Hybrid Physical-Deep Learning Model for Astronomical Inverse Problems

François Lanusse
Berkeley Center for Cosmological Physics
Berkeley Institute for Data Science
University of California, Berkeley
Berkeley, CA 94709
flanusse@berkeley.edu

Peter Melchior
Department of Astrophysical Sciences
Center for Statistics and Machine Learning
Princeton University
Princeton, NJ 08544, USA
peter.melchior@princeton.edu

Fred Moolekamp
LSST Project Management Office
Tucson, AZ, USA
Department of Astrophysical Sciences

Princeton University
Princeton, NJ 08544, USA
fredem@astro.princeton.edu

Abstract

We present a Bayesian machine learning architecture that combines a physically motivated parametrization and an analytic error model for the likelihood with a deep generative model providing a powerful data-driven prior for complex signals. This combination yields an interpretable and differentiable generative model, allows the incorporation of prior knowledge, and can be utilized for observations with different data quality without having to retrain the deep network. We demonstrate our approach with an example of astronomical source separation in current imaging data, yielding a physical and interpretable model of astronomical scenes.

1 Introduction

Deep Learning is extremely efficient at solving a wide range of inverse problems, but what is gained in performance is often lost in interpretability due the black box nature of deep neural networks. In scientific applications however, the interpretability of the solution is often paramount, as is the ability to imbue pre-existing physical knowledge directly into the model. To this end, we propose an hybrid model based on solving for the Maximum A Posteriori (MAP) solution of an inverse problem, where a physical model is used to describe the forward data acquisition process, and a deep generative model with explicit likelihood is used to provide a complex data-driven signal prior.

Related Work With the success of deep learning in many classical imaging problems [e.g. 1], a significant amount of effort has been aimed at linking deep learning successes back to the classical inverse problems literature. Several avenues have been explored, e.g. learning a denoiser as an implicit proximal operator [2]; using a generative model as deep image prior [3]; learning a convolutional dictionary and proximal operator [4]; or learning implicitly both the prior and the inference algorithm itself [5]. Finally, most closely related to our work, [6] introduces the idea of modeling the prior explicitly at the pixel-level using a deep generative model.

33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.

2 Problem statement

We consider a general linear inverse problems of the form:

$$\boldsymbol{y} = \mathbf{A}\boldsymbol{x} + \boldsymbol{n} \,, \tag{1}$$

where $\boldsymbol{y}$ are the observations, $\boldsymbol{x}$ is the unknown signal to recover, $\mathbf{A}$ is a linear degradation operator, $\boldsymbol{n}$ is some observational noise. This problem describes a wide range of applications from MRI to radio-interferometry through different choices of the operator $\mathbf{A}$.

In the Bayesian approach to inverse problems, the posterior $p(\boldsymbol{x}|\boldsymbol{y})$ can be expressed as

$$p(\boldsymbol{x}|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{x})\; p(\boldsymbol{x}). \tag{2}$$

- The *data likelihood* term $p(\boldsymbol{y}|\boldsymbol{x})$ encodes our physical understanding of the forward process that leads to the observations. We assume here that for a given $\boldsymbol{x}$, this term can be evaluated explicitly given our physical model.
- The *prior* term $p(\boldsymbol{x})$ encodes our prior knowledge on the solution we seek to recover. This prior can be informed by prior experiments, complementary data, or physical considerations. A tractable expression of this prior can often be obtained only for simple signal classes.

While the Bayesian solution to such inverse problem is the full posterior $p(\boldsymbol{x}|\boldsymbol{y})$, in many practical applications the full distribution is typically reduced to a single point estimate, i.e. the Maximum A Posteriori (MAP) solution

$$\boldsymbol{x}_{MAP} = \arg\max_{x} \; \log p(\boldsymbol{y}|\boldsymbol{x}) + \log p(\boldsymbol{x}) \,. \tag{3}$$

We will exploit this separable representation, limiting the use of deep learning to the prior term $p(\boldsymbol{x})$.

Model for the data likelihood With a given operator $\mathbf{A}$, the likelihood term is completely characterized by the noise model for $\boldsymbol{n}$. We assume a Gaussian noise model, i.e. $\boldsymbol{n} \sim \mathcal{N}(0, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma}$ is the noise covariance. In this case $\log p(\boldsymbol{y}|\boldsymbol{x}) = -\frac{1}{2}\|\boldsymbol{y} - \mathbf{A}\boldsymbol{x}\|^2_{\boldsymbol{\Sigma}^{-1}} + cst$, where $\|\boldsymbol{x}\|^2_{\mathbf{M}} = \boldsymbol{x}^\top \mathbf{M}\boldsymbol{x}$.

Deep generative models as complex data priors Models like Variational AutoEncoders (VAEs) [7] and Generative Adversarial Networks (GANs) [8] have been extremely successful, but they do not provide an explicit likelihood $p(\boldsymbol{x})$. Instead we choose to rely on pixel autoregressive models [9, 10, 11] which provide an explicit likelihood, factorized into separate conditional distributions $p_\theta(\boldsymbol{x}) = \Pi_i p_\theta(x_i|x_0 \dots x_{i-1})$ where θ are the weights of the model. These models achieve state of the art performance, are stable during training, and do not suffer from mode collapse (contrary to GANs). Our prior model is trained on uncorrupted examples of data $\boldsymbol{x}$, which may come from simulations or from high-fidelity observations.

