Towards solving model bias in cosmic shear forward modeling

Benjamin Remy[*]**, François Lanusse and Jean-Luc Starck**
Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, 91191, Gif-sur-Yvette, France

Abstract

As the volume and quality of modern galaxy surveys increase, so does the difficulty of measuring the cosmological signal imprinted in galaxy shapes. Weak gravitational lensing sourced by the most massive structures in the Universe generates a slight shearing of galaxy morphologies called cosmic shear, key probe for cosmological models. Modern techniques of shear estimation based on statistics of ellipticity measurements suffer from the fact that the ellipticity is not a well-defined quantity for arbitrary galaxy light profiles, biasing the shear estimation. We show that a hybrid physical and deep learning Hierarchical Bayesian Model, where a generative model captures the galaxy morphology, enables us to recover an unbiased estimate of the shear on realistic galaxies, thus solving the model bias.

1   Introduction

Over the next couple of years, a new generation of wide-field optical galaxy surveys will be coming online, starting with the ESA Euclid mission [1] and the Vera C. Rubin Observatory Legacy Survey of Space and Time [2], soon followed by the Roman Space Telescope [3]. At the core of the cosmology science that these surveys will enable is the measurement of the weak gravitational lensing effect, a very subtle shearing of the apparent images of distant galaxies in the presence of massive structures along the line of sight. On cosmological scales, this effect, also known as cosmic shear, is an invaluable probe of the total matter content of the Universe, and will play a key role in further constraining and testing our cosmological model over the next decade.

Despite over 20 years of research in this field, unbiased measurements of gravitational shear from observed galaxy images remains one of the most challenging aspects of the scientific exploitation of these surveys. The prevailing strategy for measuring shear from data has been to build estimators of the ellipticity of galaxy light profiles, due to the fact that one can in theory relate how a galaxy ellipticity e changes under a gravitational shear γ, as $e_{obs} \simeq e_i + \gamma$. Under the assumption that galaxies should have no preferential intrinsic ellipticity, a simple estimator for γ can therefore be constructed from $\hat{\gamma} = \langle e_{obs} \rangle = \langle e_i \rangle + \gamma$. However, to this day, no shear measurement algorithm has proven to be unbiased.

In fact, state-of-the-art approaches for shear measurement [4, 5] no longer attempt to build an unbiased shear estimator from images, but rather calibrate the shape measurement methods estimating a multiplicative and an additive bias. There are many factors contributing to these biases, but a fundamental issue comes from the fact that the ellipticity e is not a well defined quantity for arbitrary galaxy light profiles.

In this paper we propose a new paradigm for cosmic shear estimation where we no longer rely on ellipticities anymore. We propose to forward model the galaxy images at the pixel level, similarly to [6], but using non-parametric galaxy light profiles. The main idea is to train a deep generative model

---

[*]Corresponding author: benjamin.remy@cea.fr

to learn galaxy morophologies, which can then be used as part of a hierarchical Bayesian model, a hybrid deep learning and physical model, simulating the observations and used to infer the shear.

The purpose of this paper is first to demonstrate that using parametric light profiles as part of the forward model induces a bias in the shear estimation, which we call a *model bias*, and then to show that this bias can be removed by using a deep generative model for the morphology. We show in this preliminary study, that using a hybrid deep learning-physical Bayesian forward model can not only solve the model bias, but also provide uncertainty quantification to the shear estimation.

