

Accurate cosmological emulator for the probability distribution function of gravitational lensing of point sources

Tunç Turker^{1,2,*}, Valerio Marra^{3,1,4}, Tiago Castro^{1,4,5,7}, Miguel Quartin^{2,8,9}, and Stefano Borgani^{1,4,5,6,7}

¹ INAF – Osservatorio Astronomico di Trieste, Via Tiepolo 11, 34131 Trieste, Italy

² PPGCosmo, Universidade Federal do Espírito Santo, 29075-910 Vitória, ES, Brazil

³ Departamento de Física, Universidade Federal do Espírito Santo, 29075-910 Vitória, ES, Brazil

⁴ IFPU – Institute for Fundamental Physics of the Universe, Via Beirut 2, 34151 Trieste, Italy

⁵ INFN – Sezione di Trieste, Via Valerio 2, 34127 Trieste, Italy

⁶ Dipartimento di Fisica, Sezione di Astronomia, Università di Trieste, 34143 Trieste, Italy

⁷ ICSC – Centro Nazionale di Ricerca in High Performance Computing, Big Data e Quantum Computing, Bologna, Italy

⁸ Centro Brasileiro de Pesquisas Físicas, 22290-180 Rio de Janeiro, Brazil

⁹ Observatório do Valongo, Universidade Federal do Rio de Janeiro, 20080-090 Rio de Janeiro, RJ, Brazil

Received 8 December 2025 / Accepted 6 February 2026

ABSTRACT

Aims. We developed an accurate and computationally efficient emulator to model the gravitational lensing magnification probability distribution function (PDF), enabling robust cosmological inference of point sources such as supernovae and gravitational-wave observations.

Methods. We constructed a pipeline utilizing cosmological N -body simulations, creating past light cones to compute convergence and shear maps. Principal component analysis (PCA) was employed for dimensionality reduction, followed by an extreme gradient boosting (XGBoost) machine learning model to interpolate magnification PDFs across a broad cosmological parameter space (Ω_m , σ_8 , w , h) and redshift range ($0.2 \leq z \leq 6$). We identified the optimal number of PCA components to balance accuracy and stability.

Results. Our emulator, publicly released as `ace_lensing`, accurately reproduces lensing PDFs with a median Kullback–Leibler divergence of 0.007. Validation on the test set confirms that the model reliably reproduces the detailed shapes and statistical properties of the PDFs across the explored parameter range, showing no significant degradation for specific parameter combinations or redshifts. Future work focuses on incorporating baryonic physics through hydrodynamical simulations and expanding the training set to further enhance model accuracy and generalizability.

Key words. gravitational lensing: weak – methods: numerical – methods: statistical – large-scale structure of Universe

1. Introduction

Gravitational lensing magnification, caused by the intervening large-scale structure, encodes valuable cosmological information by altering the observed brightness and angular sizes of distant sources. This effect both impacts and enhances cosmological and astrophysical inferences drawn from, for example, supernovae and gravitational-wave observations.

In this work, we focused on the one-point statistics of lensing magnification, which has emerged as a valuable complement to traditional analyses based on two-point statistics (see, e.g., Friedrich et al. 2026). Unlike the latter, which only exploits the variance and covariance of observables, the one-point statistics captures the full non-Gaussian structure of the magnification probability distribution function (PDF) (see, e.g., Patton et al. 2017; Liu & Madhavacheril 2019; Uhlemann et al. 2020, 2022). Accurate modeling of the magnification PDF is essential for unbiased cosmological inference from magnification data, enabling not only accurate corrections to the observed luminosity distances of sources, which can otherwise bias supernovae (Pierel et al. 2021) or gravitational wave distances (Shan et al. 2021; Canevarolo & Chisari 2024; Ferri et al. 2025),

but also the extraction of cosmological information directly from the measured PDF, as in the method of the moments (MeMo, Quartin et al. 2014). However, existing analytic approximations and simplified methods lack the precision required to capture nonlinear structures, particularly in the high-magnification regime.

Fully nonlinear cosmological simulations have become essential tools for obtaining accurate lensing PDFs, as demonstrated, for example, by Alfradique et al. (2024). Simplified semi-analytic approaches suffer from biases in lensing moments due to limitations in capturing small-scale nonlinearities and halo internal structures, ultimately impacting cosmological inference. Fully analytical approaches also exist, but are limited to the weak-lensing limit and to the convergence PDF (Thiele et al. 2020). Hydrodynamical simulations incorporating baryonic physics significantly alter the lensing PDF at high magnifications (Castro et al. 2018). Nevertheless, due to computational constraints, dark-matter-only (DMO) simulations remain widely used. While this approximation introduces some inaccuracies, they are expected to be minor in most practical scenarios, such as applications to supernova and gravitational-wave observations.

In this work, we constructed a comprehensive pipeline to generate accurate lensing PDFs using a suite of 70 N -body

* Corresponding author: tuncturker.work@gmail.com

simulations spanning broad ranges in cosmological parameters and redshift. The pipeline, summarized in Appendix A, includes the construction of past light cones (PLCs), computation of convergence and shear maps, principal component analysis (PCA) for dimensionality reduction, and machine learning (ML) modeling using the eXtreme gradient boosting algorithm¹ (XGBoost, Chen & Guestrin 2016). The final product is an efficient emulator capable of interpolating magnification PDFs across the studied parameter space, providing a powerful and computationally inexpensive tool for cosmological analyses. The `ace_lensing` emulator is publicly available².

This work is organized as follows. In Section 2, we describe the part of the pipeline responsible for generating the lensing PDF, from numerical simulations to dataset construction, including numerical tests. Section 3 addresses dimensionality reduction and feature extraction, while Section 4 presents the machine learning model employed. The emulator’s performance is evaluated in Section 5, and conclusions are drawn in Section 6.

2. PDF generation

2.1. Simulations

Alfradique et al. (2024) concluded that fully nonlinear cosmological simulations are necessary to obtain an accurate lensing PDF. Approximate methods such as PINOCCHIO (Monaco et al. 2002; Taffoni et al. 2002; Munari et al. 2017) and `turboGL` (Kainulainen & Marra 2009, 2011a,b) currently fail to model lensing accurately at large values of magnification ($\mu \gtrsim 1.5$). This is due to their reduced precision in representing small-scale nonlinear matter fields, stemming from oversimplified assumptions about halo internal structure and a reliance on perturbation theory. These limitations introduce biases in the lensing moments, which would in turn propagate as biases in cosmological parameter estimates when using, for example, the method of the moments (Quartin et al. 2014; Castro & Quartin 2014).

Castro et al. (2018) showed that baryonic feedback significantly affects the lensing PDF at $\mu \gtrsim 3$, and that full-physics (FP) hydrodynamical simulations – including the key astrophysical processes governing galaxy formation and evolution – are recommended for fully accurate modeling. Here, however, we adopt dark-matter-only (DMO) simulations due to the high computational cost of FP simulations, deferring to future work the task of calibrating the lensing PDF emulator using a suitable suite of hydrodynamical simulations. We expect the bias introduced by this approximation to be small for applications to supernova and gravitational-wave observations. This is because full-physics and dark-matter-only simulations differ most strongly in the high-magnification tail of the distribution – corresponding to the multiple-imaging regime – where local discrepancies have little impact on global statistical properties. To quantify this, Alfradique et al. (2024) introduced the \mathcal{L}_1 norm, defined as the weighted relative difference between a surrogate and a reference PDF. For example, a local discrepancy of 50% in the PDF’s high magnification tail corresponds to a weighted relative difference of only 2% in the magnification statistics.

For our DMO simulations, we employed the Tree-PM gravity solver of the `Opengadget3` code (Dolag et al., in prep.). We adopted a box size of $250 \text{ Mpc } h^{-1}$, which is sufficiently large to suppress significant cosmic variance and mode-coupling effects (Castro et al. 2018) and, with a computationally inexpen-

Table 1. Cosmological parameters and simulation setup.

Cosmological parameters	
Ω_m	[0.2, 0.4]
σ_8	[0.7, 0.9]
w	[-1.5, -0.7]
h	[0.6, 0.8]
Simulation parameters	
Box size L	$250 \text{ Mpc } h^{-1}$
Number of particles N	4×640^3
Particle-Mesh grid	2048^3
Initial redshift	49
Softening length	$6 \text{ kpc } h^{-1}$ (comoving)
Particle mass	$4.1 \Omega_m \times 10^9 M_\odot h^{-1}$

sive setup of 4×640^3 particles³, yields a particle mass resolution of $\sim 10^9 M_\odot h^{-1}$. This resolution ensures an effective modeling of small-scale nonlinear structures relevant for lensing (Castro et al. 2018). The chosen resolution guarantees an accurate reproduction of the matter clustering down to the scales where baryonic feedback becomes relevant. The inclusion of baryonic effects will be addressed in a forthcoming work.