Physical constraints as proximal regularization terms In this approach, most of the prior information stems from the data-driven deep generative model $p_\theta(\boldsymbol{x})$. However, additional physical constraints on the solution can applied by adding regularization terms to the prior: $\log p(\boldsymbol{x}) = \log p_\theta(\boldsymbol{x}) + \sum_j R_j(\boldsymbol{x})$. The regularizers may be non-differentiable as long as they can be expressed by their proximal operators $\mathrm{prox}_{R_j}(x)$. As an example, the flux of astronomical sources is a positive quantity, we may therefore want to impose a non-negativity constraint $\iota_{>0}(\boldsymbol{x})$ on the solution, using the associated proximal operator $\mathrm{prox}_{\iota_{>0}}(x) = \max(0, x)$.

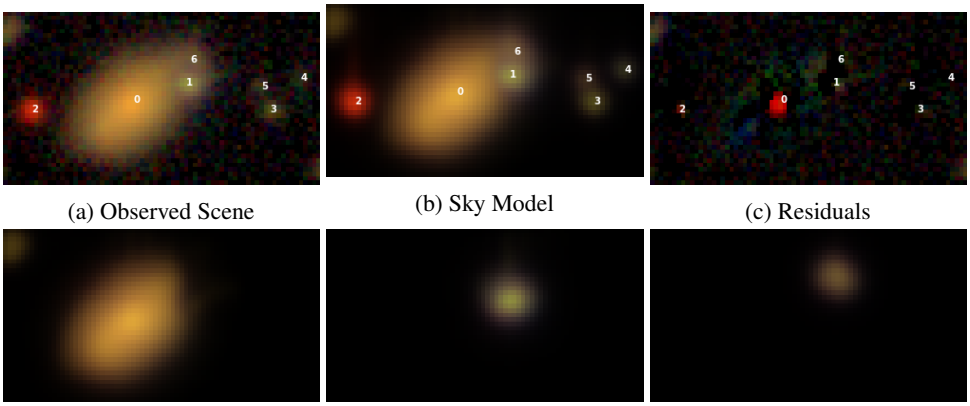
Combining all these elements, we can characterize our x_{MAP} solution as the minimum of the following loss function:

$$\mathcal{L} = \frac{1}{2}\|\boldsymbol{y} - \mathbf{A}\boldsymbol{x}\|^2_{\boldsymbol{\Sigma}^{-1}} - \log p_\theta(\boldsymbol{x}) + \sum_i R_i(\boldsymbol{x}) \equiv f(\boldsymbol{x}) + g(\boldsymbol{x}) + r(\boldsymbol{x}) \,. \tag{4}$$

The two first components of this loss are differentiable and therefore amenable to gradient descent. Due to the presence of non-differentiable regularizers, the optimization makes use of the iterative Proximal Gradient Method (also known as forward-backward splitting [12])

$$\boldsymbol{x}_{t+1} = \underset{\lambda_t\, r}{\mathrm{prox}}\left(\boldsymbol{x}_t - \lambda_t \nabla(f+g)(\boldsymbol{x}_t)\right), \tag{5}$$

which converges to a minimum of $\mathcal{L}$ if the step size λ is smaller than $2/L$, where L is the Lipschitz constant of the gradient term.



(a) Observed Scene (b) Sky Model (c) Residuals

(d) Individual components of the model

Figure 1: Deblending of a scene from the HSC imaging survey using the proposed model. The sky model (b), composed of individual sources (marked in white) with their own morphology and SED, is fitted to the observations (a). A subset of the individual recovered components are shown in (d).

3 Application: deblending galaxy images

Modern wide-field cosmological surveys cover large areas of the sky with ever increasing imaging quality [e.g. 13, 14]. They often seek to simultaneously address many scientific goals, e.g. mapping the large scale structure of the Universe to answer fundamental questions on the nature of Dark Matter and Dark Energy. One outstanding challenge faced by all modern imaging surveys is the overlap of several sources (stars and galaxies) on the sky, so-called "blending". It complicates the measurement of properties of individual members of the blend. This problem constitutes a typical instance of blind source separation, which attempts to separate all components of a blended scene without *a priori* knowledge of their nature (see Figure 1). A GAN-based approach to deblending was introduced in [15] but suffers from the typical limitations of black-box deep learning models, i.e. it cannot account for different observing conditions or noise levels without retraining the full model, is limited to separating two components, and has no test-time flux preservation.