(a) Realistic galaxies fitted by a parametric profile   (b) Realistic galaxies fitted by a deep morphology model
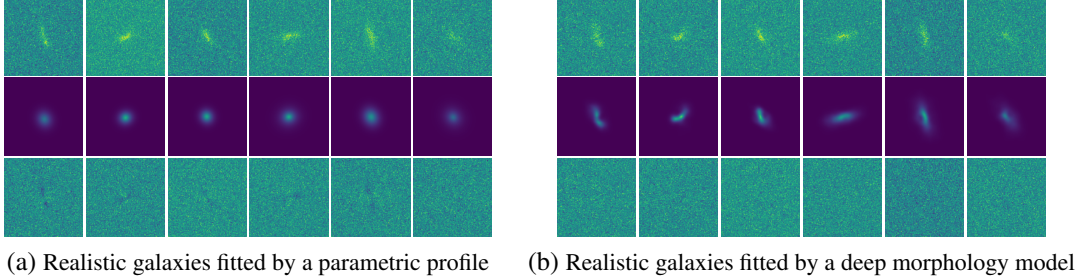
Figure 1: The first row of images correspond to observed galaxies. The second row corresponds to the light profiles Maximum a Posteriori fits, convolved with their PSFs. The third row shows residual images between the observations and the model fits. Column (a) is using Sérsic model as light profile and column (b) is using `deepgal` model as light profile.

2 Forward modeling weak lensing observations at the pixel level

From a Bayesian perspective, making the joint inference of cosmic shear and galaxy morphologies means estimating the *joint posterior* distribution $p(\mathcal{G}, \gamma | \mathcal{D})$, where $\mathcal{G} = \{z_i\}_{i=1...N}$ is the set of morphology parameters and $\mathcal{D} = \{d_i\}_{i=1...N}$ is the set of observed data. The Bayes identity tells us:

$$p(\mathcal{G}, \gamma | \mathcal{D}) \propto p(\mathcal{D} | \mathcal{G}, \gamma) p(\mathcal{G}, \gamma), \tag{1}$$

where $p(\mathcal{D} | \mathcal{G}, \gamma)$ is the data *likelihood* given by the forward model described bellow. Since the intrinsic morphology of galaxies are independent of the cosmic shear, then $p(\mathcal{G}, \gamma) = p(\mathcal{G})p(\gamma)$, where $p(\mathcal{G})$ and $p(\gamma)$ are respectively the *prior* distributions of the morphological parameters and of the cosmic shear.

We show the probabilistic graphical model in Figure 2. Nodes denote random variables and arrows conditional relation between them. White nodes corresponds to random variable on which we do the inference on, conditioned on the shaded nodes corresponding to observed variables. Dot nodes correspond to fixed variables, e.g. neural network weights θ. In this paper we compare two forward models: (1) using Sérsic law and (2) using the `deepgal` model to generate the galaxy light profile. Once the galaxy light profiles are drawn, we shear images with a constant value γ shared by all the galaxies. Each sheared galaxy i is then convolved by its Point Spread Function (PSF) Π_i. Assuming Gaussian noise corruption, the resulting image is therefore evaluated by a Gaussian likelihood of variance σ_{pix}^2.

Sérsic model The light profile of the Sérsic model is given by the following parametrization: $I(r) = F \times I_0 \exp(-b_n[(\frac{r}{r_{\text{hlr}}})^{\frac{1}{n}} - 1])$, where r_{hlr} is the half-light-radius, I_0 is a normalization factor such that the integration of I over all the space is F, which corresponds to the flux of the galaxy. n is called the Sérsic index and b_n is a dimensionless constant depending on n. The galaxy is then sheared with parameters $e = (e_1, e_2)$ corresponding to its ellipticity. Finally, since the real galaxy may not be perfectly centered in the postage stamp, we add two shift parameters s_x and s_y. Therefore, each galaxy i has a set of parameter $z_i = \{n, r_{\text{hlr}}, F, e_1, e_2, s_x, s_y\}_i$. Prior
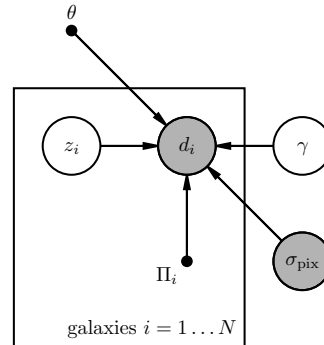

Figure 2: Probabilistic Graphical Model.

distributions and parameters follow empirical distributions of the COSMOS shape catalog and are described in subsection A.1.

