


Galaxy Dust Maps with Conditional Score Models

Jared C. Siegel 
Department of Astrophysical Sciences
Princeton University
siegeljc@princeton.edu

Peter Melchior 
Department of Astrophysical Sciences
Center for Statistics & Machine Learning
Princeton University
peter.melchior@princeton.edu

Abstract

A notorious problem in observational astrophysics is the recovery of the true shape and spectral energy distribution (SED) of a galaxy despite absorption by interstellar dust embedded in the same galaxy. It has been solved only for a few hundred nearby galaxies with exquisite data coverage, but these techniques will not be applicable to the billions of galaxies in upcoming large wide-field surveys like LSST and Euclid. We present a method to infer the SEDs and spatial distribution of both the galaxy and its interstellar dust from multi-band imaging. To stabilize this massively underconstrained inverse problem, we utilize two score-matching models as data-driven priors: the first informs our inference of the galaxy’s underlying shape, the second informs the galaxy’s dust morphology conditioned on the current estimate of the galaxy’s shape. We believe that this is the first time a set of coupled score-matching models have been utilized to solve a complex data-analysis challenge. We demonstrate with realistic simulations that we can accurately measure the parameters of the underlying host and its dust content. In addition to providing galaxy SEDs unbiased by dust attenuation for subsequent analyses, such as photometric redshifts, our dust maps will allow the study of the interplay between star formation and dust production and destruction.

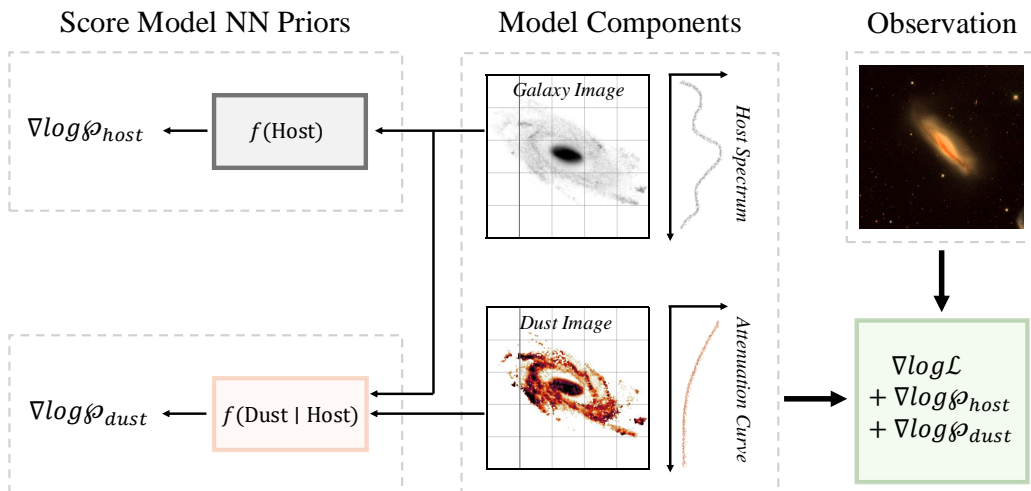


Figure 1: Our method for modeling a dust-attenuated galaxy. Score-matching neural network priors inform our inference of the galaxy and dust attenuation morphologies. The dust prior is conditional on the host galaxy morphology.

1 Introduction

In astronomy, inverse problems are often heavily under-constrained, because it is difficult or even impossible to gather enough observational data. As a result, multiple degenerate solutions cannot be distinguished with the likelihood alone, necessitating additional regularization. In the last few years, neural networks have provided the most successful avenue to stabilize such inverse problems, from super-resolution and inpainting to 3D depth reconstruction and source separation. Often, these methods internally form an implicit prior over the data space, learned from the available training data. Alternatively, one can use a neural network to directly represent an explicit prior distribution [e.g. van den Oord et al., 2016]. Particularly effective for higher-dimensional parameters are score-matching models that underpin diffusion models [Song and Ermon, 2019, 2020]. Given a dataset $\{\mathbf{x}_n\}$ of independent samples from a data distribution $p(\mathbf{x})$, a network $s_\theta(\mathbf{x})$ with parameters θ is trained to approximate the score function $\nabla \log p(\mathbf{x})$.