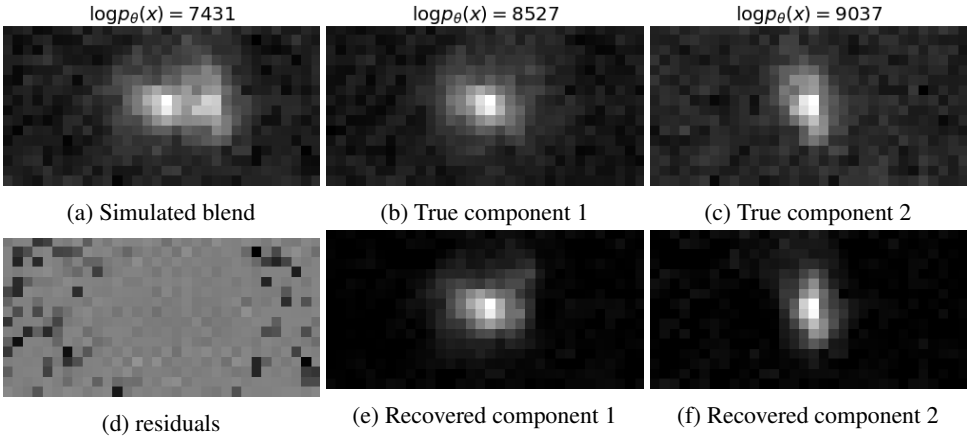
We propose to address the deblending problem using the analytic model of the scene introduced by [16]. Given an astronomical scene $\boldsymbol{y} \in \mathbb{R}^{B\times N\times N}$ observed in B bands (i.e. using multiple filters), each source k in the scene is modeled with a non-parametric shape $S_k \in \mathbb{R}^{N\times N}$ and an amplitude $A_k \in \mathbb{R}^B$, the so-called Spectral Energy Distribution (SED), which determines how bright the object will appear in each band. Multiple sources contribute additively to the scene, which is correct in the absence of absorbers, e.g. inter-stellar dust. The forward model also needs to account for degradation of the image caused by the atmosphere and the instrumental optics. This can be described as a band-wise 2D convolution by a Point Spread Function (PSF). We denote $\mathbf{P}$ as block-diagonal linear operator implementing the convolution in each band by the appropriate PSF. It acts as the operator $\mathbf{A}$ in Equation 1. Our full physical model for the scene can now be expressed as:

$$\boldsymbol{y} = \mathbf{P}\sum_{k=1}^{K} A_k^T \times S_k + \boldsymbol{n} \tag{6}$$

where $\boldsymbol{n} \in \mathbb{R}^{B\times N\times N}$ is typically assumed to be Gaussian noise with covariance Σ. The deblending problem is to recover an estimate of both the morphology S_k and the SED A_k of each component of the blend, subject to additional constraints such as positivity of the source emission ($S_k > 0$ and $A_k > 0$). Applying the framework described in section 2, we solve the optimization problem

$$\underset{S_k, A_k}{\arg\min}\, \frac{1}{2}\|\boldsymbol{y} - \mathbf{P}\sum_{k=1}^{K} A_k^T \times S_k\|^2_{\boldsymbol{\Sigma}^{-1}} + \sum_{k}^{K} \log p_\theta(S_k) + \iota_{>0}(A_k) + \iota_{>0}(S_k), \tag{7}$$

by a block-wise application of Equation 5 to every optimization variable. We base the morphology prior p_θ on the PixelCNN++ model [10], which we adjust for continuous signals by using a simple Gaussian model for the conditional distribution $p_\theta(x_i|x_0 \dots x_{i-1})$. The prior p_θ is trained on an existing set of high-resolution images of isolated galaxies [17, 18, 19] from the Hubble Space Telescope (HST)/Advanced Camera for Surveys (ACS) COSMOS survey [20]. These single-band



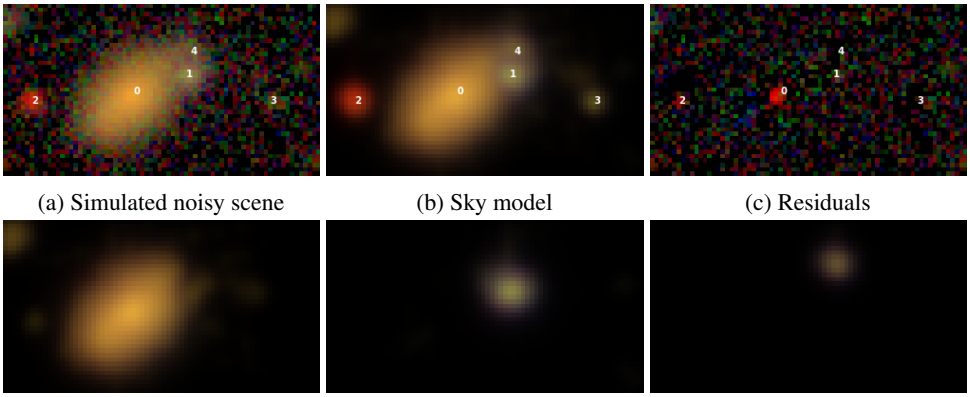
(a) Simulated blend (b) True component 1 (c) True component 2

(d) residuals (e) Recovered component 1 (f) Recovered component 2

Figure 2: Artificial blend experiment, simulated from HST observations, demonstrating the ability of the prior to disentangle sources from their morphologies alone. The log likelihood evaluated on the trained prior for the blended and isolated sources is reported in the top row: a blended object (a) has lower log likelihood than isolated objects (b, c) under the prior. This allows us to separate (e) and (f) when given (a).