`Deepgal` **model** `deepgal` is a VAE-based model, proposed in [7], generating complex galaxy light profiles. In this model, the observations are generated following a random process mapping a latent space representation $z_i \in \mathbb{R}^{16}$ to the observed image d_i following the probabilistic model illustrated in Figure 2. The neural network weights θ of the VAE are trained maximizing the Evidence Lower Bound (ELBO) of the data. Moreover, another generative model was trained on top of the VAE to learn latent space distribution using Masked Auto Regressive Flows [8]. This model was trained on the HST/ACS COSMOS survey postage stamps.

In the weak lensing regime, the shear is very small, therefore we can assume that it does not change the morphological statistics of the training sample. Thus, the model can be trained once, a priori, over a large catalog of galaxies and be used to perform the joint inference of the latent variables and the cosmic shear. For both light profile models, we denote the generated image $\mathcal{I}_0$.

Cosmic shear   In this model every galaxy image is sheared by a constant value $\gamma = \gamma_1 + i\gamma_2$ shared by all the galaxies. The image shearing transform, denoted $\mathcal{I}_0 \otimes \gamma$, is given by the following pixel coordinate change: $\left(\begin{smallmatrix} x' \\ y' \end{smallmatrix} \right) = S(\gamma)\left(\begin{smallmatrix} x \\ y \end{smallmatrix} \right)$, such that $S(\gamma) = \frac{1}{\sqrt{1-|\gamma|^2}}\left(\begin{smallmatrix} 1+\gamma_1 & \gamma_2 \\ \gamma_2 & 1-\gamma_1 \end{smallmatrix} \right)$.

Convolution with the PSF   The Point Spread Function (PSF) model is assumed to be known and exact for each galaxy. We draw the PSF evaluated at the galaxy i position, denoted by Π_i, in Fourier space and then multiply it to the Fourier image of the light profile before returning in the real space, resulting in an image $\mathcal{I}$, such that $\tilde{\mathcal{I}} = (\tilde{\mathcal{I}}_0 \otimes \gamma) \cdot \tilde{\Pi}$.

Noise corruption   We assume a known fixed gaussian noise identical for all the galaxies. For simplicity, all of the galaxies in our observations are corrupted by a constant Gaussian noise of variance σ_{pix}^2, as well as for our forward model. In a more realistic scenario where we would have information about the noise correlation properties for each postage stamp, one could consider a noise model per galaxy. The likelihood is therefore given by

$$p(\mathcal{D}|\mathcal{G}, \gamma) \propto \prod_{i=1,\ldots,N} \mathcal{N}\left(d_i \middle| \mathcal{F}^\dagger \left[\left(\tilde{\mathcal{I}}_{0,i} \otimes \gamma \right) \cdot \tilde{\Pi}_i \right], \sigma^2 \right), \quad \mathcal{F}^\dagger \text{ is the inverse Fourier operator.} \quad (2)$$
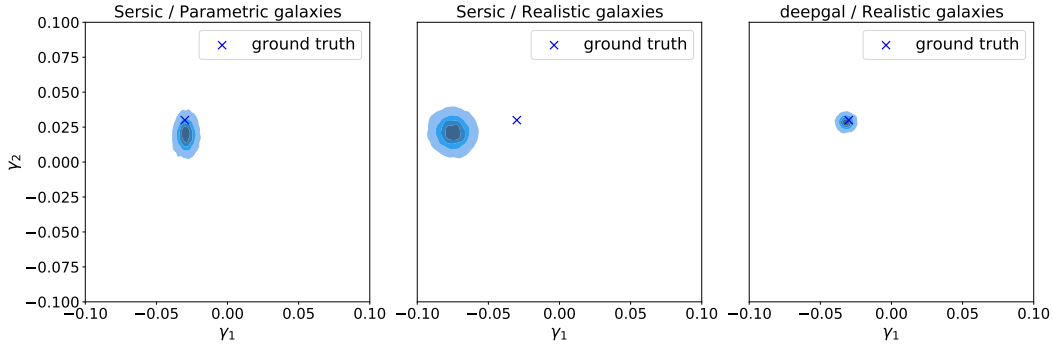
3   Illustration on realistic data

Figure 3: Shear marginal posterior contours against ground truth $\gamma = (-0.03, 0.03)$. The left panel shows the posterior using the Sérsic light profile over the parametric observations. The middle panel shows the posterior using the Sérsic light profile over the realistic galaxies. The right panel shows the posterior using the `deepgal` model over the realistic galaxies.