The initial condition snapshots were generated using `MonoFonIC`⁴, using third-order Lagrangian perturbation theory, and subsequently provided to `Opengadget3` at a redshift of $z = 49$. To reduce the effect of cosmic variance we fixed the initial power spectrum amplitudes (Angulo & Pontzen 2016). We conducted a total of 70 simulations with the parameters reported in Table 1. In all simulations, the spatial-curvature density parameter was set to zero, and the dark energy equation of state was assumed to be non-dynamical, with no time evolution. We adopted a fixed baryon density parameter of 0.0494 and a primordial scalar spectral index of 0.9661, defined at the pivot scale 0.05 Mpc^{-1} .

Neutrinos were modeled as three species, with two assumed to be massless and one massive species with mass 0.06 eV. For each simulation, we stored approximately 30 snapshots covering the redshift range from 10 to 0, yielding a total of about 2000 snapshots and, as explained below, corresponding lensing PDFs. To systematically sample the parameter space, we employed the Latin Hypercube Sampling (LHS) strategy. We employ an LHS technique based on DeRose et al. (2019), but instead of maximizing the distance between the points based on their 2D projections, we maximize their full 4D Euclidean distance, which results in a better 4D point spread. This ensures an efficient and representative coverage of the multidimensional parameter space for our simulations. This parameter space is visualized in the Figure 1.

Figure 2 compares the matter power spectrum measured from our simulations with the nonlinear prediction of `baccoemu`⁵ (Angulo et al. 2021), which is calibrated on N -body runs with larger volumes and higher mass resolution. The simulation results closely follow the emulator prediction.

To assess the limitations of our simulations, we examined the scale at which matter perturbations enter the nonlinear regime and compared it to the mean inter-particle distance, $L/N^{1/3}$. The nonlinear scale is defined as the scale R at which the variance

³ FCC lattice `MonoFonIC` configuration (Michaux et al. 2020).

⁴ <https://bitbucket.org/ohahn/monofonic>

⁵ <https://baccoemu.readthedocs.io>

¹ <https://github.com/dmlc/xgboost>

² <https://github.com/Turkero/ACE-Lensing>

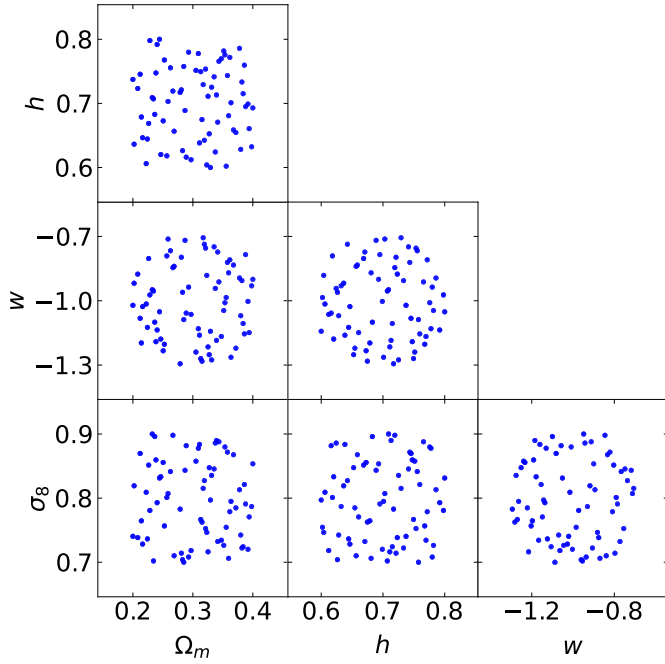


Fig. 1. Distribution of the values of the cosmological parameters for the 70 cosmologies here employed to build the lensing PDF emulator. These values were generated via Latin hypercube sampling.

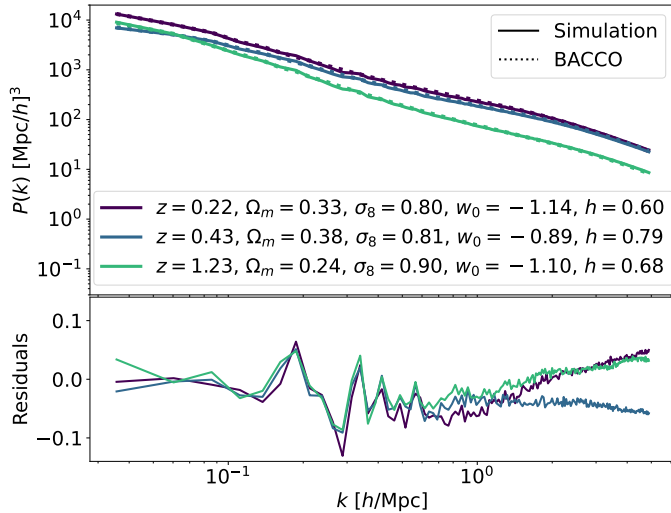


Fig. 2. Comparison between the matter power spectrum measured from the simulation snapshots and the nonlinear prediction from baccoemu.

of matter fluctuations satisfies $\sigma_R^2 = 1$, using a top-hat window function. For the computation of σ_R , we use the actual matter power spectrum measured from the snapshots of simulation C10, which has parameters $\Omega_m = 0.331$, $\sigma_8 = 0.835$, $w = -1.28$, and $h = 0.725$. Figure 3 shows that for $z \gtrsim 6$ the nonlinear scale drops below the interparticle distance, indicating that nonlinear clustering is suppressed. We thus restricted our PLCs to the redshift interval $0.2 \leq z \leq 6$ to guarantee that the analysis remains within the regime of sufficient simulation resolution⁶.

⁶ Since lensing efficiency approaches zero at the source position, contributions from the matter field at $z \approx 6$ have negligible impact on our results.

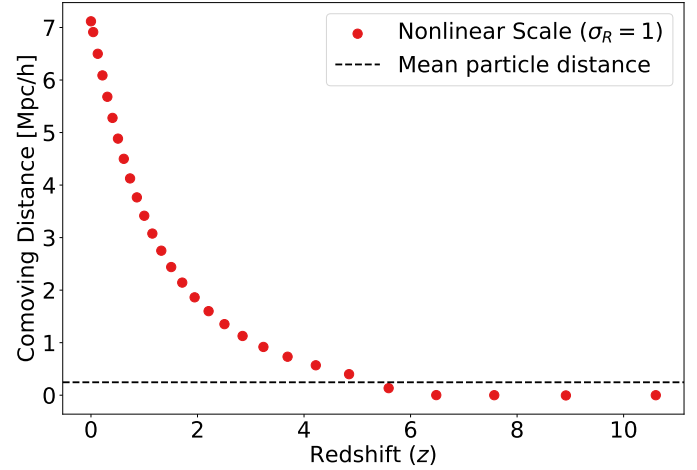


Fig. 3. Redshift dependence of resolvable matter fluctuations. The dots show the scale at which density fluctuations reach unity ($\sigma_R = 1$), and the black dashed line indicates the simulation’s mean particle separation. Regions below the dashed line are not reliably captured by the simulation due to resolution limits.

Consequently, we utilized approximately 20 out of the 30 available snapshots for each simulation⁷.

2.2. PLC construction technique

To construct past light cones (PLCs) for each simulation, we sequentially stack the snapshots, applying random rotations and translations before placing them along the line of sight to minimize correlations and boundary effects. Subsequently, the lensing PDF was computed by tracing light rays through the PLC. To this end, we divided the resulting image into pixels, the size of which introduces an additional resolution scale, distinct from the intrinsic softening scale and particle mass of the simulation. Each snapshot was projected onto a single lens plane that is perpendicular to the observer’s line of sight and passes through the center of the randomized simulation box. Consequently, the total number of lens planes coincides with the number of snapshots, approximately 30.

For the construction of past light cones (PLCs), we employed two complementary computational approaches, namely SLICER⁸ and PAINLESS⁹. These methods differ primarily in how they handle spatial sampling across redshift.

The SLICER method constructs PLCs by tracing light rays along the observer’s past light cone, maintaining a constant angular resolution θ_{grid} at all redshifts. This technique closely mimics observational conditions – such as those of wide-field imaging surveys – by preserving a fixed angular pixel scale, although the corresponding physical scale increases with distance. By construction, this approach samples only portions of the simulation boxes at each redshift, necessitating the generation of multiple PLC realizations (with varying random seeds) to fully exploit the simulation boxes.

Conversely, the PAINLESS method constructs PLCs by stacking slices from successive simulation snapshots, projected onto planes with a fixed spatial (comoving) grid size r_{grid} . This ensures a consistent physical resolution across redshift and uti-

⁷ The precise number depends on the cosmological parameters specific to each simulation.

⁸ <https://github.com/TiagoBsCastro/SLICER>

⁹ PAINLESS is a sub-routine of SLICER.

lizes all simulation particles, thus simplifying lensing calculations and eliminating the need for multiple PLC realizations. Although this method does not trace particles along the true past light cone, it provides a computationally efficient approximation well suited for theoretical analyses involving one-point statistics, such as the PDF of convergence κ or shear γ . However, due to its inherent loss of angular coherence across different redshifts, it is not appropriate for analyses based on two-point statistics, such as the angular power spectrum.