high-resolution images from HST are reconvolved with a uniform reference PSF and resampled to match the pixel scale of the test survey, the Hyper Suprime-Cam (HSC) survey from the Subaru 8.2 meter telescope on Maunakea, Hawaii [21]. Figure 2b and Figure 2c show examples of two isolated galaxies obtained by this procedure, annotated with the log likelihood $\log p_\theta(x)$ of the trained prior. Figure 2a shows a simulated blend obtained by adding these two isolated galaxies. The log likelihood under the prior for both galaxies combined is lower than for each isolated galaxy, demonstrating that the morphology prior provides information that can be leveraged for deblending.

Equipped with the prior on single-galaxy morphology, we tackle the blended scene in Figure 1 from the first public HSC data. The observations, made in $B = 5$ different filters, are modelled with Equation 6, where the number of sources K and a first guess of their positions is provided by an external detection algorithm. We infer the parameters of the scene by solving Equation 7, yielding a sky model (Figure 1b) which can be separated into its individual components (Figure 1d). The model creates an excellent fit to morphologies and SEDs despite the strong overlap of several sources. The residuals are dominated by an undetected nuclear component in the brightest galaxy, which could be modeled by adding another component there. Finally, our model analytically accounts for different observing conditions. We provide in Appendix A a comparison to the state-of-the-art SCARLET deblender [16] on this scene to highlight the benefits of the deep morphology prior. In Figure 3 we show the same blended scene, but we artificially increased the noise RMS by a factor of 3. By adjusting the noise covariance in Equation 7, the methods recovers a very similar result without the



(a) Simulated noisy scene (b) Sky model (c) Residuals

(d) Individual components of the model

Figure 3: Same as Figure 1, but with artificially increased noise (RMS $\times$ 3).

need to retrain the deep learning prior. We could also address e.g. additional blurring from a wider PSF or resampling to lower resolution with the same approach.

4 Conclusion

We have presented a hybrid Bayesian framework for inverse problems that combines analytic forward modeling for the likelihood with deep generative models for complex data-driven signal priors. This approach makes explicit use of physically motivated problem structure and prior knowledge from high-quality observations. When applied to the blind-source separation problem of galaxy blending, we can retrieve multi-components models of astronomical scenes that are by construction robust to changes in observational conditions.

References

[1] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image Super-Resolution Using Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2):295–307, 2016.

[2] Tim Meinhardt, Michael Möller, Caner Hazirbas, and Daniel Cremers. Learning Proximal Operators: Using Denoising Networks for Regularizing Inverse Imaging Problems. (October), 2017.

[3] Victor Lempitsky, Andrea Vedaldi, and Dmitry Ulyanov. Deep Image Prior. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 9446–9454, 2018.

[4] Yan Yang, Jian Sun, Huibin Li, and Zongben Xu. Deep ADMM-Net for compressive sensing MRI. *Advances in Neural Information Processing Systems*, (Nips):10–18, 2016.

[5] Patrick Putzky and Max Welling. Recurrent Inference Machines for Solving Inverse Problems. (Nips), 2017.

[6] Akshat Dave, Anil Kumar Vadathya, Ramana Subramanyam, Rahul Baburajan, and Kaushik Mitra. Solving Inverse Computational Imaging Problems Using Deep Pixel-Level Prior. *IEEE Transactions on Computational Imaging*, 5(1):37–51, 2018.

[7] Diederik P Kingma and Max Welling. Auto-Encoding Variational Bayes. 587(Ml), 12 2013.

[8] Ij Goodfellow, J Pouget-Abadie, and Mehdi Mirza. Generative Adversarial Networks. *arXiv preprint arXiv:* . . . , pages 1–9, 2014.

[9] Aaron van den Oord, Nal Kalchbrenner, and Koray Kavukcuoglu. Pixel Recurrent Neural Networks. *International Conference on Machine Learning (ICML)*, 48, 2016.

[10] Tim Salimans, Andrej Karpathy, Xi Chen, and Diederik P. Kingma. PixelCNN++: Improving the PixelCNN with Discretized Logistic Mixture Likelihood and Other Modifications. pages 1–10, 2017.

[11] Xi Chen, Nikhil Mishra, Mostafa Rohaninejad, and Pieter Abbeel. PixelSNAIL: An improved autoregressive generative model. *35th International Conference on Machine Learning, ICML 2018*, 2:1364–1372, 2018.

[12] Patrick L. Combettes and Valérie R. Wajs. Signal Recovery by Proximal Forward-Backward Splitting. *Multiscale Modeling & Simulation*, 4(4):1168–1200, 2005.

