Cosmology with Only the Photometry of Galaxies

Anonymous Authors[1]

Abstract

We present the first cosmological constraints from only the observed photometry of galaxies. Villaescusa-Navarro et al. (2022a) recently demonstrated that the internal physical properties of a single galaxy contain a significant amount of cosmological information. These physical properties, however, cannot be directly measured from observations. In this work, we present how we can go beyond theoretical demonstrations to infer cosmological constraints from actual galaxy observables (*e.g.* optical photometry) using neural density estimation and the CAMELS suite of hydrodynamical simulations. We find that the cosmological information in the photometry of a single galaxy is limited. However, we can combine the constraining power of photometry from many galaxies using hierarchical population inference and place significant cosmological constraints. With the observed photometry of ∼15,000 NASA-Sloan Atlas galaxies, we constrain $\Omega_m = 0.310^{+0.080}_{-0.098}$ and $\sigma_8 = 0.792^{+0.099}_{-0.090}$.

1. Introduction

In recent work, Villaescusa-Navarro et al. (2022a) showed that it is possible to place cosmological constraints from only the internal properties of a single galaxy. They used galaxies from 2,000 state-of-the-art hydrodynamical simulations with different cosmologies and astrophysical models from CAMELS (Villaescusa-Navarro et al., 2021; 2022b) to train moment networks (Jeffrey & Wandelt, 2020) that predict cosmological parameters from galaxy properties. With only a handful of galaxy properties, including stellar mass (M_*), stellar metallicity (Z_*), and maximum circular velocity ($V_{\rm max}$), they were able to constrain Ω_m to 10% precision with a single galaxy. They found similar constraining power for galaxies simulated using the subgrid physics models of the IllustrisTNG (Pillepich et al., 2018; Weinberger et al., 2018) and SIMBA (Davé et al., 2019). Since then, follow-up works have found consistent results for other hydrodynamicl models: (Echeverri et al., 2023).

According to Villaescusa-Navarro et al. (2021), the cosmological information is derived from the imprint of Ω_m on the dark matter content of galaxies that affects galaxy properties in a distinct way than astrophysical processes. Also, since Ω_b is fixed in CAMELS, which is justified by the tight constraints from Big Bang Nucleosynthesis, the galaxy properties are effectively measuring the baryon fraction, Ω_b/Ω_m. For instance, $V_{\rm max}$ measures the depth of the total matter gravitational potential while other properties like M_* and Z_* measures the mass in baryons so together they can constrain the ratio Ω_b/Ω_m. In fact, a similar approach was used three deacdes ago in White et al. (1993) to constrain Ω_b/Ω_m using galaxy clusters.

Despite the promising signs that they may be useful cosmological probes, galaxy properties themselves are *not* actual observable. They are derived quantities that are typically inferred from photometry or spectra of galaxies and require modeling the spectral energy distribution (SED) or emission lines (Conroy, 2013). In this work, we go beyond the theoretical considerations of previous works and infer cosmological parameters from actual galaxy observables — optical photometry. We leverage a simulation-based inference method that employs neural density estimation to estimate the posterior of cosmological parameters given galaxy photometry, similar to the approach of (Hahn & Melchior, 2022). Furthermore, since we expect a limited amount of cosmological information from the photometry of a single galaxy, we present a hierarhical population inference approach for inferring the posterior of cosmological parameters from the photometry of multiple galaxies. Lastly, we present the cosmological constraints derived from applying this approach to the photometry of ∼15,000 SDSS galaxies from the NASA-Sloan Atlas (Sec. A).

2. Forward Model: CAMELS

We use simulated galaxies from 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 constructed using

[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.

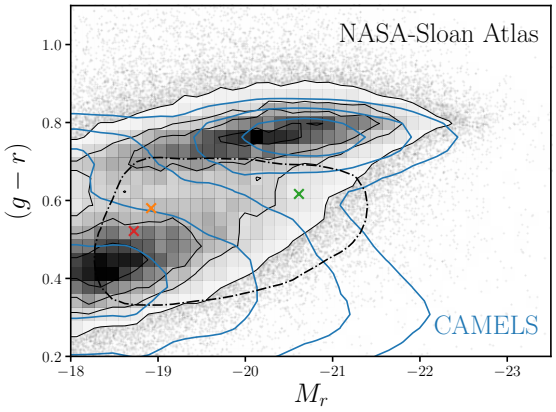

Figure 1. Color-magnitude distrubtion, $(g-r)-M_r$, of observed galaxies from the NSA (black) and simulated galaxies from CAMELS-TNG (blue). Overall, the distributions of NSA and CAMELS-TNG galaxies are in good agreement. The distribution for CAMELS-TNG is significantly broader since its simulated galaxies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