In Appendix B we compare the SLICER and PAINLESS methods, finding a good agreement between both methods. We choose PAINLESS for the analysis presented below, as it offers better control over spatial sampling. In SLICER, the angular resolution is fixed, so the comoving pixel size decreases toward low redshift. As a result, each pixel contains fewer particles, enhancing particle discreteness effects in the projected maps (since shot noise scales as $1/n$, with n the number of particles per pixel) that are not associated with physical lensing. By adopting a fixed comoving grid at all redshifts, PAINLESS keeps n more uniform across the light cone, giving better control over these numerical fluctuations. While the PAINLESS approach is less suited for mimicking observational conditions, it is more appropriate for theoretical studies in cosmological lensing, where full resolution control is critical.

2.3. Lensing maps

We constructed convergence and shear maps by implementing the gravitational lensing formalism (Bartelmann & Schneider 2001). The PLC construction yields maps of the surface mass density Σ_i on successive lens planes, each representing the projected mass distribution of the corresponding simulation snapshot.

First, we computed the effective comoving lens distance D_l and its corresponding redshift z_l by performing a distance-weighted average over the redshift interval spanned by each lens slice:

$$z_l = \frac{\int_{z_{\text{low}}}^{z_{\text{up}}} z D_M(z) dz}{\int_{z_{\text{low}}}^{z_{\text{up}}} D_M(z) dz}, \quad D_l = D_M(z_l), \quad (1)$$

where z_{low} and z_{up} delimit the redshift boundaries of the box within the PLC, $D_M(z) = \int_0^z c dz' / H(z')$ is the transverse comoving distance, and c is the speed of light.

Next, for a given source redshift z_s , we identified all lens planes at redshifts $z_i < z_s$ and processed their associated projected mass density maps Σ_i to obtain the convergence map $\kappa(\mathbf{x})$:

$$\kappa(\mathbf{x}) = \frac{4\pi G}{c^2} \sum_i W(z_i, z_s) [\Sigma_i(\mathbf{x}) - \langle \Sigma_i \rangle] (1 + z_i)^2 \frac{G(z_i)}{G(z_{\text{snap}})}, \quad (2)$$

where $W(z_i, z_s) = D_l D_{l,s} / D_s$ is the lensing efficiency kernel, computed from the comoving distances between the observer, lens, and source, and the factor $(1 + z_i)^2$ accounts for the proper-to-comoving scaling in the Poisson equation. The last term corrects for the mismatch between the lens-plane redshift and the snapshot redshift via the ratio of the linear growth functions G . This correction is small since the snapshots were saved at redshifts very close to those of the lens planes.

Finally, we computed the lensing potential ϕ from the convergence maps by solving the Poisson equation in Fourier space:

$$\tilde{\phi}(\mathbf{k}) = -\frac{2}{|\mathbf{k}|^2} \tilde{\kappa}(\mathbf{k}), \quad (3)$$

and transformed it back to real space. This procedure takes into account the correct periodic boundary conditions as PAINLESS uses the full box. The shear fields were derived as second-order derivatives of the potential:

$$\gamma_1 = \frac{1}{2}(\partial_{xx}\phi - \partial_{yy}\phi), \quad \gamma_2 = \partial_{xy}\phi. \quad (4)$$

These maps are crucial for modeling gravitational lensing by large-scale structure, encapsulating how matter distributions influence the propagation of light. The convergence field (κ) characterizes isotropic magnification, while the shear field (γ) quantifies the anisotropic distortions arising from differential gravitational deflection. From these convergence and shear maps, we computed magnification maps (μ), which serve as the foundation for our emulator framework. The magnification factor, describing the amplification or de-amplification of background sources due to gravitational lensing, is computed through:

$$\mu = \frac{1}{(1 - \kappa)^2 - \gamma^2}, \quad (5)$$

which follows directly from the Jacobian determinant of the lensing transformation and quantifies the total modification in the apparent brightness and angular size of background sources. Since magnification represents a flux ratio, which is intrinsically positive, we consider the absolute value $|\mu|$.

The magnification field is especially relevant for weak-lensing analyses, as it affects the observed luminosity of distant point-like objects (e.g., supernovae) and modifies source number counts at various flux thresholds. Furthermore, the statistical distribution of magnification encodes essential cosmological information regarding matter density fluctuations and the evolution of large-scale structure.

After generating the maps, we computed the magnification probability distribution function (PDF), which captures the statistical properties of the magnification field across various cosmological models and source redshifts. Our analysis spans 1447 distinct configurations of cosmological parameters and redshift values ($\sim 70 \times 20$)¹⁰, enabling a comprehensive investigation of lensing-induced magnification effects throughout the parameter space. These PDFs form the input to our emulator, which interpolates lensing statistics across a broad range of cosmologies without the need for additional costly simulations.

It is important to mention that all of our calculations are performed in the source plane. The transformation of the simulation outputs from the lens plane to the source plane is carried out by scaling with the magnification weight (Takahashi et al. 2011). This is necessary because the results in the source plane differ from those in the lens plane due to magnification or de-magnification effects caused by the matter content of the intervening region. As a sanity check, we verified at the map level, before PDF construction, that:

$$\langle \mu \rangle_{\text{source}} = \frac{\sum_i \mu_i \cdot (1/\mu_i)}{\sum_i (1/\mu_i)} = \frac{N}{\sum_i (1/\mu_i)} \approx 1. \quad (6)$$

2.4. Lensing map resolution

Finally, we determined the optimal grid size r_{grid} to effectively exploit the nonlinear structures probed by the simulations without entering the shot-noise dominated regime. Following the

¹⁰ As discussed in Section 2.1 and Figure 3, we restrict the analysis to the PDFs relative to $z < 6$.

approach of Takahashi et al. (2011), we model the convergence variance as:

$$\langle k^2 \rangle = \frac{9}{8\pi} H_0^4 \Omega_m^2 \int_0^{z_s} \frac{dz}{H(z)} (1+z)^2 W^2(z_l, z_s) \times \int dk k \left[P(k, z) + \frac{1}{n} \right] e^{-(k/k_{\text{cut}})^2}, \quad (7)$$

where $P(k, z)$ is the nonlinear matter power spectrum obtained from CAMB (HaloFit, mead2020 version)¹¹. The shot-noise contribution is modeled by the term $1/n$, where n denotes the particle number density in the simulations. The Gaussian smoothing introduced by the finite grid size r_{grid} is modeled by the exponential factor $e^{-(k/k_{\text{cut}})^2}$, where the cutoff wavenumber is defined as $k_{\text{cut}} = 2\pi\alpha/r_{\text{grid}}$. We adopted $\alpha = 0.2$, following Takahashi et al. (2011), which provides an empirical mapping between the smoothing effects of a cubic top-hat filter and a Gaussian kernel.

From Figure 4, we conclude that $r_{\text{grid}} = 12 \text{ kpc } h^{-1}$ is an optimal grid size, as it marks the transition to the shot-noise-dominated regime. The convergence test shows that reducing r_{grid} below this value mainly amplifies particle shot noise, rather than adding physically meaningful small-scale information. As a result, magnification PDFs computed at smaller r_{grid} become increasingly dominated by numerical artifacts. Any additional sensitivity to unresolved structure is expected to affect mainly the extreme high-magnification tail. Moreover, baryonic effects dominate the PDF at $\mu \gtrsim 3$ (Castro et al. 2018), and since baryons are not included in our simulations, exploring smaller grid sizes is beyond the scope of this first work. For $r_{\text{grid}} = 12 \text{ kpc } h^{-1}$, the resulting two-dimensional density, convergence, and shear maps have a resolution of 20833×20833 pixels.

We also stress that sources are treated as pointlike. In this framework, r_{grid} sets the comoving pixel size used to project the density field, and thus the effective angular resolution of the lensing maps probed by a point-like beam. The emulator should therefore be interpreted as predicting an effective, resolution-limited point-source magnification PDF. Sub-grid small-scale structure is not captured by our simulations and is therefore not included in the present modeling; for discussions and possible prescriptions to account for effects such as microlensing by compact objects, we refer to Boscá et al. (2022) and references therein. While baryonic and sub-grid effects can be important in the extreme tail, they are not expected to substantially affect the bulk of the PDF, which is the primary target of the present emulation.

2.5. Dataset generation and availability

At this stage, we produced a dataset consisting of approximately 2000 lensing PDFs spanning 70 cosmological models, publicly available at github.com/Turker/A&A-Lensing. The lensing PDFs were constructed by binning the magnification maps into 5000 logarithmic bins over the range $0.1 < \mu < 6$.