[13] LSST Science Collaboration, Paul A. Abell, Julius Allison, Scott F Anderson, John R Andrew, J Roger P Angel, Lee Armus, David Arnett, S J Asztalos, Tim S Axelrod, Stephen Bailey, D R Ballantyne, Justin R Bankert, Wayne A Barkhouse, Jeffrey D Barr, L Felipe Barrientos, Aaron J Barth, James G Bartlett, Andrew C Becker, Jacek Becla, Timothy C Beers, Joseph P Bernstein, Rahul Biswas, Michael R Blanton, Joshua S Bloom, John J Bochanski, Pat Boeshaar, Kirk D Borne, Marusa Bradac, W N Brandt, Carrie R Bridge, Michael E Brown, Robert J Brunner, James S Bullock, Adam J Burgasser, James H Burge, David L Burke, Phillip A Cargile, Srinivasan Chandrasekharan, George Chartas, Steven R Chesley, You-Hua Chu, David

Cinabro, Mark W Claire, Charles F Claver, Douglas Clowe, A J Connolly, Kem H Cook, Jeff Cooke, Asantha Cooray, Kevin R Covey, Christopher S Culliton, Roelof de Jong, Willem H de Vries, Victor P Debattista, Francisco Delgado, Ian P Dell'Antonio, Saurav Dhital, Rosanne Di Stefano, Mark Dickinson, Benjamin Dilday, S G Djorgovski, Gregory Dobler, Ciro Donalek, Gregory Dubois-Felsmann, Josef Durech, Ardis Eliasdottir, Michael Eracleous, Laurent Eyer, Emilio E Falco, Xiaohui Fan, Christopher D Fassnacht, Harry C Ferguson, Yanga R Fernandez, Brian D Fields, Douglas Finkbeiner, Eduardo E Figueroa, Derek B Fox, Harold Francke, James S Frank, Josh Frieman, Sebastien Fromenteau, Muhammad Furqan, Gaspar Galaz, A Gal-Yam, Peter Garnavich, Eric Gawiser, John Geary, Perry Gee, Robert R Gibson, Kirk Gilmore, Emily A Grace, Richard F Green, William J Gressler, Carl J Grillmair, Salman Habib, J S Haggerty, Mario Hamuy, Alan W Harris, Suzanne L Hawley, Alan F Heavens, Leslie Hebb, Todd J Henry, Edward Hileman, Eric J Hilton, Keri Hoadley, J B Holberg, Matt J Holman, Steve B Howell, Leopoldo Infante, Zeljko Ivezic, Suzanne H Jacoby, Bhuvnesh Jain, R, Jedicke, M James Jee, J Garrett Jernigan, Saurabh W Jha, Kathryn V Johnston, R Lynne Jones, Mario Juric, Mikko Kaasalainen, Styliani, Kafka, Steven M Kahn, Nathan A Kaib, Jason Kalirai, Jeff Kantor, Mansi M Kasliwal, Charles R Keeton, Richard Kessler, Zoran Knezevic, Adam Kowalski, Victor L Krabbendam, K Simon Krughoff, Shrinivas Kulkarni, Stephen Kuhlman, Mark Lacy, Sebastien Lepine, Ming Liang, Amy Lien, Paulina Lira, Knox S Long, Suzanne Lorenz, Jennifer M Lotz, R H Lupton, Julie Lutz, Lucas M Macri, Ashish A Mahabal, Rachel Mandelbaum, Phil Marshall, Morgan May, Peregrine M McGehee, Brian T Meadows, Alan Meert, Andrea Milani, Christopher J. Miller, Michelle Miller, David Mills, Dante Minniti, David Monet, Anjum S Mukadam, Ehud Nakar, Douglas R Neill, Jeffrey A Newman, Sergei Nikolaev, Martin Nordby, Paul O'Connor, Masamune Oguri, John Oliver, Scot S Olivier, Julia K. Olsen, Knut Olsen, Edward W Olszewski, Hakeem Oluseyi, Nelson D Padilla, Alex Parker, Joshua Pepper, John R Peterson, Catherine Petry, Philip A Pinto, James L Pizagno, Bogdan Popescu, Andrej Prsa, Veljko Radcka, M Jordan Raddick, Andrew Rasmussen, Arne Rau, Jeonghee Rho, James E Rhoads, Gordon T Richards, Stephen T Ridgway, Brant E Robertson, Rok Roskar, Abhijit Saha, Ata Sarajedini, Evan Scannapieco, Terry Schalk, Rafe Schindler, Samuel Schmidt, Sarah Schmidt, Donald P Schneider, German Schumacher, Ryan Scranton, Jacques Sebag, Lynn G Seppala, Ohad Shemmer, Joshua D Simon, M Sivertz, Howard A. Smith, J. Allyn Smith, Nathan Smith, Anna H Spitz, Adam Stanford, Keivan G Stassun, Jay Strader, Michael A Strauss, Christopher W Stubbs, Donald W Sweeney, Alex Szalay, Paula Szkody, Masahiro Takada, Paul Thorman, David E Trilling, Virginia Trimble, Anthony Tyson, Richard Van Berg, Daniel Vanden Berk, Jake VanderPlas, Licia Verde, Bojan Vrsnak, Lucianne M Walkowicz, Benjamin D Wandelt, Sheng Wang, Yun Wang, Michael Warner, Risa H Wechsler, Andrew A West, Oliver Wiecha, Benjamin F Williams, Beth Willman, David Wittman, Sidney C Wolff, W Michael Wood-Vasey, Przemek Wozniak, Patrick Young, Andrew Zentner, and Hu Zhan. LSST Science Book, Version 2.0. *arXiv.org*, astro-ph.I(November):201, 2009.

