Cosmology with Only the Photometry of Many Galaxies

Anonymous Authors[1]

Abstract

We present the first cosmological constraints from photometric observations of galaxies. Villaescusa-Navarro+(2022) recently demonstrated that a galaxy's internal physical properties contain a significant amount of cosmological information: *i.e.* cosmology with one galaxy. These physical properties, however, cannot be directly measured from observations. In this talk, I will present how we can go beyond theoretical demonstrations to infer cosmological constraints from actual galaxy observables (*e.g.* optical SDSS photometry) using simulation-based inference and neural density estimation on the CAMELS dataset. While the cosmological information from photometric observations of a single galaxy is limited, I will present preliminary results that place significant cosmological constraints from hierarchical population inference of photometry from thousands of SDSS galaxies. Furthermore, I will discuss domain adaptation (TNG versus SIMBA) and challenges in SED modeling, which underpins the connection between the simulations and observations. Lastly, I will discuss the potential of extending this work to upcoming surveys (*e.g.* DESI, PFS, Rubin).

1. Introduction

2. Data

2.1. NASA-Sloan Atlas

For our observations, we use galaxy photometry from the NASA-Sloan Atlas[1] (hereafter NSA). The NSA provides photometry of $z < 0.05$ galaxies observed by the Sloan Digital Sky Survey Data Release 8 (SDSS; Aihara et al., 2011) with improved background subtraction (Blanton et al., 2011). In particular, we use optical g, r, i, z band absoluate magnitudes derived using KCORRECT (Blanton & Roweis, 2007), assuming a (Chabrier, 2003) initial mass function. In Fig. 1, we present the $(g-r)-M_r$ color-magnitude distribution of $\sim 120,000$ NSA galaxies (black), which reveals the modes that correspond to blue star-forming and red quiescent galaxies.

Among the full NSA sample, we focus on luminous galaxies with $-18 > M_r > -22$. We exclude galaxies brighter than $M_r > -22$ since our simulated galaxy sample does not include a large number of the most luminous galaxies. In addition, we only select galaxies with precisely measured photometry: magnitude uncertainties below $\sigma_g, \sigma_r, \sigma_i < 0.022$ and $\sigma_z < 0.04$. Lastly, we impose the color cuts to exclude galaxies within the central 68 percentile of the $g-r$, $g-i$, $g-z$, $r-i$, $r-z$, $i-z$ color distributions. The color cuts remove NSA galaxies that potentially have observational artifacts or problematic photometry. They also ensure that the NSA galaxies are within the photometric distribution (*i.e.* support) of the simulated galaxies used in this work. We mark the 95^{th} percentile contour of our NSA subsample in Fig. 1 (black dot-dashed). In total, we use 14,736 NSA galaxies.

2.2. Forward Model: CAMELS

In this work, we also use simulated galaxies from the Cosmology and Astrophysics with MachinE Learning Simulations (Villaescusa-Navarro et al., 2021; 2022, CAMELS;), a suite of hydrodynamical simulations constructed over a wide range of cosmological and hydrodynamical parameters. In particular, we use the 1,000 hydrodynamical simulations (CAMELS-TNG) constructed using the subrid physics model of the state-of-the-art IllustrisTNG (**?**). The simulations are generated with different cosmoloigcal parameters, $\Omega = \{\Omega_m, \sigma_8\}$, and baryonic feedback parameters, $\mathcal{B} = \{A_{\rm SN1}, A_{\rm SN2}, A_{\rm AGN1}, A_{\rm AGN2}\}$, arranged in a latin hypercube. $A_{\rm SN1}$ and $A_{\rm SN2}$ represent the normalization factors for the galactic wind flux and speed; $A_{\rm AGN1}$ and $A_{\rm AGN2}$ represent the normalization factors for the energy output and specific energy of AGN feedback.