The choice of the maximum magnification is somewhat arbitrary. In tests, setting a value greater than 6 degraded the performance of the PCA dimensionality reduction because the high- μ tail disproportionately weighted the variance, downweighting the body of the PDF. Conversely, adopting a smaller μ_{max} prematurely truncated the high-magnification tail relative to the low-magnification wing. Although modeling is less reliable at very high magnifications, we retain support up to $\mu = 6$ to ensure a

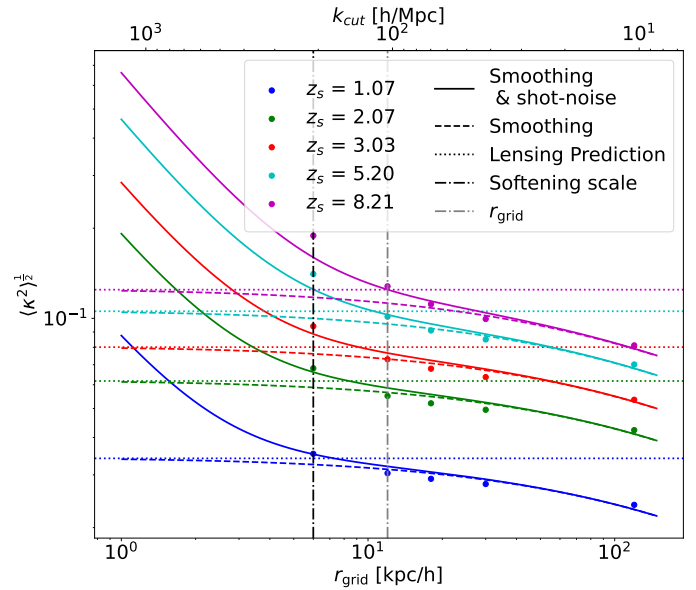


Fig. 4. Convergence root mean square as a function of grid size, used to determine the optimal resolution before entering the shot-noise dominated regime. The chosen optimal value (12 kpc/h) balances noise and resolution.

sufficiently wide domain. For supernova and gravitational-wave applications, the residual inaccuracy near the upper bound is expected to introduce at most a small bias.

3. PDF feature extraction

3.1. Standardization of the PDFs

We adopted Principal Component Analysis (PCA) to efficiently represent the PDFs. Before applying PCA, we standardized the PDFs to place them on a comparable scale and shape, which improves the performance of the PCA decomposition and, consequently, the training of our models. We standardized the PDFs as follows:

$$\mu_{\text{std}} = \frac{\mu - \bar{\mu}}{\sigma}, \quad P(\mu_{\text{std}}) = P(\mu) \cdot \sigma. \quad (8)$$

Here, μ denotes the lensing magnification, with mean $\bar{\mu}$ and standard deviation σ . The variable μ_{std} is the standardized magnification, and $P(\mu_{\text{std}})$ is its corresponding probability distribution function.

After standardization, we trimmed the PDF tails for values smaller than $\max(\text{PDF})/10^4$ to remove noisy artifacts. This operation does not affect the PDF region relevant for emulation. We then interpolated all PDFs onto a common magnification grid of 5000 evenly spaced points, spanning from -2 to 8 standard deviations ($\sigma_{\text{std}} = 1$). This procedure ensures a fixed, shared grid across all PDFs, eliminating the need for explicit emulation of the grid itself. While no PDF extends below -2 standard deviations due to intrinsic skewness, some may exhibit tails beyond 8 standard deviations. However, considering the intended applications of our emulator, the finite resolution of the simulations, and the absence of baryonic physics, extending the standardized PDFs beyond 8σ would not provide additional meaningful cosmological information.

¹¹ <https://camb.readthedocs.io>

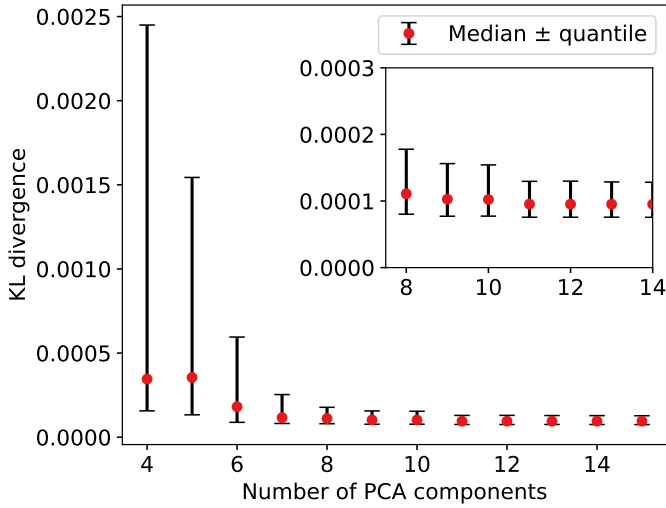


Fig. 5. KL divergence between original and reconstructed PDFs as a function of the number of PCA components. Error bars indicate the 16th–84th percentile range across 1447 PDFs.

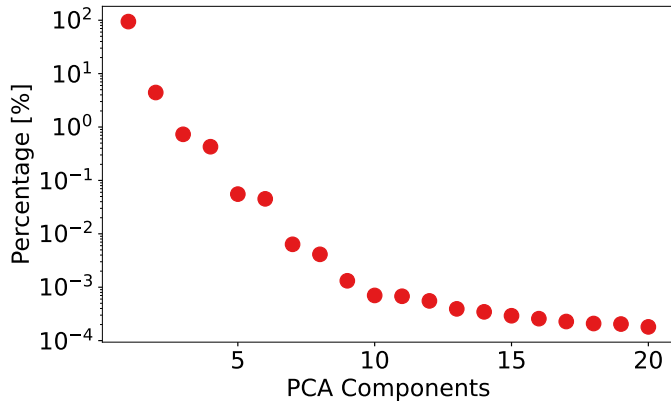


Fig. 6. Non-cumulative explained variance ratio for each PCA component. The first component captures over 90% of the total variance.

3.2. Dimensionality reduction with PCA

After standardizing the PDFs, we performed several tests to determine the optimal number of PCA components. The first test involved computing the Kullback–Leibler (KL) divergence between the original and reconstructed PDFs, evaluated over the 1447 PDFs corresponding to redshifts $z < 6$, as a function of the number of PCA components. The KL divergence is defined as:

$$D_{\text{KL}}(P \parallel Q) = \sum P(x) \log \left(\frac{P(x)}{Q(x)} \right), \quad (9)$$

where P is the original PDF and Q is the PCA-reconstructed PDF. Figure 5 shows that reconstruction fidelity improves with the number of components, reaching a plateau around 10 components, beyond which no significant gain is observed.

The relative importance of each PCA component is shown in Figure 6, where the first component alone interestingly accounts for approximately 91% of the total variance. Sharp drops in explained variance are observed after the 4th, 6th, and 8th components. Once again, we notice a plateau after the 10th component, indicating diminishing returns beyond this point. These thresholds provide natural cutoffs for retaining the most infor-

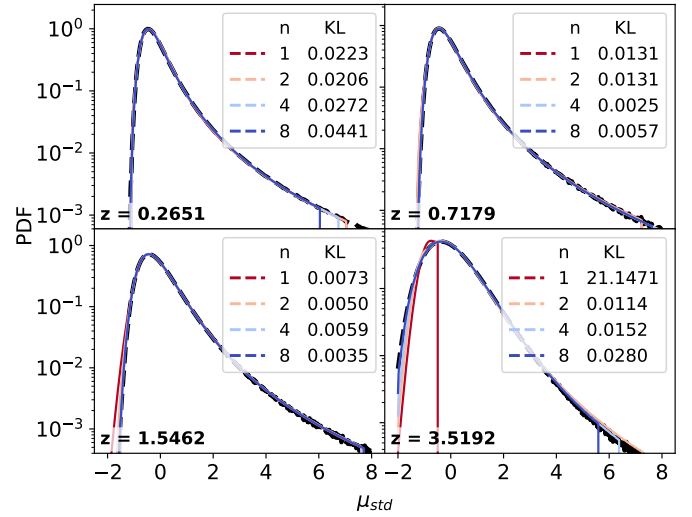


Fig. 7. Visual comparison of randomly selected standardized PDFs at representative redshifts. The black line in the background represents the original PDF. The legend shows the KL divergence values corresponding to the PDFs reconstructed with a given number of PCA components.

native features of the PDFs while minimizing redundancy and the risk of overfitting.

Figure 7 provides a visual comparison of four randomly selected standardized PDFs at representative redshifts. We note that one PCA component indeed captures the main features of each PDF, broadly reproducing their overall shape, and that it is hard to spot by eye the differences after two PCA components in all redshifts. Additional components are nevertheless required to achieve an accurate reconstruction.

Finally, we analyzed the correlation between the PCA components and the cosmological parameters, including redshift. Figure 8 shows that the first few PCA components display strong non-linear correlations with the input features and among themselves, indicating that they encode meaningful cosmological information relevant for reconstructing the PDFs. As expected, the correlation strength progressively decreases for higher-order components and becomes negligible beyond the tenth, consistent with the trends observed in Figures 5 and 6.