[14] R. Laureijs, J. Amiaux, S. Arduini, J. L. Augueres, J. Brinchmann, R. Cole, M. Cropper, C. Dabin, L. Duvet, A. Ealet, B. Garilli, P. Gondoin, L. Guzzo, J. Hoar, H. Hoekstra, R. Holmes, T. Kitching, T. Maciaszek, Y. Mellier, F. Pasian, W. Percival, J. Rhodes, G. Saavedra Criado, M. Sauvage, R. Scaramella, L. Valenziano, S. Warren, R. Bender, F. Castander, A. Cimatti, O. Le Fevre, H. Kurki-Suonio, M. Levi, P. Lilje, G. Meylan, R. Nichol, K. Pedersen, V. Popa, R. Rebolo Lopez, H. W. Rix, H. Rottgering, W. Zeilinger, F. Grupp, P. Hudelot, R. Massey, M. Meneghetti, L. Miller, S. Paltani, S. Paulin-Henriksson, S. Pires, C. Saxton, T. Schrabback, G. Seidel, J. Walsh, N. Aghanim, L. Amendola, J. Bartlett, C. Baccigalupi, J. P. Beaulieu, K. Benabed, J. G. Cuby, D. Elbaz, P. Fosalba, G. Gavazzi, A. Helmi, I. Hook, M. Irwin, J. P. Kneib, M. Kunz, F. Mannucci, L. Moscardini, C. Tao, R. Teyssier, J. Weller, G. Zamorani, M. R. Zapatero Osorio, O. Boulade, J. J. Foumond, A. Di Giorgio, P. Guttridge, A. James, M. Kemp, J. Martignac, A. Spencer, D. Walton, T. Blumchen, C. Bonoli, F. Bortoletto, C. Cerna, L. Corcione, C. Fabron, K. Jahnke, S. Ligori, F. Madrid, L. Martin, G. Morgante, T. Pamplona, E. Prieto, M. Riva, R. Toledo, M. Trifoglio, F. Zerbi, F. Abdalla, M. Douspis, C. Grenet, S. Borgani, R. Bouwens, F. Courbin, J. M. Delouis, P. Dubath, A. Fontana, M. Frailis, A. Grazian, J. Koppenhofer, O. Mansutti, M. Melchior, M. Mignoli, J. Mohr, C. Neissner, K. Noddle, M. Poncet, M. Scodeggio, S. Serrano, N. Shane, J. L. Starck, C. Surace, A. Taylor, G. Verdoes-Kleijn, C. Vuerli, O. R. Williams, A. Zacchei, B. Altieri, I. Escudero Sanz, R. Kohley, T. Oosterbroek, P. Astier, D. Bacon, S. Bardelli, C. Baugh, F. Bellagamba, C. Benoist, D. Bianchi, A. Biviano, E. Branchini, C. Carbone, V. Cardone, D. Clements, S. Colombi, C. Conselice, G. Cresci,

N. Deacon, J. Dunlop, C. Fedeli, F. Fontanot, P. Franzetti, C. Giocoli, J. Garcia-Bellido, J. Gow, A. Heavens, P. Hewett, C. Heymans, A. Holland, Z. Huang, O. Ilbert, B. Joachimi, E. Jennins, E. Kerins, A. Kiessling, D. Kirk, R. Kotak, O. Krause, O. Lahav, F. van Leeuwen, J. Lesgourgues, M. Lombardi, M. Magliocchetti, K. Maguire, E. Majerotto, R. Maoli, F. Marulli, S. Maurogordato, H. McCracken, R. McLure, A. Melchiorri, A. Merson, M. Moresco, M. Nonino, P. Norberg, J. Peacock, R. Pello, M. Penny, V. Pettorino, C. Di Porto, L. Pozzetti, C. Quercellini, M. Radovich, A. Rassat, N. Roche, S. Ronayette, E. Rossetti, B. Sartoris, P. Schneider, E. Semboloni, S. Serjeant, F. Simpson, C. Skordis, G. Smadja, S. Smartt, P. Spano, S. Spiro, M. Sullivan, A. Tilquin, R. Trotta, L. Verde, Y. Wang, G. Williger, G. Zhao, J. Zoubian, and E. Zucca. Euclid Definition Study Report. *ArXiv e-prints*, page arXiv:1110.3193, 10 2011.

[15] David M. Reiman and Brett E. Göhre. Deblending galaxy superpositions with branched generative adversarial networks. *Monthly Notices of the Royal Astronomical Society*, 485(2):2617–2627, 2019.

[16] P. Melchior, F. Moolekamp, M. Jerdee, R. Armstrong, A.-L. Sun, J. Bosch, and R. Lupton. scarlet : Source separation in multi-band images by Constrained Matrix Factorization. *Astronomy and Computing*, 24:129–142, 7 2018.

