Improving Posterior Inference of Galaxy Properties with Image-Based Conditional Flow Matching

Mikaeel Yunus
Department of Physics and Astronomy
Johns Hopkins University
Baltimore, MD 21218
myunus1@jh.edu

John F. Wu
Space Telescope Science Institute
Baltimore, MD 21218
jowu@stsci.edu

Benne W. Holwerda
Department of Physics and Astronomy
University of Louisville
Louisville, KY 40208
benne.holwerda@louisville.edu

Abstract

Estimating physical properties of galaxies from wide-field surveys remains a central challenge in astrophysics. While spectroscopy provides precise measurements, it is observationally expensive, and photometry discards morphological information that correlates with mass, star formation history, metallicity, and dust. We present a conditional flow matching (CFM) framework that leverages pixel-level imaging alongside photometry to improve posterior inference of galaxy properties. Using $\sim 10^5$ SDSS galaxies, we compare models trained on photometry alone versus photometry plus images. The image+photometry model outperforms the photometry-only model in posterior inference and more reliably recovers known scaling relations. Morphological information also helps mitigate the dust–age degeneracy. Our results highlight the potential of integrating morphology into photometric SED fitting pipelines, opening a pathway towards more accurate and physically informed constraints on galaxy properties.

1 Introduction

Modern galaxy surveys capture millions of objects with unprecedented depth and detail, but turning these observations into reliable physical properties can be difficult. The gold standard is spectroscopy, where a galaxy's light is dispersed into a detailed spectrum across many wavelengths, and absorption and emission features pin down the physics of the galaxy with high precision. However, acquiring spectra for every galaxy in a wide survey is observationally expensive.[1]

Broadband photometry – measuring the integrated flux of a galaxy through a small number of wide filters – offers a cheaper and more scalable alternative for measuring galaxy properties. With only a relatively small number of flux measurements across different filters, stellar population synthesis (SPS) methods can recover many physical parameters with high accuracy [7, 17]. However, this photometry-based approach discards valuable morphological information (e.g., spatial structure, color gradients, and overall visual appearance) that encodes properties such as stellar mass, star-formation history, and metallicity (e.g., [27, 2, 22]).

[1]The advent of wide-field, complex spectroscopic surveys such as DESI [6, 10] and PFS [14] is rapidly increasing the availability of spectra. Nevertheless, spectroscopy will remain limited in scale relative to imaging, making photometric surveys the primary avenue for estimating galaxy properties at scale.

Machine Learning and the Physical Sciences Workshop, NeurIPS 2025.

Since galaxy images contain a wealth of morphological information, it is natural to ask whether images can help improve galaxy property estimation. Recent work on this subject [11] has demonstrated that images can be used to generate optical spectra via generative models, from which galaxy properties can then be inferred. However, this approach relies on artificial spectra as an intermediate step before any galaxy physics can be constrained.

Alternatively, simulation-based inference (SBI) provides a framework to *directly* incorporate galaxy morphology into physical property estimation (e.g., [8]). While SBI has been applied to galaxy property inference [3, 15], only recently has imaging data been incorporated into such pipelines (e.g., [16]).

In this work, we test the hypothesis that explicitly including morphology from images in an SBI framework can sharpen galaxy property inference. To do so, we train two complementary models: a photometry-only baseline and a model augmented with latent representations of images. By comparing their posteriors, we directly assess the value of morphology for improving galaxy property estimates and potentially breaking astrophysical degeneracies.

2 Data

We use galaxies drawn from the Sloan Digital Sky Survey [28, 1] Main Galaxy Sample ($r < 17.78$; [24]). Starting from the SDSS `galSpecExtra` catalog of spectroscopically confirmed galaxies [18, 5, 26, 23], we remove objects with unphysical values and discard systems with model magnitudes brighter than $r < 16$ (see Figure A.1). All objects in our sample are bright, star-forming galaxies (based on BPT classification [4]). After filtering, our working sample contains $106{,}800$ galaxies, which we split $80/10/10$ into training, validation, and test sets ($85{,}440/10{,}680/10{,}680$).

We work with five physical property variables: stellar mass $M_\star$, star formation rate (SFR), gas-phase metallicity $Z_{\rm gas}$, the narrow 4000 Å break index $D_n(4000)$ (a proxy for stellar population age), and the V-band dust attenuation A_V. Figure A.1 shows the distributions of these properties in our sample.

We train models conditioned on SDSS *ugriz* photometry and 128×128 image cutouts. We download these *gri*-band cutouts from SDSS SkyServer at the native pixel scale ($0.396''$ pixel^{-1}). Each image is centered on the corresponding galaxy.

3 Methods

We aim to infer posterior distributions of five galaxy properties – stellar mass, star formation rate, metallicity, $D_n(4000)$, and dust attenuation – from SDSS observations using Conditional Flow Matching (CFM) [20, 25]. We train two complementary models: (i) a baseline that conditions only on *ugriz* photometry, and (ii) a model that augments photometry with morphological information extracted from images. Throughout, we refer to these as the **photometry model** and the **image model**, respectively.

Conditional Flow Matching. Let θ represent the galaxy properties and $\mathcal{D}$ the corresponding observational data (photometry and/or images). CFM learns a time-dependent velocity field $v_\phi(t, \theta, \mathcal{D})$ that transports a simple prior to the posterior $p(\theta \mid \mathcal{D})$. We use a linear interpolation path – i.e. we sample $\theta_0 \sim \mathcal{N}(0, I)$, $t \sim \mathcal{U}(0, 1)$, $\epsilon \sim \mathcal{N}(0, I)$ and form

$$\theta_t = (1 - t)\,\theta_0 + t\,\theta_1 + \sigma\,\epsilon,$$

where θ_0 is drawn from a Gaussian prior, θ_1 denotes the target properties for each training example, and σ is chosen to be 0.05. This results in the probability path $p_t(\theta \mid \mathcal{D}, \theta_0, \theta_1) = \mathcal{N}\big(\theta;\ (1-t)\theta_0 + t\theta_1,\ \sigma^2 I\big)$ [13]. We train with an MSE loss to the time-independent target velocity $\theta_1 - \theta_0$. At test time we integrate $\dot\theta = v_\phi(t, \theta, \mathcal{D})$ from t=0 to 1 using the fourth-order Runge-Kutta method (RK4) with 100 steps, drawing 1000 trajectories per object to approximate the posterior.

