

---

# Bridging Simulators with Conditional Optimal Transport

---

Justine Zeghal<sup>1 2 3</sup> Benjamin Remy<sup>4</sup> Yashar Hezaveh<sup>1 2 3 5 6</sup> François Lanusse<sup>7 8</sup>  
Laurence Perreault-Levasseur<sup>1 2 3 5 6 9</sup>

## Abstract

We propose a new field-level emulator that bridges two simulators using unpaired simulation datasets. Our method leverages a flow-based approach to learn the likelihood transport from one simulator to the other. Since multiple transport maps exist, we employ Conditional Optimal Transport Flow Matching (COT-FM) to ensure that the transformation minimally distorts the underlying structure of the data. We demonstrate the effectiveness of this approach by bridging weak lensing simulators: a Lagrangian Perturbation Theory (LPT) to a N-body Particle-Mesh (PM). We demonstrate that our emulator captures the full correction between the simulators, by showing that it enables full-field inference to accurately recover the true posterior, validating its accuracy beyond traditional summary statistics.

## 1. Introduction

In recent years, there has been a growing shift from traditional analytic inference methods to simulation-based inference (SBI) (e.g. Fluri et al., 2022; Porqueres et al., 2023; Jeffrey et al., 2024). This shift is driven by SBI’s ability to provide more precise constraints on cosmological parameters, particularly when analyzing data at the pixel level. However, it comes with the downside of relying solely on simulation accuracy. Any mismatches between the simulated and true underlying physical processes can result in biased parameter estimates (e.g. Filipp et al., 2024; Bayer

et al., 2025). Although significant effort has been dedicated to improving the realism of simulations, a major challenge remains: high-fidelity simulations are computationally expensive, limiting both the number of samples and the size of each simulation. This constraint makes full-field inference difficult to apply without compromising simulation quality.

Emulators have emerged as a way to accelerate the generation of high-fidelity simulations. Neural network-based emulators are trained on simulations to either correct fast approximate simulations (e.g. Dai et al., 2018; Kodi Ramanah et al., 2020; Lanzieri et al., 2022; Payot et al., 2023; Bartlett et al., 2025) or emulate the full simulations evolution (e.g. Lucie-Smith et al., 2018; He et al., 2019; Jamieson et al., 2023). As learning a small correction between low-fidelity (LF) and high-fidelity (HF) simulations should require fewer simulations, we focus on the formal approach. Specifically, we propose an emulator based on flow models (Lipman et al., 2023; Pooladian et al., 2023; Albergo et al., 2023; Tong et al., 2024; Lipman et al., 2024) that learn mappings between two probability distributions: the HF distribution and LF distribution. Since many mappings can transform one distribution into another, additional structure is needed to ensure that the correction minimally distorts the data manifold. To achieve this, we leverage optimal transport (OT) theory, which seeks to find minimal-effort mappings between distributions according to a chosen cost function. Specifically, we use COT-FM (Kerrigan et al., 2024), an approach that aligns conditional distributions (i.e., likelihoods rather than marginals), ensuring that the mapping between input parameters and observables is preserved. Meanwhile, Flow Matching (FM) provides a framework for learning deterministic mappings by minimizing a regression loss on velocity fields, remaining efficient even in high dimensions. A key distinction of our approach compared to standard field emulators is that it does not require paired data (e.g., simulations with shared initial conditions), a flexibility enabled by our OT-based framework (Tong et al., 2024).

We apply our COT-FM emulator to learn the transformation that maps second-order Lagrangian Perturbation Theory (LPT) convergence maps to their N-body Particle-Mesh (PM) counterparts. Emulators are usually validated at the summary statistics level. In this study, we assess the emulator performance by performing full-field inference. We first

---

<sup>1</sup>Department of Physics, Université de Montréal, Montréal, Canada <sup>2</sup>Mila – Quebec Artificial Intelligence Institute, Montréal, Canada <sup>3</sup>Ciela – Montreal Institute for Astrophysical Data Analysis and Machine Learning, Montréal, Canada <sup>4</sup>Department of Astrophysical Sciences, Princeton University, Princeton, NJ, USA <sup>5</sup>Center for Computational Astrophysics, Flatiron Institute, New York, USA <sup>6</sup>Trottier Space Institute, McGill University, Montreal, Canada <sup>7</sup>Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, Gif-sur-Yvette, France <sup>8</sup>Flatiron Institute, New York, NY, USA <sup>9</sup>Perimeter Institute for Theoretical Physics, Waterloo, Canada. Correspondence to: Justine Zeghal <justine.zeghal@umontreal.ca>.

show that analyzing PM convergence maps using the LPT forward model results in biased posterior estimates, highlighting simulation mismatch. Then, we demonstrate that the transformed simulations can recover the true posterior.

## 2. Conditional Optimal Transport Flow Matching

### 2.1. Flow Matching

FM (Lipman et al., 2023; Albergo et al., 2023) is a method for training continuous normalizing flows (CNFs) (Chen et al., 2019). CNF aims to continuously transport a source distribution  $p_0$  into a target distribution  $p_1$  over time through a diffeomorphic map  $\phi : [0, 1] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ . This map is defined as the solution to the ordinary differential equation (ODE):

$$dx = v_t(x)dt, \quad \phi_0(x) = x_0, \quad (1)$$

where  $x_0 \sim p_0$  are samples from the source distribution, and  $v : [0, 1] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is the time-dependent velocity field guiding the transport. Intuitively, the velocity field tells each particle  $x$  in the distribution how to move at every time  $t$  to gradually reshape  $p_0$  into  $p_1$ . The intermediate distribution at any time  $t$ , denoted  $p : [0, 1] \times \mathbb{R}^d \rightarrow \mathbb{R}$ , and known as the marginal probability path, is the pushforward<sup>1</sup> of  $p_0$  by the map  $\phi_t$ :  $p_t = [\phi_t]_{\#}(p_0)$ . As long as the velocity field satisfies the continuity equation, the evolving distribution is a valid probability density function.

