High-resolution weak lensing mass mapping from DES-Y3 data using diffusion-based prior

Supranta S. Boruah
Department of Physics and Astronomy, University of Pennsylvania,
Philadelphia, PA 19103 `supranta@sas.upenn.edu`

Michael Jacob
Department of Physics and Astronomy, University of Pennsylvania,
Philadelphia, PA 19103 `mgjacob@sas.upenn.edu`

Bhuvnesh Jain
Department of Physics and Astronomy, University of Pennsylvania,
Philadelphia, PA 19103 `bjain@physics.upenn.edu`

Riya Maiya
Department of Physics and Astronomy, University of Pennsylvania,
Philadelphia, PA 19103 `rmaiya@sas.upenn.edu`

Raghav Venkataramanan
Department of Physics and Astronomy, University of Pennsylvania,
Philadelphia, PA 19103 `krvenkat@seas.upenn.edu`

Abstract

High-resolution mapping of cosmic mass distribution is essential for a variety of astrophysical applications including understanding cosmic structure formation, and galaxy formation and evolution. However dark matter is not directly observed and therefore we need advanced methods for solving inverse problems to reconstruct the underlying cosmic matter distribution. Here, we train a generative diffusion model and use it in the Diffusion Posterior Sampling (DPS) framework to reconstruct mass maps from Dark Energy Survey-Year 3 (DES-Y3) weak gravitational lensing data at high (1 arcminute) resolution. We show that the standard DPS results are biased, but they can be easily corrected by scaling the log-likelihood score during the diffusion process, yielding unbiased results with proper uncertainty quantification. The resulting mass maps reveal cosmic structures with enhanced detail, opening the door for improved astrophysical studies using the obtained mass maps.

1 Introduction

When light from distant galaxies passes near massive structures, it gets bent, subtly distorting the galaxy shapes we observe. This effect, called weak gravitational lensing, allows us to map the invisible dark matter distribution in the Universe without relying on luminous tracers. These mass maps are fundamental for understanding dark matter structures [1, 2], constraining cosmological parameters [3, 4] and understanding the physics of galaxy formation and evolution in different astrophysical environments [5, 6].

Traditional mass mapping approaches face significant limitations in extracting the full potential from modern surveys. The widely-used Kaiser-Squires reconstruction [7] results in highly noisy reconstructions, while Wiener filtering [8, 9] assumes Gaussian priors that are inadequate for capturing the highly non-Gaussian nature of late-time cosmic structure. Although lognormal models [10–12] offer some improvement over Gaussian assumptions, they remain insufficient for the precision requirements of Stage IV surveys [13]. These limitations restrict our ability to utilize small-scale structures from weak lensing mass maps for astrophysical studies.

Machine learning offers a promising path forward to solve the inverse problem of reconstructing cosmic matter distribution. Specifically, we can use generative machine learning models to learn complex, non-Gaussian priors directly from N-body simulations and then use this prior to reconstruct the matter distribution in the Universe with a high resolution. Several works have applied generative models to reconstruct dark matter maps from weak lensing data. [14] used conditional diffusion models for CMB lensing mass maps, learning the posterior directly, while [15] employed a similar method as us by using score-based diffusion models to learn dark matter map priors, and then used a convolved likelihood approach for the mass map reconstruction. Other generative methods include [16, 17], with [16] being the only prior work to apply generative machine learning methods to Stage-III survey data to the best of our knowledge. In this paper, we use diffusion models [18, 19] to learn the cosmic matter distribution from N-body simulations and then use the Diffusion Posterior Sampling [DPS, 20] approach to reconstruct mass maps from Dark Energy Survey Year 3 (DES-Y3) weak gravitational lensing data. We have previously demonstrated the application of this method to simulated weak lensing maps in [21]. We show that the standard DPS results are biased, but they can be corrected by using hyperparameters for likelihood scaling. Our method achieves accurate mass map reconstruction with well-calibrated uncertainties. The resulting mass maps reveal smaller-scale structures beyond those captured in earlier reconstructions, presenting new opportunities for astrophysical studies. Our work provides the first diffusion-based reconstruction of weak lensing mass maps from Stage-III surveys with comprehensive validation.