Architectures. For our **photometry** model, the velocity network is an MLP (three layers, width 256) whose input is the concatenation $[t; \theta; f_{\rm phot}]$, where $f_{\rm phot} \in \mathbb{R}^5$ is the normalized photometry. The input dimensionality is $1 + d_\theta + 5 = 11$ with $d_\theta = 5$. For our **image** model, we encode a 128×128 RGB cutout with a CNN comprising four stride-1 convolutional blocks with average pooling, followed by global average pooling, yielding a 256-D representation $f_{\rm img}$. We use average pooling to preserve

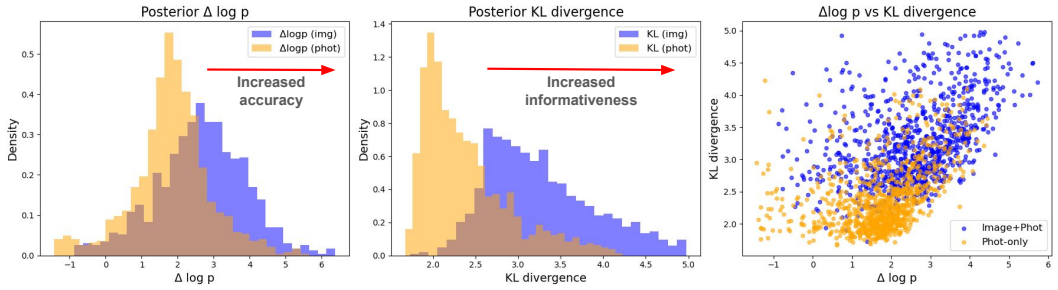

Figure 1: Left: distributions of $\Delta \log p(\theta_*; \mathcal{D})$ for the image model (purple, $\mu = 2.17,\ \sigma = 3.30$) and the photometry model (yellow, $\mu = 1.26,\ \sigma = 3.98$). Middle: distributions of $D_{\mathrm{KL}}[p(\theta \,|\, \mathcal{D}) \,\|\, p(\theta)]$ for the image model (purple, $\mu = 3.41,\ \sigma = 0.95$) and the photometry model (yellow, $\mu = 2.55,\ \sigma = 0.97$). Right: per-object scatter plot of $\Delta \log p$ vs. D_{KL} (color indicates model). Panels show $N = 1000$ galaxies from the test set. For these objects, the image model attains **higher** $\Delta \log p$ **for 81.5%** of objects and **higher** D_{KL} **for 96.5%** of objects compared to the photometry model.

extended light distributions, which are informative for galaxy property estimation. We then condition the velocity MLP (three layers, width 256) on $[t; \theta; f_{\mathrm{img}}; f_{\mathrm{phot}}]$ (input size $1{+}d_\theta{+}256{+}5 = 267$).

Training Details. Models are implemented in PyTorch and trained with AdamW (learning rate 5×10^{-5}, batch size 64) with early stopping on validation loss. All continuous variables (photometry and properties) are standardized to zero mean and unit variance; inverse transformations are applied to samples for evaluation. We hold out a separate test set, used only for posterior evaluation. Training uses PyTorch `DataParallel` on four NVIDIA V100 GPUs.

Posterior Performance Metrics. We evaluate our posteriors using two complementary statistics.

Accuracy: For each object we compute $\Delta \log p(\theta_*; \mathcal{D}) = \log p(\theta_* | \mathcal{D}) - \log p(\theta_*)$, where θ_* are the target parameters for that object, and $\mathcal{D}$ is the conditioning data. Positive values mean the posterior assigns higher density to the target than the prior does, i.e. a per-object Bayes factor gain. (Note that the "prior" here is the empirical distribution of properties in the spectroscopic dataset, not the Gaussian prior used in CFM training.)

Informativeness: We also measure the Kullback-Leibler divergence $D_{\mathrm{KL}}[p(\theta|\mathcal{D}) \,\|\, p(\theta)]$, quantifying how different the posterior distribution is from the prior distribution. Averaging D_{KL} over $\mathcal{D}$ yields the mutual information $I(\theta; \mathcal{D})$, a population-level summary of information gain.

4 Results

We compare posterior quality between the **image** and **photometry** models using the two metrics defined in Section 3: (i) the *accuracy* statistic $\Delta \log p(\theta_*; \mathcal{D})$, evaluated at the target θ_*, and (ii) the *informativeness* statistic $D_{\mathrm{KL}}[p(\theta \,|\, \mathcal{D}) \,\|\, p(\theta)]$, which measures departure from the empirical (dataset) prior. Figure 1 summarizes these comparisons.

The image model shifts $\Delta \log p$ rightward relative to the photometry baseline, indicating higher posterior density at the target for more galaxies (higher per-object Bayes-factor gain). Simultaneously, D_{KL} increases, showing that the image-conditioned posteriors move further from the dataset prior. The joint trend in the right panel demonstrates that informativeness and accuracy both increase: objects with larger D_{KL} typically also exhibit larger $\Delta \log p$. A small number of outliers with negative $\Delta \log p$ remain, but since these appear in both models, they likely reflect CFM architecture limitations rather than the inclusion of morphological information in the image model.

In addition to our per-object posterior accuracy metrics, we evaluate how well each model reproduces the *population-level* distribution of physical parameters. Specifically, for each galaxy property variable, we compute the Wasserstein distance (WD) between (i) its one-dimensional marginal distribution in the test set and (ii) the distribution of posterior means obtained from each model across the test set. This metric evaluates whether the model recovers the global population distribution. Smaller distances indicate that the model's predictions, aggregated across the test set, match the true population more closely. Table 1 reports these Wasserstein distances for both the image model and the photometry model.