3.1   Simulating observations

We designed a toy model where galaxies are assumed to be already detected, deblended and centered in postage stamps, so that we only need to simulate the light profile, shearing transform and convolution with the PSF. In order to demonstrate the model bias, we distinguish two sets of observations.

(1) generated with parametric galaxy light profiles using a Sérsic model and (2) using realistic galaxies based on the galaxy postage stamps from the COSMOS HST Advanced Camera for Surveys (ACS) field [9–11]. We used the postage stamps from [12] and Sérsic fit parameters compiled in [13]. Because the unconvolved galaxies are very noisy, we used the Auto-Encoder from [7] to get denoised version of unconvolved galaxies, therefore this is not a demonstration on real galaxies yet. However, the purpose of this paper is to demonstrate the model bias, so we find that realistic galaxy morphologies that cannot be captured by analytic profiles are sufficient to illustrate the argument.

Each set of observation is a field of 256 galaxies, generated from the COSMOS catalog, with the following cuts: $n > 0.4$, $r_{hlr} > 0.2$ and $23.5 < i^+ < 25.2$ so rather faint galaxies. That way the entire galaxy is in the postage stamp, is not too peaky and not too resolved to be close to a Sérsic model. Every postage stamps were corrupted by a Gaussian noise of standard deviation $\sigma_{pix} = 0.01$. In order to simulate shearing, we removed any shear present in the selected sample from the COSMOS catalog, by adopting the same strategy than in the GREAT3 Challenge [14], i.e. doubling each galaxy, and rotating the doubled galaxies by 90°, as well as their PSF, to cancel the overall sample shear and then apply the same constant shear to every galaxy $\gamma = (-0.03, 0.03)$.

3.2 Approximate inference with mean-field VI

In order to estimate the posterior distribution $p(\mathcal{G}, \gamma | \mathcal{D})$, we relied on approximate inference using mean-field Variational Inference (VI), where we approximate the posterior distribution using a parametric distribution $q_\phi(\mathcal{G}, \gamma) = q_{\phi_\gamma}(\gamma) \prod_i \prod_j q_{\phi_{i,j}}(z_{i,j})$, parameterized by ϕ. We used Normal distributions $q_{\phi_{i,j}}$ to approximate the posteriror of parameters i of the jth galaxy, so the parameters $\phi_{i,j} = (\mu_{i,j}, \sigma_{i,j})$ correspond to means and standard deviations. Similarly, q_{ϕ_γ} is a 2d-diagonal multivariate Gaussian. The parameters $\phi_{i,j}$ and ϕ_γ are then optimized maximizing the ELBO:

$$\phi^* = \arg\max_\phi \ \mathbb{E}_{(\mathcal{G}, \gamma) \sim q_\phi} [\log p(\mathcal{G}, \gamma, \mathcal{D}) - \log q_\phi(\mathcal{G}, \gamma)] \tag{3}$$

This remains a very simplistic approach, which does not assume any correlation between variables. A more accurate model would condition the morphological parameters on the shear. In principle, the most accurate method would be to run a Markov Chain Monte Carlo, but several experiments showed that sampling based-methods are too computationally expensive for such high-dimensional problem.