The CAMELS-TNG simulation all have a comoving volume of $(25\ h^{-1}{\rm Mpc})^3$. They each evolve 256^3 dark matter particles and 256^3 fluid elements from $z = 127$ to $z = 0$ from

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.

[1]http://www.nsatlas.org/

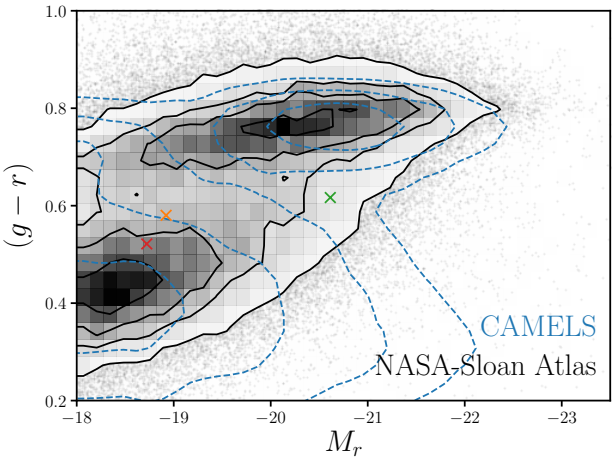

Figure 1. Color-magnitude distrubtion, $(g-r)-M_r$, of observed galaxies from the NSA (black) and simulated galaxies from CAMELS-TNG (blue). The distribution reveals the bimodality distribution of blue star-forming and red quiescent galaxies. Overall, the distributions of NSA and CAMELS-TNG galaxies are in good agreement. The distribution for CAMELS-TNG is significantly broader since its simulated galaxies are generated using a wide range of cosmological and baryonic feedback parameters.

initial conditions generated using second order perturbation theory. Each simulation has 34 saved snapshots from $z = 6$ to $z = 0$. Since we target $z < 0.05$ NSA galaxies, we only use the $z = 0$ snapshot. All the cosmological parameter besides Ω_m and σ_8 are fixed: $\Omega_b = 0.049$, $h = 0.6711$, $n_s = 0.9624$, $\sum m_\nu = 0.0$eV, and $w = -1$. The SUBFIND (Springel et al., 2001; Dolag et al., 2009) algorithm is run on each simulation to identify halos and subhalos. SUBFIND computes physical quantities of the subhalos and the galaxies that resides in them. For additional details on CAMELS, we refer readers to (Villaescusa-Navarro et al., 2021; 2022).

In all 1,000 CAMELS-TNG simulations, there are ∼700,000 galaxies with more than 20 star particles. The galaxies, however, are not evenly distributed across the simulations and have a significant dependence on the CAMELS parameters. For instance, simulations constructed at higher Ω_m values have more galaxies. Since the goal of this work is to conduct cosmological inference on a per galaxy basis, we must correct for this implicit prior on the CAMELS parameters. We do this by randomly selecting 100 galaxies from each simulation. This imposes a uniform prior: $p(\Omega, \mathcal{B}) = 1$. Thus, we use a total of 100,000 CAMELS-TNG galaxies.

Since the goal of this work is to analyze the observed photometry of NSA galaxies, we forward model observed photometry for the simulated galaxies. CAMELS-TNG already provides synthetic dust attenuated stellar photometry for each simulated galaxy calculated using the same procedure as (Nelson et al., 2018). The unattenuated spectral energy distribution (SED) of a galaxy is computed by combining the SEDs of every member star particle of the host subhalo. Each star particle SED is modeled as a single-burst simple stellar population using stellar population synthesis (SPS) based on the recorded birth time, metallicity, and mass. The SPS uses FSPS (Conroy et al., 2009; **?**), Padova isochrones, MILES stellar library, and assumes a Chabrier initial mass function. The unattenuated galaxy SEDs are then dust attenuated using the "Model C" dust model of (Nelson et al., 2018), which is based on the metal content of the neutral gas distribution in and around each galaxy. Afterwards, each galaxy SED is convolved with the SDSS g, r, i, z-band photometric bandpasses to produce the photometry: X_i. For additional details on the synthetic photometry, we refer readers to (Nelson et al., 2018).