Once the velocity field  $v_t$  is learned, generating samples from the target distribution simply requires integrating the ODE forward in time starting from samples of  $p_0$ . Conversely, one can reverse the process by integrating backward in time. If the density  $p_1$  is required, it can be recovered via  $p_1 = [\phi_1]_{\#}(p_0)$ .

In practice, FM turns learning  $v_t$  into the following regression task

$$\mathcal{L}_{\text{FM}} = \mathbb{E}_{t \sim \mathcal{U}[0,1], p_t(x)} \|v_t(x) - v_{\varphi}(t, x)\|^2, \quad (2)$$

where  $v_{\varphi}$  is a neural network (NN). Because neither  $p_t$  nor  $v_t$  is known a priori, Lipman et al., 2023 proposed an alternative *conditional flow matching* (CFM) loss which shares the same gradients:

$$\mathcal{L}_{\text{CFM}} = \mathbb{E}_{t \sim \mathcal{U}[0,1], q(z), p_t(x|z)} \|v_t(x|z) - v_{\varphi}(t, x)\|^2, \quad (3)$$

with  $z$  a conditioning variable,  $p_t(x|z)$  the conditional probability path such that  $p_t(x) = \int p_t(x|z)q(z)dz$ , and  $v_t(x|z)$  the corresponding conditional velocity field. As explain in

<sup>1</sup>The pushforward operator  $\#$  is defined by the change of variables formula  $[\phi_t]_{\#}(p_0) = p_0(\phi_t^{-1}(x)) \det[\partial_x \phi_t^{-1}(x)]$ .

Tong et al., 2024, to bridge two arbitrary distributions, one can define  $q(z)$  to generate pairs of samples, for instance  $q(z) = p_0(x_0)p_1(x_1)$ . Then the conditional distribution  $p_t(x|z)$  can be taken as a simple interpolating distribution, such as a Gaussian with mean along the straight path between  $x_0$  and  $x_1$ :  $\mu_t(z) = (1-t)x_0 + tx_1$ , with a small variance  $\sigma^2$ . Specifically, as  $\sigma \rightarrow 0$ ,  $p_t$  at  $t = 0$  and  $t = 1$  approaches the true source distribution  $p_0$  and the true target distribution  $p_1$ , respectively.

Note that multiple velocity fields can transform  $p_0$  into  $p_1$ , and enforcing only straight-line conditional probability paths between unpaired samples does not ensure that the velocity field minimizes the transport cost. Additionally, FM only transports marginal to marginal distributions.

### 2.2. Conditional Optimal Transport

FM enables us to learn a deterministic transport map  $\phi_1$  such that a new sample  $x_1$  can be obtained from  $x_0$  via  $x_1 = \phi_1(x_0)$ . In this work, however, we seek the transport map that minimizes a specific cost, ensuring that the transformation from  $x_0$  to  $x_1$  retains essential information without unnecessary alterations. Moreover, this transformation must preserve the conditioning: each generated PM map should share the cosmological parameters of its corresponding LPT input. Conditional Optimal Transport (COT) offers a rigorous framework for this goal by extending classical OT to the setting where one seeks to bridge conditional distributions across different parameters in an amortized way (Hosseini et al., 2024; Kerrigan et al., 2024).

Formally, let  $\Theta$  denotes the parameter space (e.g., cosmological parameters) and  $X$  the observation space (e.g., convergence maps). We denote by  $p_0(\theta, x)$  with  $\theta \in \Theta$  and  $x \in X$ , the joint source distribution (LPT forward model), and by  $p_1(\theta, x)$  that of the target (PM forward model), with both sharing the same prior  $p(\theta)$  over parameters. To preserve conditioning during transport, we define a triangular map  $\phi : \Theta \times X \rightarrow \Theta \times X$  of the form

$$\phi(\theta, x) = (\theta, \phi_X(\theta, x)), \quad (4)$$

where  $\phi_X : \Theta \times X \rightarrow X$ . A key result from Baptista et al., 2023, shows that if  $p_1 = [\phi]_{\#}(p_0)$ , then for almost every  $\theta$ ,  $[\phi_X(\theta, \cdot)]_{\#}(p_0(\cdot|\theta)) = p_1(\cdot|\theta)$ . That is, a single triangular map transports all likelihoods.

COT seeks such a triangular map  $\phi$  that minimizes a displacement cost. In our case, we consider the cost  $c(x_0, \theta_0, x_1, \theta_1) = \|x_0 - \phi_X(\theta_0, x_0)\|^2$ , leading to the following conditional Monge problem:

$$\inf_{\phi_X} \int \|x - \phi_X(\theta, x)\|^2 p_0(\theta, x) dx d\theta, \quad (5)$$

subject to the constraint that  $p_1 = [\phi]_{\#}(p_0)$ . If there exists a unique transport map  $\phi^*$  that solves this problem, then

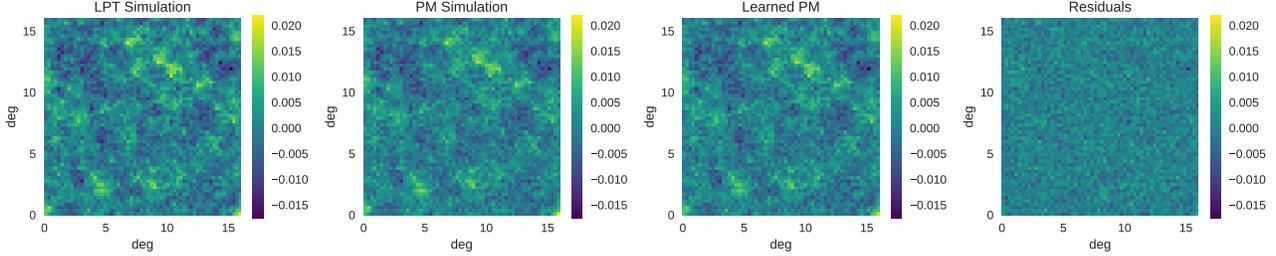


Figure 1. From left to right: An example of LPT simulations with noise; The corresponding PM simulation sharing the same noise, cosmological parameters and initial conditions as the LPT simulation; The LPT simulations optimally transporting onto the PM space; the residuals, i.e., the difference of the learned PM map with the true PM map without noise.

