Sparse Bayesian mass-mapping with uncertainties

Matthew A. Price*, Jason D. McEwen*, Xiaohao Cai*, Thomas D. Kitching*, Christopher G. R. Wallis*, and Marcelo Pereyra†.
* Mullard Space Science Laboratory, University College London, RH5 6NT, UK.
† Maxwell Institute for Mathematical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom

***Abstract*—Mass-mapping *via* weak gravitational lensing has until recently lacked principled statistical consideration of uncertainties introduced during the reconstruction process – solving of an often seriously ill-posed inverse problem. In recent work we posed the mass-mapping inverse problem as an unconstrained Bayesian inference problem with Laplace-type ℓ_1-norm sparsity-promoting prior, which we solve via convex optimization. Formulating the problem in this way allows us to exploit recent developments in probability concentration theory to infer tightly bound, theoretically conservative uncertainties $\mathcal{O}(10^6)$ times faster than traditional MCMC techniques. Building on these new fast Bayesian inference techniques we have developed several uncertainty quantification techniques primarily aimed towards the gravitational lensing paradigm, though entirely generalizable to other settings. The uncertainty quantification techniques reviewed here are: *knock-out* hypothesis testing of structure, local credible regions (*cf.* pixel-level Bayesian error bars), and Bayesian locational uncertainty of structure. Additionally, these conservative Bayesian inferences can be leveraged to aggregate uncertainties which are often computed by the weak lensing community (*e.g.* peak statistics).**

I. Forward model

A mapping relation can be drawn between the two first order lensing fields κ and γ giving the *planar forward model* in Fourier space,

$$\hat{\gamma}(k_x, k_y) = \mathbf{D}_{k_x,k_y}\hat{\kappa}(k_x, k_y), \tag{1}$$

for Fourier space mapping $\mathbf{D}_{k_x,k_y} = (k_x^2 - k_y^2 + 2ik_xk_y)/(k_x^2 + k_y^2)$.

Due to the so called *mass-sheet degeneracy* the convergence κ is not observable and so typically measurements of the shearing field γ are taken and inverted to form estimators for κ.

II. Sparse hierachical Bayesian inference

We formulate this lensing inversion as a hierachical Bayesian inference problem. Bayes' theorem for the posterior distribution is given by $p(\kappa|\gamma) \propto p(\gamma|\kappa)p(\kappa)$, where $p(\gamma|\kappa)$ is the likelihood function representing data fidelity and $p(\kappa)$ is a prior on the statistical nature of κ. Suppose the pixel-level noise n on the shear field γ is i.i.d. Gaussian noise, such that measurements of γ are obtained by $\gamma = \mathbf{\Phi}\kappa + n$, where the measurement operator $\mathbf{\Phi} = \mathbf{F}^{-1}\mathbf{D}\mathbf{F}$ for forward (inverse) Fourier transforms $\mathbf{F}(\mathbf{F}^{-1})$. Then our likelihood term $p(\gamma|\kappa)$ is given by $p(\gamma|\kappa) \propto \exp\big((-\|\mathbf{\Phi}\kappa - \gamma\|_2^2)/(2\sigma_n^2)\big)$ which we choose to regularize with a sparsity promoting Laplace type ℓ_1-norm wavelet prior $p(\kappa) \propto \exp\big(-\mu\|\mathbf{\Psi}^\dagger\kappa\|_1\big)$, where $\mathbf{\Psi}$ is a wavelet dictionary and the regularization parameter μ is drawn from a gamma-type hyper-prior distribution [1]. Due to the monotonicity of the logarithm function, maximizing the posterior is equivalent to minimizing the log-posterior, thus the problem may be recast as a convex optimization problem,

$$\kappa^{\text{map}} = \underset{\kappa}{\operatorname{argmin}} \left\{ \mu\|\mathbf{\Psi}^\dagger\kappa\|_1 + \frac{\|\mathbf{\Phi}\kappa - \gamma\|_2^2}{2\sigma_n^2} \right\}, \tag{2}$$

where κ^{map} is the *maximum a posteriori* (MAP) convergence field.