The synthetic photometry in CAMELS-TNG does not include any uncertainties. However, since we have measured uncertainties, σ_X, for NSA galaxies, we can construct a realistic noise model. For each CAMELS-TNG galaxy, we randomly sample $\sigma_{X,i}$ from the range of uncertainties measured in NSA. Afterwards, we apply the uncertainty using a Gaussian with standard deviation $\sigma_{X,i}$: $\hat{X}_i \sim \mathcal{N}(X_i, \sigma_{X,i})$. Our noise model does not include correlations between the magnitudes and the uncertainties. However, as we later discuss, this is not an issue in our approach beacuse the posteriors we ultimately evaluate are conditioned on the uncertainties. In Fig. 1, we present the color-magnitude distribution of the forward modeled CAMELS-TNG galaxies in blue.

3. Methods

3.1. Hierarchical Bayesian Inference

Our goal in this paper is to infer the posterior of cosmological parameters $\Omega = \{\Omega_m, \sigma_8\}$ and baryonic feedback parameters $\mathcal{B} = \{A_{\rm SN1}, A_{\rm SN2}, A_{\rm AGN1}, A_{\rm AGN2}\}$ from the observed photometry of galaxies in the NSA catalog, $\{\boldsymbol{X}_i\}$: $p(\Omega, \mathcal{B} \,|\, \{\boldsymbol{X}_i\})$. $\boldsymbol{X}_i$ represents both the measured absolute magnitudes and uncertainties: $\{\hat{X}_i, \sigma_{X,i}\}$. With our forward model, based on CAMELS-TNG, we can simulate noisy galaxy photometry from Ω and $\mathcal{B}$. Hence, the cosmological inference from photometry can be reformulated as a hierarchical population inference problem.

To illustrate this, we graphically represent our forward model in Figure 2. Circles, shaded circles, and dots represent random variables, observed quantities, and random variables that are deterministic. θ_i^g represents the physical properties of galaxies (*e.g.* M_*, star-formation history), which are determined from Ω and $\mathcal{B}$ through the CAMELS-

TNG simulation. Then the noisy photometry $\hat{X}_i$ is determined from θ_i^g through SPS and our noise model.

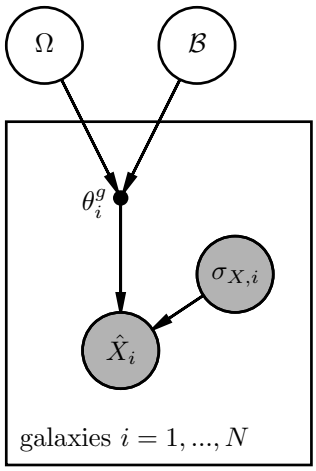

Figure 2. Graphical representation of our hierarchical approach that illustrate the relationship among the parameters of our model. Circles are inferred random variables, shaded circles are observed quantities, and dots indicate random variables that are deterministic. The physical properties of galaxies, θ_i^g, are determined from the cosmological and hydrodynamical parameters Ω and $\mathcal{B}$ through the CAMELS-TNG simulation. Then the noisy optical photometry, $\hat{X}_i$, is derived from θ_i^g using SPS and our noise model.