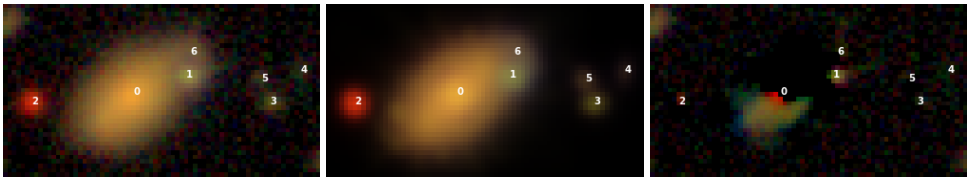
[17] Rachel Mandelbaum, Claire Lackner, Alexie Leauthaud, and Barnaby Rowe. COSMOS real galaxy dataset, 1 2012.

[18] Rachel Mandelbaum, Barnaby Rowe, James Bosch, Chihway Chang, Frederic Courbin, Mandeep Gill, Mike Jarvis, Arun Kannawadi, Tomasz Kacprzak, Claire Lackner, Alexie Leauthaud, Hironao Miyatake, Reiko Nakajima, Jason Rhodes, Melanie Simet, Joe Zuntz, Bob Armstrong, Sarah Bridle, Jean Coupon, Jrthrg P. Dietrich, Marc Gentile, Catherine Heymans, Alden S. Jurling, Stephen M. Kent, David Kirkby, Daniel Margala, Richard Massey, Peter Melchior, John Peterson, Aaron Roodman, and Tim Schrabback. The Third Gravitational Lensing Accuracy Testing (GREAT3) Challenge Handbook. *The Astrophysical Journal Supplement Series*, 212(1):5, 5 2014.

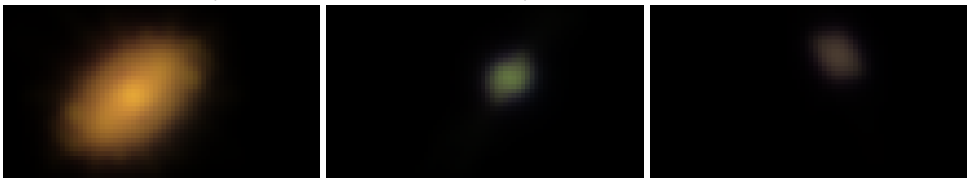
[19] B. T P Rowe, M. Jarvis, R. Mandelbaum, G. M. Bernstein, J. Bosch, M. Simet, J. E. Meyers, T. Kacprzak, R. Nakajima, J. Zuntz, H. Miyatake, J. P. Dietrich, R. Armstrong, P. Melchior, and M. S S Gill. GalSim: The modular galaxy image simulation toolkit. *Astronomy and Computing*, 10:121–150, 2015.

[20] A.M. Koekemoer, H Aussel, D Calzetti, P Capak, M Giavalisco, J.-P. Kneib, A Leauthaud, O. Le Fèvre, H.J. Mccracken, R Massey, B Mobasher, J Rhodes, N Scoville, and P.L. Shopbell. The COSMOS survey: Hubble Space Telescope Advanced Camera for Surveys observations and data processing. *Astrophysical Journal, Supplement Series*, 172(1):196–202, 2007.

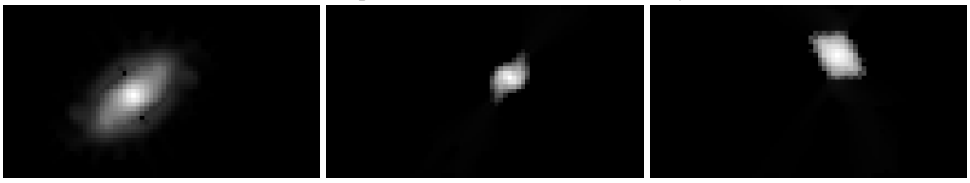
[21] Hiroaki Aihara, Nobuo Arimoto, Robert Armstrong, Stéphane Arnouts, Neta A. Bahcall, Steven Bickerton, James Bosch, Kevin Bundy, Peter L. Capak, James H. H. Chan, Masashi Chiba, Jean Coupon, Eiichi Egami, Motohiro Enoki, Francois Finet, Hiroki Fujimori, Seiji Fujimoto, Hisanori Furusawa, Junko Furusawa, Tomotsugu Goto, Andy Goulding, Johnny P. Greco, Jenny E. Greene, James E. Gunn, Takashi Hamana, Yuichi Harikane, Yasuhiro Hashimoto, Takashi Hattori, Masao Hayashi, Yusuke Hayashi, Krzysztof G. Hełminiak, Ryo Higuchi, Chiaki Hikage, Paul T. P. Ho, Bau-Ching Hsieh, Kuiyun Huang, Song Huang, Hiroyuki Ikeda, Masatoshi Imanishi, Akio K. Inoue, Kazushi Iwasawa, Ikuru Iwata, Anton T. Jaelani, Hung-Yu Jian, Yukiko Kamata, Hiroshi Karoji, Nobunari Kashikawa, Nobuhiko Katayama, Satoshi Kawanomoto, Issha Kayo, Jin Koda, Michitaro Koike, Takashi Kojima, Yutaka Komiyama, Akira Konno, Shintaro Koshida, Yusei Koyama, Haruka Kusakabe, Alexie Leauthaud, Chien-Hsiu Lee, Lihwai Lin, Yen-Ting Lin, Robert H. Lupton, Rachel Mandelbaum, Yoshiki Matsuoka, Elinor Medezinski, Sogo Mineo, Shoken Miyama, Hironao Miyatake, Satoshi Miyazaki, Rieko Momose, Anupreeta More, Surhud More, Yuki Moritani, Takashi J. Moriya, Tomoki Morokuma, Shiro Mukae, Ryoma Murata, Hitoshi Murayama, Tohru Nagao, Fumiaki Nakata, Mana Niida, Hiroko Niikura, Atsushi J. Nishizawa, Yoshiyuki Obuchi, Masamune Oguri, Yukie Oishi, Nobuhiro Okabe, Sakurako Okamoto, Yuki Okura, Yoshiaki Ono, Masato Onodera, Masafusa Onoue, Ken Osato, Masami Ouchi, Paul A. Price, Tae-Soo Pyo, Masao Sako, Marcin Sawicki, Takatoshi Shibuya, Kazuhiro Shimasaku, Atsushi Shimono, Masato Shirasaki, John D. Silverman, Melanie Simet, Joshua Speagle, David N. Spergel, Michael A. Strauss, Yuma Sugahara, Naoshi Sugiyama, Yasushi Suto, Sherry H. Suyu, Nao Suzuki, Philip J. Tait, Masahiro