Exploiting recent developments in probability concentration theory, a conservative approximation of the *highest posterior density* (HPD) credible region has been proposed [2] such that the approximate credible set is given by $C'_\alpha := \{\kappa : f(\kappa) + g(\kappa) \leq \epsilon'_\alpha\}$ where the approximate isocontour ϵ'_α is given by $\epsilon'_\alpha = f(\kappa^{\text{map}}) + g(\kappa^{\text{map}}) + \tau_\alpha\sqrt{N} + N$, with constant $\tau_\alpha = \sqrt{16\log(3/\alpha)}$ and dimensionality N.

Crucially, this approximate credible region of the posterior can be readily computed from the MAP solution alone, and so avoids the high dimensional computationally taxing integrals present in the true HPD credible region.

III. Hypothesis testing of structure

Mechanically, one first removes a feature of interest to create a surrogate convergence field κ^{sgt}. If $\kappa^{\text{sgt}} \notin C'_\alpha$ then the data (noisy measurements of γ) is sufficient to reject the null hypothesis which implies that the feautre is physical at $100(1-\alpha)\%$ confidence. If $\kappa^{\text{sgt}} \in C'_\alpha$ then the null hypothesis cannot be rejected by the data and so the physicality is inconclusive [1], [3].

IV. Local credible intervals

Suppose the i^{th} pixel is selected within the MAP convergence field κ^{MAP} and has intensity j. Let us iteratively increase (decrease) the intensity $j \to j + \delta j$ at each step creating a new surrogate solution κ^{sgt} until $\kappa^{\text{sgt}} \notin C'_\alpha$. At this point the local pixel (or in practise group of pixels – *'super-pixel'*) has reached the highest (lowest) intensity value which is supported by the data at some confidence level $100(1-\alpha)\%$. One can therefore imagine constructing a complete map of Bayesian pixel-level error bars for a given reconstruction [3], [4].

V. Bayesian locational uncertainty

Suppose one iteratively perturbs the location of a feature of interest within the convergence domain, creating surrogate images κ^{sgt}, for each determining whether $\kappa^{\text{sgt}} \in C'_\alpha$. If $\kappa^{\text{sgt}} \in C'_\alpha$ then the data cannot reject the hypothesis that the feature was located at a given location and so the location is accepted. If $\kappa^{\text{sgt}} \notin C'_\alpha$ then the data rejects the hypothesis that the feature could have been observed at this location and so the location is rejected. One can then create a convex set of accepted positions which we take to be the Bayesian locational uncertainty [5].

References

[1] M. A. Price, X. Cai, J. D. McEwen, T. D. Kitching, and C. G. R. Wallis, "Sparse Bayesian mass-mapping with uncertainties: hypothesis testing of structure," *submitted to MNRAS*, 2018.

[2] M. Pereyra, "Maximum-a-posteriori estimation with bayesian confidence regions," *SIAM Journal on Imaging Sciences*, vol. 10, no. 1, pp. 285–302, 2017. [Online]. Available: https://doi.org/10.1137/16M1071249

[3] X. Cai, M. Pereyra, and J. D. McEwen, "Uncertainty quantification for radio interferometric imaging: II. MAP estimation," *ArXiv e-prints*, Nov. MNRAS in press, 2017.

[4] M. A. Price, X. Cai, J. D. McEwen, M. Pereyra, and T. D. Kitching, "Sparse Bayesian mass-mapping with uncertainties: local credible intervals," *submitted to MNRAS*, 2018.

[5] M. A. Price, X. Cai, J. D. McEwen, and T. D. Kitching, "Sparse Bayesian mass-mapping with uncertainties: peak statistics and feature locations," *submitted to MNRAS*, 2018.