Given this hierarchical model, we can rewrite the posterior as:

$$p(\Omega, \mathcal{B} \,|\, \{\boldsymbol{X}_i\}) = \frac{p(\Omega, \mathcal{B})\; p(\{\boldsymbol{X}_i\} \,|\, \Omega, \mathcal{B})}{p(\{\boldsymbol{X}_i\})} \tag{1}$$

$$= \frac{p(\Omega, \mathcal{B})}{p(\{\boldsymbol{X}_i\})} \prod_{i=1}^{N} p(\boldsymbol{X}_i \,|\, \Omega, \mathcal{B}) \tag{2}$$

$$= \frac{p(\Omega, \mathcal{B})}{p(\{\boldsymbol{X}_i\})} \prod_{i=1}^{N} \frac{p(\boldsymbol{X}_i)\, p(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i)}{p(\Omega, \mathcal{B})} \tag{3}$$

$$= \frac{1}{p(\Omega, \mathcal{B})^{N-1}} \prod_{i=1}^{N} p(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i) \tag{4}$$

We downsampled the galaxies in the CAMELS-TNG simulations to impose $p(\Omega, \mathcal{B}) = 1$ (Sec. 2.2), so the posterior simply becomes:

$$= \prod_{i=1}^{N} p(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i). \tag{5}$$

Hence, we can evaluate $p(\Omega, \mathcal{B} \,|\, \{\boldsymbol{X}_i\})$, and sample it, as long as we can accurately estimate $p(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i)$, the posterior for photometry from an individual galaxy.

3.2. Neural Density Estimation

One way to accurately estimate $p(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i)$ is by applying neural density estimation (NDE) to the CAMELS-TNG. The CAMELS-TNG forward model provides a training dataset of 100,000 parameter-photometry pairs: $\{(\Omega, \mathcal{B}, \boldsymbol{X}_i)\}$. With NDE, we can use this data to train a neural network q with parameters ϕ to estimate $p(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i) \approx q_\phi(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i)$. This type of simulaiton-based inference using NDE has now been applied to a broad range of astronomical applications: *e.g.* analyzing gravitational waves (Wong et al., 2020; Dax et al., 2021), binary microlensing lensing (**?**), galaxy SEDs (Hahn & Melchior, 2022), and galaxy clustering (Hahn et al., 2022; 2023).

In this work, our NDE is based on "normalizing flow" models (Tabak & Vanden-Eijnden, 2010; Tabak & Turner, 2013), which use neural networks to learn a flexible and bijective transformation, f, that maps a complex target distribution to a simple base distribution that is fast to evaluate. f is defined to be invertible and have a tractable Jacobian so that the target distribution can be evaluated with change of variables. Since $\pi(\boldsymbol{z})$ is easy to evaluate, this enables us to also easily evaluate the target distribution. In our case, the target distribution is $p(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i)$ and we set the base distribution to be a multivariate Gaussian. Among different flow architectures, we use Masked Autoregressive Flow (MAF; Papamakarios et al., 2017) models implemented in the `sbi` Python package[2] (Greenberg et al., 2019; Tejero-Cantero et al., 2020).

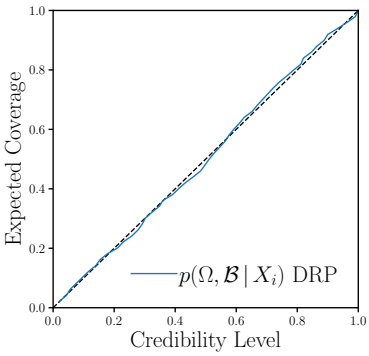

Figure 3. DRP coverage test validating the accuracy of our $q_\phi(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i)$ posterior estimate trained using the forward modeled CAMELS-TNG data (blue). The black-dashed line represents an optimal estimate of the posterior. The coverage test demonstrates that q_ϕ provides a near optimal estimate of the true posterior.

Our goal in training the flow is to determine q_ϕ that best approximates $p(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i)$. We can reformulate this into an optimization problem for determining ϕ that minimizes the KL

[2] https://github.com/mackelab/sbi/

divergence between $p(\Omega, \mathcal{B}, \boldsymbol{X}_i) = p(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i) p(\boldsymbol{X}_i)$ and $q_\phi(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i) p(\boldsymbol{X}_i)$. In practice, we split the $\{(\Omega, \mathcal{B}, \boldsymbol{X}_i\}$ data from CAMELS-TNG into a training and validation set with a 90/10 split. Then, we maximize the total log-likelihood $\sum_i \log q_\phi(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i)$ over the training set, which is equivalent to the KL divergence, using the ADAM optimizer (Kingma & Ba, 2017) with a learning rate of 5×10^{-4}. To prevent overfitting, we evaluate the total log-likelihood on the validation data at every training epoch and stop the training when the validation log-likelihood fails to increase after 20 epochs.