2 Weak lensing mass mapping using diffusion posterior sampling

Diffusion models [18, 19] are generative models that progressively add noise to training data until it becomes pure Gaussian noise, then use a neural network to reverse this process. The forward diffusion process gradually adds noise over timesteps, while the reverse process uses the gradient of the log-probability density, $\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t)$ (the 'score' function). A neural network estimates this score function: $\boldsymbol{s_\theta}(\boldsymbol{x}_t) \approx \nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t)$. Once trained, the model generates samples by starting from Gaussian noise and iteratively denoising.

2.1 Diffusion Posterior Sampling and likelihood scaling

For weak lensing mass mapping, we reconstruct the convergence field $\boldsymbol{\kappa}$, from noisy shear observations $\boldsymbol{D}$. To sample from the posterior, we need the gradient of the log-posterior [22] which is given using the Bayes' theorem as:

$$\nabla_{\boldsymbol{\kappa}_t} \log p(\boldsymbol{\kappa}_t|\boldsymbol{D}) = \nabla_{\boldsymbol{\kappa}_t} \log p(\boldsymbol{D}|\boldsymbol{\kappa}_t) + \nabla_{\boldsymbol{\kappa}_t} \log p(\boldsymbol{\kappa}_t), \tag{1}$$

where $\boldsymbol{\kappa}_t$ is the latent convergence map at timestep t. The likelihood gradient $\nabla_{\boldsymbol{\kappa}_t} \log p(\boldsymbol{D}|\boldsymbol{\kappa}_t)$ is intractable [20]. Diffusion Posterior Sampling (DPS) [20] approximates this using Tweedie's formula [23, 24], evaluating the likelihood at the posterior mean which we denote as $\hat{\boldsymbol{\kappa}}_{0|t}$. We refer the readers to [20] for details of DPS. DPS produces biased inference due to accumulated errors during early reverse diffusion [25]. We address this with likelihood scaling, multiplying the likelihood gradient by a time-dependent factor $A(t)$:

$$\nabla_{\boldsymbol{\kappa}_t} \log p(\boldsymbol{\kappa}_t|\boldsymbol{D}) \approx A(t) \cdot \nabla_{\boldsymbol{\kappa}_t} \log p(\boldsymbol{D}|\hat{\boldsymbol{\kappa}}_{0|t}) + \boldsymbol{s_\theta}(\boldsymbol{\kappa}_t), \tag{2}$$

where $A(t) = \text{Sigmoid}((\bar{t} - t)/\sigma_t)$ with $\bar{t} = 500$ and $\sigma_t = 20$. This down-weights the likelihood at early timesteps when samples are prone to going out-of-distribution (Figure 1, left), while maintaining full weighting at final steps. To determine $\bar{t}$ and σ_t, we run the TARP test (section 3.1) and search for hyperparameters that provide a good coverage test. The prior score is estimated using a neural network as $\boldsymbol{s_\theta}(\boldsymbol{\kappa}_t)$. The details of the neural network and its training can be found in [21].
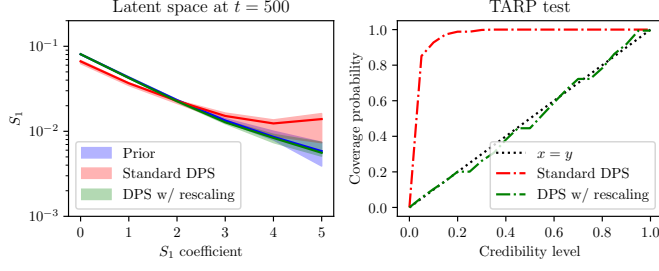
Figure 1: *Left:* Standard DPS (*red*) makes the latent maps go out-of-distribution. We resolve this with our likelihood scaling method (*green*). *Right:* TARP coverage test showing well-calibrated uncertainty quantification (following the $x = y$ line) with likelihood scaling, whereas the standard DPS gives biased results.

