Accelerated Galaxy SED Modeling using Amortized Neural Posterior Estimation

ChangHoon Hahn [1] **Peter Melchior** [1]

Abstract

State-of-the-art spectral energy distribution (SED) analyses use Bayesian inference to derive physical properties of galaxies from observed photometry or spectra. They require sampling from a high-dimensional space of model parameters and take $> 10 - 100$ CPU hours per galaxy. This renders them practically infeasible for analyzing the *billions* of galaxies that will be observed by upcoming galaxy surveys (*e.g.* DESI, PFS, Rubin, Webb, and Roman). In this work, we present an alternative approach using Amortized Neural Posterior Estimation (ANPE). ANPE is a likelihood-free inference method that employs neural networks to estimate the posterior over the full range of observations. Once trained, it requires no additional model evaluations to estimate the posterior. We present SEDFLOW, an ANPE method to produce posteriors of the recent (Hahn et al., 2022) SED model from optical photometry and spectra. SEDFLOW takes $\sim$1 *second per galaxy* to obtain the posterior distributions of 12 model parameters, all of which are in excellent agreement with traditional Markov Chain Monte Carlo sampling results.

1. Introduction

Physical properties of galaxies are the building blocks of our understanding of galaxies and their evolution. We can determine properties such as stellar mass, star formation rate, metallicity, and age of a galaxy by analyzing its spectral energy distribution (SED). Theoretical modeling of galaxy SEDs is currently based on stellar population synthesis (SPS) and describes the SED as a composite stellar population constructed from isochrones, stellar spectra, an initial mass function, a star formation and chemical evolution history, and dust attenuation (see Walcher et al., 2011; Conroy, 2013, for a review). In state-of-the-art SED modeling, theoretical SPS models are compared to observed SEDs using Bayesian inference (Acquaviva et al., 2011; Chevallard & Charlot, 2016; Leja et al., 2017; Carnall et al., 2018; Johnson et al., 2021; Hahn et al., 2022).

However, current Bayesian SED modeling methods, which use Markov Chain Monte Carlo (MCMC) sampling techniques, take $10 - 100$ CPU hours per galaxy (*e.g.* Carnall et al., 2019; Tacchella et al., 2021). While this is merely very expensive with current data sets of hundreds of thousands of galaxy SEDs, observed by current surveys (*e.g.*, SDSS, DEEP2, GAMA), it is prohibitive for the next generation of surveys. Over the next decade, surveys with DESI, PFS, Rubin, JWST, and Roman will observe *billions* of galaxy SEDs. The task of SED modeling alone for these surveys would amount to tens or hundreds of billions of CPU hours.

But Bayesian inference does not require MCMC sampling. Likelihood-free inference (LFI) is a rapidly developing class of inference methods that offers alternatives for many applications (Papamakarios et al., 2017; Alsing et al., 2019; Cranmer et al., 2020; Huppenkothen & Bachetti, 2021; Zhang et al., 2021). Of particular interest for SED modeling is a technique called Amortized Neural Posterior Estimation (ANPE). Instead of using MCMC to sample the posterior for every single galaxy separately, ANPE uses neural density estimators (NDE) to build a model of the posterior for *all* observable galaxies. Once the NDE is trained, generating the posterior requires only the observed SED and no additional model evaluations. In this work, we present SEDFLOW, a method that applies ANPE to Bayesian galaxy SED modeling using the recent (Hahn et al., 2022) SED model. We demonstrate that we can derive accurate posteriors with SEDFLOW and make Bayesian SED modeling fully scalable for the billions of galaxies that will be observed by upcoming surveys.

2. Amortized Neural Posterior Estimation

The goal of Bayesian SED modeling is to infer the posterior probability distributions $p(\theta \,|\, x)$ of galaxy properties, θ, given observations, x. For a specific θ and x, we typically evaluate the posterior using Bayes' rule, $p(\theta \,|\, x) \propto p(\theta)\, p(x \,|\, \theta)$, where $p(\theta)$ denotes the prior distribution and $p(x \,|\, \theta)$ the likelihood, which is typically assumed to have a

ICML 2022 Workshop on Machine Learning for Astrophysics, Baltimore, Maryland, USA, 2022.

Gaussian functional form.