We determine the architecture of our flow by training a large number of flows with architectures determined using the (Akiba et al., 2019) hyperparameter optimization framework. Afterwards, we define our final flow as an equally weighted ensemble of five flows with the lowest validation losses: $q_\phi(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i) = \sum_{j=1}^{5} q_{\phi,j}(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i)/5$. Ensembling flows with different initializations and architectures improves the overall robustness of our normalizing flow (Alsing et al., 2019). To validate the accuracy of q_ϕ, we use the "distance to random point" (DRP) coverage test from (Lemos et al., 2023), which is necessary and sufficient to show that a posterior estimator is optimal. In Fig 3, we present the DRP coverage of q_ϕ (blue) and find that it provides a near optimal estimate of the true posterior.

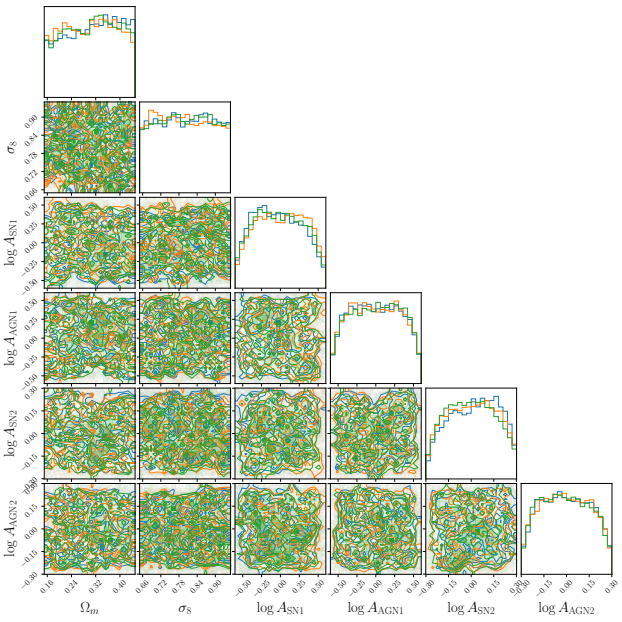

Figure 4. Estimated posteriors of the cosmological and hydrodynamical parameters (Ω, $mathcalB$) from photometry, $q_\phi(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i)$,for three arbitrarily selected NSA galaxies. The posteriors correspond to the galaxies marked in Fig. 1.

In Fig. 4, we present $q_\phi(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i)$ for three arbitrarily selected NSA galaxies. The selected galaxies are also marked in Fig. 1. The individual posteriors reveal that there is very little cosmological information in the photometry of a single galaxy. However, with Eq. 5, we can extract the cosmological information from *thousands* of galaxies.

4. Results

Our posterior on Ω_m and σ_8 is inferred assuming a galaxy formation model and a SED model. This is illustrated in Eq. **??** where $p(\Omega, B \,|\, \theta_i^g)$ is determined by the galaxy formation model and $p(\theta_i^g \,|\, \boldsymbol{X}_i)$ is determined by the SED model. Below we discuss the caveats of these assumptions and how they can be accounted for in future research.