the state-of-the-art IllustrisTNG model (CAMELS-TNG). The simulations are generated with different cosmological parameters, $\Omega = \{\Omega_m, \sigma_8\}$, and baryonic feedback parameters, $\mathcal{B} = \{A_{\rm SN1}, A_{\rm SN2}, A_{\rm AGN1}, A_{\rm AGN2}\}$, 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.

In the 1,000 simulations, there are ∼700,000 galaxies with $M_* > 2\times10^8 M_\odot/h$. The galaxies, however, are not evenly distributed across them and have a significant dependence on the parameters. For instance, simulations constructed at higher Ω_m values have more galaxies. We correct for this implicit prior on the CAMELS parameters by randomly selecting 100 galaxies from each simulation. This imposes a uniform prior: $p(\Omega, \mathcal{B}) = 1$ and leaves us with a total of 100,000 CAMELS-TNG galaxies.

Since our aim is to analyze the observed photometry of NSA galaxies, we forward model photometry for the simulated galaxies. CAMELS-TNG already provides synthetic dust attenuated stellar photometry for each simulated galaxy (Nelson et al., 2018). The unattenuated SED of a galaxy is computed by combining the SEDs of every star particle of the host subhalo. Each 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 attenuated using a dust model based on the metal content of the neutral gas distribution in and around each galaxy. Afterwards, the SED is convolved with SDSS g, r, i, z-band photometric bandpasses to produce the photometry: X_i.

Next, we add noise to the synthetic photometry based on the measured uncertainties, σ_X, of NSA galaxies. For each 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 $\sigma_{X,i}$: $\hat{X}_i \sim \mathcal{N}(X_i, \sigma_{X,i})$. Although our noise model is simplistic, this is not an issue in our approach beacuse the posteriors we ultimately evaluate are conditioned on the uncertainties (Hahn & Melchior, 2022). In Fig. 1, we present the color-magnitude distribution of the forward modeled CAMELS-TNG galaxies in blue. We reserve a random 10% of these galaxies as a test dataset for validation later Sec. B.

3. Methods

3.1. Hierarchical Bayesian Inference

Our goal is to infer the posterior of cosmological parameters $\Omega = \{\Omega_m, \sigma_8\}$ and baryonic feedback parameters $\mathcal{B}$ 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 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 CAMELS-TNG. Then the noisy photometry $\hat{X}_i$ is determined from θ_i^g through SPS and our noise model.

We can rewrite the posterior as:

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

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

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

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

Since $p(\Omega, \mathcal{B}) = 1$ (Sec. 2):

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

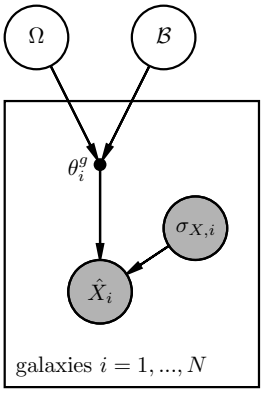

Figure 2. Graphical representation of our hierarchical approach that illustrate the relationship among the parameters of our model. Circles are inferred random variables, shaded circles are observed quantities, and dots indicate random variables that are deterministic. The physical properties of galaxies,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

Hence, we can evaluate $p(\Omega, \mathcal{B} \,|\, \{\boldsymbol{X}_i\})$ if we can accurately estimate $p(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i)$, the posterior for a single 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, which 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 simulation-based inference using NDE has been applied to a broad range of astronomical applications (*e.g.* Wong et al., 2020; Dax et al., 2021; Zhang et al., 2021; Hahn et al., 2022a).

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. 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[1] (Greenberg et al., 2019; Tejero-Cantero et al., 2020).

In training, we want to determine q_ϕ that best approximates $p(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i)$. We reformulate this into an optimization problem of determining ϕ that minimizes the KL 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 CAMELS-TNG data 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 minimizing the KL divergence. 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 train a large number of flows with architectures determined by the Akiba et al. (2019) hyperparameter optimization. Then, we construct 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. In Appendix B, we extensively validate the accuracy of q_ϕ, using Simulation-Based Calibration (Talts et al., 2020) and the Lemos et al. (2023) coverage test. Our final flow is a near optimal estimate of the true posterior.