The discussion above focused solely on determining the optimal number of PCA components based on the reconstruction accuracy of the PDFs, without accounting for the additional noise introduced by the machine learning model. As we will show in the next section, it is this model-induced noise that ultimately determines the truly optimal number of components.

4. Machine learning model

As discussed in the previous section, we modeled the lensing PDF via its mean and standard deviation (whose information is absent in the standardized PDF used for the PCA analysis), together with the first n PCA components. Although the mean is very close to the theoretical value of unity, we found that to improve numerical accuracy it is useful to model it along side the variance. We remind that the emulator outputs the lensing PDF in the source plane.

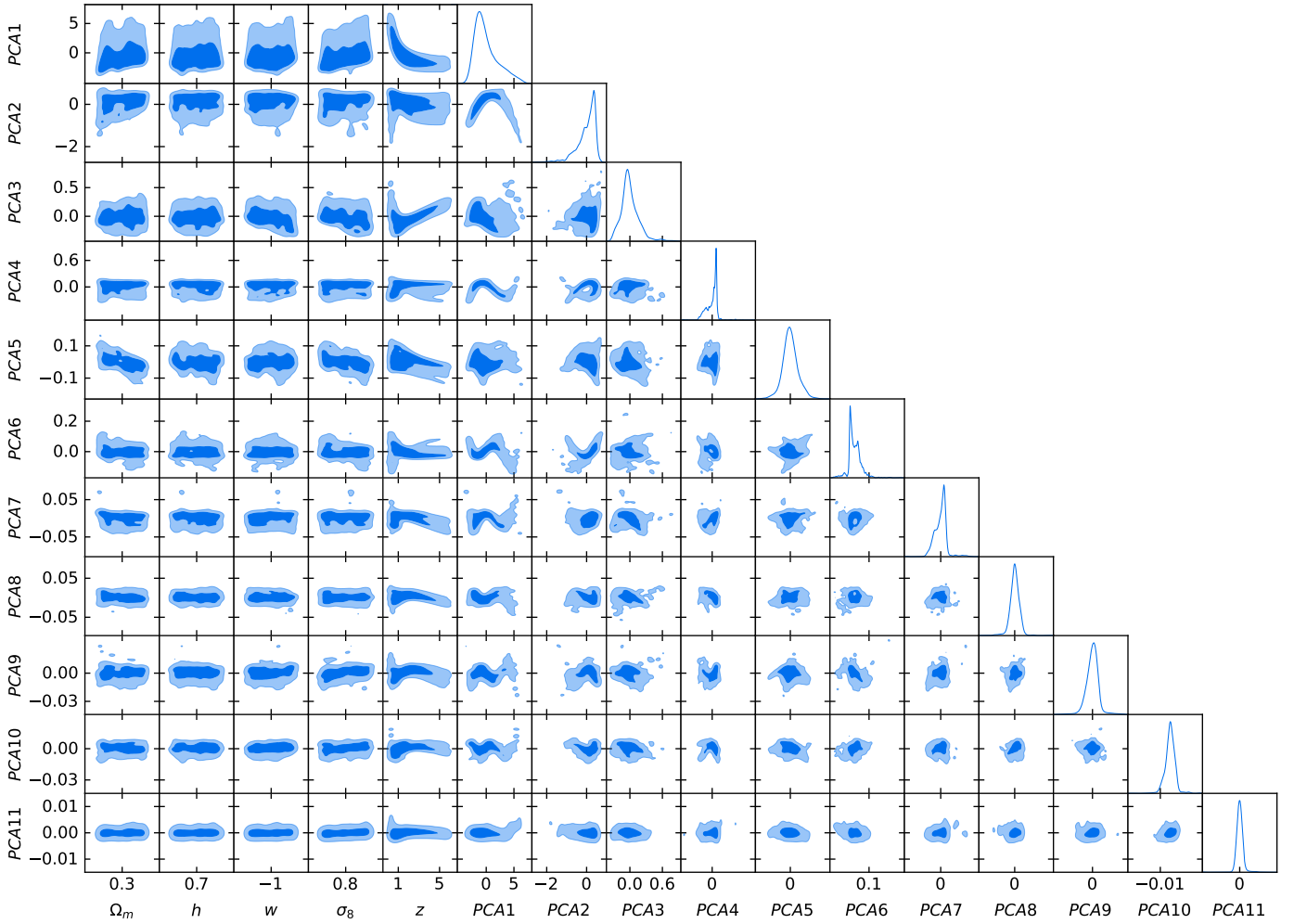


Fig. 8. Triangular correlation matrix between PCA components and input parameters, summarizing the relationships between the principal modes and the cosmological parameters. Numerical values indicate the nonlinear Spearman correlation coefficients.

4.1. XGBoost

To complete the emulator, we require a model that maps the five input features – the cosmological parameters (Ω_m , σ_8 , w , h) and redshift – to the $n+2$ output quantities describing the PDF. Given the size of the training set, we employ XGBoost, which builds an ensemble of gradient-boosted decision trees to model nonlinear relations efficiently. While neural networks may outperform tree-based methods on substantially larger datasets, they are unlikely to provide significant advantages in our setting.

Consequently, we built $n+2$ XGBoost models and perform an individual hyperparameter search for each of them over a parameter grid to optimize performance. The grid includes variations in the number of estimators, learning rate, tree depth, and sampling ratios, as summarized in Table 2. We split our set of 1447 PDFs into 80% for training (1157 PDFs) and 20% for testing (290 PDFs).

4.2. Choosing the optimal number of PCA components

In Section 3.2, we concluded that up to 10 PCA components were sufficient to accurately reconstruct the lensing PDF. However, that analysis considered only the PCA reconstruction step, ignoring the additional prediction noise introduced by the ML modeling stage. Here, we refine this analysis by accounting explicitly for the full pipeline, including the ML-induced uncer-

Table 2. Hyperparameter search grid for the XGBoost model.

Parameter	Values
n_estimators	50, 100, 200
learning_rate	0.01, 0.1, 0.2
max_depth	3, 5, 7
subsample	0.7, 1.0
colsample_bytree	0.7, 1.0

tainty. We expect the conclusions of this Section to evolve as the training set expands and ML techniques improve.

Figure 9 compares the KL divergence as a function of the number of PCA components, contrasting the scenario that includes only PCA dimensionality reduction with the full pipeline of Figure A.1, which incorporates ML prediction errors. We ultimately selected 4 PCA components, as this choice provides an optimal balance between reconstruction accuracy and model stability, yielding greater robustness against outliers. Indeed, Figure 9 presents median KL divergence values, which are robust measures of central tendency but do not fully capture the spread or presence of outliers. We observed that the number of outliers decreases when using 4 PCA components.

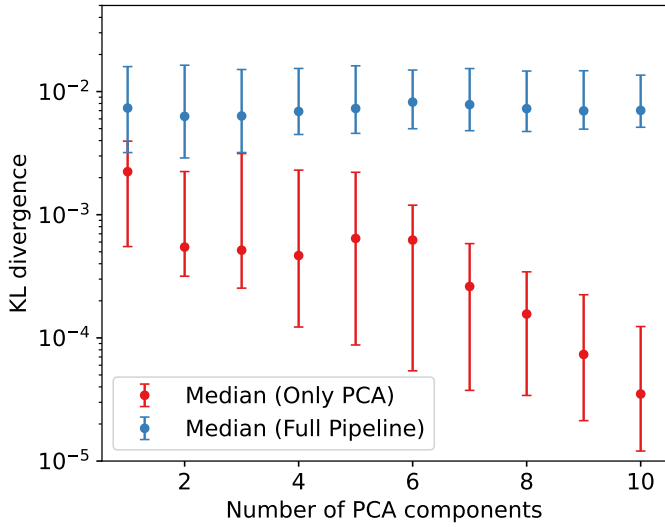


Fig. 9. Median KL divergence as a function of the number of PCA components. Error bars indicate the 16th–84th percentile range across the test set (290 PDFs). Red points correspond to an analysis similar to that of Figure 5, accounting only for information loss due to dimensionality reduction. Blue points include the full pipeline, incorporating ML prediction errors.

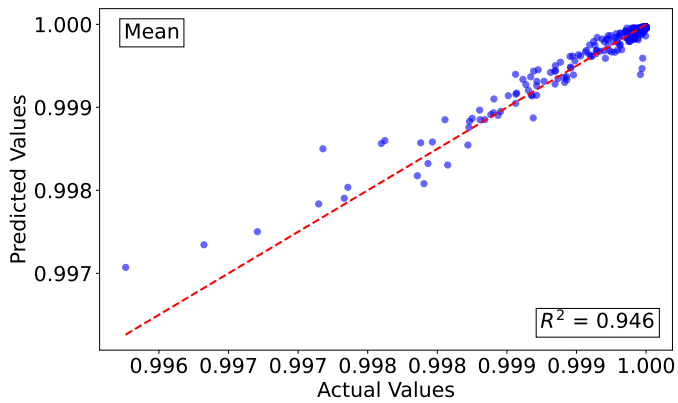


Fig. 10. Performance of the XGBoost regressor in predicting the mean of the magnification PDFs.