4.1. Galaxy Formation Model

Some interpretation of cosmology with one galaxy here and what that may actually imply for $p(\Omega, B \,|\, \theta_i^g)$

4.2. SED Modeling

The forward modeled photometry for the simulated galaxies in our training data are constructed using SED modeling (Section 2.2.

dust modeling is unrealistic, which could bias the inferred θ^g

emission lines are crap

hard assumptions on the IMF

Despite these serious limitations, we can identify subpopulations of galaxies that are the least sensitive to them. For example, quiescent galaxies do not have emission lines and we can identify ones without dust attenuation from XXXX (galaxy zoo? emission lines?). Furthermore, recent studies using both lensing and IR spectral features also find no IMF variations in the bulk populations of nearby quiescent galaxies. Hence, although $p(\theta_i^g \,|\, \boldsymbol{X}_i)$ is not robust over the full space of $\boldsymbol{X}$, it is robust for dustless quiescent galaxies.

In principle, $p(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i) = \int p(\Omega, \mathcal{B} \,|\, \theta_i^g)\, p(\theta_i^g \,|\, \boldsymbol{X}_i)\, \mathrm{d}\theta_i^g$, where θ_i^g is a latent variable that is marginalized.

References

Aihara, H., Allende Prieto, C., An, D., Anderson, S. F., Aubourg, É., Balbinot, E., Beers, T. C., Berlind, A. A., Bickerton, S. J., Bizyaev, D., Blanton, M. R., Bochanski, J. J., Bolton, A. S., Bovy, J., Brandt, W. N., Brinkmann, J., Brown, P. J., Brownstein, J. R., Busca, N. G., Campbell, H., Carr, M. A., Chen, Y., Chiappini, C., Comparat, J., Connolly, N., Cortes, M., Croft, R. A. C., Cuesta, A. J., da Costa, L. N., Davenport, J. R. A., Dawson, K., Dhital, S., Ealet, A., Ebelke, G. L., Edmondson,

E. M., Eisenstein, D. J., Escoffier, S., Esposito, M., Evans, M. L., Fan, X., Femenía Castellá, B., Font-Ribera, A., Frinchaboy, P. M., Ge, J., Gillespie, B. A., Gilmore, G., González Hernández, J. I., Gott, J. R., Gould, A., Grebel, E. K., Gunn, J. E., Hamilton, J.-C., Harding, P., Harris, D. W., Hawley, S. L., Hearty, F. R., Ho, S., Hogg, D. W., Holtzman, J. A., Honscheid, K., Inada, N., Ivans, I. I., Jiang, L., Johnson, J. A., Jordan, C., Jordan, W. P., Kazin, E. A., Kirkby, D., Klaene, M. A., Knapp, G. R., Kneib, J.-P., Kochanek, C. S., Koesterke, L., Kollmeier, J. A., Kron, R. G., Lampeitl, H., Lang, D., Le Goff, J.-M., Lee, Y. S., Lin, Y.-T., Long, D. C., Loomis, C. P., Lucatello, S., Lundgren, B., Lupton, R. H., Ma, Z., MacDonald, N., Mahadevan, S., Maia, M. A. G., Makler, M., Malanushenko, E., Malanushenko, V., Mandelbaum, R., Maraston, C., Margala, D., Masters, K. L., McBride, C. K., McGehee, P. M., McGreer, I. D., Ménard, B., Miralda-Escudé, J., Morrison, H. L., Mullally, F., Muna, D., Munn, J. A., Murayama, H., Myers, A. D., Naugle, T., Neto, A. F., Nguyen, D. C., Nichol, R. C., O'Connell, R. W., Ogando, R. L. C., Olmstead, M. D., Oravetz, D. J., Padmanabhan, N., Palanque-Delabrouille, N., Pan, K., Pandey, P., Pâris, I., Percival, W. J., Petitjean, P., Pfaffenberger, R., Pforr, J., Phleps, S., Pichon, C., Pieri, M. M., Prada, F., Price-Whelan, A. M., Raddick, M. J., Ramos, B. H. F., Reylé, C., Rich, J., Richards, G. T., Rix, H.-W., Robin, A. C., Rocha-Pinto, H. J., Rockosi, C. M., Roe, N. A., Rollinde, E., Ross, A. J., Ross, N. P., Rossetto, B. M., Sánchez, A. G., Sayres, C., Schlegel, D. J., Schlesinger, K. J., Schmidt, S. J., Schneider, D. P., Sheldon, E., Shu, Y., Simmerer, J., Simmons, A. E., Sivarani, T., Snedden, S. A., Sobeck, J. S., Steinmetz, M., Strauss, M. A., Szalay, A. S., Tanaka, M., Thakar, A. R., Thomas, D., Tinker, J. L., Tofflemire, B. M., Tojeiro, R., Tremonti, C. A., Vandenberg, J., Vargas Magaña, M., Verde, L., Vogt, N. P., Wake, D. A., Wang, J., Weaver, B. A., Weinberg, D. H., White, M., White, S. D. M., Yanny, B., Yasuda, N., Yeche, C., and Zehavi, I. The Eighth Data Release of the Sloan Digital Sky Survey: First Data from SDSS-III. *The Astrophysical Journal Supplement Series*, 193:29, April 2011. ISSN 0067-0049. doi: 10.1088/0067-0049/193/2/29.

Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M. Optuna: A Next-generation Hyperparameter Optimization Framework, July 2019.

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.

Blanton, M. R. and Roweis, S. K-corrections and filter transformations in the ultraviolet, optical, and near infrared. *The Astronomical Journal*, 133(2):734–754, February 2007. ISSN 0004-6256, 1538-3881. doi: 10.1086/510127.

Blanton, M. R., Kazin, E., Muna, D., Weaver, B. A., and Price-Whelan, A. Improved Background Subtraction for the Sloan Digital Sky Survey Images. *The Astronomical Journal*, 142:31, July 2011. ISSN 0004-6256. doi: 10.1088/0004-6256/142/1/31.

Chabrier, G. Galactic Stellar and Substellar Initial Mass Function. *Publications of the Astronomical Society of the Pacific*, 115:763–795, July 2003. ISSN 0004-6280. doi: 10.1086/376392.

Conroy, C., Gunn, J. E., and White, M. The propagation of uncertainties in stellar population synthesis modeling I: The relevance of uncertain aspects of stellar evolution and the IMF to the derived physical properties of galaxies. *The Astrophysical Journal*, 699(1):486–506, July 2009. ISSN 0004-637X, 1538-4357. doi: 10.1088/0004-637X/699/1/486.

Dax, M., Green, S. R., Gair, J., Macke, J. H., Buonanno, A., and Schölkopf, B. Real-Time Gravitational Wave Science with Neural Posterior Estimation. *Physical Review Letters*, 127:241103, December 2021. ISSN 0031-9007. doi: 10.1103/PhysRevLett.127.241103.

Dolag, K., Borgani, S., Murante, G., and Springel, V. Substructures in hydrodynamical cluster simulations. *Monthly Notices of the Royal Astronomical Society*, 399:497–514, October 2009. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2009.15034.x.

Greenberg, D. S., Nonnenmacher, M., and Macke, J. H. Automatic Posterior Transformation for Likelihood-Free Inference, May 2019.

Hahn, C. and Melchior, P. Accelerated Bayesian SED Modeling using Amortized Neural Posterior Estimation, March 2022.

Hahn, C., Eickenberg, M., Ho, S., Hou, J., Lemos, P., Massara, E., Modi, C., Moradinezhad Dizgah, A., Régaldo-Saint Blancard, B., and Abidi, M. M. ${\rm S{\scriptsize IM}BIG}$: A Forward Modeling Approach To Analyzing Galaxy Clustering, November 2022.

Hahn, C., Eickenberg, M., Ho, S., Hou, J., Lemos, P., Massara, E., Modi, C., Moradinezhad Dizgah, A., Régaldo-Saint Blancard, B., and Abidi, M. M. SIMBIG: Mock challenge for a forward modeling approach to galaxy clustering. *Journal of Cosmology and Astroparticle Physics*, 2023:010, April 2023. ISSN 1475-7516. doi: 10.1088/1475-7516/2023/04/010.