3.3 Identifying and solving the model bias

We ran three sets of experiments to demonstrate the model bias. First fitting the fully parametric model to the parametric observations. Then fitting the same parametric model to the realistic galaxies and then fitting the hybrid model to the realistic galaxies. We used a Sérsic law as light profile as proposed in [6]. One could argue that there exist more complex parametric profiles such as the bulge+disk model which could reduce the aforementioned bias, but we propose here a model which does not rely on ellipticities and in principle can capture any complexity of galaxy morphologies.

First, as illustrated in Figure 3 (a), the inferred shear using the parametric model on the parametric observations is unbiased. This was indeed expected since the observations and the model are exactly the same. Secondly, Figure 3 (b) shows that the parametric model induces a bias in the shear estimation when the galaxy morphologies are no longer elliptical. Besides, one can observe the non-zero residuals in the Figure 1 fit to real galaxies, illustrating the model discrepancy when applied to complex morphologies. Finally, Figure 3 (c) shows that when we substitute the parametric light profile model by the deepgal model, the shear estimation becomes unbiased. One can notice that the uncertainties using the generative model are smaller than for the Sérsic model, but these contours are not representative of the error bars due to the mean-field approximation.

4 Discussion and conclusion

We have presented in this work the necessity of an accurate light profile model for forward modeling of the cosmic shear. We have shown that a hybrid physical and deep learning HBM has the potential to solve the model bias, being able to capture the complex morphology of galaxies and simulating physical processes such as shearing transform and convolution by the PSF. We identify many directions of research towards a real cosmological analysis by forward modeling. On the one

hand, many assumptions could be relaxed by building a more complex hierarchical model, e.g. parametrizing the PSF model and jointly inferring it along with the shear. Concerning the detection and deblending of galaxies, similar work have already been proposed, e.g. BLISS [15]. On the other hand, accurate uncertainty quantification could be enabled by scaling MCMCs for HBMs or using neural density estimator to capture the shape of the joint posterior distribution.

5   Acknowledgments

6   Broader impact

The methodology described here finds many useful applications within and outside of cosmology. Hierarchical Bayesian Models are very efficient tools for statistical inference but can suffer form model misspecification. Combining deep generative models and physical models seems to be a promising approach to solve this problem. We believe this work does not entail any negative consequences or ethical issues.

References

[1] Rene Laureijs, J Amiaux, S Arduini, J-L Augueres, J Brinchmann, R Cole, M Cropper, C Dabin, L Duvet, A Ealet, et al. Euclid definition study report. *arXiv preprint arXiv:1110.3193*, 2011.

[2] Zeljko Ivezić et al. LSST: From Science Drivers to Reference Design and Anticipated Data Products. *The Astrophysical Journal*, 873(2):111, 3 2019.

[3] D Spergel, N Gehrels, C Baltay, D Bennett, J Breckinridge, M Donahue, A Dressler, BS Gaudi, T Greene, O Guyon, et al. Wide-field infrared survey telescope-astrophysics focused telescope assets wfirst-afta 2015 report. *arXiv preprint arXiv:1503.03757*, 2015.

[4] Eric Huff and Rachel Mandelbaum. Metacalibration: direct self-calibration of biases in shear measurement. *arXiv preprint arXiv:1702.02600*, 2017.

[5] Erin S Sheldon and Eric M Huff. Practical weak-lensing shear measurement with metacalibration. *The Astrophysical Journal*, 841(1):24, 2017.

[6] Michael D Schneider, David W Hogg, Philip J Marshall, William A Dawson, Joshua Meyers, Deborah J Bard, and Dustin Lang. Hierarchical probabilistic inference of cosmic shear. *The Astrophysical Journal*, 807(1):87, 2015.

[7] François Lanusse, Rachel Mandelbaum, Siamak Ravanbakhsh, Chun-Liang Li, Peter Freeman, and Barnabás Póczos. Deep generative models for galaxy image simulations. *Monthly Notices of the Royal Astronomical Society*, 504(4):5543–5555, 2021.

[8] George Papamakarios, Theo Pavlakou, and Iain Murray. Masked autoregressive flow for density estimation. *Advances in neural information processing systems*, 30, 2017.