$p_t = [\phi_t^*]_{\#}(p_0)$ , with  $\phi_t^* = (1-t)\text{Id} + t\phi^*$ , is the conditional McCann interpolation (Kerrigan et al., 2024). This probability path and the corresponding triangular velocity field  $v_t(\phi_t^*(\theta, x)) = (0, \phi_X^*(\theta, x) - x)$ , solve the COT problem. This result, known as the conditional Benamou–Brenier theorem, establishes the equivalence between the static and dynamic COT problem (Kerrigan et al., 2024).

Hence, to learn the velocity that solves the COT problem, we use the loss introduced in Equation 3 with the same straight-line conditional probability paths and the same corresponding conditional velocity field  $v_t(x|z) = x_1 - x_0$ . However, to ensure that the resulting velocity field corresponds to the solution of the COT problem, the sample pairs must be built as  $x_1 = \phi_X^*(\theta_0, x_0)$  with  $\phi_X^*$  the optimal conditional Monge map. Moreover, to preserve the conditioning during transport, the learned velocity field must be triangular  $v_t(x, \theta) = (0, v_\varphi(t, x, \theta))$ , with  $v_\varphi(t, x, \theta)$  the NN, so that  $\theta$  remains unchanged.

In practice, solving the static COT problem at the dataset scale is computationally intractable. We therefore use a minibatch approximation, solving COT separately on each minibatch (Fatraş et al., 2021; Tong et al., 2024). Finally, to further ease optimization under finite sample constraints, we use a relaxed cost of the form  $c(x_0, \theta_0, x_1, \theta_1) = |\theta_0 - \theta_1| + \epsilon|x_0 - x_1|$ . As  $\epsilon \rightarrow 0$ , Hosseini et al., 2024 show that we recover the triangular map solving the conditional Monge problem. Note that when paired simulations from both models are available, finding  $\phi^*$  becomes unnecessary.

### 3. Experiments

#### 3.1. Simulation Models

We generate two suites of simulations: second-order LPT and PM N-body, using JaxPM<sup>2</sup>. Both sets of simulations are run in a cubic box of comoving size  $500 \times 500 \times 5000$  Mpc  $h^{-1}$  discretized into a  $60 \times 60 \times 200$

<sup>2</sup><https://github.com/DifferentiableUniverseInitiative/JaxPM>

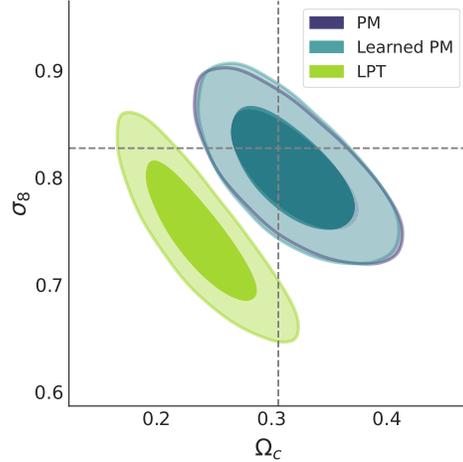


Figure 2. Posterior distributions of the cosmological parameters evaluated on the PM map shown in Figure 1, and learned from LPT simulations (green), PM simulations (purple), and emulated simulations (blue).

voxel grid, corresponding to a transverse resolution of  $8.3\text{Mpc } h^{-1}$  and a radial resolution of  $25\text{Mpc } h^{-1}$ . Displacements of the PM simulations are first computed using first-order LPT up to scale factor 0.1, and then according to the Poisson equation. The particle displacement is interpolated onto a 3D density grid using the Cloud-in-Cell (CIC) scheme, and we treat the resulting volume as our light cone. Convergence maps are then computed under the Born approximation by integrating density fluctuations along the line of sight. The resulting maps span an area of  $16^\circ \times 16^\circ$  on the sky, discretized into  $60 \times 60$  pixels (pixel scale  $\approx 0.27^\circ$ ). We assume a single source redshift distribution given by  $n(z) \propto z^{0.5} e^{-(z/2.0)^{1.0}}$  with a galaxy density of  $n_{\text{gal}} = 30 \text{ arcmin}^{-2}$ . Shape noise is added per pixel using  $\sigma_{\text{pix}} = \sigma_e / \sqrt{n_{\text{gal}} A_{\text{pix}}}$ , where  $\sigma_e = 0.26$  and  $A_{\text{pix}} = 256 \text{ arcmin}^2$ . Cosmological parameters  $\theta = (\Omega_c, \sigma_8)$  are drawn from independent Gaussians,  $\Omega_c \sim \mathcal{N}(0.3, 0.05^2)$  and  $\sigma_8 \sim \mathcal{N}(0.8, 0.05^2)$ . An example of the different convergence maps is shown in Figure 1.