The score can be leveraged in an optimization or gradient-based sampling scheme as a regularizer to yield the Bayesian log-posterior

$$\log p(\mathbf{x} | \mathcal{D}) = \log \mathcal{L}(\mathcal{D} | \mathbf{x}) + \log p(\mathbf{x}). \quad (1)$$

If the solution has multiple components, $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_K)$, that are assumed to be independent from each other, the prior term factorizes accordingly. However, this is not necessarily true, and this assumption weakens the inference if the conditional dependence of one component on another is strong. In such a case, it would be easier to learn the conditional than the unconditional prior, and one should choose a factorization of the joint distribution that exploits dependencies, representing the log-prior as

$$\log p(\mathbf{x}) = \log p(\mathbf{x}_1) + \sum_{k=1}^K \log p(\mathbf{x}_{k+1} | \mathbf{x}_1, \dots, \mathbf{x}_k), \quad (2)$$

where the components of \mathbf{x} are ordered, e.g., by decreasing mutual independence. Regularizing such problems requires several score-matching models, including conditional ones. If we limit the dependencies to two components, we get gradients of the log-posterior of

$$\nabla \log p(\mathbf{x} | \mathcal{D}) = \nabla \log \mathcal{L}(\mathcal{D} | \mathbf{x}) + s(\mathbf{x}_1) + s(\mathbf{x}_2 | \mathbf{x}_1). \quad (3)$$

Below we will show that this problem arises in modeling galaxies with dust attenuation and that Equation 3 yields stable and accurate solutions of the multi-component inverse problem. Utilizing multiple score models has been applied to conditional generation of images [Nair et al., 2022] and for astronomical inverse problems to model signal and noise separately [Adam et al., 2023], but to the best of our knowledge this is the first time score-matching models are conditionally coupled to represent dependencies between different components of the solution.

2 Galaxy and Dust Model

A key problem in observational astronomy is determining a galaxy's shape and SED, the emitted energy as a function of wavelength, from observations taken in multiple filters. These measurements encode the formation history and composition of a galaxy. However, the light emitted by the stars that make up the galaxy can be scattered or absorbed by dust grains before reaching the observer. The effective loss of photons is called dust attenuation. Not accounting for dust attenuation will lead to biases in inferred galaxy properties and the studies that make use of them.

At a given wavelength, dust attenuation reduces the emitted intensity of light $F_{\text{host}}(\lambda)$ to the observed intensity $F(\lambda)$, which is commonly expressed as

$$F(\lambda) = F_{\text{host}}(\lambda) 10^{-0.4A(\lambda)}, \quad (4)$$

where $A(\lambda)$ is the wavelength dependence of dust attenuation. We adopt the empirical relation of Calzetti et al. [2000] with modifications by Noll et al. [2009] and Kriek and Conroy [2013],

$$A(\lambda) = A_0 \frac{2.659 \left(\frac{1.509}{\lambda} - \frac{0.198}{\lambda^2} + \frac{0.011}{\lambda^3} - 2.156 \right) + 4.05}{4.05} \left(\frac{\lambda}{\lambda_0} \right)^\delta \equiv A_0 k(\lambda, \delta), \quad (5)$$

where A_0 sets the attenuation amplitude at a reference wavelength λ_0 , and $k(\lambda, \delta)$ is a polynomial with a single free parameter δ . Throughout this work, we set $\lambda_0 = 0.4754$ microns.