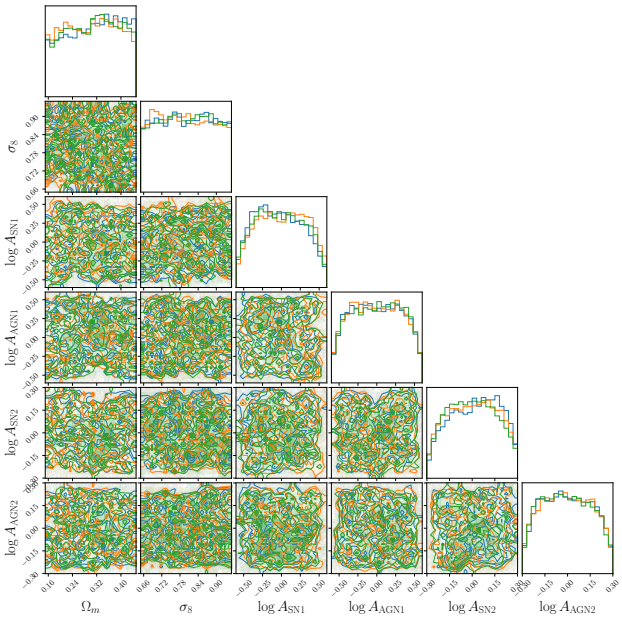

Figure 3. Estimated posteriors of the cosmological and hydrodynamical parameters (Ω, $\mathcal{B}$) from photometry, $q_\phi(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i)$, for three arbitrarily selected NSA galaxies. The posteriors correspond to the galaxies marked in Fig. 1. The photometry of a single galaxy contains limited cosmological information.

In Fig. 3, 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 limited cosmological information in the photometry of a single galaxy. However, with Eq. 5, we can extract the cosmological information from *thousands* of galaxies.

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

4. Results & Discussion

With our trained NDE, $q_\phi(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i)$, we can now evaluate the posterior $p(\Omega, \mathcal{B} \,|\, \{\boldsymbol{X}_i\})$ for multiple galaxies using Eq. 5. In Fig. 4, we present $p(\Omega, \mathcal{B} \,|\, \{\boldsymbol{X}_i\})$ for the 14,736 observed galaxies in our NSA sample. The contours mark the 68 and 95 percentiles of the distribution and we list the median and $\pm 1\sigma$ marginalized constraints on Ω_m, σ_8, and $S_8 = \sigma_8 \sqrt{\Omega_m / 0.3}$. The samples from $p(\Omega \,|\, \{\boldsymbol{X}_i\})$ are derived using Markov Chain Monte Carlo (MCMC). We use the `emcee` sampler (Foreman-Mackey et al., 2013) with 1,000 walkers evaluated 36,000 iterations, discarding the first 5,000 iterations for burn in.

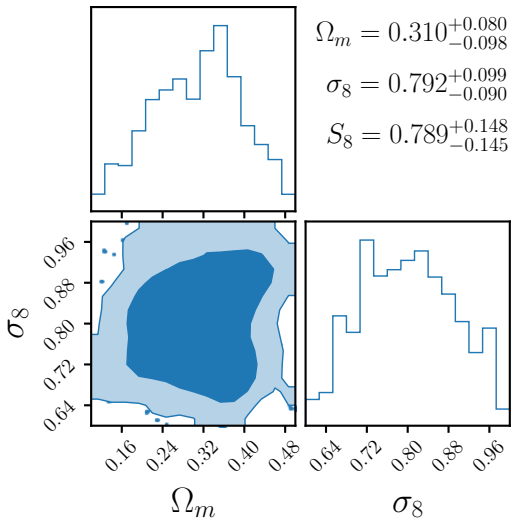

Figure 4. The posterior on Ω_m and σ_8 from the observed photometry of 14,736 NSA galaxies. The contours mark the 68 and 95 percentiles. We place significant cosmological constraints from only the photometry of galaxies.