Likelihood-free inference (LFI; also known as "simulation-based" inference) offers an alternative that requires no assumptions on the form of the likelihood. Instead, LFI uses a generative forward model, *i.e.* a simulation, to generate mock observations x' given parameters θ'. It uses simulated pairs (θ', x') to directly estimate either the posterior, the likelihood, or $p(\theta, x)$. Hence, it provides an inference framework that can exploit high-fidelity simulations. LFI has already been successfully applied to a wide range of inference problems in astronomy (Cameron & Pettitt, 2012; Weyant et al., 2013; Hahn et al., 2017; Huppenkothen & Bachetti, 2021; Zhang et al., 2021).

In this work, we utilize neural density estimation based LFI, where a neural network with parameters ϕ is trained using the simulated pairs (θ', x') to estimate the density $p_\phi(\theta \,|\, x')$. In particular, SEDFLOW uses normalizing flows as density estimators (Tabak & Vanden-Eijnden, 2010; Tabak & Turner, 2013).

2.1. SEDflow

A key ingredient in LFI is a forward model that is capable of generating mock observations that are practically indistinguishable from observations. In our case, our forward model must be able to generate realistic galaxy photometry and spectroscopy. To do this, we use the state-of-the-art PROVABGS SED model from (Hahn et al., 2022) and realistic noise models for photometry and spectroscopy.

In PROVABGS, the SED of a galaxy is modeled as a composite of stellar populations defined by stellar evolution theory (in the form of isochrones, stellar spectral libraries, and an initial mass function) and its star formation and chemical enrichment histories (SFH and ZH), attenuated by dust (see Walcher et al., 2011; Conroy, 2013, for a review). The PROVABGS model, in particular, utilizes a non-parametric SFH with a starburst, a non-parametric ZH that varies with time, and a flexible dust attenuation prescription.

Using the PROVABGS SED model, we can forward model the training data for SEDFLOW. First, we sample $N_{\rm train}$ model parameters from a prior: $\theta' \sim p(\theta)$. We use the same priors as (Hahn et al., 2022). For each θ', we also uniformly sample a redshift $z' \sim \mathcal{U}(0., 0.2)$. Next, we forward model mock observables. We calculate the rest-frame galaxy SED from PROVABGS and redshift it: $F(\lambda; \theta', z)$.

For photometry, we convolve F with optical broadband filters, R_X, to generate noiseless photometric fluxes, $f_X(\theta', z')$. Then, we assign photometric uncertainties, σ'_X, by sampling an empirical estimate of the observed $p(\sigma_X | f_X)$ of NSA galaxies. Afterwards, we apply Gaussian noise $\hat{f}_X(\theta', z', \sigma'_x) = f_X(\theta', z') + n_X$ where $n_X \sim \mathcal{N}(0, \sigma'_X)$ to derive the forward modeled photometric flux.

To forward model spectra, we apply a noise model to the noiseless SEDs, F_λ. To assign uncertainties, σ'_λ, we first sample the distribution $p(h_i | F_\lambda)$, where h_i represents the compressed inverse variance of spectra. The compression is performed using a variational autoencoder (VAE) similar to the one used in (Portillo et al., 2020) and trained on SDSS spectral noise measurements. $p(h_i | F_\lambda)$ is also estimated from SDSS observations using normalizing flows. After h'_i is assigned, we use the decoder of our inverse variance VAE to obtain inverse variance as a function wavelength and then σ'_λ. Lastly, we apply σ'_λ to F_λ as Gaussian noise to get realistic SDSS-like spectra: f_λ.

f_λ has ~3,600 spectral elements. Applying ANPE directly to f_λ would require estimating a $\gtrsim$3,600 dimensional distribution with an NDE. To avoid this complication, we use another VAE, trained on observed SDSS spectra, to compress f_λ to a 10 dimensional latent variable space g. We apply SEDFLOW to this compressed spectra.

2.2. Training SEDflow

Among the various normalizing flow density estimators now available in the literature, we use a Masked Autoregressive Flow (MAF; Papamakarios et al., 2017). MAF combines normalizing flows with an autoregressive design (Uria et al., 2016), which is well-suited for estimating conditional probability distributions such as a posterior. A MAF model is built by stacking multiple Masked Autoencoder for Distribution Estimation (MADE; Germain et al., 2015) models so that it has the autoregressive structure of MADE but with more flexibility to describe complex probability distributions. We use the MAF implementation from (Greenberg et al., 2019; Tejero-Cantero et al., 2020).