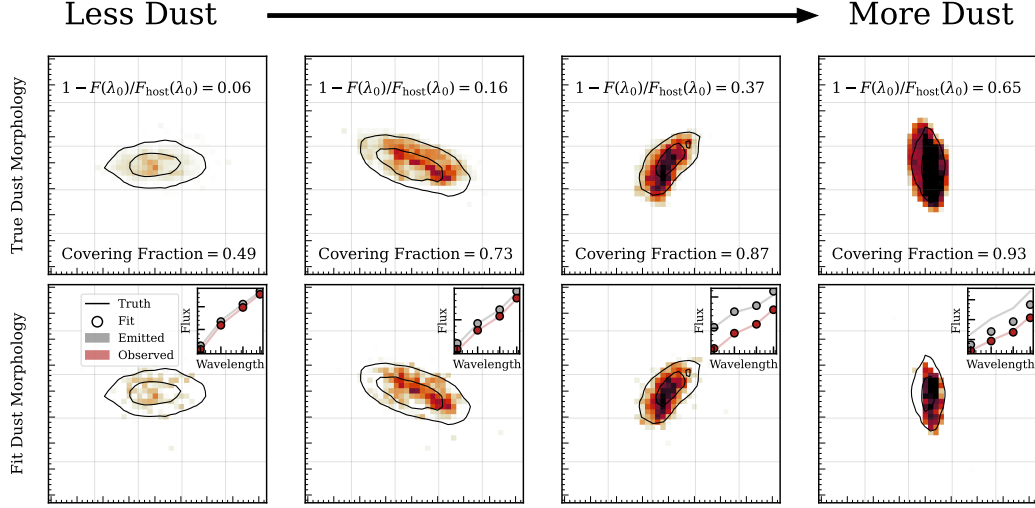


Figure 2: Ground truth (top row) and inferred (bottom row) dust morphologies for four representative simulated galaxies. The color scale covers $\mathbf{D} \in [0, 0.75]$. The contours present the 50 and 84-th percentiles of the unattenuated galaxy light. The inset panels show the spectrum of the galaxy before (grey) and after (red) dust attenuation.

For observed multi-band images, the spatial and spectral variation of the galaxy are integrated and discretized. If we stack these images and assume that spatial and spectral variation are independent, we can model the image cube as

$$\mathbf{Y} = \mathbf{F}^T \times \mathbf{S}, \quad (6)$$

where $\mathbf{F} = (\int d\lambda F(\lambda) c_i(\lambda), \dots)$ is the vectorized spectrum for a list of filter curves $c_i(\lambda)$, and \mathbf{S} is a monochrome image of the pixel-integrated spatial distribution, respectively. \mathbf{F} is one-dimensional, with a length equal to the number of observed filters N_{band} ; \mathbf{S} is two-dimensional (64×64); \mathbf{Y} is three-dimensional ($N_{\text{band}} \times 64 \times 64$). This parameterization has been found to be quite effective for galaxy images [Melchior et al., 2018], but it does not account for dust attenuation. With dust, Equation 6 needs to be modified according to Equation 4:

$$\mathbf{Y} = (\mathbf{F}^T \times \mathbf{S}) \odot 10^{-0.4\mathbf{A}}, \quad (7)$$

where \mathbf{A} denotes the multi-band image cube of dust attenuation. If we make the same assumption of spatial-spectral separability, we can express the attenuation cube as

$$\mathbf{A} = \mathbf{K}^T \times [-2.5 \log_{10}(\mathbf{D})], \quad (8)$$

where $\mathbf{K} = (k(\lambda_i, \delta), \dots)$ is the attenuation curve evaluated for filters i , and \mathbf{D} is the image of the spatial distribution of $\frac{F}{F_{\text{host}}}(\lambda_0)$.

Fitting Equation 7 to multi-band observations of a galaxy faces severe degeneracies. Because of the shape of the attenuation curve in Equation 4, the presence of dust is revealed as a reddening of the remaining light, but in areas without light, any amount of attenuation is consistent with the data; the likelihood alone cannot distinguish between physical and non-physical galaxy or dust morphologies. This inverse problem thus requires priors on at least \mathbf{S} and \mathbf{D} . Because the dust is produced by stars, we expect dust only in places with starlight, but it is entirely plausible to find starlight without dust. We therefore posit a dependency structure for the prior in the form of

$$p(\mathbf{S}, \mathbf{D}) = p(\mathbf{S}) p(\mathbf{D} | \mathbf{S}). \quad (9)$$

For the host morphologies \mathbf{S} , we adopt the publicly available score-matching network of Sampson et al. [2024]. That model was trained on 600,000 galaxy images from the Subaru Hyper-Suprime Cam catalogue [Bosch et al., 2018]. We build the dust morphology prior as a score-matching model that is conditional on the host morphology: $s(\mathbf{D} | \mathbf{S})$. We adopt a U-Net architecture for the neural network [Ronneberger et al., 2015]. We train it with simulated galaxies from Faucher et al. [2023], which include hydrodynamical and the radiative transfer physics. Unlike observed objects, these