Overall, we derive significant constraint on both Ω_m and σ_8: $\Omega_m = 0.310^{+0.080}_{-0.098}$ and $\sigma_8 = 0.792^{+0.099}_{-0.090}$. Although, we also derive significant constraints on the hydrodynamical parameters, we do not include them for clarity. Our cosmological constraints demonstrate that although the photometry of a single galaxy does not contain a significant amount of cosmological information, we can place significant cosmological constraints by combining the information from 14,736 galaxies.

Interestingly, our constraints can inform the latest tension between growth of structure S_8 constraints from large-scale structure (LSS) and CMB experiments. Low redshift LSS experiments find a significantly lower value of S_8 than CMB experiments and has placed renewed scrutiny on the standard ΛCDM cosmological model. Our $S_8 = 0.789^{+0.148}_{-0.145}$ constraint is consistent with the lower LSS S_8 values; however, it is not precise enough to competitive inform the tension.

A major caveat of our constraint is that our posterior assumes a galaxy formation and a SED models. This assumption is more clearly expressed if we rewrite

$$p(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i) \propto p(\Omega, \mathcal{B}) \int p(\theta_i^g \,|\, \boldsymbol{X}_i)\, p(\Omega, \mathcal{B} \,|\, \theta_i^g)\, \mathrm{d}\theta_i^g. \quad (6)$$

$p(\theta_i^g \,|\, \Omega, \mathcal{B})$, in our case, is the TNG simulation. Although TNG is a state-of-the-art hydrodynamical simulation, it makes specific choices and assumptions for its subgrid approximations. Consequently, an alternative galaxy formation model would produce a significantly different $p(\theta_i^g \,|\, \Omega, \mathcal{B})$. For instance, Hahn et al. (2019) found significantly discrepant M_*-star formation rate relations in three state-of-the-art hydrodynamical simulations.

Similarly, $p(\boldsymbol{X}_i \,|\, \theta_i^g)$ is given by the SED model, which also include a number of choices and assumptions. For instance, we use a dust model that assumes that dust is cospatial and uniformly mixed. Yet, recent studies suggest that the dust-star geometry can be significantly more complex and, thus, produce a wider range of attenuation curves (Narayanan et al., 2018; Hahn et al., 2021). The SED model also does not accurately model nebular emission lines, which may be significantly impact of photometry of emission ilne galaxies. Furthermore, the SED model assumes a fixed Chabrier IMF.

Despite these major limitations, a key advantage of our approach is that we can systematically choose subpopulations of galaxies that are the most robust to the assumptions in the galaxy formation and SED models. For example, we can compare $p(\theta_i^g \,|\, \Omega, \mathcal{B})$ across multiple galaxy formation models and choose the galaxies that are consistently predicted by the models. We can take a similar approach for the SED modeling assumptions by selecting galaxies without emission lines and with limited dust attenuation (*e.g.* by using infrared photometry). We can also identify galaxy populations with observationally constrained IMFs (Myers et al., 2013; Smith et al., 2015).

Another key advantage is that our cosmological constraining power depends on the number of galaxies we include. In this work, we selected ∼15,000 NSA galaxies using conservative noise and color cuts. They were from the NSA catalog, which only has ∼120,000 galaxies total. Upcoming spectroscopic surveys such as the Dark Energy Spectroscopic Survey (DESI; DESI Collab. et al., 2022) Bright Galaxy Survey (BGS; Hahn et al., 2022b) will observe *orders of magnitude* larger galaxy samples. With the millions of galaxies that they will observe, these surveys will provide an enormous number of galaxies that we can analyze even if we restrict our samples to galaxy populations robust to our modeling assumptions. Furthermore, these upcoming galaxy samples will have spectra, in addition to photometry, which will also provide more information per galaxy.

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.

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. 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.

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.

Davé, R., Anglés-Alcázar, D., Narayanan, D., Li, Q., Rafieferantsoa, M. H., and Appleby, S. SIMBA: Cosmological simulations with black hole growth and feedback. *Monthly Notices of the Royal Astronomical Society*, 486:2827–2849, June 2019. ISSN 0035-8711. doi: 10.1093/mnras/stz937.

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.