[9] Anton M Koekemoer, H Aussel, D Calzetti, P Capak, M Giavalisco, J-P Kneib, A Leauthaud, O Le Fevre, HJ McCracken, R Massey, et al. The cosmos survey: Hubble space telescope advanced camera for surveys observations and data processing. *The Astrophysical Journal Supplement Series*, 172(1):196, 2007.

[10] Nick Scoville, H Aussel, M Brusa, Peter Capak, C Marcella Carollo, M Elvis, M Giavalisco, L Guzzo, G Hasinger, C Impey, et al. The cosmic evolution survey (cosmos): overview. *The Astrophysical Journal Supplement Series*, 172(1):1, 2007.

[11] Nick Scoville, RG Abraham, H Aussel, JE Barnes, A Benson, AW Blain, D Calzetti, A Comastri, P Capak, C Carilli, et al. Cosmos: Hubble space telescope observations. *The Astrophysical Journal Supplement Series*, 172(1):38, 2007.

[12] Alexie Leauthaud, Richard Massey, Jean-Paul Kneib, Jason Rhodes, David E Johnston, Peter Capak, Catherine Heymans, Richard S Ellis, Anton M Koekemoer, Oliver Le Fevre, et al. Weak gravitational lensing with cosmos: Galaxy selection and shape measurements. *The Astrophysical Journal Supplement Series*, 172(1):219, 2007.

[13] Rachel Mandelbaum, Claire Lackner, Alexie Leauthaud, and Barnaby Rowe. Cosmos real galaxy dataset. *Zenodo*, January 2012.

[14] Rachel Mandelbaum, Barnaby Rowe, James Bosch, Chihway Chang, Frederic Courbin, Mandeep Gill, Mike Jarvis, Arun Kannawadi, Tomasz Kacprzak, Claire Lackner, et al. The third gravitational lensing accuracy testing (great3) challenge handbook. *The Astrophysical Journal Supplement Series*, 212(1):5, 2014.

[15] Derek Hansen, Ismael Mendoza, Runjing Liu, Ziteng Pang, Zhe Zhao, Camille Avestruz, and Jeffrey Regier. Scalable bayesian inference for detection and deblending in astronomical images. *arXiv preprint arXiv:2207.05642*, 2022.

Checklist

1. For all authors...

    (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]

    (b) Did you describe the limitations of your work? [Yes]

    (c) Did you discuss any potential negative societal impacts of your work? [Yes]

    (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]

2. If you are including theoretical results...

    (a) Did you state the full set of assumptions of all theoretical results? [N/A]

    (b) Did you include complete proofs of all theoretical results? [N/A]

3. If you ran experiments...

    (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [No]

    (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [No] Because of a lack of space but notebooks reproducing the results will be released once the paper is accepted. We mentioned the prior parameters but not the optimization of hyperparameters.

    (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [No] For this short paper, we show the experiments ran once.

    (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [No] because of a lack of space. Moreover, we did not intend to benchmark computation time here.

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...

    (a) If your work uses existing assets, did you cite the creators? [Yes]

    (b) Did you mention the license of the assets? [No]

    (c) Did you include any new assets either in the supplemental material or as a URL? [No]

    (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [No]

    (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [No]

5. If you used crowdsourcing or conducted research with human subjects...

(a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]

(b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]

(c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]

A Appendix

A.1 Prior distributions and parameters for Sérsic light profile

The following distributions were fitted from the COSMOS shape catalog [12, 13].

- $\log_{10} r_{\mathrm{hlr}} \sim \mathcal{N}(-0.68, 0.30)$
- $\log_{10} n \sim \mathcal{N}(0.1, 0.39)$
- $\log_{10} F \sim \mathcal{N}(-1.97, 0.53)$
- $e_{1/2} \sim \mathcal{N}(0, 0.28)$
- $\gamma_{1/2} \sim \mathcal{N}(0, 0.09)$
- $s_{x/y} \sim \mathcal{N}(0, 1)$