2.2 Training data and likelihood

We train our diffusion model on 50 full-sky ray-traced convergence maps from PKDGRAV3 simulations [26], creating 4 tomographic maps matching DES-Y3 source redshift distributions [27]. These simulations are all produced at a fixed cosmology and the diffusion model learns the dark matter distributions in these simulations as the prior. From each of these simulations, we extract 2,146 flat 256×256 pixel maps per full-sky map at 1 arcminute resolution (resulting in a total of $> 100k$ flat maps). We assume Gaussian likelihood for shear observations with independent shape noise per pixel:

$$p(\mathbf{D}|\boldsymbol{\kappa}) \propto \prod_{\alpha=1}^{2} \prod_{i=1}^{N_{\text{pix}}} \exp\left[-\frac{(\gamma_{\text{obs},\alpha}^{i} - \gamma_{\alpha}^{i}(\boldsymbol{\kappa}))^2}{2\sigma_{\epsilon,i}^2} \right] \qquad (3)$$

where $\gamma_{\text{obs},\alpha}^{i}$ are observed ellipticity, $\gamma_{\alpha}^{i}(\boldsymbol{\kappa})$ is shear calculated from the convergence field, and $\sigma_{\epsilon,i}$ is shape noise in the i-th pixel. While DES-Y3 shape noise is non-Gaussian [28], our framework can incorporate non-Gaussian likelihoods without retraining the diffusion model. We will assess the impact of non-Gaussian noise in a future work.

3 Results

3.1 Validation on simulation

We validate our method using mock DES-Y3 observations generated from simulated convergence maps. We create realistic mock shear data that includes a survey mask and Gaussian shape noise with $\sigma_\epsilon = 0.26$ per galaxy, using number densities corresponding to the DES-Y3 survey. Pixels outside the survey footprint are masked during likelihood evaluation Equation (3). Sampling 1 posterior map requires around 1 minute on a NVIDIA A40 GPU.

We demonstrate that the standard DPS method produces biases when early steps go out of distribution (Figure 1, left panel). However, our modified DPS algorithm with likelihood scaling successfully corrects this bias and reconstructs the underlying mass distribution.

To assess the uncertainty quantification of our method, we employ the Tests of Accuracy with Random Points [TARP, 29] framework. TARP provides a sampling-based calibration test that evaluates whether generated posterior samples accurately represent the true marginal posterior distribution. Through hyperparameter tuning of the likelihood scaling, we achieve well-calibrated posteriors with accurate uncertainty quantification (right panel of Figure 1). We refer the readers to [29] for more details on the TARP test.

Figure 2 demonstrates the performance of our method on mock simulations. Visual inspection of the maps confirms that we successfully reconstruct prominent dark matter features. We compute the cross-correlation between the reconstructed maps and the true map, showing that we correctly reconstruct structure on scales up to $\ell \lesssim 1000$. This represents an improvement of almost a factor of 5 over state-of-the-art mass mapping methods. For example, [30] achieves $\rho_c(\ell) \approx 0.5$ at $\ell \sim 200$, whereas our maps achieve the same correlation level at $\ell \sim 1000$. We further validate our method's performance by computing the power spectrum and other non-Gaussian summary statistics. Our reconstructed maps show excellent agreement with the true underlying maps across all these statistics.
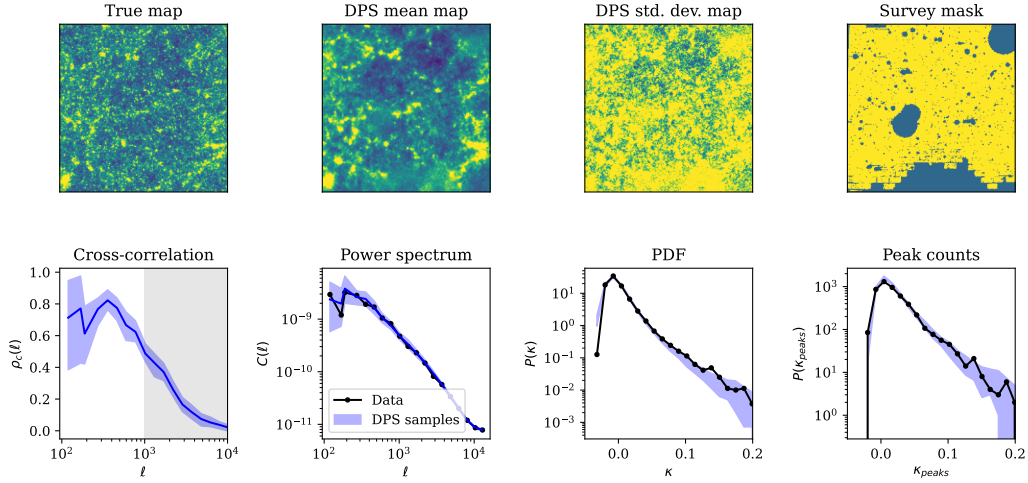
Figure 2: Validation of our method on simulations. **Top row**: True underlying mass map (left), mean of reconstructed mass map samples (center left), standard deviation of reconstructed samples (center right), and survey mask (right). **Bottom row**: Cross-correlation between reconstructed samples and true map (left), followed by comparisons of power spectrum, PDF, and peak counts between reconstructed samples (blue) and true map (black).
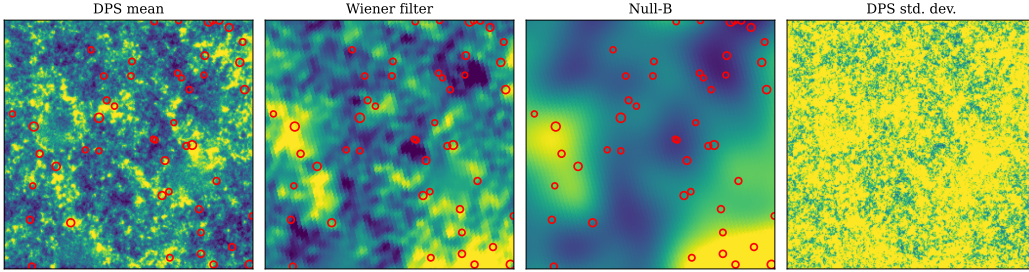