In training, our goal is to determine the parameters, ϕ, of our MAF model, $p_\phi(\theta \,|\, x)$, so that it accurately estimates the posterior, $p(\theta \,|\, x)$. We do this by minimizing the KL divergence between $p_\phi(\theta \,|\, x)$ and $p(\theta \,|\, x)$. We split the training data into a training and validation set with a 90/10 split, then maximize the total log likelihood $\sum_i \log p_\phi(\theta_i | x_i)$ over training set, which is equivalent to minimizing the KL divergence. We use the ADAM optimizer (Kingma & Ba, 2017) with a learning rate of 5×10^{-4}. We prevent overfitting by stopping the training when the validation log likelihood fails to increase after 20 epochs. We determine the architecture of our MAF model through experimentation. Our final trained MAF model has 15 MADE blocks, each with 2 hidden layers and 500 hidden units.

3. Results

Now that we have trained SEDFLOW, we can estimate the posterior, $p(\theta \,|\, x_i)$, for any $x_i = \{f_{X,i}, \sigma_{X,i}, z_i\}$. In practice, we do this by drawing samples from the SED-

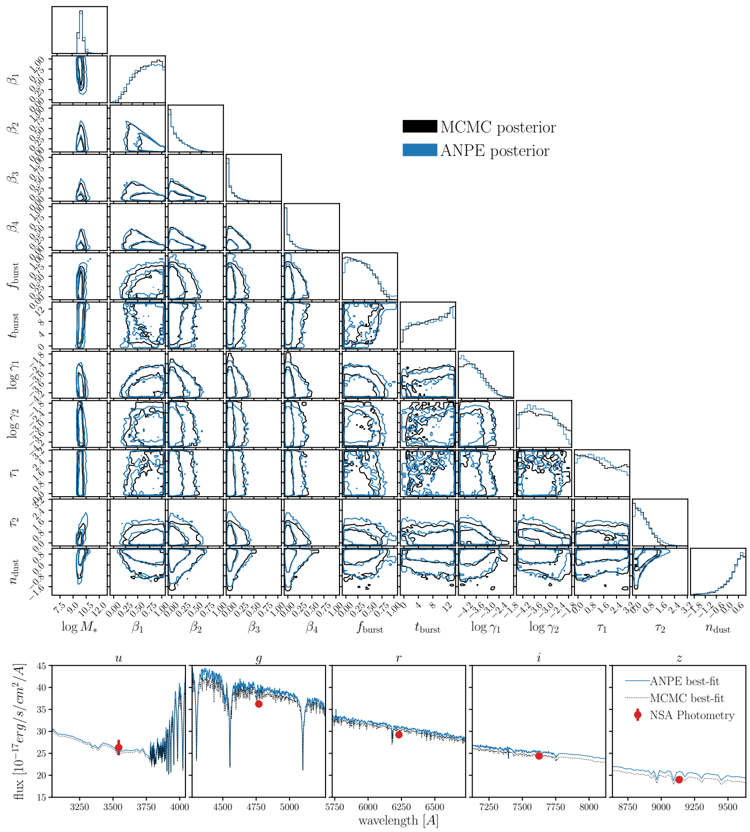

Figure 1. Posteriors of the 12 SED model parameters derived from standard MCMC sampling (black) and our SEDFLOW (blue) for an arbitrarily selected NSA galaxy. The posteriors are in excellent agreement for all of the SED parameters. In the bottom panel, we present the SEDs of the the best-fit parameter values from the SEDFLOW (blue) and MCMC posteriors (black dotted), which are both in good agreement with the observed NSA photometric flux (red). Estimating the posterior using MCMC sampling requires ∼10 CPU hours. Even using neural emulators to accelerate likelihood evaluations, MCMC sampling requires ∼10 CPU minutes. *With SEDFLOW, inferring the full posterior takes 1 second per galaxy.*

FLOW NDE model. Since we use a normalizing flow, this is trivial: we generate samples from the target distribution of the normalizing flow, a multivariate Gaussian distribution in our case, then we transform the samples using the bijective transformation that we trained. The transformed samples follow $p_\phi(\theta \,|\, x_i)$ and estimate the posterior, $p(\theta \,|\, x_i)$.

