbf{y} p(\mathbf{y}) \log \left[\int d\mathbf{x} p_{\theta_\alpha}(\mathbf{x}|\mathbf{y}) \frac{p_{\theta_{\alpha+1}}(\mathbf{x})}{p_{\theta_\alpha}(\mathbf{x})}\right] \tag{13}$$

$$\geq \int\int d\mathbf{y} p(\mathbf{y}) d\mathbf{x} p_{\theta_\alpha}(\mathbf{x} \mid \mathbf{y}) \left[\log p_{\theta_{\alpha+1}}(\mathbf{x}) - \log p_{\theta_\alpha}(\mathbf{x})\right] \tag{14}$$

Here, to go from line (12) to line (13), we have multiplied by $1 = p_\alpha(\mathbf{x}|\mathbf{y})/p_\alpha(\mathbf{x}|\mathbf{y})$ inside the $\mathbf{x}$ integral and used that the likelihood $p(\mathbf{y}|\mathbf{x})$ is the same for all α's, and have used that line (14) follows from Jensen's inequality. Now, since the way we define $\theta_{\alpha+1}$ in our iterative update is by finding the values of θ that maximize the first terms in Eqn. (14) (see Appendix E), and that, at worse, we could have $\theta_{\alpha+1} = \theta_\alpha$, we conclude that Eqn. (9) follows.

G Posterior sampling SDE solver and coverage test

Across all experiments, the posterior sampling procedure $\mathcal{F}$ is a Predictor-Corrector SDE solver, with a different number of steps. When doing posterior sampling, we use the convolved likelihood approximation. Since approximations are involved, and the problem is discretized with a certain number of steps, it is important to test the correctness of $\mathcal{F}$.

We choose to use TARP [62], a sample-based method to estimate coverage probabilities of generative posterior estimators. It has been shown that passing this test is a necessary and sufficient condition for the accuracy of $\mathcal{F}$. We perform several tests for experiments with both MNIST and galaxies. These tests can be conducted either using the test set of the dataset used to train the true distribution $p_{\theta^\star}$, or with samples from the true distribution since it is defined to be the target.

For the experiment with galaxies, when using a Predictor-Corrector solver with $1\,024$ steps (one corrector step per predictor step) as $\mathcal{F}$, posterior sampling is exact, as shown in Figure 7. This uses the correct prior (SBM) and observational noise of $\sigma_{\boldsymbol{\eta}} = 1$. However, increasing the level of observational noise to $\sigma_{\boldsymbol{\eta}} = 3$ makes the procedure biased, indicating the limit of the approximations used for $\mathcal{F}$. Furthermore, when using the test set instead of samples from the prior, again with $\sigma_{\boldsymbol{\eta}} = 3$, we obtain more biased results. This could be because the SBM is not in distribution with the test set, possibly due to model capacity or imperfect learning.

When using $\mathcal{F}$ with a misspecified prior, the test shows an important bias, as it is sensitive to the correct prior. Nonetheless, we use $\mathcal{F}$ as is for the updates. To run TARP, we simulate 256 observations and obtain 256 posterior samples from each one.

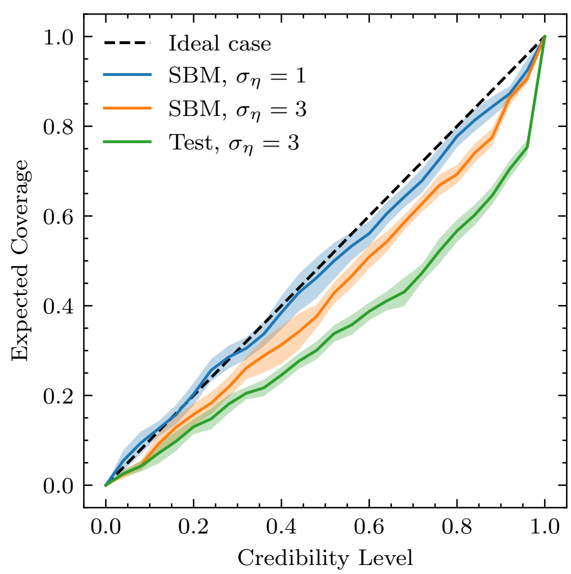

Figure 7: Coverage test for the posterior sampling procedure $\mathcal{F}$ using TARP [62] for the galaxy experiment. Credibility levels are plotted against expected coverage, comparing results from mock observations created with samples from the Score-Based Model (SBM) prior under different noise levels: $\sigma_{\boldsymbol{\eta}} = 1$ (blue) and $\sigma_{\boldsymbol{\eta}} = 3$ (orange), along with mock observations from the test set with noise $\sigma_{\boldsymbol{\eta}} = 3$ (green). The dashed line represents the ideal case where expected coverage matches credibility levels perfectly. Posterior sampling is exact when we have the correct prior (here enforced by creating mock observations with samples from the prior), and with a specific noise level. Results vary with different noise levels or a misspecified prior, indicating $\mathcal{F}$'s sensitivity to these factors.