4.3. PDF reconstruction and post-processing

Using the n PCA components predicted by XGBoost, we first reconstructed the standardized PDF by applying the inverse PCA transformation. Artifacts occasionally appear in the reconstructed PDFs and so we applied a systematic post-processing step. Since the magnification PDFs are expected to be monotonic on both sides of the mode, we searched for local maxima and minima, and then smoothed these anomalies to restore monotonicity. Next, we applied a Fourier transformation and filtered out high-frequency components to remove residual noise.

To recover the original (unstandardized) PDF, we applied the inverse transformation using the predicted values for the mean and standard deviation:

$$\mu = \sigma_{\text{pred}} \mu_{\text{std}} + \bar{\mu}_{\text{pred}}, \quad P(\mu) = \frac{1}{\sigma_{\text{pred}}} P_{\text{pred}}(\mu_{\text{std}}), \quad (10)$$

where $P_{\text{pred}}(\mu_{\text{std}})$ is the standardized PDF reconstructed from the predicted PCA components. We renormalized the PDF to preserve unit integral.

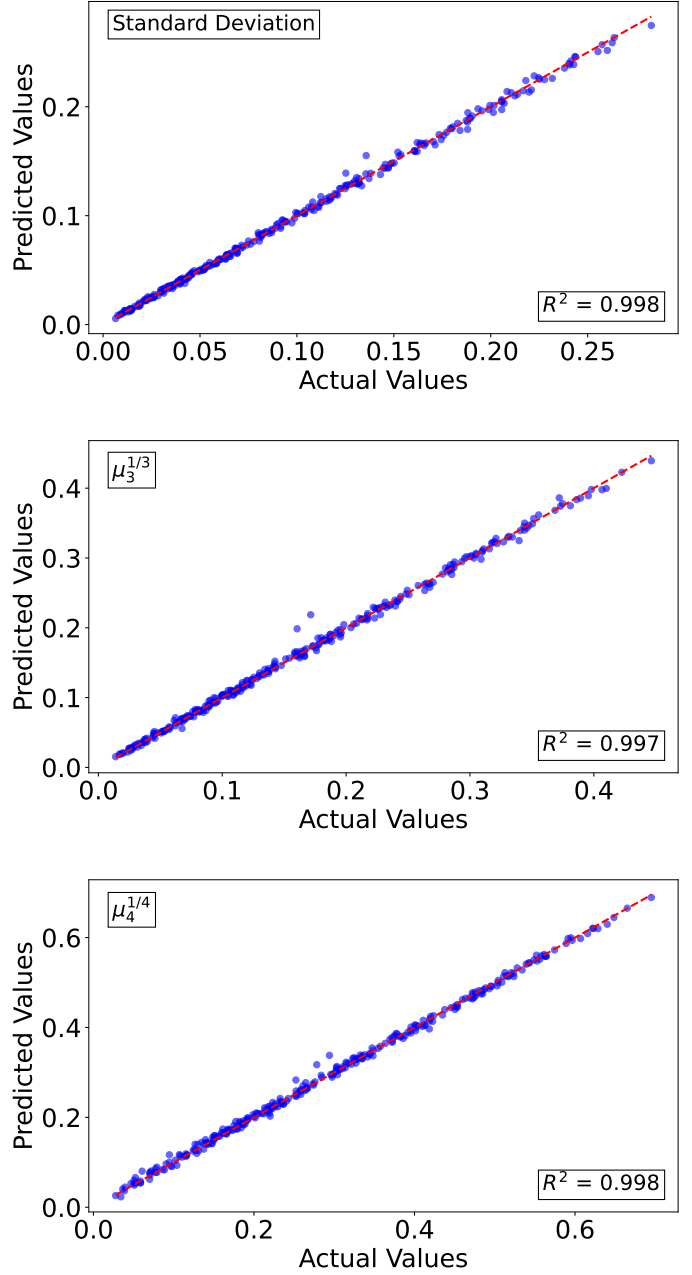


Fig. 11. Performance of the XGBoost regressor in predicting the second-to-fourth moments of the magnification PDFs.

Due to flux conservation in the source plane, the mean of the magnification PDF is expected to be unity, see Eq. (6). To exactly enforce this, we rescaled the emulator output as follows. We defined the rescaled magnification and corresponding PDF by

$$\mu' = \frac{\mu}{m_1}, \quad P'(\mu') = m_1 P(\mu), \quad m_1 = \int \mu P(\mu) d\mu, \quad (11)$$

which ensures that

$$\int P'(\mu') d\mu' = 1, \quad \int \mu' P'(\mu') d\mu' = 1. \quad (12)$$

Although this correction is very small, it ensures that the theoretical mean is recovered, yielding an unbiased measurement of the luminosity distances.

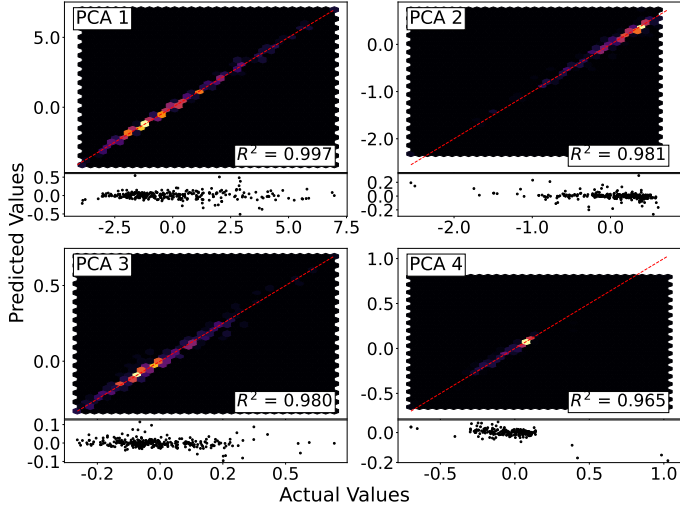


Fig. 12. Performance of the XGBoost regressor in predicting the PCA component values.

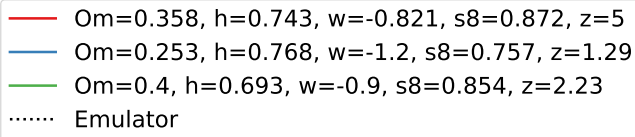


Fig. 13. Comparison between emulated PDFs (dotted lines) and actual PDFs from the test set (solid lines) at representative redshifts.

4.4. ML performance

As discussed earlier in Eq. (6), our maps yield a mean value very close to the theoretical expectation of unity. However, as mentioned in Section 2.5, we bin the PDF within the magnification range $0.1 < \mu < 6$, thereby excluding highly magnified events for which our physical model is unreliable and that are not relevant for the emulator's purpose. This truncation leads to a mean magnification slightly below unity, since only the high-magnification tail is systematically removed. The resulting bias, however, is well correlated with redshift and cosmology, as evidenced by the good performance of the XGBoost model in reconstructing the mean (Figure 10), allowing for an accurate modeling and a posteriori correction of the PDF mean, as implemented through Eq. (11).

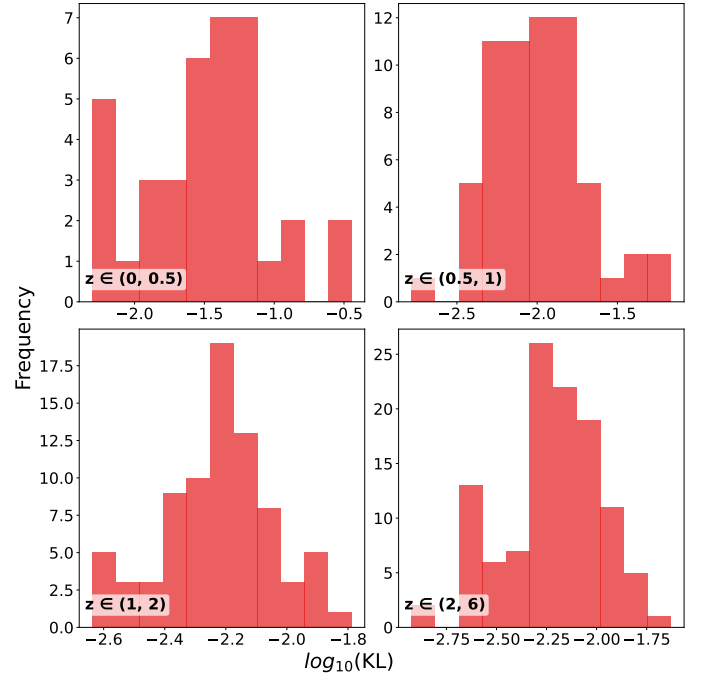


Fig. 14. Histogram of KL divergence values over the test set (290 PDFs) across the full pipeline using 4 PCA components.