Property	WD (Image Model)	WD (Photometry Model)	Difference
$M_\star$	0.0264	0.0547	0.0283
SFR	0.0639	0.1119	0.0480
$Z_{\rm gas}$	0.0156	0.0302	0.0146
$D_n(4000)$	0.0103	0.0131	0.0028
A_V	0.1937	0.2565	0.0628

Table 1: Wasserstein distances between the marginal test-set distribution of each variable and the distribution of posterior means predicted by the image model and the photometry model. Lower values indicate better agreement with the true population distribution. The rightmost column shows the absolute improvement (reduction in WD) achieved by adding morphology.

Across all parameters, the image model achieves substantially smaller Wasserstein distances than the photometry model, indicating that morphological information improves not only per-object posterior accuracy and informativeness, but also fidelity to population-level marginal distributions in the test set.

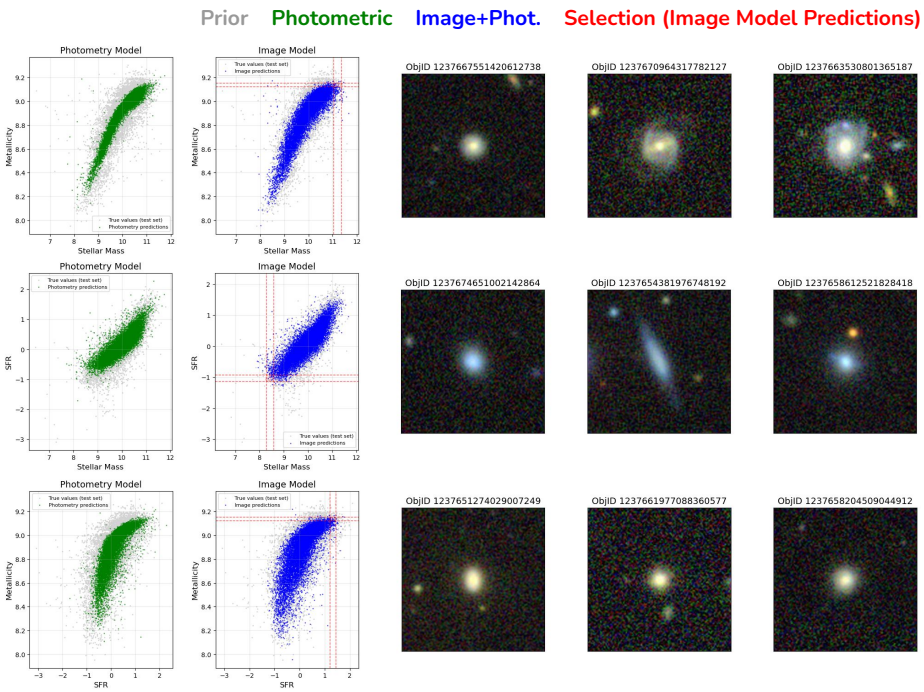

Figure 2: Image-model predictions (blue) recover known scaling relations in SDSS data (gray) more faithfully than the photometry model (green). Each row shows a different relation: $M_\star$–$Z_{\rm gas}$ (top), $M_\star$–SFR (middle), and SFR–$Z_{\rm gas}$ (bottom). Red boxes mark selections defined on the image-model predictions, with three representative galaxy cutouts shown on the right.

In Figure 2, we directly compare scaling relations (see [21, 9, 12, 26, 19]) recovered by the two models. The top, middle, and bottom rows show photometry-model and image-model predictions on our test set in the $M_\star$–$Z_{\rm gas}$, $M_\star$–SFR, and SFR–$Z_{\rm gas}$ planes, respectively. These predictions are obtained for each galaxy by drawing 1000 samples from the appropriate CFM model, finding the mean, and displaying it as a point in the appropriate scatterplot. In each case, the image model (blue) better reproduces known galaxy scaling relations represented by SDSS data (gray) than the photometry model (green).

In each row of Figure 2, we also highlight selection boxes (red) defined on the image model predictions and display three galaxy cutouts from each selection. These examples show that the image model yields visually coherent samples aligned with astrophysical expectations – for instance, low-$M_\star$, low-SFR galaxies appear blue and diffuse, as seen in the middle row.

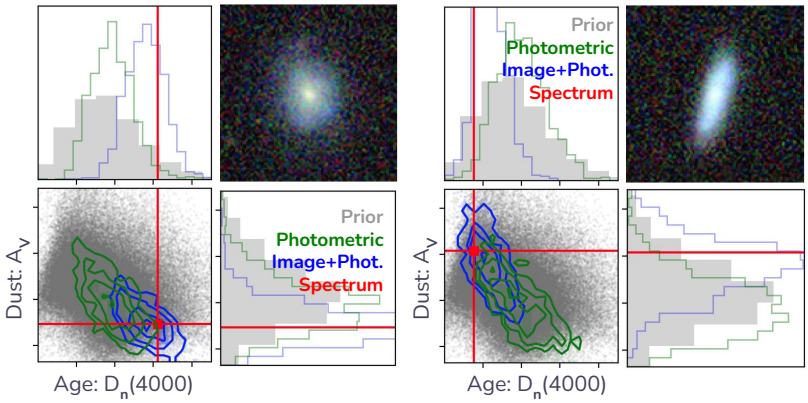

Figure 3: Corner plots of A_V (dust attenuation) versus $D_n(4000)$ (stellar age proxy) for two representative galaxies. *Left:* old, dust-poor galaxy. *Right:* young, dust-rich galaxy. Contours show posterior predictions for the photometry model (green) and image model (blue); red crosses mark spectroscopic targets. The image model moves closer to the spectroscopic target than the photometry model does in both cases, indicating potential to weaken the dust–age degeneracy.