Next, we validate the accuracy of the SEDFLOW posterior estimates. As a first test, in Fig. 1, we compare the posterior from SEDFLOW to the posterior derived from MCMC sampling for a single arbitrarily chosen galaxy from the NASA-Sloan Atlas[1] (NSA). In the top, we present the the posterior distribution of the 12 SED model parameters for the SEDFLOW posterior (blue) and MCMC posterior (black). The SEDFLOW posterior is in excellent agreement with the MCMC posterior for all of the SED parameters.

In the bottom of Fig. 1, we compare the SEDs of the best-fit parameter values from the SEDFLOW (blue) and MCMC posteriors (black dotted). We also include the NSA photometric flux of the selected galaxy (red). The best-fit SED from the SEDFLOW posterior is in good agreement with both the MCMC best-fit SED and the NSA photometry.

The key advantage of ANPE is that it enables accurate Bayesian inference orders of magnitude faster than conventional methods. We derive the MCMC posterior using the (Karamanis & Beutler, 2020) ensemble slice-sampler with 30 walkers and 10,000 iterations. In total, the MCMC posterior requires >100,000 SED model evaluations — *i.e.* ∼10 CPU hours. Meanwhile, after training, he SEDFLOW posterior takes 1 second — > $10^4\times$ faster than MCMC.

The posteriors from SEDFLOW and MCMC are overall in excellent for NSA galaxies, besides the one in Figure 1. However, we do not know the true SED parameters for these galaxies so to further validate SEDFLOW, we use test

[1] http://nsatlas.org/

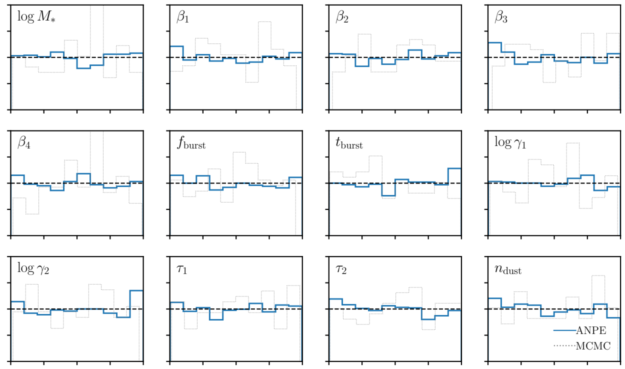

Figure 2. Simulation-based calibration of the SEDFLOW posteriors for 1000 synthetic test observations. The histogram in each panel represents the distribution of the rank statistic of the true value within the marginalized SEDFLOW posterior (blue) for each SED parameter. For the true posterior, the rank statistics will have a uniform distribution (black dashed). The rank statistic distribution of SEDFLOW is nearly uniform for all of the SED parameters. SEDFLOW provides unbiased and accurate estimates of the true posteriors.

synthetic photometry, where we know the truth. We sample 1000 SED parameters from the prior, $\{\theta_i^{\rm test}\} \sim p(\theta)$, and forward model synthetic NSA observations, $\{x_i^{\rm test}\}$, for them in the same way as the training data. Afterwards, we generate posteriors for each of $\boldsymbol{x}_i^{\rm test}$ using SEDFLOW.

In Fig. 2, we present simulation-based calibration (SBC; Talts et al., 2020) of each SED parameter for the SEDFLOW posteriors (blue) using the 1000 test observations. SBC examines the distribution of the rank statistics of the true parameter values within the marginalized posteriors. For true posteriors, the rank statistics are uniformly distributed (black dashed). For comparison, we include the SBC for the MCMC posteriors (gray dotted). The rank statistic distribution for the SEDFLOW posteriors are nearly uniform for all SED parameters. Hence, we confirm that the SEDFLOW posteriors are in excellent agreement with the true posterior.

We repeat the validation tests above for SEDFLOW applied to spectra. For synthetic test observations, we accurately recover the true posterior based on SBC. Meanwhile, when we compare the SEDFLOW posterior to a posterior derived from full spectral fitting with MCMC, we find that the SEDFLOW posterior is consistent but significantly broader than the MCMC posterior. This is due to the fact that some constraining power in the spectra is erased in the spectral compression procedure. We are currently exploring more efficient compression schemes that will reduce this loss of information.