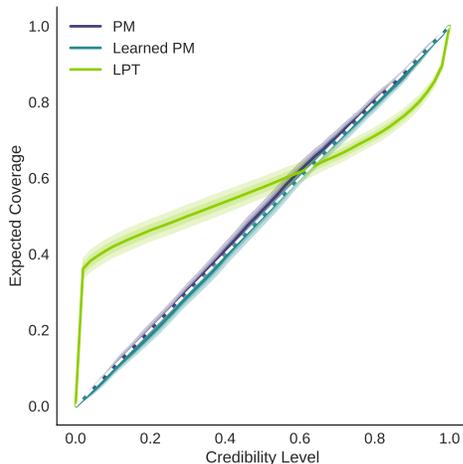


Figure 3. TARP expected coverage test over 500 simulations, with error bars from 100 bootstrap iterations. The dashed white line shows ideal coverage. Posteriors from PM maps analyzed with the correct model (purple) closely match this curve, while using the LPT model (green) leads to bias. COT-FM restores accurate coverage (blue), demonstrating effective emulation of PM maps.

Although we acknowledge that this is a simplistic setup, it is sufficient for initial testing purposes. A more complete study, with more realistic simulations will be presented in a forthcoming paper.

### 3.2. Implicit full-field inference

We perform full-field inference using an implicit inference approach that allows for amortized inference across various simulations, which is useful for assessing the robustness of our correction method. However, sampling schemes such as Hamiltonian Monte Carlo (Neal et al., 2011; Betancourt, 2017) can also be applied to our corrected forward model.

Convergence maps are first compressed into a two-dimensional summary statistic by training a convolutional neural network (CNN) under the Variational Mutual Information Maximization (VMIM) loss (Jeffrey et al., 2020), which has been shown to build near-sufficient statistics for parameter inference (Lanzieri et al., 2025). Then, we perform Neural Posterior Estimation (NPE) (Papamakarios & Murray, 2018) by training a Normalizing Flow (NF) on the compressed summaries to learn the posterior distribution over cosmological parameters. All training details, including the number of simulations, are provided in Appendix A.

### 3.3. Results

We begin by highlighting the simulation mismatch at the pixel level. To this end, we apply our implicit inference framework: For the LPT simulations, we train the CNN compressor and NF posterior estimator using only LPT sam-

ples. For the PM simulations, we repeat the same procedure but using PM samples. Once trained, we evaluate the two inferred posterior distributions on a fixed PM fiducial convergence map (shown in Figure 1). The results in Figure 2 illustrate a significant bias when analyzing PM data with a model trained on LPT simulations.

To correct this bias, we apply COT-FM. As explained in subsection 2.2, COT-FM learns a COT map that transforms LPT simulations into PM-like simulations, preserving the associated cosmological parameters and finding the minimal correction to the LPT simulation needed to match its PM counterpart, a process that is effective even when only two unpaired sets of simulations are available. Our mapping is parameterized using a 2D conditional U-Net (Ronneberger et al., 2015) trained on two noisy datasets: the set of LPT simulations and the set of PM simulations. At each training step, we solve the COT problem on the mini-batch to effectively pair each LPT and PM simulation, allowing us to learn the velocity field that solves the dynamic COT problem. Once the velocity field is learned, we integrate the forward ODE to transport LPT samples into their corrected PM-like counterparts. An example of a transported convergence map is shown in Figure 1. Training details, including the number of simulations, can be found in Appendix A.

As a first validation, we project the original LPT, learned PM, and true PM simulations into summary space using the PM-trained compressor. As shown in Figure 4, the learned PM simulations closely align with the true PM simulations, while differing from the original LPT simulations. We complete the evaluation by performing implicit full-field inference, i.e., training the CNN compressor to build sufficient statistics and training the NF posterior estimator using the learned PM convergence maps. We evaluate the posterior distribution on the same fixed convergence maps as for LPT and PM posteriors. The results in Figure 2 show that COT-FM effectively corrects the inference bias, demonstrating the emulator accuracy. To further validate the posterior distributions across multiple simulations, we use the TARP (Lemos et al., 2023) expected coverage probability (ECP) test. The ECP corresponds to the expected number of posterior samples that fall in a credible region of credibility level  $\alpha$ . TARP provides a method to estimate this ECP such that it is precisely equal  $\alpha$  for all  $\alpha \in [0, 1]$  if and only if the posterior distribution is calibrated. In Figure 3, we apply TARP to the posterior distributions obtained from the three simulation models and evaluated on 500 PM observations. We show that, unlike the LPT maps, the learn maps recover the true posterior distributions, as indicated by the ECP falling along the diagonal. Finally, additional posterior distributions are displayed in Figure 5 with their corresponding convergence maps displayed in Figure 6.

## 4. Conclusion and Discussion

We have presented a new method to accurately emulate simulations at the pixel level. Instead of modeling complex simulation dynamics end-to-end, our approach focuses on transforming simulation likelihoods. This significantly reduces the complexity of the learning task, requiring fewer training simulations and lowering the risk of generating unphysical outputs. In addition, by framing the problem as a transport between distributions under the COT constraint, the method applies minimal yet targeted corrections to align forward models. We also note that because the transport operates on likelihood distributions, the emulator is naturally conditioned on cosmological parameters. The learned transformation is differentiable, making it suitable for explicit full-field inference that relies on gradient sampling schemes. Finally, the COT-FM framework allows training on unpaired datasets, which increases flexibility when exact simulation pairs are not available. Applied to a toy weak lensing example, we demonstrated the effectiveness of this method by successfully recovering the full-field posterior distribution using our emulated convergence maps.

In future work, we plan to more rigorously investigate the limitations of this approach, for instance, when the LF and HF distributions differ significantly or in high-noise regimes. We also aim to extend our experiments to more realistic convergence maps from Stage-IV surveys, for example, by mapping PM simulations to hydrodynamic simulations, incorporating baryonic effects, and modeling observational systematics. Additionally, we will explore transporting full simulation volumes directly. We note that, for this method to be effective, it is crucial that COT-FM training works well, even with a limited number of high-quality simulations. This study will be conducted in future work. Another potential direction is to condition the transport not only on cosmological parameters  $\theta$ , but also on the initial conditions. This would ensure that paired simulations share the same IC, potentially yielding even more minimal transports. We believe that this framework also offers a promising avenue for future extensions, in which it could be used as a correction tool for bridging simulations to real data. However, it will require rethinking assumptions made in this study, since our Universe has a fixed but unknown set of cosmological parameters.