Figure 3: Comparison of the reconstructed maps from the DES-Y3 data in a $4.26° \times 4.26°$ representative patch. The left panel shows the mean of the DPS reconstructions, the center left panel shows the Weiner filtered reconstruction, and the third panel shows the Null-B reconstruction. The right panel shows the standard deviations computed from the DPS reconstructions. DES-Y3 RedMapper clusters are overplotted with red circles.

3.2 Mass map reconstruction of DES-Y3 data

After validating our method on simulated convergence maps, we apply it to DES-Y3 shear data. We created pixelized maps of mean galaxy ellipticity at 1 arcminute resolution. While we present in this section our results for only one of the patches, we have used our method to create the mass maps for the full DES-Y3 footprint by dividing the data in 410 patches of the same size. Our DPS maps are reconstructed independently on each flat patch. Figure 3 shows the reconstruction for one representative patch. We compare our DPS reconstruction (left) to Wiener filter and Null-B maps of the same data produced in [9]. From the figure, we see that: *i)* while our maps show overdensities and underdensities in the same sky regions as the other maps, they reconstruct structures on much smaller scales, and *ii)* DES-Y3 RedMaPPer clusters (red circles) correlate strongly with high-convergence regions in our maps.

We plan to perform additional validation of our reconstructed maps by cross-correlating them with external datasets such as galaxy catalogs and Compton y-maps. The enhanced small-scale recon-

structions will provide a powerful tool for probing galaxy formation and evolution at unprecedented resolution. The rightmost panel in Figure 3 shows the pixel-wise uncertainty in our reconstruction, which is essential for downstream analyses.

4   Conclusion

In this paper, we present a diffusion-based reconstruction of mass maps from DES-Y3 weak lensing data. This represents one of the first applications of generative AI-based mass map reconstruction to Stage-III survey data. By applying our diffusion-based approach to DES-Y3, we have produced high-resolution mass maps with robust uncertainty quantification. Our maps improve the resolution of mass map reconstruction by more than a factor of 5 compared to previous methods applied on the same data.

These high-resolution maps reveal significantly more small-scale structures than earlier reconstructions, opening new avenues for astrophysical studies. This enhanced resolution enables novel scientific applications including the identification of filaments and voids, and investigations of galaxy formation and evolution in different dark matter environments.

Future work will carefully examine the impact of various assumptions made in this study. The plug-and-play nature of DPS makes it straightforward to incorporate these improvements without retraining our diffusion model. One limitation of our current method is the need to carefully tune likelihood scaling hyperparameters. We plan to implement more robust methods for solving inverse problems with diffusion models to address this limitation [e.g, 31, 32]. Another potential limitation is that because our simulations cannot capture every aspect of the real data, the observed data may not lie within the training distribution. Understanding how such out-of-distribution (OoD) data impact the inference remains an open question. We expect these effects to introduce biases, and we need to explore strategies to quantify and reduce them is left for future work. Finally, we have not compared our results with other ML-based methods, as DES-Y3 mass maps generated using such approaches are not publicly available. A detailed comparison across different ML-based weak-lensing mass-mapping techniques is planned for future work.