4. Conclusions

By analyzing the SED of a galaxy, we can infer detailed physical properties such as its stellar mass, star formation rate, metallicity, and dust content. These properties serve as the building blocks of our understanding of how galaxies form and evolve. State-of-the-art SED modeling methods use MCMC sampling to perform Bayesian statistical inference. For the dimensionality of current SED models, deriving a posterior requires $\gtrsim$100, 000 model evaluations and take $\gtrsim$10 $-$ 100 CPU hours per galaxy. Upcoming galaxy surveys, however, will observe *billions* of galaxies using *e.g.*DESI, PFS, Rubin, JWST, and Roman. Analyzing all of these galaxies with current Bayesian SED models is infeasible and would require hundreds of billions of CPU hours.

We demonstrate in this work that Amortized Neural Posterior Estimation (ANPE) provides an alternative *scalable* approach for Bayesian inference in SED modeling. We present SEDFLOW, a galaxy SED modeling method using ANPE and PROVABGS, a flexible SED model that uses a compact non-parametric SFH and ZH prescriptions (Hahn et al., 2022). SEDFLOW is based on a MAF normalizing flow and trained using a data set of $\sim$1 million SED model parameters and forward model synthetic SEDs. Once trained, deriving posteriors of galaxy properties for a galaxy takes $\sim$1 second, $10^5\times$ faster than traditional MCMC sampling. Furthermore, SEDFLOW posteriors show excellent agreement with MCMC posteriors and accurately estimate the true posterior, based on validation using synthetic observations. For SED modeling of spectra, SEDFLOW accurately estimate the true posterior but are broader than posteriors from full spectral fitting.

This work highlights the advantages of using an ANPE approach to Bayesian SED modeling. Our approach can easily be extended beyond this application. For instance, we can include multi-wavelength photometry at ultra-violet or infrared wavelengths. We can also modify SEDFLOW to infer redshift from photometry to infer more physically motivated photometric redshifts, where we marginalize over our understanding of galaxies rather than using templates.

References

Acquaviva, V., Gawiser, E., and Guaita, L. Spectral Energy Distribution Fitting with Markov Chain Monte Carlo: Methodology and Application to z = 3.1 Ly-emitting Galaxies. *The Astrophysical Journal*, 737:47, August 2011. ISSN 0004-637X. doi: 10.1088/0004-637X/737/2/47. URL https://ui.adsabs.harvard.edu/abs/2011ApJ...737...47A. ADS Bibcode: 2011ApJ...737...47A.

Alsing, J., Charnock, T., Feeney, S., and Wandelt, B. Fast likelihood-free cosmology with neural density estimators and active learning. *Monthly Notices of the Royal Astronomical Society*, 488:4440–4458, September 2019. ISSN 0035-8711. doi: 10.1093/mnras/stz1960. URL https://ui.adsabs.harvard.edu/abs/2019MNRAS.488.4440A. ADS Bibcode: 2019MNRAS.488.4440A.

Cameron, E. and Pettitt, A. N. Approximate Bayesian Computation for astronomical model analysis: a case study in galaxy demographics and morphological transformation at high redshift. *Monthly Notices of the Royal Astronomical Society*, 425:44–65, September 2012. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2012.21371.x. URL http://adsabs.harvard.edu/abs/2012MNRAS.425...44C.

Carnall, A. C., McLure, R. J., Dunlop, J. S., and Davé, R. Inferring the star formation histories of massive quiescent galaxies with BAGPIPES: evidence for multiple quenching mechanisms. *Monthly Notices of the Royal Astronomical Society*, 480:4379–4401, November 2018. ISSN 0035-8711. doi: 10.1093/mnras/sty2169. URL https://ui.adsabs.harvard.edu/abs/2018MNRAS.480.4379C. ADS Bibcode: 2018MNRAS.480.4379C.

Carnall, A. C., Leja, J., Johnson, B. D., McLure, R. J., Dunlop, J. S., and Conroy, C. How to Measure Galaxy Star Formation Histories. I. Parametric Models. *The Astrophysical Journal*, 873:44, March 2019. ISSN 0004-637X. doi: 10.3847/1538-4357/ab04a2. URL https://ui.adsabs.harvard.edu/abs/2019ApJ...873...44C. ADS Bibcode: 2019ApJ...873...44C.