## Acknowledgements

The authors would like to thank Nikolay Malkin and Alexandre Adam for their valuable discussions. We also thank the reviewers of the ML4Astro 2025 workshop for their constructive feedback. This work is partially supported by Schmidt Futures, a philanthropic initiative founded by Eric and Wendy Schmidt as part of the Virtual Institute for Astrophysics (VIA). The work is in part supported by computa-

tional resources provided by Calcul Quebec and the Digital Research Alliance of Canada. Y.H. and L.P. acknowledge support from the Canada Research Chairs Program, the National Sciences and Engineering Council of Canada through grants RGPIN-2020-05073 and 05102.

## References

- Albergo, M. S., Boffi, N. M., and Vanden-Eijnden, E. Stochastic interpolants: A unifying framework for flows and diffusions, 2023. URL <https://arxiv.org/abs/2303.08797>.
- Baptista, R., Hosseini, B., Kovachki, N. B., and Marzouk, Y. Conditional sampling with monotone gans: from generative models to likelihood-free inference, 2023. URL <https://arxiv.org/abs/2006.06755>.
- Bartlett, D. J., Chiarenza, M., Doerer, L., and Leclercq, F. Comoving computer acceleration (coca): N-body simulations in an emulated frame of reference. *Astronomy & Astrophysics*, 694:A287, February 2025. ISSN 1432-0746. doi: 10.1051/0004-6361/202452217. URL <http://dx.doi.org/10.1051/0004-6361/202452217>.
- Bayer, A. E., Villaescusa-Navarro, F., Sharief, S., Teyssier, R., Garrison, L. H., Perreault-Levasseur, L., Bryan, G. L., Gatti, M., and Visbal, E. Field-level comparison and robustness analysis of cosmological n-body simulations, 2025. URL <https://arxiv.org/abs/2505.13620>.
- Betancourt, M. A conceptual introduction to hamiltonian monte carlo. *arXiv preprint arXiv:1701.02434*, 2017.
- Chen, R. T. Q., Rubanova, Y., Bettencourt, J., and Duvenaud, D. Neural ordinary differential equations, 2019. URL <https://arxiv.org/abs/1806.07366>.
- Dai, B., Feng, Y., and Seljak, U. A gradient based method for modeling baryons and matter in halos of fast simulations. *Journal of Cosmology and Astroparticle Physics*, 2018(11):009–009, November 2018. ISSN 1475-7516. doi: 10.1088/1475-7516/2018/11/009. URL <http://dx.doi.org/10.1088/1475-7516/2018/11/009>.
- Dinh, L., Sohl-Dickstein, J., and Bengio, S. Density estimation using real nvp, 2017. URL <https://arxiv.org/abs/1605.08803>.
- Dormand, J. R. and Prince, P. J. A family of embedded Runge–Kutta formulae. *J. Comp. Appl. Math.*, 6:19–26, 1980.
- Fatras, K., Zine, Y., Majewski, S., Flamary, R., Gribonval, R., and Courty, N. Minibatch optimal transport dis-