DESI Collab., Abareshi, B., Aguilar, J., Ahlen, S., Alam, S., Alexander, D. M., Alfarsy, R., Allen, L., Prieto, C. A., Alves, O., Ameel, J., Armengaud, E., Asorey, J., Aviles, A., Bailey, S., Balaguera-Antolínez, A., Ballester, O., Baltay, C., Bault, A., Beltran, S. F., Benavides, B., Ben-Zvi, S., Berti, A., Besuner, R., Beutler, F., Bianchi, D., Blake, C., Blanc, P., Blum, R., Bolton, A., Bose, S., Bramall, D., Brieden, S., Brodzeller, A., Brooks, D., Brownewell, C., Buckley-Geer, E., Cahn, R. N., Cai, Z., Canning, R., Rosell, A. C., Carton, P., Casas, R., Castander, F. J., Cervantes-Cota, J. L., Chabanier, S., Chaussidon, E., Chuang, C., Circosta, C., Cole, S., Cooper,

A. P., da Costa, L., Cousinou, M.-C., Cuceu, A., Davis, T. M., Dawson, K., de la Cruz-Noriega, R., de la Macorra, A., de Mattia, A., Della Costa, J., Demmer, P., Derwent, M., Dey, A., Dey, B., Dhungana, G., Ding, Z., Dobson, C., Doel, P., Donald-McCann, J., Donaldson, J., Douglass, K., Duan, Y., Dunlop, P., Edelstein, J., Eftekharzadeh, S., Eisenstein, D. J., Enriquez-Vargas, M., Escoffier, S., Evatt, M., Fagrelius, P., Fan, X., Fanning, K., Fawcett, V. A., Ferraro, S., Ereza, J., Flaugher, B., Font-Ribera, A., Forero-Romero, J. E., Frenk, C. S., Fromenteau, S., Gänsicke, B. T., Garcia-Quintero, C., Garrison, L., Gaztañaga, E., Gerardi, F., Gil-Marín, H., Gontcho, S. G. A., Gonzalez-Morales, A. X., Gonzalez-de-Rivera, G., Gonzalez-Perez, V., Gordon, C., Graur, O., Green, D., Grove, C., Gruen, D., Gutierrez, G., Guy, J., Hahn, C., Harris, S., Herrera, D., Herrera-Alcantar, H. K., Honscheid, K., Howlett, C., Huterer, D., Iršič, V., Ishak, M., Jelinsky, P., Jiang, L., Jimenez, J., Jing, Y. P., Joyce, R., Jullo, E., Juneau, S., Karaçaylı, N. G., Karamanis, M., Karcher, A., Karim, T., Kehoe, R., Kent, S., Kirkby, D., Kisner, T., Kitaura, F., Koposov, S. E., Kovács, A., Kremin, A., Krolewski, A., L'Huillier, B., Lahav, O., Lambert, A., Lamman, C., Lan, T.-W., Landriau, M., Lane, S., Lang, D., Lange, J. U., Lasker, J., Guillou, L. L., Leauthaud, A., Van Suu, A. L., Levi, M. E., Li, T. S., Magneville, C., Manera, M., Manser, C. J., Marshall, B., McCollam, W., McDonald, P., Meisner, A. M., Mezcua, J. M.-F. M., Miller, T., Miquel, R., Montero-Camacho, P., Moon, J., Martini, J. P., Meneses-Rizo, J., Moustakas, J., Mueller, E., Muñoz-Gutiérrez, A., Myers, A. D., Nadathur, S., Najita, J., Napolitano, L., Neilsen, E., Newman, J. A., Nie, J. D., Ning, Y., Niz, G., Norberg, P., Noriega, H. E., O'Brien, T., Obuljen, A., Palanque-Delabrouille, N., Palmese, A., Zhiwei, P., Pappalardo, D., Peng, X., Percival, W. J., Perruchot, S., Pogge, R., Poppett, C., Porredon, A., Prada, F., Prochaska, J., Pucha, R., Pérez-Fernández, A., Pérez-Ráfols, I., Rabinowitz, D., Raichoor, A., Ramirez-Solano, S., Ramírez-Pérez, C., Ravoux, C., Reil, K., Rezaie, M., Rocher, A., Rockosi, C., Roe, N. A., Roodman, A., Ross, A. J., Rossi, G., Ruggeri, R., Ruhlmann-Kleider, V., Sabiu, C. G., Safonova, S., Said, K., Saintonge, A., Catonga, J. S., Samushia, L., Sanchez, E., Saulder, C., Schaan, E., Schlafly, E., Schlegel, D., Schmoll, J., Scholte, D., Schubnell, M., Secroun, A., Seo, H., Serrano, S., Sharples, R. M., Sholl, M. J., Silber, J. H., Silva, D. R., Sirk, M., Siudek, M., Smith, A., Sprayberry, D., Staten, R., Stupak, B., Tan, T., Tarlé, G., Tie, S. S., Tojeiro, R., Ureña-López, L. A., Valdes, F., Valenzuela, O., Valluri, M., Vargas-Magaña, M., Verde, L., Walther, M., Wang, B., Wang, M. S., Weaver, B. A., Weaverdyck, C., Wechsler, R., Wilson, M. J., Yang, J., Yu, Y., Yuan, S., Yèche, C., Zhang, H., Zhang, K., Zhao, C., Zhou, R., Zhou, Z., Zou, H., Zou, J., Zou, S., and Zu, Y. Overview of the Instrumentation for the Dark Energy Spectroscopic Instrument, May 2022.