5 Discussion

We have developed an image-conditioned flow matching model that outperforms a photometry-only model, and more accurately reproduces known galaxy scaling relations. This supports our aim to leverage morphological information to constrain galaxy properties.

Long-standing degeneracies in galaxy physics, such as the dust–age degeneracy, provide another motivation for incorporating morphology into posterior inference of galaxy properties. Figure 3 depicts two illustrative cases: an old, dust-poor system (left) and a young, dust-rich system (right). In both examples, the image model (blue) shifts closer to the spectroscopic target (red) compared to the photometry-only model (green), suggesting that morphology can provide improvements even when A_V is difficult to constrain. However, overall constraints on A_V remain weak, and we therefore only partially disentangle A_V from $D_n(4000)$. Nonetheless, Figure 3 highlights the potential of our approach to break dust–age degeneracies once integrated into a more comprehensive pipeline.

Looking ahead, we plan to combine physics-based photometric SED fitting with our flow-based framework in order to incorporate morphological information into galaxy property inference. Traditional SED fitting encodes physics-based priors and captures parameter extremes well, but it has not yet offered a statistical pathway to leverage morphology. Our work can provide such a pathway. Future pipelines that merge SED models with SBI can retain the physical interpretability of SED fitting while using morphology to supply missing information, thereby yielding more accurate and comprehensive galaxy property estimates.

References

[1] Kevork N. Abazajian, Jennifer K. Adelman-McCarthy, Marcel A. Agüeros, Sahar S. Allam, Carlos Allende Prieto, Deokkeun An, Kurt S. J. Anderson, Scott F. Anderson, James Annis, Neta A. Bahcall, C. A. L. Bailer-Jones, J. C. Barentine, Bruce A. Bassett, Andrew C. Becker, Timothy C. Beers, Eric F. Bell, Vasily Belokurov, Andreas A. Berlind, Eileen F. Berman, Mariangela Bernardi, Steven J. Bickerton, Dmitry Bizyaev, John P. Blakeslee, Michael R. Blanton, John J. Bochanski, William N. Boroski, Howard J. Brewington, Jarle Brinchmann, J. Brinkmann, Robert J. Brunner, Tamás Budavári, Larry N. Carey, Samuel Carliles, Michael A. Carr, Francisco J. Castander, David Cinabro, A. J. Connolly, István Csabai, Carlos E. Cunha, Paul C. Czarapata, James R. A. Davenport, Ernst de Haas, Ben Dilday, Mamoru Doi, Daniel J. Eisenstein, Michael L. Evans, N. W. Evans, Xiaohui Fan, Scott D. Friedman, Joshua A. Frieman, Masataka Fukugita, Boris T. Gänsicke, Evalyn Gates, Bruce Gillespie, G. Gilmore, Belinda Gonzalez, Carlos F. Gonzalez, Eva K. Grebel, James E. Gunn, Zsuzsanna György, Patrick B. Hall, Paul Harding, Frederick H. Harris, Michael Harvanek, Suzanne L. Hawley, Jeffrey J. E. Hayes, Timothy M. Heckman, John S. Hendry, Gregory S. Hennessy, Robert B. Hindsley, J. Hoblitt, Craig J. Hogan, David W. Hogg, Jon A. Holtzman, Joseph B. Hyde, Shin-ichi Ichikawa, Takashi Ichikawa, Myungshin Im, Željko Ivezić, Sebastian Jester, Linhua Jiang, Jennifer A. Johnson, Anders M. Jorgensen, Mario Jurić, Stephen M. Kent, R. Kessler, S. J. Kleinman, G. R. Knapp, Kohki Konishi, Richard G. Kron, Jurek Krzesinski, Nikolay Kuropatkin, Hubert Lampeitl, Svetlana Lebedeva, Myung Gyoon Lee, Young Sun Lee, R. French Leger, Sébastien Lépine, Nolan Li, Marcos Lima, Huan Lin, Daniel C. Long, Craig P. Loomis, Jon Loveday, Robert H. Lupton, Eugene Magnier, Olena Malanushenko, Viktor Malanushenko, Rachel Mandelbaum, Bruce Margon, John P. Marriner, David Martínez-Delgado, Takahiko Matsubara, Peregrine M. McGehee, Timothy A. McKay, Avery Meiksin, Heather L. Morrison, Fergal Mullally, Jeffrey A. Munn, Tara Murphy, Thomas Nash, Ada Nebot, Jr. Neilsen, Eric H., Heidi Jo Newberg, Peter R. Newman, Robert C. Nichol, Tom Nicinski, Maria Nieto-Santisteban, Atsuko Nitta, Sadanori Okamura, Daniel J. Oravetz, Jeremiah P. Ostriker, Russell Owen, Nikhil Padmanabhan, Kaike Pan, Changbom Park, George Pauls, Jr. Peoples, John, Will J. Percival, Jeffrey R. Pier, Adrian C. Pope, Dimitri Pourbaix, Paul A. Price, Norbert Purger, Thomas Quinn, M. Jordan Raddick, Paola Re Fiorentin, Gordon T. Richards, Michael W. Richmond, Adam G. Riess, Hans-Walter Rix, Constance M. Rockosi, Masao Sako, David J. Schlegel, Donald P. Schneider, Ralf-Dieter Scholz, Matthias R. Schreiber, Axel D. Schwope, Uroš Seljak, Branimir Sesar, Erin Sheldon, Kazu Shimasaku, Valena C. Sibley, A. E. Simmons, Thirupathi Sivarani, J. Allyn Smith, Martin C. Smith, Vernesa Smolčić, Stephanie A. Snedden, Albert Stebbins, Matthias Steinmetz, Chris Stoughton, Michael A. Strauss, Mark SubbaRao, Yasushi Suto, Alexander S. Szalay, István Szapudi, Paula Szkody, Masayuki Tanaka, Max Tegmark, Luis F. A. Teodoro, Aniruddha R. Thakar, Christy A. Tremonti, Douglas L. Tucker, Alan Uomoto, Daniel E. Vanden Berk, Jan Vandenberg, S. Vidrih, Michael S. Vogeley, Wolfgang Voges, Nicole P. Vogt, Yogesh Wadadekar, Shannon Watters, David H. Weinberg, Andrew A. West, Simon D. M. White, Brian C. Wilhite, Alainna C. Wonders, Brian Yanny, D. R. Yocum, Donald G. York, Idit Zehavi, Stefano Zibetti, and Daniel B. Zucker. The Seventh Data Release of the Sloan Digital Sky Survey. *The Astrophysical Journal Supplement Series*, 182(2):543–558, June 2009.