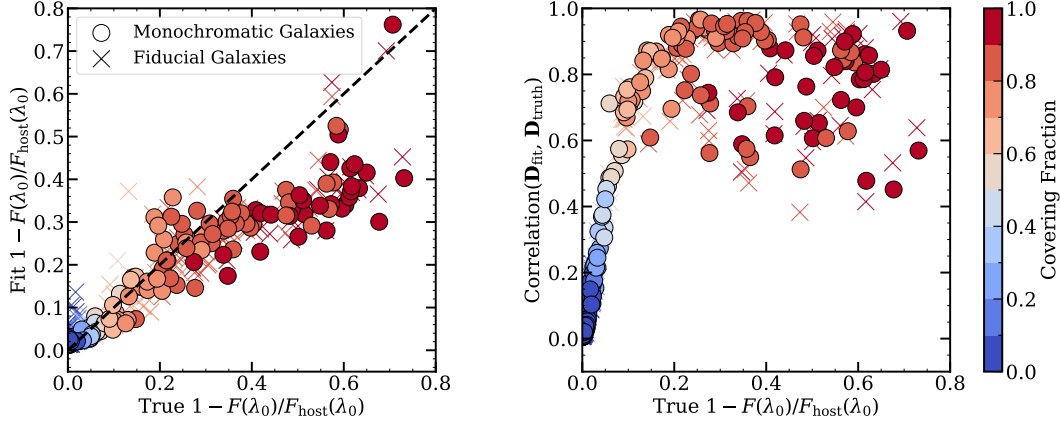


Figure 3: Reconstructed attenuation strength (left) and correlation of the inferred dust map to the true dust distribution (right) for 65 simulated galaxies with 3 orientation angles each. Monochromatic host galaxies are indicated by circles, fiducial hosts with intrinsic color variations by crosses.

simulations allow us to calculate the galaxies’ ground-truth host and dust attenuation morphologies. Multi-band image cubes are publicly available for 65 galaxies, with 10 viewing angles for each. Such a low number of training samples is only acceptable because the dependency structure, i.e. the dust distribution given the light distribution, is highly informative.

3 Results

We run gradient-descent with the Adam optimizer [Kingma and Ba, 2015] to determine the MAP estimates of \mathbf{F} , \mathbf{S} , \mathbf{D} , δ ; optimization is terminated based on iteration-to-iteration convergence of the host spectrum parameter. Our method accurately models dusty galaxies for a wide range of attenuation levels. In Figure 2, we visualize our fits to four representative simulated galaxies from Faucher et al. [2023]. The galaxies span a wide range of attenuation levels and covering fractions (the fraction of the galaxy light covered by non-zero dust attenuation). For galaxies with covering fractions $< 90\%$, we accurately recover the host galaxy’s emitted spectrum, as well as the amplitude and morphology of dust attenuation. As the covering fraction approaches unity (i.e., the entire galaxy is covered by dust), the results deteriorate because there is a fundamental degeneracy between the host’s emitted spectrum and the level of dust attenuation; if everything gets attenuated, one can compensate for more attenuation with a bluer host spectrum \mathbf{F} . We apply our method to the entire suite of simulated galaxies from Faucher et al. [2023]; the results are shown in Figure 3. Our reconstructions are accurate at low to intermediate dust attenuation levels. In addition to the failure mode at high covering fractions described above, the spatial distribution of low levels of dust is also hard to infer because they suffer from very weak evidence (seen in the drop of the correlation of the inferred to the true dust maps in the right panel).

As an additional complication, galaxies also exhibit color variations in absence of dust, most notably from differences in stellar age across the galaxy. This can result in erroneously inferring dust attenuation where none takes place. To determine the impacts of that complication, we create “monochromatic” galaxies by removing non-attenuation color gradients from the host data cubes. The results are shown as circles in Figure 3. Compared to the fiducial galaxy images, we see a much tighter relation in very low attenuation cases, indicating that color variations for them stem primarily from the stellar populations and not from dust.

4 Conclusions

We apply data-driven priors to a complex inverse problem: inferring the spectra and morphologies of galaxies as well as the spectra and morphologies of the interstellar dust embedded in those galaxies. We solve this highly degenerate problem with a combination of score-matching neural network priors, one of which is conditional to reflect the causal dependency structure of the problem. This choice

allows us to operate with a very small training sample for the conditional prior. From testing with simulated galaxies, our model successfully measured the level and morphology of dust attenuation and produces accurate estimates of the galaxies' unattenuated properties.