Echeverri, N., Villaescusa-Navarro, F., Chawak, C., Ni, Y., Hahn, C., Hernandez-Martinez, E., Teyssier, R., Angles-Alcazar, D., Dolag, K., and Castro, T. Cosmology with one galaxy? – The ASTRID model and robustness, April 2023.

Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J. Emcee: The MCMC Hammer. *Publications of the Astronomical Society of the Pacific*, 125:306, March 2013. ISSN 0004-6280. doi: 10.1086/670067.

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., Starkenburg, T. K., Choi, E., Davé, R., Dickey, C. M., Geha, M. C., Genel, S., Hayward, C. C., Maller, A. H., Mandyam, N., Pandya, V., Popping, G., Rafieferantsoa, M., Somerville, R. S., and Tinker, J. L. IQ-Collaboratory 1.1: The Star-forming Sequence of Simulated Central Galaxies. *The Astrophysical Journal*, 872 (2):160, February 2019. doi: 10.3847/1538-4357/aafedd.

Hahn, C., Starkenburg, T. K., Anglés-Alcázar, D., Choi, E., Davé, R., Dickey, C., Iyer, K. G., Maller, A. H., Somerville, R. S., Tinker, J. L., and Yung, L. Y. A. IQ Collaboratory III: The Empirical Dust Attenuation Framework – Taking Hydrodynamical Simulations with a Grain of Dust, June 2021.

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 2022a.

Hahn, C., Wilson, M. J., Ruiz-Macias, O., Cole, S., Weinberg, D. H., Moustakas, J., Kremin, A., Tinker, J. L., Smith, A., Wechsler, R. H., Ahlen, S., Alam, S., Bailey, S., Brooks, D., Cooper, A. P., Davis, T. M., Dawson, K., Dey, A., Dey, B., Eftekharzadeh, S., Eisenstein, D. J., Fanning, K., Forero-Romero, J. E., Frenk, C. S., Gaztañaga, E., Gontcho, S. G. A., Guy, J., Honscheid, K., Ishak, M., Juneau, S., Kehoe, R., Kisner, T., Lan, T.-W., Landriau, M., Le Guillou, L., Levi, M. E., Magneville, C., Martini, P., Meisner, A., Myers, A. D., Nie, J., Norberg, P., Palanque-Delabrouille, N., Percival, W. J., Poppett, C., Prada, F., Raichoor, A., Ross, A. J., Safonova, S., Saulder, C., Schlafly, E., Schlegel, D., Sierra-Porta, D., Tarle, G., Weaver, B. A., Yèche, C., Zarrouk, P., Zhou, R., Zhou, Z.,

and Zou, H. DESI Bright Galaxy Survey: Final Target Selection, Design, and Validation, August 2022b.

Jeffrey, N. and Wandelt, B. D. Solving high-dimensional parameter inference: Marginal posterior densities & Moment Networks. *arXiv:2011.05991 [astro-ph, stat]*, November 2020.

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.

Myers, A. T., McKee, C. F., Cunningham, A. J., Klein, R. I., and Krumholz, M. R. The Fragmentation of Magnetized, Massive Star-forming Cores with Radiative Feedback. *The Astrophysical Journal*, 766:97, April 2013. ISSN 0004-637X. doi: 10.1088/0004-637X/766/2/97.

Narayanan, D., Conroy, C., Davé, R., Johnson, B. D., and Popping, G. A Theory for the Variation of Dust Attenuation Laws in Galaxies. *The Astrophysical Journal*, 869:70, December 2018. ISSN 0004-637X. doi: 10.3847/1538-4357/aaed25.