- tances; analysis and applications, 2021. URL <https://arxiv.org/abs/2101.01792>.
- Filipp, A., Hezaveh, Y., and Perreault-Levasseur, L. Robustness of neural ratio and posterior estimators to distributional shifts for population-level dark matter analysis in strong gravitational lensing, 2024. URL <https://arxiv.org/abs/2411.05905>.
- Flamary, R., Courty, N., Gramfort, A., Alaya, M. Z., Boissunon, A., Chambon, S., Chapel, L., Corenflos, A., Fatras, K., Fournier, N., Gautheron, L., Gayraud, N. T., Janati, H., Rakotomamonjy, A., Redko, I., Rolet, A., Schutz, A., Seguy, V., Sutherland, D. J., Tavenard, R., Tong, A., and Vayer, T. Pot: Python optimal transport. *Journal of Machine Learning Research*, 22(78):1–8, 2021. URL <http://jmlr.org/papers/v22/20-451.html>.
- Fluri, J., Kacprzak, T., Lucchi, A., Schneider, A., Refregier, A., and Hofmann, T. Full  $w$  cdm analysis of kids-1000 weak lensing maps using deep learning. *Physical Review D*, 105(8):083518, 2022.
- He, S., Li, Y., Feng, Y., Ho, S., Ravanbakhsh, S., Chen, W., and Póczos, B. Learning to predict the cosmological structure formation. *Proceedings of the National Academy of Sciences*, 116(28):13825–13832, June 2019. ISSN 1091-6490. doi: 10.1073/pnas.1821458116. URL <http://dx.doi.org/10.1073/pnas.1821458116>.
- Hosseini, B., Hsu, A. W., and Taghvaei, A. Conditional optimal transport on function spaces, 2024. URL <https://arxiv.org/abs/2311.05672>.
- Jamieson, D., Li, Y., de Oliveira, R. A., Villaescusa-Navarro, F., Ho, S., and Spergel, D. N. Field-level neural network emulator for cosmological  $n$ -body simulations. *The Astrophysical Journal*, 952(2):145, July 2023. ISSN 1538-4357. doi: 10.3847/1538-4357/acdb6c. URL <http://dx.doi.org/10.3847/1538-4357/acdb6c>.
- Jeffrey, N., Alsing, J., and Lanusse, F. Likelihood-free inference with neural compression of des sv weak lensing map statistics. *Monthly Notices of the Royal Astronomical Society*, 501(1):954–969, November 2020. ISSN 1365-2966. doi: 10.1093/mnras/staa3594. URL <http://dx.doi.org/10.1093/mnras/staa3594>.
- Jeffrey, N., Whiteway, L., Gatti, M., Williamson, J., Alsing, J., Porredon, A., Prat, J., Doux, C., Jain, B., Chang, C., Cheng, T. Y., Kacprzak, T., Lemos, P., Alarcon, A., Amon, A., Bechtol, K., Becker, M. R., Bernstein, G. M., Campos, A., Rosell, A. C., Chen, R., Choi, A., DeRose, J., Drlica-Wagner, A., Eckert, K., Everett, S., Ferté, A., Gruen, D., Gruendl, R. A., Herner, K., Jarvis, M., McCullough, J., Myles, J., Navarro-Alsina, A., Pandey, S., Raveri, M., Rollins, R. P., Rykoff, E. S., Sánchez, C., Secco, L. F., Sevilla-Noarbe, I., Sheldon, E., Shin, T., Troxel, M. A., Tutusaus, I., Varga, T. N., Yanny, B., Yin, B., Zuntz, J., Aguena, M., Allam, S. S., Alves, O., Bacon, D., Bocquet, S., Brooks, D., da Costa, L. N., Davis, T. M., Vicente, J. D., Desai, S., Diehl, H. T., Ferrero, I., Frieman, J., García-Bellido, J., Gaztanaga, E., Giannini, G., Gutierrez, G., Hinton, S. R., Hollowood, D. L., Honscheid, K., Huterer, D., James, D. J., Lahav, O., Lee, S., Marshall, J. L., Mena-Fernández, J., Miquel, R., Pieres, A., Malagón, A. A. P., Roodman, A., Sako, M., Sanchez, E., Cid, D. S., Smith, M., Suchyta, E., Swanson, M. E. C., Tarle, G., Tucker, D. L., Weaverdyck, N., Weller, J., Wiseman, P., and Yamamoto, M. Dark energy survey year 3 results: likelihood-free, simulation-based  $w$ cdm inference with neural compression of weak-lensing map statistics, 2024. URL <https://arxiv.org/abs/2403.02314>.
- Kerrigan, G., Migliorini, G., and Smyth, P. Dynamic conditional optimal transport through simulation-free flows, 2024. URL <https://arxiv.org/abs/2404.04240>.
- Kidger, P. *On Neural Differential Equations*. PhD thesis, University of Oxford, 2021.
- Kodi Ramanah, D., Charnock, T., Villaescusa-Navarro, F., and Wandelt, B. D. Super-resolution emulator of cosmological simulations using deep physical models. *Monthly Notices of the Royal Astronomical Society*, 495(4):4227–4236, May 2020. ISSN 1365-2966. doi: 10.1093/mnras/staa1428. URL <http://dx.doi.org/10.1093/mnras/staa1428>.
- Lanzieri, D., Lanusse, F., and Starck, J.-L. Hybrid physical-neural odes for fast  $n$ -body simulations, 2022. URL <https://arxiv.org/abs/2207.05509>.
- Lanzieri, D., Zeghal, J., Makinen, T. L., Boucaud, A., Starck, J.-L., and Lanusse, F. Optimal neural summarisation for full-field weak lensing cosmological implicit inference, 2025. URL <https://arxiv.org/abs/2407.10877>.
- Lemos, P., Coogan, A., Hezaveh, Y., and Perreault-Levasseur, L. Sampling-based accuracy testing of posterior estimators for general inference, 2023. URL <https://arxiv.org/abs/2302.03026>.
- Lipman, Y., Chen, R. T. Q., Ben-Hamu, H., Nickel, M., and Le, M. Flow matching for generative modeling, 2023. URL <https://arxiv.org/abs/2210.02747>.
- Lipman, Y., Havasi, M., Holderrieth, P., Shaul, N., Le, M., Karrer, B., Chen, R. T. Q., Lopez-Paz, D., Ben-Hamu, H., and Gat, I. Flow matching guide and code, 2024. URL <https://arxiv.org/abs/2412.06264>.

- Lucie-Smith, L., Peiris, H. V., Pontzen, A., and Lochner, M. Machine learning cosmological structure formation. *Monthly Notices of the Royal Astronomical Society*, 479(3):3405–3414, June 2018. ISSN 1365-2966. doi: 10.1093/mnras/sty1719. URL <http://dx.doi.org/10.1093/mnras/sty1719>.
- Neal, R. M. et al. Mcmc using hamiltonian dynamics. *Handbook of markov chain monte carlo*, 2(11):2, 2011.
- Papamakarios, G. and Murray, I. Fast  $\epsilon$ -free inference of simulation models with bayesian conditional density estimation, 2018. URL <https://arxiv.org/abs/1605.06376>.
- Payot, N., Lemos, P., Perreault-Levasseur, L., Cuesta-Lazaro, C., Modi, C., and Hezaveh, Y. Learning an effective evolution equation for particle-mesh simulations across cosmologies, 2023. URL <https://arxiv.org/abs/2311.18017>.
- Pooladian, A.-A., Ben-Hamu, H., Domingo-Enrich, C., Amos, B., Lipman, Y., and Chen, R. T. Q. Multisample flow matching: Straightening flows with minibatch couplings, 2023. URL <https://arxiv.org/abs/2304.14772>.
- Porqueres, N., Heavens, A., Mortlock, D., Lavaux, G., and Mäkinen, T. L. Field-level inference of cosmic shear with intrinsic alignments and baryons, 2023. URL <https://arxiv.org/abs/2304.04785>.
- Ronneberger, O., Fischer, P., and Brox, T. U-net: Convolutional networks for biomedical image segmentation, 2015. URL <https://arxiv.org/abs/1505.04597>.
- Shampine, L. F. Some practical Runge-Kutta formulas. *Mathematics of Computation*, 46(173):135–150, 1986. doi: <https://doi.org/10.2307/2008219>.
- Tong, A., Fatras, K., Malkin, N., Huguet, G., Zhang, Y., Rector-Brooks, J., Wolf, G., and Bengio, Y. Improving and generalizing flow-based generative models with minibatch optimal transport, 2024. URL <https://arxiv.org/abs/2302.00482>.