Acknowledgements

References

[1] Moews B., et al., 2021, MNRAS, 500, 859

[2] Kovács A., et al., 2022, MNRAS, 510, 216

[3] Gatti M., et al., 2024, arXiv e-prints, p. arXiv:2405.10881

[4] Jeffrey N., et al., 2025, MNRAS, 536, 1303

[5] Eckert D., et al., 2015, Nature, 528, 105

[6] Yang T., Hudson M. J., Afshordi N., 2020, MNRAS, 498, 3158

[7] Kaiser N., Squires G., 1993, ApJ, 404, 441

[8] Wiener N., 1949, Extrapolation, Interpolation, and Smoothing of Stationary Time Series: With Engineering Applications. The MIT Press (https://direct.mit.edu/book-pdf/2313079/book_9780262257190.pdf), doi:10.7551/mitpress/2946.001.0001, https://doi.org/10.7551/mitpress/2946.001.0001

[9] DES Collaboration 2021, MNRAS, 505, 4626

[10] Boruah S. S., Rozo E., Fiedorowicz P., 2022, MNRAS, 516, 4111

[11] Fiedorowicz P., Rozo E., Boruah S. S., Chang C., Gatti M., 2022, MNRAS, 512, 73

[12] Zhou A. J., Li X., Dodelson S., Mandelbaum R., 2024, PRD, 110, 023539

[13] Boruah S. S., Fiedorowicz P., Garcia R., Coulton W. R., Rozo E., Fabbian G., 2024, arXiv e-prints, p. arXiv:2406.05867

[14] Flöss T., Coulton W. R., Duivenvoorden A. J., Villaescusa-Navarro F., Wandelt B. D., 2024, MNRAS, 533, 423

[15] Remy B., Lanusse F., Jeffrey N., Liu J., Starck J. L., Osato K., Schrabback T., 2023, A&A, 672, A51

[16] Shirasaki M., Moriwaki K., Oogi T., Yoshida N., Ikeda S., Nishimichi T., 2021, MNRAS, 504, 1825

[17] Whitney J. J., Liaudat T. I., Price M. A., Mars M., McEwen J. D., 2024, arXiv e-prints, p. arXiv:2410.24197

[18] Ho J., Jain A., Abbeel P., 2020, arXiv e-prints, p. arXiv:2006.11239

[19] Song Y., Sohl-Dickstein J., Kingma D. P., Kumar A., Ermon S., Poole B., 2020, arXiv e-prints, p. arXiv:2011.13456

[20] Chung H., Kim J., Mccann M. T., Klasky M. L., Ye J. C., 2022, arXiv e-prints, p. arXiv:2209.14687

[21] Boruah S. S., Jacob M., Jain B., 2025, PRD, 111, 083542

[22] Anderson B. D., 1982, Stochastic Processes and their Applications, 12, 313

[23] Robbins H. E., 1992, in , Breakthroughs in Statistics: Foundations and basic theory. Springer, pp 388–394

[24] Kim K., Ye J. C., 2021, arXiv e-prints, p. arXiv:2106.07009

[25] Zhang B., Chu W., Berner J., Meng C., Anandkumar A., Song Y., 2024, arXiv e-prints, p. arXiv:2407.01521

[26] Potter D., Stadel J., Teyssier R., 2017, Computational Astrophysics and Cosmology, 4, 2

[27] Collaboration D., 2021, MNRAS, 505, 4249

[28] DES Collaboration 2024, MNRAS, 527, L115

[29] Lemos P., Coogan A., Hezaveh Y., Perreault-Levasseur L., 2023, 40th International Conference on Machine Learning, 202, 19256

[30] Boruah S. S., Fiedorowicz P., Rozo E., 2024, PRD, 110, 023524

[31] Tolooshams B., Chandrashekar A., Zirvi R., Mammadov A., Yao J., Wang C., Anandkumar A., 2025, arXiv e-prints, p. arXiv:2505.22973

[32] Karan A., Shah K., Chen S., 2025, arXiv e-prints, p. arXiv:2506.10955