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.

Pillepich, A., Springel, V., Nelson, D., Genel, S., Naiman, J., Pakmor, R., Hernquist, L., Torrey, P., Vogelsberger, M., Weinberger, R., and Marinacci, F. Simulating galaxy formation with the IllustrisTNG model. *Monthly Notices of the Royal Astronomical Society*, 473:4077–4106, January 2018. ISSN 0035-8711. doi: 10.1093/mnras/stx2656.

Smith, R. J., Lucey, J. R., and Conroy, C. The SINFONI Nearby Elliptical Lens Locator Survey: Discovery of two new low-redshift strong lenses and implications for the initial mass function in giant early-type galaxies. *Monthly Notices of the Royal Astronomical Society*, 449:3441–3457, June 2015. ISSN 0035-8711. doi: 10.1093/mnras/stv518.

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.

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.

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., Ding, J., Genel, S., Tonnesen, S., La Torre, V., Spergel, D. N., Teyssier, R., Li, Y., Heneka, C., Lemos, P., Anglés-Alcázar, D., Nagai, D., and Vogelsberger, M. Cosmology with One Galaxy? *The Astrophysical Journal*, 929:132, April 2022a. ISSN 0004-637X. doi: 10.3847/1538-4357/ac5d3f.

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 2022b.

Weinberger, R., Springel, V., Pakmor, R., Nelson, D., Genel, S., Pillepich, A., Vogelsberger, M., Marinacci, F., Naiman, J., Torrey, P., and Hernquist, L. Supermassive black holes and their feedback effects in the IllustrisTNG simulation. *Monthly Notices of the Royal Astronomical Society*, 479: 4056–4072, September 2018. ISSN 0035-8711. doi: 10.1093/mnras/sty1733.

White, S. D. M., Navarro, J. F., Evrard, A. E., and Frenk, C. S. The baryon content of galaxy clusters: A challenge to cosmological orthodoxy. *Nature*, 366(6454): 429–433, December 1993. ISSN 1476-4687. doi: 10.1038/366429a0.

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.

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. *The Astronomical Journal*, 161:262, June 2021. ISSN 0004-6256. doi: 10.3847/1538-3881/abf42e.

A. Observations: NASA-Sloan Atlas

In this work, we analyze observed galaxy photometry from the NASA-Sloan Atlas[2] (hereafter NSA). The NSA provides photometry of $z < 0.05$ galaxies observed by the Sloan Digital Sky Survey (SDSS; Aihara et al., 2011) Data Release 8 with improved background subtraction (Blanton et al., 2011). We use optical g, r, i, z band absolute magnitudes derived using KCORRECT (Blanton & Roweis, 2007), assuming a (Chabrier, 2003) initial mass function (IMF).

Out of the full NSA sample, we focus on luminous galaxies with $-18 > M_r > -22$. 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 outside 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 our simulated galaxies. 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.

B. Validating the Neural Density Estimator

Our posterior (Eq. 5) requires an accurate estimate of the individual posterior from NDE: $p(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i) \approx q_\phi(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i)$ (Sec. 3.2). To validate q_ϕ, we use 10% of the CAMELS-TNG data reserved for testing and two validation methods: (1) Simulation-Based Calibration (SBC) and (2) the "distance to random point" (DRP) coverage test.