[2] Juan Pablo Alfonzo, Kartheik G. Iyer, Masayuki Akiyama, Greg L. Bryan, Suchetha Cooray, Eric Ludwig, Lamiya Mowla, Kiyoaki C. Omori, Camilla Pacifici, Joshua S. Speagle, and John F. Wu. Katachi: Decoding the Imprints of Past Star Formation on Present-day Morphology in Galaxies with Interpretable CNNs. *The Astrophysical Journal*, 967(2):152, June 2024.

[3] Justin Alsing, Hiranya Peiris, Joel Leja, ChangHoon Hahn, Rita Tojeiro, Daniel Mortlock, Boris Leistedt, Benjamin D. Johnson, and Charlie Conroy. Speculator: Emulating stellar population synthesis for fast and accurate galaxy spectra and photometry. *The Astrophysical Journal Supplement Series*, 249(1):5, June 2020.

[4] J. A. Baldwin, M. M. Phillips, and R. Terlevich. Classification parameters for the emission-line spectra of extragalactic objects. *Publications of the Astronomical Society of the Pacific*, 93:5–19, February 1981.

[5] J. Brinchmann, S. Charlot, S. D. M. White, C. Tremonti, G. Kauffmann, T. Heckman, and J. Brinkmann. The physical properties of star-forming galaxies in the low-redshift Universe. *Monthly Notices of the Royal Astronomical Society*, 351(4):1151–1179, July 2004.

[6] DESI Collaboration, B. Abareshi, J. Aguilar, S. Ahlen, Shadab Alam, David M. Alexander, R. Alfarsy, L. Allen, C. Allende Prieto, O. Alves, J. Ameel, E. Armengaud, J. Asorey, Alejandro Aviles, S. Bailey, A. Balaguera-Antolínez, O. Ballester, C. Baltay, A. Bault, S. F. Beltran, B. Benavides, S. BenZvi, A. Berti, R. Besuner, Florian Beutler, D. Bianchi, C. Blake, P. Blanc, R. Blum, A. Bolton, S. Bose, D. Bramall, S. Brieden, A. Brodzeller, D. Brooks, C. Brownewell, E. Buckley-Geer, R. N. Cahn, Z. Cai, R. Canning, R. Capasso, A. Carnero Rosell, P. Carton, R. Casas, F. J. Castander, J. L. Cervantes-Cota, S. Chabanier, E. Chaussidon, C. Chuang, C. Circosta, S. Cole, A. P. Cooper, L. da Costa, M.-C. Cousinou, A. Cuceu, T. M. Davis, K. Dawson, R. de la Cruz-Noriega, A. de la Macorra, A. de Mattia, J. Della Costa, P. Demmer, M. Derwent, A. Dey, B. Dey, G. Dhungana, Z. Ding, C. Dobson, P. Doel, J. Donald-McCann, J. Donaldson, K. Douglass, Y. Duan, P. Dunlop, J. Edelstein, S. Eftekharzadeh, D. J. Eisenstein, M. Enriquez-Vargas, S. Escoffier, M. Evatt, P. Fagrelius, X. Fan, K. Fanning, V. A. Fawcett, S. Ferraro, J. Ereza, B. Flaugher, A. Font-Ribera, J. E. Forero-Romero, C. S. Frenk, S. Fromenteau, B. T. Gänsicke, C. Garcia-Quintero, L. Garrison, E. Gaztañaga, F. Gerardi, H. Gil-Marín, S. Gontcho A Gontcho, Alma X. Gonzalez-Morales, G. Gonzalez-de Rivera, V. Gonzalez-Perez, C. Gordon, O. Graur, D. Green, C. Grove, D. Gruen, G. Gutierrez, J. Guy, C. Hahn, S. Harris, D. Herrera, Hiram K. Herrera-Alcantar, K. Honscheid, C. Howlett, D. Huterer, V. Iršič, M. Ishak, P. Jelinsky, L. Jiang, J. Jimenez, Y. P. Jing, R. Joyce, E. Jullo, S. Juneau, N. G. Karaçaylı, M. Karamanis, A. Karcher, T. Karim, R. Kehoe, S. Kent, D. Kirkby, T. Kisner, F. Kitaura, S. E. Koposov, A. Kovács, A. Kremin, Alex Krolewski, B. L'Huillier, O. Lahav, A. Lambert, C. Lamman, Ting-Wen Lan, M. Landriau, S. Lane, D. Lang, J. U. Lange, J. Lasker, L. Le Guillou, A. Leauthaud, A. Le Van Suu, Michael E. Levi, T. S. Li, C. Magneville, M. Manera, Christopher J. Manser, B. Marshall, Paul Martini, W. McCollam, P. McDonald, Aaron M. Meisner, J. Mena-Fernández, J. Meneses-Rizo, M. Mezcua, T. Miller, R. Miquel, P. Montero-Camacho, J. Moon, J. Moustakas, E. Mueller, Andrea Muñoz-Gutiérrez, Adam D. Myers, S. Nadathur, J. Najita, L. Napolitano, E. Neilsen, Jeffrey A. Newman, J. D. Nie, Y. Ning, G. Niz, P. Norberg, Hernán E. Noriega, T. O'Brien, A. Obuljen, N. Palanque-Delabrouille, A. Palmese, P. Zhiwei, D. Pappalardo, X. PENG, W. J. Percival, S. Perruchot, R. Pogge, C. Poppett, A. Porredon, F. Prada, J. Prochaska, R. Pucha, A. Pérez-Fernández, I. Pérez-Ràfols, D. Rabinowitz, A. Raichoor, S. Ramirez-Solano, César Ramírez-Pérez, C. Ravoux, K. Reil, M. Rezaie, A. Rocher, C. Rockosi, N. A. Roe, A. Roodman, A. J. Ross, G. Rossi, R. Ruggeri, V. Ruhlmann-Kleider, C. G. Sabiu, S. Gaines, K. Said, A. Saintonge, Javier Salas Catonga, L. Samushia, E. Sanchez, C. Saulder, E. Schaan, E. Schlafly, D. Schlegel, J. Schmoll, D. Scholte, M. Schubnell, A. Secroun, H. Seo, S. Serrano, Ray M. Sharples, Michael J. Sholl, Joseph Harry Silber, D. R. Silva, M. Sirk, M. Siudek, A. Smith, D. Sprayberry, R. Staten, B. Stupak, T. Tan, Gregory Tarlé, Suk Sien Tie, R. Tojeiro, L. A. Ureña-López, F. Valdes, O. Valenzuela, M. Valluri, M. Vargas-Magaña, L. Verde, M. Walther, B. Wang, M. S. Wang, B. A. Weaver, C. Weaverdyck, R. Wechsler, Michael J. Wilson, J. Yang, Y. Yu, S. Yuan, Christophe Yèche, H. Zhang, K. Zhang, Cheng Zhao, Rongpu Zhou, Zhimin Zhou, H. Zou, J. Zou, S. Zou, and Y. Zu. Overview of the instrumentation for the dark energy spectroscopic instrument. *The Astronomical Journal*, 164(5):207, October 2022.