References

- 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. *arXiv [astro-ph.IM]*, November 2023. URL <http://arxiv.org/abs/2311.18002>.
- James Bosch, Robert Armstrong, Steven Bickerton, Hisanori Furusawa, Hiroyuki Ikeda, Michitaro Koike, Robert Lupton, Sogo Mineo, Paul Price, Tadafumi Takata, Masayuki Tanaka, Naoki Yasuda, Yusra AlSayyad, Andrew C. Becker, William Coulton, Jean Coupon, Jose Garmilla, Song Huang, K. Simon Krughoff, Dustin Lang, Alexie Leauthaud, Kian-Tat Lim, Nate B. Lust, Lauren A. MacArthur, Rachel Mandelbaum, Hironao Miyatake, Satoshi Miyazaki, Ryoma Murata, Surhud More, Yuki Okura, Russell Owen, John D. Swinbank, Michael A. Strauss, Yoshihiko Yamada, and Hitomi Yamanoi. The Hyper Suprime-Cam software pipeline. , 70:S5, January 2018. doi: 10.1093/pasj/psx080.
- Daniela Calzetti, Lee Armus, Ralph C. Bohlin, Anne L. Kinney, Jan Koornneef, and Thaisa Storchi-Bergmann. The Dust Content and Opacity of Actively Star-forming Galaxies. , 533(2):682–695, April 2000. doi: 10.1086/308692.
- Nicholas Faucher, Michael R. Blanton, and Andrea V. Macciò. Panchromatic Simulated Galaxy Observations from the NIHAO Project. , 957(1):7, November 2023. doi: 10.3847/1538-4357/acf9f0.
- Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015. URL <http://arxiv.org/abs/1412.6980>.
- Mariska Kriek and Charlie Conroy. THE DUST ATTENUATION LAW IN DISTANT GALAXIES: EVIDENCE FOR VARIATION WITH SPECTRAL TYPE. *The Astrophysical Journal Letters*, 775(1):L16, September 2013. ISSN 2041-8205. doi: 10.1088/2041-8205/775/1/L16. URL <https://iopscience.iop.org/article/10.1088/2041-8205/775/1/L16/meta>.
- P Melchior, F Moolekamp, M Jerdee, R Armstrong, A-L Sun, J Bosch, and R Lupton. scarlet: Source separation in multi-band images by constrained matrix factorization. *Astronomy and Computing*, 24:129–142, July 2018. ISSN 2213-1337. doi: 10.1016/j.ascom.2018.07.001. URL <http://www.sciencedirect.com/science/article/pii/S2213133718300301>.
- Nithin Gopalakrishnan Nair, Wele Gedara Chaminda Bandara, and Vishal M Patel. Image generation with multimodal priors using denoising diffusion probabilistic models. *arXiv [cs.CV]*, June 2022. URL <http://arxiv.org/abs/2206.05039>.
- S Noll, D Pierini, A Cimatti, E Daddi, J D Kurk, M Bolzonella, P Cassata, C Halliday, M Mignoli, L Pozzetti, A Renzini, S Berta, M Dickinson, A Franceschini, G Rodighiero, P Rosati, and G Zamorani. GMASS ultradeep spectroscopy of galaxies at $z \sim 2$: IV. the variety of dust populations. *Astronomy and astrophysics*, 499(1):69–85, May 2009. ISSN 0004-6361,1432-0746. doi: 10.1051/0004-6361/200811526. URL <https://www.aanda.org/articles/aa/abs/2009/19/aa11526-08/aa11526-08.html>.
- Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv e-prints*, art. arXiv:1505.04597, May 2015. doi: 10.48550/arXiv.1505.04597.
- Matt L. Sampson, Peter Melchior, Charlotte Ward, and Sufia Birmingham. Score-matching neural networks for improved multi-band source separation. *arXiv e-prints*, art. arXiv:2401.07313, January 2024. doi: 10.48550/arXiv.2401.07313.

- Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper_files/paper/2019/file/3001ef257407d5a371a96dcd947c7d93-Paper.pdf.
- 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. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/92c3b916311a5517d9290576e3ea37ad-Paper.pdf.
- Aaron van den Oord, Nal Kalchbrenner, Oriol Vinyals, Lasse Espeholt, Alex Graves, and Koray Kavukcuoglu. Conditional image generation with PixelCNN decoders. *arXiv [cs.CV]*, June 2016. URL <http://arxiv.org/abs/1606.05328>.