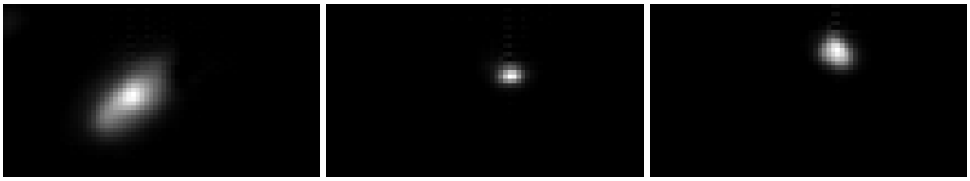
(a) Simulated noisy scene (b) SCARLET sky model (c) Residuals



(d) Individual components of the model recovered by SCARLET



(e) Individual morphology components recovered with SCARLET



(f) Individual morphology components recovered by our proposed method

Figure 4: Similar to Figure 1, but using the deblender SCARLET with default settings. We also compare in (e) and (f) the deconvolved morphology components (S_k in Equation 6) recovered using the strict monotonicity and symmetry constraints in SCARLET with the deep morphology priors from this work.

Takada, Tadafumi Takata, Naoyuki Tamura, Manobu M Tanaka, Masaomi Tanaka, Masayuki Tanaka, Yoko Tanaka, Tsuyoshi Terai, Yuichi Terashima, Yoshiki Toba, Nozomu Tominaga, Jun Toshikawa, Edwin L. Turner, Tomohisa Uchida, Hisakazu Uchiyama, Keiichi Umetsu, Fumihiro Uraguchi, Yuji Urata, Tomonori Usuda, Yousuke Utsumi, Shiang-Yu Wang, Wei-Hao Wang, Kenneth C. Wong, Kiyoto Yabe, Yoshihiko Yamada, Hitomi Yamanoi, Naoki Yasuda, Sherry Yeh, Atsunori Yonehara, and Suraphong Yuma. The Hyper Suprime-Cam SSP Survey: Overview and survey design. *Publications of the Astronomical Society of Japan*, 70(SP1):1–5, 1 2018.

A Comparison to the SCARLET deblender

In this appendix, we compare the proposed method with the baseline deblending algorithm SCARLET, which constitutes a state-of-the-art deblender for ground-based images [16]. SCARLET uses the same parameterization and loss function; in fact the work described here uses the same code base and only differs in the assumptions about galaxy morphologies.

In its default configuration, SCARLET assumes every S_k to be non-negative, symmetric under rotation of 180° and monotonically decreasing away from the center. These hard constraints can directly be enforced through proximal mappings in Equation 4 and have been found successful as regularizers of the deblending problem for ground-based images. They do not perform well on complex and irregular galaxies, which is the original motivation for the present work: replacing analytic, heuristic constraints by a data-driven deep morphology prior.

Figure 4 show the result of running baseline SCARLET on the same data set. The residuals are significantly larger than in Figure 2 due to the inadequacies of the strict symmetry and monotonicity assumptions. This can also be seen by directly comparing the recovered deconvolved morphologies (lower panels of Figure 4). The symmetry constraint can lead to artifacts in the direction of a nearby source, in this example the model of source 0 is influenced by source 1. The morphologies recovered under the deep morphology prior are by construction realistic and do not exhibit such obvious artifacts.

While this comparison remains qualitative, it illustrates that the deep morphology prior addresses one main limitation of the SCARLET algorithm. A thorough study of our extension for the science cases of the upcoming LSST survey [13] will be the main focus of an upcoming science paper geared towards the astronomical community.