[7] Charlie Conroy. Modeling the panchromatic spectral energy distributions of galaxies. *Annual Review of Astronomy and Astrophysics*, 51(1):393–455, aug 2013.

[8] Kyle Cranmer, Johann Brehmer, and Gilles Louppe. The frontier of simulation-based inference. *Proceedings of the National Academy of Sciences*, 117(48):30055–30062, 2020.

[9] E. Daddi, M. Dickinson, G. Morrison, R. Chary, A. Cimatti, D. Elbaz, D. Frayer, A. Renzini, A. Pope, D. M. Alexander, F. E. Bauer, M. Giavalisco, M. Huynh, J. Kurk, and M. Mignoli. Multiwavelength Study of Massive Galaxies at z$\sim$2. I. Star Formation and Galaxy Growth. *The Astrophysical Journal*, 670(1):156–172, November 2007.

[10] Arjun Dey, David J. Schlegel, Dustin Lang, Robert Blum, Kaylan Burleigh, Xiaohui Fan, Joseph R. Findlay, Doug Finkbeiner, David Herrera, Stéphanie Juneau, Martin Landriau, Michael Levi, Ian McGreer, Aaron Meisner, Adam D. Myers, John Moustakas, Peter Nugent, Anna Patej, Edward F. Schlafly, Alistair R. Walker, Francisco Valdes, Benjamin A. Weaver,

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

[11] Lars Doorenbos, Eva Sextl, Kevin Heng, Stefano Cavuoti, Massimo Brescia, Olena Torbaniuk, Giuseppe Longo, Raphael Sznitman, and Pablo Márquez-Neila. Galaxy spectroscopy without spectra: Galaxy properties from photometric images with conditional diffusion models. *The Astrophysical Journal*, 977(1):131, dec 2024.

[12] D. Elbaz, E. Daddi, D. Le Borgne, M. Dickinson, D. M. Alexander, R. R. Chary, J. L. Starck, W. N. Brandt, M. Kitzbichler, E. MacDonald, M. Nonino, P. Popesso, D. Stern, and E. Vanzella. The reversal of the star formation-density relation in the distant universe. *Astronomy & Astrophysics*, 468(1):33–48, June 2007.

[13] Anne Gagneux, Ségolène Martin, Rémi Emonet, Quentin Bertrand, and Mathurin Massias. A visual dive into conditional flow matching. In *ICLR Blogposts 2025*, 2025. https://dl.heeere.com/conditional-flow-matching/blog/conditional-flow-matching/.

[14] Jenny Greene, Rachel Bezanson, Masami Ouchi, John Silverman, and the PFS Galaxy Evolution Working Group. The prime focus spectrograph galaxy evolution survey, 2022.

[15] ChangHoon Hahn and Peter Melchior. Accelerated bayesian sed modeling using amortized neural posterior estimation. *The Astrophysical Journal*, 938(1):11, October 2022.

[16] Patricia Iglesias-Navarro, Marc Huertas-Company, Pablo Pérez-González, Johan H. Knapen, ChangHoon Hahn, Anton M. Koekemoer, Steven L. Finkelstein, Natalia Villanueva, and Andrés Asensio Ramos. Simulation-based inference of galaxy properties from jwst pixels, 2025.

[17] Kartheik G. Iyer, Camilla Pacifici, Gabriela Calistro-Rivera, and Christopher C. Lovell. The Spectral Energy Distributions of Galaxies. *arXiv e-prints*, page arXiv:2502.17680, February 2025.

[18] Guinevere Kauffmann, Timothy M. Heckman, Christy Tremonti, Jarle Brinchmann, Stéphane Charlot, Simon D. M. White, Susan E. Ridgway, Jon Brinkmann, Masataka Fukugita, Patrick B. Hall, Željko Ivezić, Gordon T. Richards, and Donald P. Schneider. The host galaxies of active galactic nuclei. *Monthly Notices of the Royal Astronomical Society*, 346(4):1055–1077, December 2003.

[19] M. A. Lara-Lopez et al. Galaxy And Mass Assembly (GAMA): A deeper view of the mass, metallicity, and SFR relationships. *Mon. Not. Roy. Astron. Soc.*, 434:451, 2013.

[20] Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling, 2023.