Chevallard, J. and Charlot, S. Modelling and interpreting spectral energy distributions of galaxies with BEAGLE. *Monthly Notices of the Royal Astronomical Society*, 462(2):1415–1443, October 2016. ISSN 0035-8711, 1365-2966. doi: 10.1093/mnras/stw1756. URL http://arxiv.org/abs/1603.03037. arXiv: 1603.03037.

Conroy, C. Modeling the Panchromatic Spectral Energy Distributions of Galaxies. *Annual Review of Astronomy and Astrophysics*, 51:393–455, August 2013. ISSN 0066-4146. doi: 10.1146/annurev-astro-082812-141017. URL http://adsabs.harvard.edu/abs/2013ARA%26A..51..393C.

Cranmer, K., Brehmer, J., and Louppe, G. The frontier of simulation-based inference. *Proceedings of the National Academy of Sciences*, 117(48):30055–30062, 2020. ISSN 0027-8424. doi: 10.1073/pnas.1912789117. URL https://www.pnas.org/content/117/48/30055. Publisher: National Academy of Sciences tex.eprint: https://www.pnas.org/content/117/48/30055.full.pdf.

Germain, M., Gregor, K., Murray, I., and Larochelle, H. MADE: Masked Autoencoder for Distribution Estimation. *arXiv:1502.03509 [cs, stat]*, June 2015. URL http://arxiv.org/abs/1502.03509. arXiv: 1502.03509.

Greenberg, D. S., Nonnenmacher, M., and Macke, J. H. Automatic Posterior Transformation for Likelihood-Free Inference. Technical report, May 2019. URL https://ui.adsabs.harvard.edu/abs/2019arXiv190507488G. Publication Title: arXiv e-prints ADS Bibcode: 2019arXiv190507488G Type: article.

Hahn, C., Vakili, M., Walsh, K., Hearin, A. P., Hogg, D. W., and Campbell, D. Approximate Bayesian Computation in Large Scale Structure: constraining the galaxy-halo connection. *Monthly Notices of the Royal Astronomical Society*, 469(3):2791–2805, August 2017. ISSN 0035-8711, 1365-2966. doi: 10.1093/mnras/stx894. URL http://arxiv.org/abs/1607.01782. arXiv: 1607.01782.

Hahn, C., Kwon, K. J., Tojeiro, R., Siudek, M., Canning, R. E. A., Mezcua, M., Tinker, J. L., Brooks, D., Doel, P., Fanning, K., Gaztañaga, E., Kehoe, R., Landriau, M., Meisner, A., Moustakas, J., Poppett, C., Tarle, G., Weiner, B., and Zou, H. The DESI PRObabilistic Value-Added Bright Galaxy Survey (PROVABGS) Mock Challenge. Technical report, February 2022. URL https://ui.adsabs.harvard.edu/abs/2022arXiv220201809H. Publication Title: arXiv e-prints ADS Bibcode: 2022arXiv220201809H Type: article.

Huppenkothen, D. and Bachetti, M. Accurate X-ray Timing in the Presence of Systematic Biases With Simulation-Based Inference. Technical report, April 2021.

URL https://ui.adsabs.harvard.edu/abs/2021arXiv210403278H. Publication Title: arXiv e-prints ADS Bibcode: 2021arXiv210403278H Type: article.

Johnson, B. D., Leja, J., Conroy, C., and Speagle, J. S. Stellar Population Inference with Prospector. *The Astrophysical Journal Supplement Series*, 254(2):22, June 2021. ISSN 0067-0049, 1538-4365. doi: 10.3847/1538-4365/abef67. URL http://arxiv.org/abs/2012.01426. arXiv: 2012.01426.

Karamanis, M. and Beutler, F. Ensemble Slice Sampling. *arXiv e-prints*, pp. arXiv:2002.06212, February 2020. URL https://ui.adsabs.harvard.edu/abs/2020arXiv200206212K.

Kingma, D. P. and Ba, J. Adam: A Method for Stochastic Optimization. *arXiv:1412.6980 [cs]*, January 2017. URL http://arxiv.org/abs/1412.6980. arXiv: 1412.6980.

Leja, J., Johnson, B. D., Conroy, C., van Dokkum, P. G., and Byler, N. Deriving Physical Properties from Broadband Photometry with Prospector: Description of the Model and a Demonstration of its Accuracy Using 129 Galaxies in the Local Universe. *The Astrophysical Journal*, 837:170, March 2017. ISSN 0004-637X. doi: 10.3847/1538-4357/aa5ffe. URL http://adsabs.harvard.edu/abs/2017ApJ...837..170L.