Figure 11 illustrates the excellent performance of the XGBoost regressor in predicting the second-to-fourth moments of the magnification PDFs, whose accurate modeling is crucial for extracting cosmological information from lensing scatter via the method of the moments (MeMo, Quartin et al. 2014).

Finally, Figure 12 shows the accuracy in predicting the first 4 PCA component values. The accuracy in predicting the n -th PCA component decreases with increasing n , as higher-order components typically capture subtler and noisier variations.

5. Emulator performance

Figure 13 visually compares selected emulated PDFs (solid lines) with the actual PDFs from the test set (dotted lines) at four representative redshifts. The plots confirm the high quality of the emulation, consistent with the low KL divergence values reported in the legend.

Figure 14 shows the histogram of KL divergence values between the reconstructed and original PDFs over the test set, after applying the full pipeline with 4 PCA components. The median KL divergence is 0.0069, with nearly all PDFs exhibiting values below 0.1. The four redshift bins were defined according to the following rationale. The first bin, $[0, 0.5]$, includes sources for which lensing has a negligible impact. The second bin, $[0.5, 1]$, corresponds to the redshift range most relevant for current Type Ia supernova samples, such as DESY5 (DES Collaboration 2024) and Pantheon+ (Scolnic et al. 2022). The third bin, $[1, 2]$, will be important for the supernovae to be discovered by the Nancy Grace Roman Space Telescope (Rose et al. 2021). Finally, the last bin, $[2, 6]$, will be particularly relevant for the high-redshift gravitational-wave events (Menote et al. 2025) detected by third-generation observatories, such as the Einstein Telescope (Maggiore et al. 2020).

Finally, Figure 15 presents the KL divergence as a function of cosmological parameters. No clear patterns are observed, indicating that the performance is primarily limited by noise in

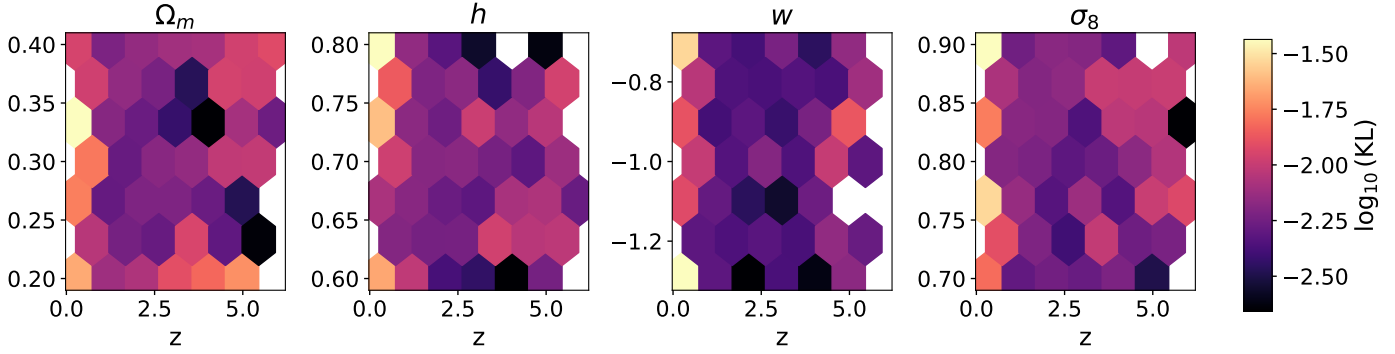


Fig. 15. KL divergence values as functions of cosmological parameters and redshift, for the full pipeline with 4 PCA components.

the ML predictions, which could be reduced with a larger training set. A mild trend of increasing KL divergence with higher σ_8 and redshift is visible, though it remains noisy.

6. Conclusions

In this work, we presented a comprehensive framework to accurately model the gravitational lensing magnification probability distribution function (PDF) using cosmological N -body simulations. Our approach addresses the limitations of previous analytic and semi-analytic approximations, particularly their inadequacies in capturing the nonlinear structures crucial for accurate cosmological inference.

We constructed a robust pipeline encompassing numerical simulations, past light cone generation, convergence and shear map computation, dimensionality reduction via PCA, and machine learning emulation employing the XGBoost algorithm. This pipeline enables efficient interpolation of lensing PDFs across a broad parameter space, covering cosmological parameters (Ω_m , σ_8 , w , h) and source redshifts ($0.2 \leq z \leq 6$).

We validated our methodology, demonstrating good agreement between the SLICER and PAINLESS approaches and identifying an optimal spatial resolution ($r_{\text{grid}} = 12 \text{ kpc } h^{-1}$) to mitigate shot noise while preserving nonlinear structure information. By performing dimensionality reduction using PCA, we found that employing 4 PCA components provides a robust balance between accuracy and stability, effectively capturing the essential features of the magnification PDFs.

Our machine-learning emulator, `ace_lensing`, achieves high accuracy, with a median KL divergence of 0.007 between emulated and original PDFs. Validation on the test set confirmed that the model reliably reproduces the detailed shapes and statistical properties of the PDFs across the explored parameter range, showing no significant degradation for specific parameter combinations or redshifts.

The emulator `ace_lensing` is publicly available, providing the cosmology and astrophysics communities with an efficient and precise tool for extracting cosmological information from gravitational lensing magnification data. The Python emulator module `ace_lensing` provides both the full lensing magnification PDF and standalone functions to directly predict its second-to-fourth moments.

In future work, we will extend the emulator on the modeling side by enlarging the training set with additional small-volume simulations to improve the stability of the ML predictions, a limited number of large-volume boxes to assess the impact of cosmic variance, and a representative set of hydrodynamical simulations to quantify and calibrate baryonic effects. This effort aims

at reducing both statistical and systematic sources of bias in the emulator.

Data availability

The `ace_lensing` code and trained machine-learning models described in this work are publicly available in the ACE-Lensing repository at <https://github.com/Turker0/ACE-Lensing>. The repository also includes processed data sets and example usage.

Acknowledgements. TT acknowledges financial support from FAPES (Brazil), CAPES (Brazil) and CNPq (Brazil). VM acknowledges partial support from CNPq (Brazil) and FAPES (Brazil). MQ is supported by FAPERJ (Brazil) project E-26/201.237/2022, CNPq (Brazil) and CAPES (Brazil). SB acknowledges support from the “Fondazione ICSC” National Recovery and Resilience Plan (PNRR) Project ID CN-00000013 “Italian Research Center on High-Performance Computing, Big Data and Quantum Computing” funded by MUR Missione 4, Componente 2 Investimento 1.4: “Potenziamento strutture di ricerca e creazione di campioni nazionali di R&S (M4C2-19)” – Next Generation EU (NGEU). We acknowledge the use of the HOTCAT computing infrastructure of the Astronomical Observatory of Trieste of the National Institute for Astrophysics (INAF, Italy) (see Bertocco et al. 2020; Taffoni et al. 2020), of the computing center of Cineca, of the Santos Dumont supercomputer of the National Laboratory of Scientific Computing (LNCC, Brazil), of the Sci-Com Lab of the Department of Physics at UFES, supported by FAPES, CAPES, and CNPq and of the computational resources of the joint CHE/Millways cluster, supported by a FAPERJ grant E-26/210.130/2023.

References

- Alfradique, V., Castro, T., Marra, V., et al. 2024, *Open J. Astrophys.*, 7
- Angulo, R. E., & Pontzen, A. 2016, *MNRAS*, 462, L1
- Angulo, R. E., Zennaro, M., Contreras, S., et al. 2021, *MNRAS*, 507, 5869
- Bartelmann, M., & Schneider, P. 2001, *Phys. Rep.*, 340, 291
- Bertocco, S., Goz, D., Tornatore, L., et al. 2020, *ASP Conf. Ser.*, 527, 303
- Boscá, V., Fleury, P., & García-Bellido, J. 2022, *JCAP*, 10, 098
- Canevarolo, S., & Chisari, N. E. 2024, *MNRAS*, 533, 36
- Castro, T., & Quartin, M. 2014, *MNRAS*, 443, L6
- Castro, T., Quartin, M., Giocoli, C., Borgani, S., & Dolag, K. 2018, *MNRAS*, 478, 1305
- Chen, T., & Guestrin, C. 2016, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785
- DeRose, J., Wechsler, R. H., Tinker, J. L., et al. 2019, *ApJ*, 875, 69
- DES Collaboration (Abbott, T. M. C., et al.) 2024, *ApJ*, 973, L14
- Ferri, J., Tashiro, I. L., Abramo, L. R., et al. 2025, *JCAP*, 04, 008
- Friedrich, O., Castiblanco, L., Halder, A., & Uhlemann, C. 2026, *MNRAS*, 545, staf2181
- Kainulainen, K., & Marra, V. 2009, *Phys. Rev. D*, 80, 123020
- Kainulainen, K., & Marra, V. 2011a, *Phys. Rev. D*, 83, 023009
- Kainulainen, K., & Marra, V. 2011b, *Phys. Rev. D*, 84, 063004
- Liu, J., & Madhavacheril, M. S. 2019, *Phys. Rev. D*, 99, 083508
- Maggiore, M., Van Den Broeck, C., Bartolo, N., et al. 2020, *JCAP*, 03, 050
- Menote, R., Marra, V., Sturani, R., Andrade-Oliveira, F., & Bom, C. R. 2025, arXiv e-prints [arXiv:2510.18727]