[21] K. G. Noeske, B. J. Weiner, S. M. Faber, C. Papovich, D. C. Koo, R. S. Somerville, K. Bundy, C. J. Conselice, J. A. Newman, D. Schiminovich, E. Le Floc'h, A. L. Coil, G. H. Rieke, J. M. Lotz, J. R. Primack, P. Barmby, M. C. Cooper, M. Davis, R. S. Ellis, G. G. Fazio, P. Guhathakurta, J. Huang, S. A. Kassin, D. C. Martin, A. C. Phillips, R. M. Rich, T. A. Small, C. N. A. Willmer, and G. Wilson. Star Formation in AEGIS Field Galaxies since z=1.1: The Dominance of Gradually Declining Star Formation, and the Main Sequence of Star-forming Galaxies. *The Astrophysical Journal*, 660(1):L43–L46, May 2007.

[22] Liam Parker, Francois Lanusse, Siavash Golkar, Leopoldo Sarra, Miles Cranmer, Alberto Bietti, Michael Eickenberg, Geraud Krawezik, Michael McCabe, Rudy Morel, Ruben Ohana, Mariel Pettee, Bruno Régaldo-Saint Blancard, Kyunghyun Cho, Shirley Ho, and Polymathic AI Collaboration. AstroCLIP: a cross-modal foundation model for galaxies. *Monthly Notices of the Royal Astronomical Society*, 531(4):4990–5011, July 2024.

[23] Samir Salim, Stéphane Charlot, R. Michael Rich, Guinevere Kauffmann, Timothy M. Heckman, Tom A. Barlow, Luciana Bianchi, Yong-Ik Byun, Jose Donas, Karl Forster, Peter G. Friedman, Patrick N. Jelinsky, Young-Wook Lee, Barry F. Madore, Roger F. Malina, D. Christopher Martin, Bruno Milliard, Patrick Morrissey, Susan G. Neff, David Schiminovich, Mark Seibert, Oswald H. W. Siegmund, Todd Small, Alex S. Szalay, Barry Y. Welsh, and Ted K. Wyder. New Constraints on the Star Formation Histories and Dust Attenuation of Galaxies in the Local Universe from GALEX. *The Astrophysical Journal Letters*, 619(1):L39–L42, January 2005.

[24] Michael A. Strauss, David H. Weinberg, Robert H. Lupton, Vijay K. Narayanan, James Annis, Mariangela Bernardi, Michael Blanton, Scott Burles, A. J. Connolly, Julianne Dalcanton, Mamoru Doi, Daniel Eisenstein, Joshua A. Frieman, Masataka Fukugita, James E. Gunn, Željko Ivezić, Stephen Kent, Rita S. J. Kim, G. R. Knapp, Richard G. Kron, Jeffrey A. Munn, Heidi Jo Newberg, R. C. Nichol, Sadanori Okamura, Thomas R. Quinn, Michael W. Richmond, David J. Schlegel, Kazuhiro Shimasaku, Mark SubbaRao, Alexander S. Szalay, Dan Vanden Berk, Michael S. Vogeley, Brian Yanny, Naoki Yasuda, Donald G. York, and Idit Zehavi. Spectroscopic Target Selection in the Sloan Digital Sky Survey: The Main Galaxy Sample. *The Astronomical Journal*, 124(3):1810–1824, September 2002.

[25] Alexander Tong, Kilian Fatras, Nikolay Malkin, Guillaume Huguet, Yanlei Zhang, Jarrid Rector-Brooks, Guy Wolf, and Yoshua Bengio. Improving and generalizing flow-based generative models with minibatch optimal transport, 2024.

[26] Christy A. Tremonti, Timothy M. Heckman, Guinevere Kauffmann, Jarle Brinchmann, Stéphane Charlot, Simon D. M. White, Mark Seibert, Eric W. Peng, David J. Schlegel, Alan Uomoto, Masataka Fukugita, and Jon Brinkmann. The Origin of the Mass-Metallicity Relation: Insights from 53,000 Star-forming Galaxies in the Sloan Digital Sky Survey. *The Astrophysical Journal*, 613(2):898–913, October 2004.

[27] John F Wu and Steven Boada. Using convolutional neural networks to predict galaxy metallicity from three-colour images. *Monthly Notices of the Royal Astronomical Society*, 484(4):4683–4694, February 2019.

[28] Donald G. York, J. Adelman, Jr. Anderson, John E., Scott F. Anderson, James Annis, Neta A. Bahcall, J. A. Bakken, Robert Barkhouser, Steven Bastian, Eileen Berman, William N. Boroski, Steve Bracker, Charlie Briegel, John W. Briggs, J. Brinkmann, Robert Brunner, Scott Burles, Larry Carey, Michael A. Carr, Francisco J. Castander, Bing Chen, Patrick L. Colestock, A. J. Connolly, J. H. Crocker, István Csabai, Paul C. Czarapata, John Eric Davis, Mamoru Doi, Tom Dombeck, Daniel Eisenstein, Nancy Ellman, Brian R. Elms, Michael L. Evans, Xiaohui Fan, Glenn R. Federwitz, Larry Fiscelli, Scott Friedman, Joshua A. Frieman, Masataka Fukugita, Bruce Gillespie, James E. Gunn, Vijay K. Gurbani, Ernst de Haas, Merle Haldeman, Frederick H. Harris, J. Hayes, Timothy M. Heckman, G. S. Hennessy, Robert B. Hindsley, Scott Holm, Donald J. Holmgren, Chi-hao Huang, Charles Hull, Don Husby, Shin-Ichi Ichikawa, Takashi