Papamakarios, G., Pavlakou, T., and Murray, I. Masked Autoregressive Flow for Density Estimation. *arXiv e-prints*, 1705:arXiv:1705.07057, May 2017. URL http://adsabs.harvard.edu/abs/2017arXiv170507057P.

Portillo, S. K. N., Parejko, J. K., Vergara, J. R., and Connolly, A. J. Dimensionality Reduction of SDSS Spectra with Variational Autoencoders. *The Astronomical Journal*, 160(1):45, June 2020. ISSN 1538-3881. doi: 10.3847/1538-3881/ab9644. URL http://arxiv.org/abs/2002.10464. arXiv: 2002.10464.

Tabak, E. G. and Turner, C. V. A Family of Nonparametric Density Estimation Algorithms. *Communications on Pure and Applied Mathematics*, 66(2):145–164, 2013. ISSN 1097-0312. doi: 10.1002/cpa.21423. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.21423. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpa.21423.

Tabak, E. G. and Vanden-Eijnden, E. Density estimation by dual ascent of the log-likelihood. *Communications in Mathematical Sciences*, 8(1):217–233, March 2010. ISSN 1945-0796. doi: 10.4310/CMS.2010.v8.n1.a11. URL https://www.intlpress.com/site/pub/pages/journals/items/cms/content/vols/0008/0001/a011/abstract.php. Publisher: International Press of Boston.

Tacchella, S., Conroy, C., Faber, S. M., Johnson, B. D., Leja, J., Barro, G., Cunningham, E. C., Deason, A. J., Guhathakurta, P., Guo, Y., Hernquist, L., Koo, D. C., McKinnon, K., Rockosi, C. M., Speagle, J. S., van Dokkum, P., and Yesuf, H. M. Fast, Slow, Early, Late: Quenching Massive Galaxies at \$z\\sim0.8\$. *arXiv e-prints*, 2102:arXiv:2102.12494, February 2021. URL http://adsabs.harvard.edu/abs/2021arXiv210212494T.

Talts, S., Betancourt, M., Simpson, D., Vehtari, A., and Gelman, A. Validating Bayesian Inference Algorithms with Simulation-Based Calibration. *arXiv:1804.06788 [stat]*, October 2020. URL http://arxiv.org/abs/1804.06788. arXiv: 1804.06788.

Tejero-Cantero, A., Boelts, J., Deistler, M., Lueckmann, J.-M., Durkan, C., Gonçalves, P. J., Greenberg, D. S., and Macke, J. H. sbi: A toolkit for simulation-based inference. *Journal of Open Source Software*, 5(52):2505, August 2020. ISSN 2475-9066. doi: 10.21105/joss.02505. URL https://joss.theoj.org/papers/10.21105/joss.02505.

Uria, B., Côté, M.-A., Gregor, K., Murray, I., and Larochelle, H. Neural Autoregressive Distribution Estimation. *arXiv:1605.02226 [cs]*, May 2016. URL http://arxiv.org/abs/1605.02226. arXiv: 1605.02226.

Walcher, J., Groves, B., Budavári, T., and Dale, D. Fitting the integrated spectral energy distributions of galaxies. *Astrophysics and Space Science*, 331:1–52, January 2011. doi: 10.1007/s10509-010-0458-z. URL http://adsabs.harvard.edu/abs/2011Ap%26SS.331....1W.

Weyant, A., Schafer, C., and Wood-Vasey, W. M. Likelihood-free Cosmological Inference with Type Ia Supernovae: Approximate Bayesian Computation for a Complete Treatment of Uncertainty. *The Astrophysical Journal*, 764:116, February 2013. ISSN 0004-637X. doi: 10.1088/0004-637X/764/2/116. URL http://adsabs.harvard.edu/abs/2013ApJ...764..116W.

Zhang, K., Bloom, J. S., Gaudi, B. S., Lanusse, F., Lam, C., and Lu, J. R. Real-Time Likelihood-Free Inference of Roman Binary Microlensing Events with Amortized Neural Posterior Estimation. February 2021. doi: 10.3847/1538-3881/abf42e. URL https://arxiv.org/abs/2102.05673v3.