## A. Training details and model architectures

**Implicit full-field inference** is performed by first compressing  $60 \times 60$  pixels convergence maps into 2-dimensional summary statistics using a 2D convolutional neural network (CNN) composed of three convolutional layers. The CNN is trained using the VMIM loss function (Jeffrey et al., 2020) on a dataset of 60 000 simulations. To compute the VMIM loss function, we use a RealNVP NF (Dinh et al., 2017) with 4 coupling layers. After training, the compressor is used to compress a new set of 40 000 simulations. A new RealNVP with 4 coupling layers is then trained on the compressed summary statistics using the NPE loss function to approximate the posterior distribution. The architectures of both the compressor and the NFs are kept fixed across all experiments.

**COT-FM** is performed using a 2D conditional U-Net (Ronneberger et al., 2015) from the `Diffusers` library<sup>3</sup>, which we adapt to put cosmological parameters as an additional conditioning input alongside the time. The U-Net is trained under the COT-FM loss to learn the marginal velocity field, using a dataset of 90 000 simulations (for both LPT and PM), different from the datasets used in implicit full-field inference. We train with a batch size of 200 simulations and use the exact linear programming Earth Mover’s Distance (EMD) algorithm from the `POT` library (Flamary et al., 2021) to solve the COT problem in each batch. For the transport cost, we set  $\epsilon = 0.1$ . Note that larger batch sizes and smaller  $\epsilon$  improve the OT approximation. Although we did not fine-tune these parameters in this study, we plan to address them in future work. Finally, we integrate the learned velocity field using the Dopri5 ODE solver (Dormand & Prince, 1980; Shampine, 1986) from the `diffraX` library (Kidger, 2021).

We trained COT-FM for seven hours on an NVidia A100SXM4.

## B. Additional figures

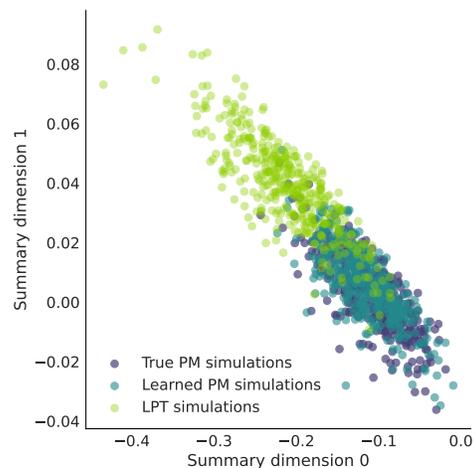


Figure 4. Simulations comparison in summary space. Using the compressor trained on PM maps, we compress PM maps (purple), learned PM maps (blue), and LPT maps (green). We show that our emulated PM maps align with the true PM ones.