Ichikawa, Željko Ivezić, Stephen Kent, Rita S. J. Kim, E. Kinney, Mark Klaene, A. N. Kleinman, S. Kleinman, G. R. Knapp, John Korienek, Richard G. Kron, Peter Z. Kunszt, D. Q. Lamb, B. Lee, R. French Leger, Siriluk Limmongkol, Carl Lindenmeyer, Daniel C. Long, Craig Loomis, Jon Loveday, Rich Lucinio, Robert H. Lupton, Bryan MacKinnon, Edward J. Mannery, P. M. Mantsch, Bruce Margon, Peregrine McGehee, Timothy A. McKay, Avery Meiksin, Aronne Merelli, David G. Monet, Jeffrey A. Munn, Vijay K. Narayanan, Thomas Nash, Eric Neilsen, Rich Neswold, Heidi Jo Newberg, R. C. Nichol, Tom Nicinski, Mario Nonino, Norio Okada, Sadanori Okamura, Jeremiah P. Ostriker, Russell Owen, A. George Pauls, John Peoples, R. L. Peterson, Donald Petravick, Jeffrey R. Pier, Adrian Pope, Ruth Pordes, Angela Prosapio, Ron Rechenmacher, Thomas R. Quinn, Gordon T. Richards, Michael W. Richmond, Claudio H. Rivetta, Constance M. Rockosi, Kurt Ruthmansdorfer, Dale Sandford, David J. Schlegel, Donald P. Schneider, Maki Sekiguchi, Gary Sergey, Kazuhiro Shimasaku, Walter A. Siegmund, Stephen Smee, J. Allyn Smith, S. Snedden, R. Stone, Chris Stoughton, Michael A. Strauss, Christopher Stubbs, Mark SubbaRao, Alexander S. Szalay, Istvan Szapudi, Gyula P. Szokoly, Anirudda R. Thakar, Christy Tremonti, Douglas L. Tucker, Alan Uomoto, Dan Vanden Berk, Michael S. Vogeley, Patrick Waddell, Shu-i. Wang, Masaru Watanabe, David H. Weinberg, Brian Yanny, Naoki Yasuda, and SDSS Collaboration. The Sloan Digital Sky Survey: Technical Summary. *The Astronomical Journal*, 120(3):1579–1587, September 2000.

A Appendix

Figure A.1 shows the univariate and bivariate distributions of SDSS galaxy properties and r-band magnitudes in our spectroscopic dataset.

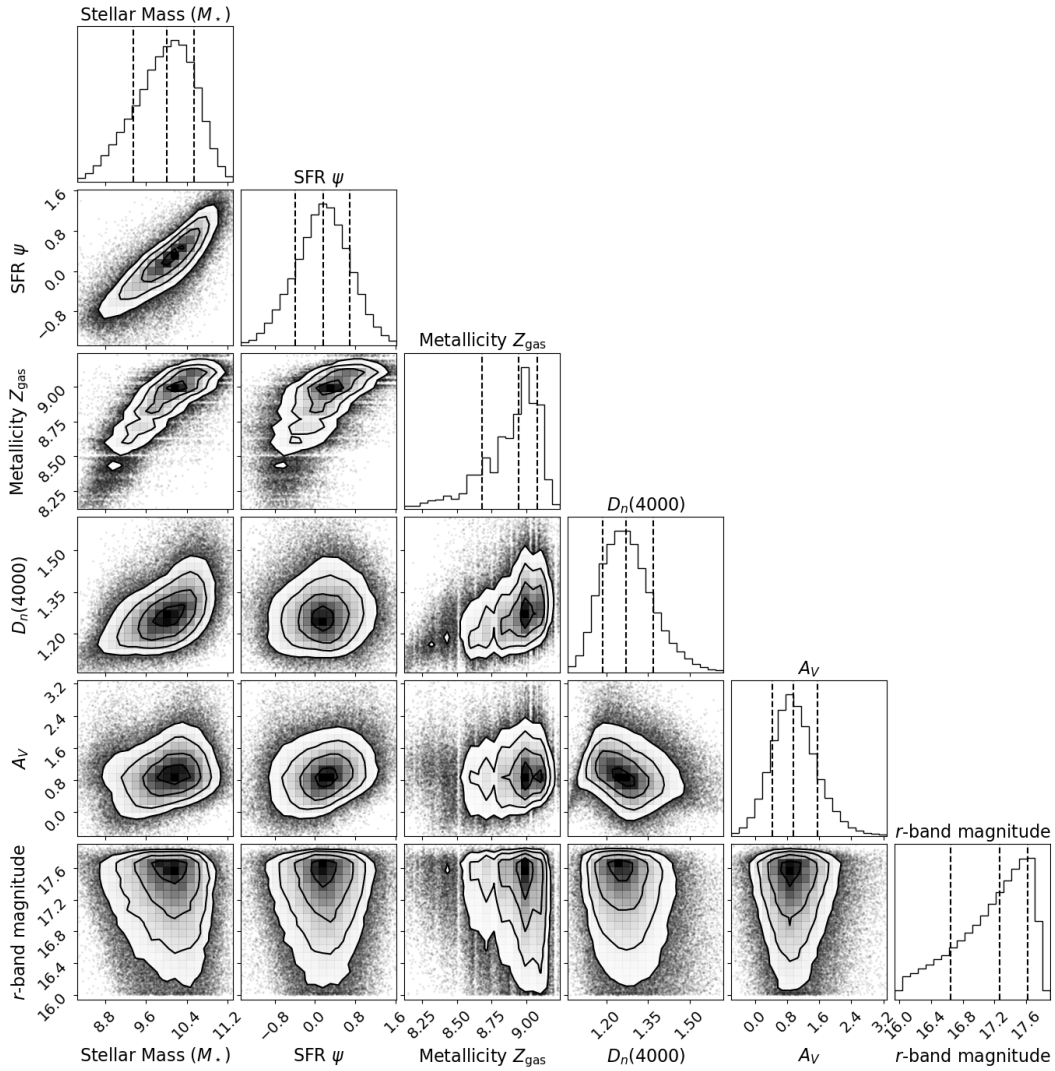

Figure A.1: Distributions of SDSS galaxy properties and r-band magnitudes. The 16th, 50th, and 84th percentiles of the univariate distributions are labeled with dashed lines.