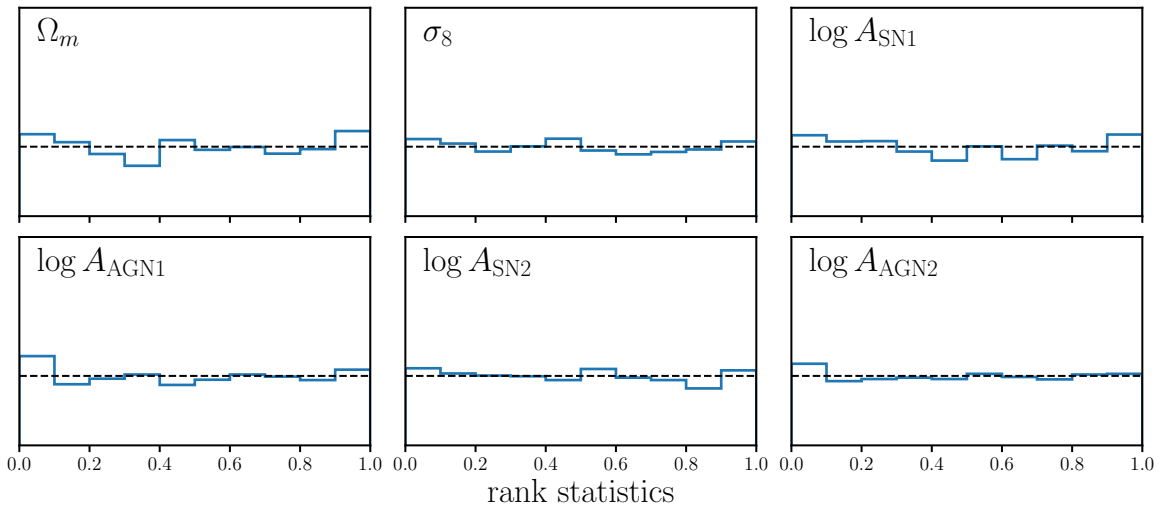

Figure 5. Simulation-based calibration plot of $q_\phi(\Omega, \mathcal{B} \,|\, \boldsymbol{X}_i)$ using 10% of the CAMELS-TNG data reserved for testing. The histogram in each panel represents the distribution of the rank statistic of the true value within the marginalized posterior (blue) for each parameter. The rank distribution is uniform for the true posterior (black dashed). The rank distribution of q_ϕ is nearly uniform for all Ω and $\mathcal{B}$ parameters. Therefore, it provides unbiased and accurate estimate of the true posterior.

Both are variations of the standard coverage test, where q_ϕ is applied to test samples not used for training. The posterior of each test sample is compared against the true parameter value its percentile score is calculated. Afterwards, cummulative distribution function (CDF) of the percentile is used to access the accuracy of q_ϕ. SBC is a modification of this standard coverage test that uses rank statistics rather than percentile score. It addresses the limitation that the CDFs only asymptotically approach the true values and that the discrete sampling of the posterior can cause artifacts in the CDFs. In Fig. 5, we present the SBC rank distributions of q_ϕ for Ω and $\mathcal{B}$ (blue). For the true posterior, rank distribution is uniform by construction (black dashed). The rank distributions are nearly uniform for all Ω and $\mathcal{B}$. Hence, we confirm that q_ϕ is in excellent agreement with the true posterior.

As additional validation, we also use the DRP coverage test from (Lemos et al., 2023). Instead of percentile scores or ranks, the DRP test assesses q_ϕ using samples drawn from q_ϕ for a test sample, the true parameter value of the test sample,

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

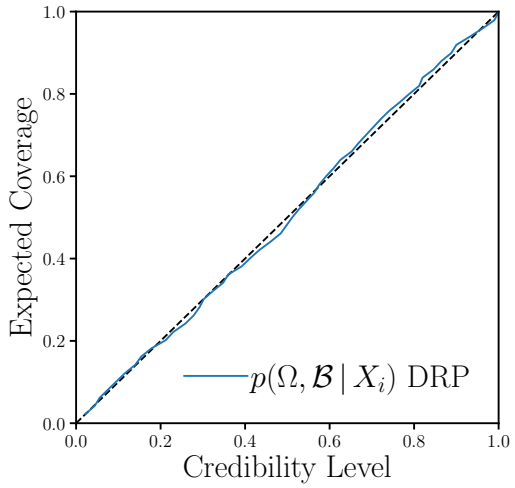

Figure 6. DRP coverage test validating the accuracy of our $q_\phi(\Omega, \mathcal{B} \mid \boldsymbol{X}_i)$ posterior estimate (blue). The DRP test is calculated using 10% of the CAMELS-TNG data reserved for testing. The black-dashed line represents an optimal estimate of the posterior. The DRP test demonstrates that q_ϕ provides a near optimal estimate of the true posterior.

and a random point in parameter space. It evalulates the distances between the q_ϕ samples and the random point. Then compares the distances to the distance between the true parameter value and the random point in order to derive an estimate of expected coverage probability. (Lemos et al., 2023) prove that this approach is necessary and sufficient to show that a posterior estimator is optimal. In Fig. 6, we present the DRP coverage test of q_ϕ (blue). Based ont he DRP test, q_ϕ provides a near optimal estimate of the true posterior (black-dashed).