<sup>3</sup><https://huggingface.co/docs/diffusers/>

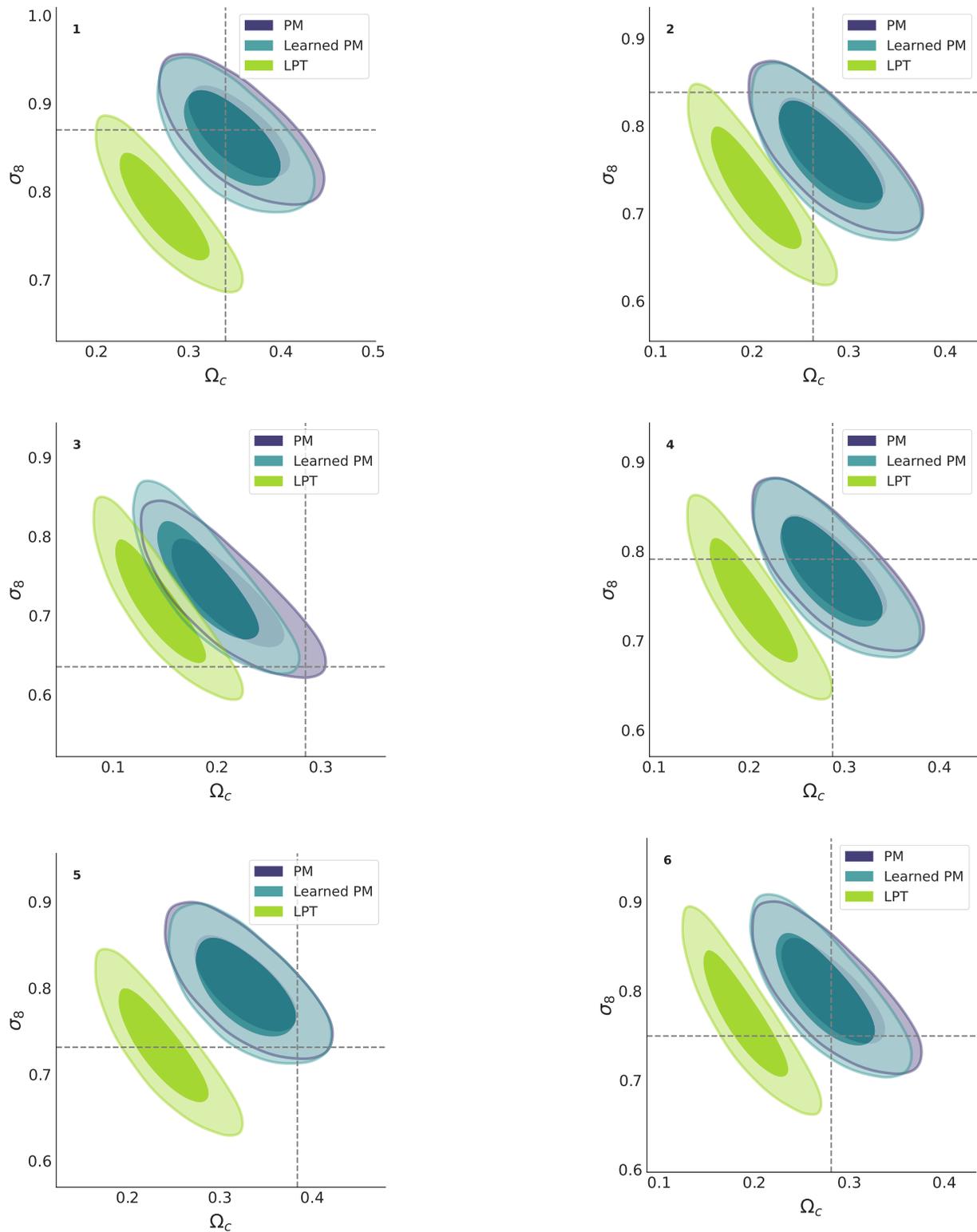


Figure 5. Additional posterior distributions of the cosmological parameters evaluated on the PM fiducial maps shown in Figure 6. The green contour corresponds to the posterior distribution learned using LPT simulations. The purple one corresponds to the posterior distribution learned using PM simulations. The blue one is the posterior distribution learned from corrected LPT simulations, i.e., PM-like simulations. This highlights the impact of model misspecification on the cosmological constraints and how COT-FM solves this problem.

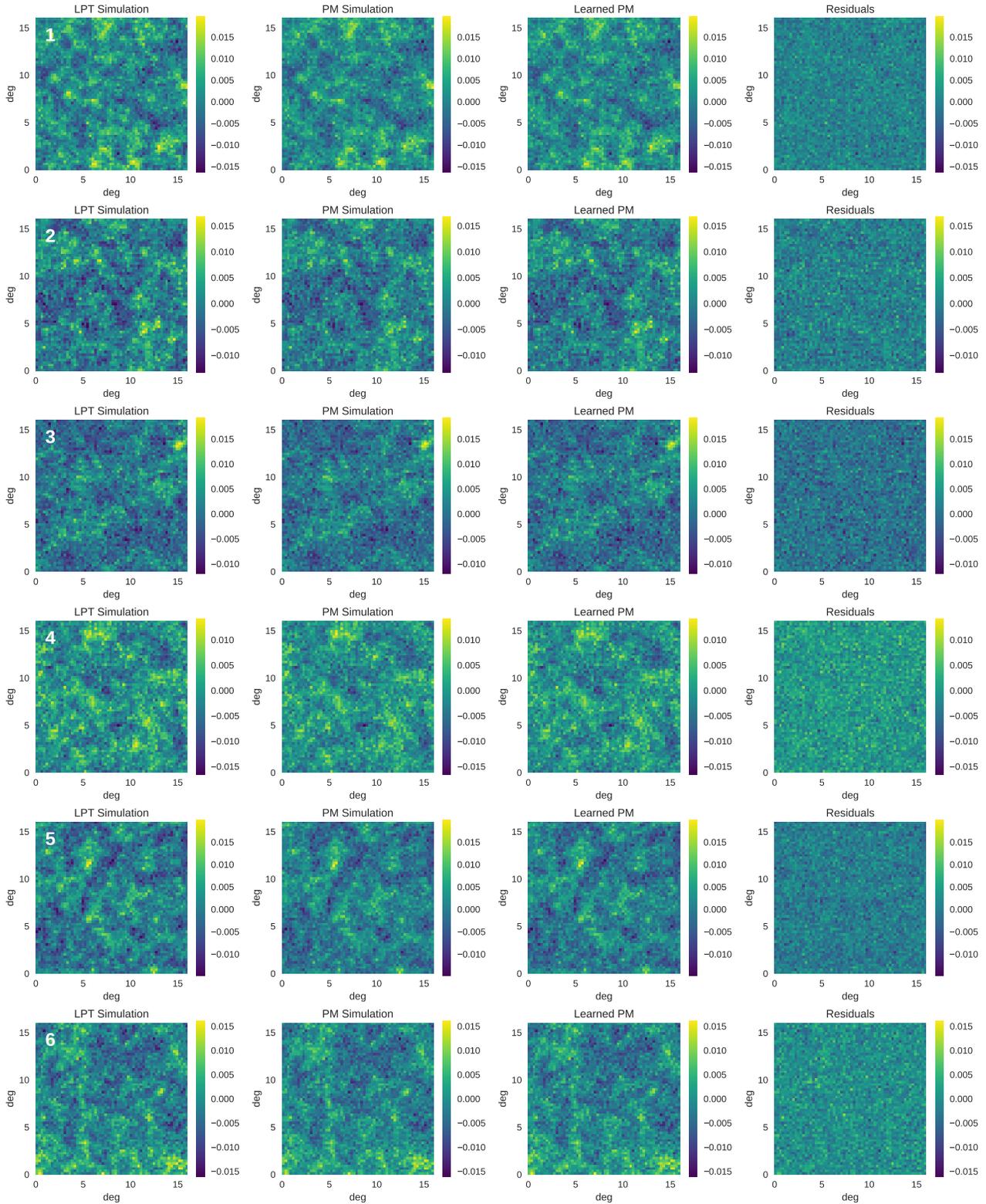


Figure 6. Additional convergence maps. From left to right: LPT simulations with noise; The corresponding PM simulation sharing the same noise, cosmological parameters and initial conditions as the LPT simulation; The LPT simulations optimally transporting onto the PM space; the residuals, i.e., the difference of the learned PM map with the true PM map without noise.