Kingma, D. P. and Ba, J. Adam: A Method for Stochastic Optimization. *arXiv:1412.6980 [cs]*, January 2017.

Lemos, P., Coogan, A., Hezaveh, Y., and Perreault-Levasseur, L. Sampling-Based Accuracy Testing of Posterior Estimators for General Inference. https://arxiv.org/abs/2302.03026v1, February 2023.

Nelson, D., Pillepich, A., Springel, V., Weinberger, R., Hernquist, L., Pakmor, R., Genel, S., Torrey, P., Vogelsberger, M., Kauffmann, G., Marinacci, F., and Naiman, J. First results from the IllustrisTNG simulations: The galaxy colour bimodality. *Monthly Notices of the Royal Astronomical Society*, 475:624–647, March 2018. ISSN 0035-8711. doi: 10.1093/mnras/stx3040.

Papamakarios, G., Pavlakou, T., and Murray, I. Masked Autoregressive Flow for Density Estimation. *arXiv e-prints*, 1705:arXiv:1705.07057, May 2017.

Springel, V., White, S. D. M., Tormen, G., and Kauffmann, G. Populating a cluster of galaxies - I. Results at [formmu2]z=0. *Monthly Notices of the Royal Astronomical Society*, 328:726–750, December 2001. ISSN 0035-8711. doi: 10.1046/j.1365-8711.2001.04912.x.

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.

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.

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.

Villaescusa-Navarro, F., Anglés-Alcázar, D., Genel, S., Spergel, D. N., Somerville, R. S., Dave, R., Pillepich, A., Hernquist, L., Nelson, D., Torrey, P., Narayanan, D., Li, Y., Philcox, O., La Torre, V., Delgado, A. M., Ho, S., Hassan, S., Burkhart, B., Wadekar, D., Battaglia, N., Contardo, G., and Bryan, G. L. The CAMELS project: Cosmology and Astrophysics with MachinE Learning Simulations. *The Astrophysical Journal*, 915 (1):71, July 2021. ISSN 0004-637X, 1538-4357. doi: 10.3847/1538-4357/abf7ba.

Villaescusa-Navarro, F., Genel, S., Anglés-Alcázar, D., Perez, L. A., Villanueva-Domingo, P., Wadekar, D., Shao, H., Mohammad, F. G., Hassan, S., Moser, E., Lau, E. T., Valle, L. F. M. P., Nicola, A., Thiele, L., Jo, Y., Philcox, O. H. E., Oppenheimer, B. D., Tillman, M., Hahn, C., Kaushal, N., Pisani, A., Gebhardt, M., Delgado, A. M., Caliendo, J., Kreisch, C., Wong, K. W. K., Coulton, W. R., Eickenberg, M., Parimbelli, G., Ni, Y., Steinwandel, U. P., La Torre, V., Dave, R., Battaglia, N., Nagai, D., Spergel, D. N., Hernquist, L., Burkhart, B., Narayanan, D., Wandelt, B., Somerville, R. S., Bryan, G. L., Viel, M., Li, Y., Irsic, V., Kraljic, K., and Vogelsberger, M. The CAMELS project: Public data release. *arXiv:2201.01300 [astro-ph]*, January 2022.

Wong, K. W. K., Contardo, G., and Ho, S. Gravitational wave population inference with deep flow-based generative network. *Physical Review D*, 101(12):123005, June 2020. ISSN 2470-0010, 2470-0029. doi: 10.1103/PhysRevD.101.123005.

A. You *can* have an appendix here.

You can have as much text here as you want. The main body must be at most 8 pages long. For the final version, one more page can be added. If you want, you can use an appendix like this one, even using the one-column format.