- Michaux, M., Hahn, O., Rampf, C., & Angulo, R. E. 2020, [MNRAS](#), **500**, 663
- Monaco, P., Theuns, T., & Taffoni, G. 2002, [MNRAS](#), **331**, 587
- Munari, E., Monaco, P., Sefusatti, E., et al. 2017, [MNRAS](#), **465**, 4658
- Patton, K., Blazek, J., Honscheid, K., et al. 2017, [MNRAS](#), **472**, 439
- Pierel, J. D. R., Rodney, S., Vernardos, G., et al. 2021, [ApJ](#), **908**, 190
- Quartin, M., Marra, V., & Amendola, L. 2014, [Phys. Rev. D](#), **89**, 023009
- Rose, B. M., Baltay, C., Hounsell, R., et al. 2021, arXiv e-prints [arXiv:[2111.03081](#)]
- Scolnic, D., Brout, D., Carr, A., et al. 2022, [ApJ](#), **938**, 113
- Shan, X., Wei, C., & Hu, B. 2021, [MNRAS](#), **508**, 1253
- Taffoni, G., Monaco, P., & Theuns, T. 2002, [MNRAS](#), **333**, 623
- Taffoni, G., Becciani, U., Garilli, B., et al. 2020, [ASP Conf. Ser.](#), **527**, 307
- Takahashi, R., Oguri, M., Sato, M., & Hamana, T. 2011, [ApJ](#), **742**, 15
- Thiele, L., Hill, J. C., & Smith, K. M. 2020, [Phys. Rev. D](#), **102**, 123545
- Uhlemann, C., Friedrich, O., Villaescusa-Navarro, F., Banerjee, A., & Codis, S. 2020, [MNRAS](#), **495**, 4006
- Uhlemann, C., Friedrich, O., Boyle, A., et al. 2022, [Open J. Astrophys.](#), **6**, 2023

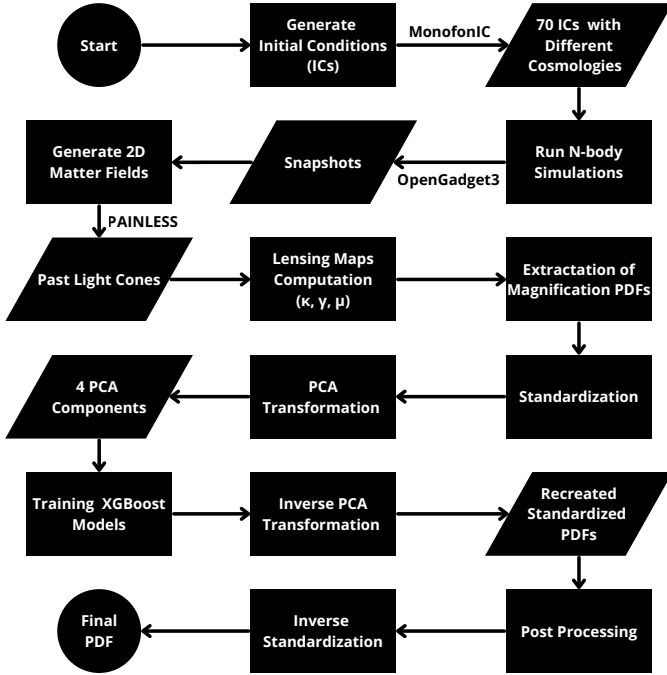


Fig. A.1. Full pipeline of the lensing PDF emulator.

Appendix A: Pipeline

Figure A.1 summarizes the pipeline adopted to build the `ace_lensing` emulator.

Appendix B: SLICER vs PAINLESS

Here we compare the SLICER and PAINLESS methods. To validate and quantify potential discrepancies between them, we define the effective grid size $r_{\text{grid}}^{\text{eff}}$ for PAINLESS that corresponds to a given angular resolution θ_{grid} used in SLICER (see Castro et al. 2018):

$$r_{\text{grid}}^{\text{eff}} = \frac{\int_0^{z_s} dz r_{\text{grid}}(z) G^2(z) \frac{(1+z)^2}{H(z)} W^2(z, z_s)}{\int_0^{z_s} dz G^2(z) \frac{(1+z)^2}{H(z)} W^2(z, z_s)}, \quad (\text{B.1})$$

where $r_{\text{grid}}(z) = \theta_{\text{grid}} D_M(z)$ is the comoving scale associated with the fixed angular resolution θ_{grid} .

We compared the convergence PDFs from SLICER with $\theta_{\text{grid}} = 14.06$ arcsec to those from PAINLESS, using $r_{\text{grid}}^{\text{eff}}$ from Eq. (B.1). Each simulation yields a single PAINLESS PDF, since all particles are used. In contrast, SLICER samples only a fraction of particles along discrete PLCs; to fully exploit the box we generated 50 PLC realizations per simulation. Figure B.1 compares the mean SLICER PDF to the corresponding PAINLESS PDF: the agreement is excellent for $\kappa \lesssim 0.1$, while systematic deviations appear in the high- κ tails, where sampling noise and rare nonlinear structures dominate. The increasing error bars at large κ reflect the smaller number of contributing pixels and the sensitivity to line-of-sight selection. Finally, Eq. (B.1) provides only an approximate mapping between the two resolution scales; residual discrepancies in the tails are therefore expected because the two methods probe the density field with different spatial samplings.

Next, in Figure B.2, we compare the n -th central moments, defined as $\kappa_n = \langle (\kappa - \langle \kappa \rangle)^n \rangle$. To facilitate this comparison, we plot the ratio $(\kappa_n^{\text{SLICER}} / \kappa_n^{\text{PAINLESS}})^{1/n}$. We obtain values close to

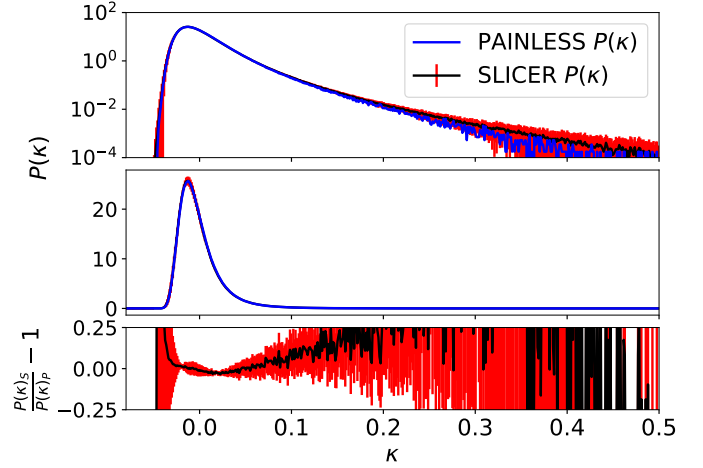


Fig. B.1. Mean of 50 SLICER realizations (with standard deviation) with $\theta_{\text{grid}} = 14.06$ arcsec compared to a single PAINLESS realization with $r_{\text{grid}} = 106.8$ kpc for a source at $z = 1$.

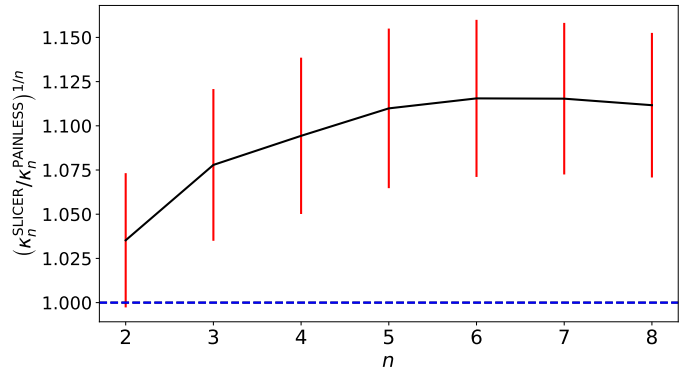


Fig. B.2. Comparison of central moments between mean SLICER PDF (50 realizations) and PAINLESS PDF.

unity for small n , exhibiting modest upward deviations at higher-order moments, consistent with the finding from Figure B.1. In general, these results reinforce the statistical consistency between the SLICER and PAINLESS methods.