# Solving high-dimensional parameter inference: marginal posterior densities & Moment Networks

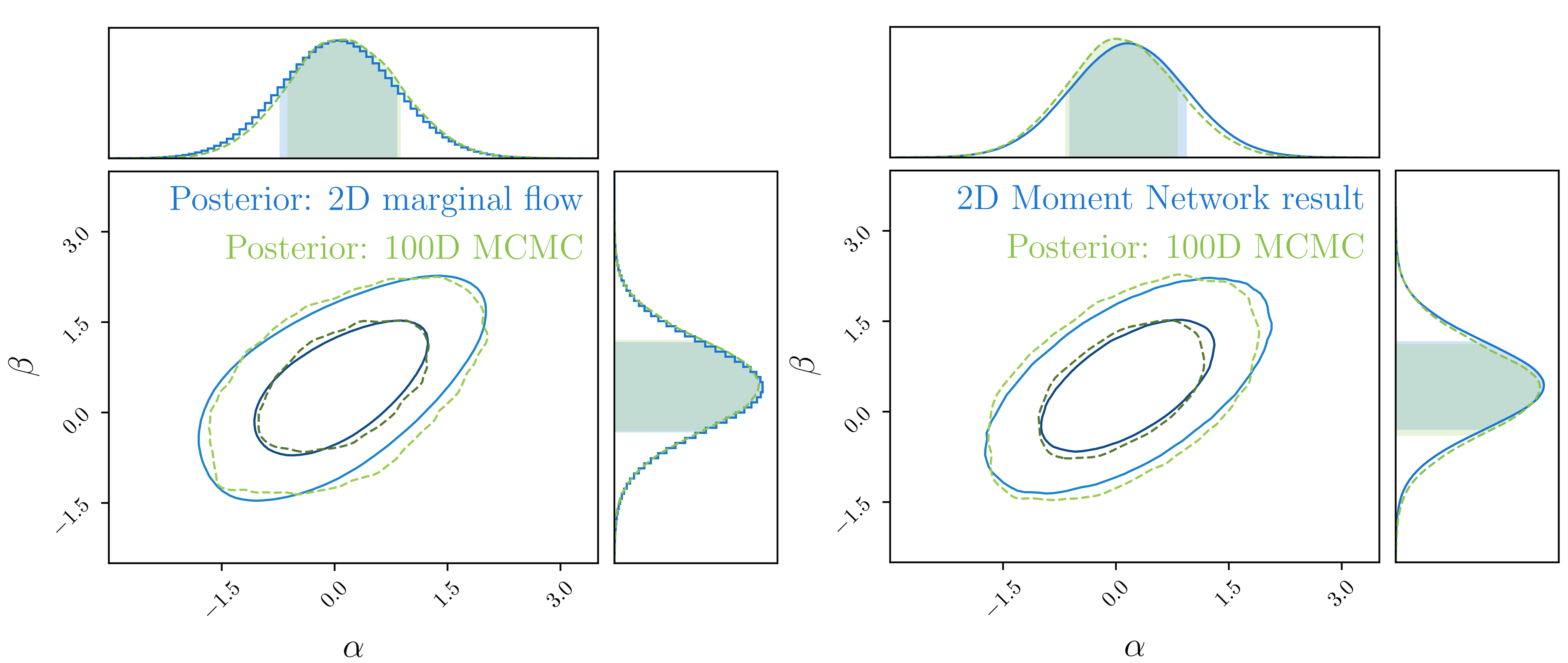
**Niall Jeffrey**[1,2], Benjamin D. Wandelt[3,4]

*1 **École Normale Supérieure, Paris***
*2 University College London*
*3 Institut d'Astrophysique de Paris (IAP), Sorbonne Université*
*4 Center for Computational Astrophysics, Flatiron Institute, New York*

ArXiv:2011.05991 [github.com/NiallJeffrey/MomentNetworks](https://github.com/NiallJeffrey/MomentNetworks) niall.jeffrey [at] phys.ens.fr

## Highlights

- High-dimensional probability density estimation suffers from "*curse of dimensionality*"
- Propose direct estimation of lower-dimensional marginal posterior distributions
- ***Marginal flow***: estimate joint-marginals of subsets of parameters
- ***Moment Networks***: hierarchy of fast regression models compute increasing moments of lower-dimensional marginal posterior density
- **Beyond likelihood-free inference**: can also efficiently solve known MCMC problems
- **Demo**: high-dimensional inference: a) MCMC reference, b) Gravitational wave data model




(A) 100-dimensional data model with known reference distribution evaluated with $10^7$ MCMC samples. Direct 2D marginal posterior estimation using a masked autoregressive flow ensemble (left panel) and representation of 2D Moment Network result (right panel) both trained with $8{\times}10^4$ simulations.

## Marginal flows

- Full $p(\boldsymbol{x}|\boldsymbol{\theta})$ density estimation for data $\boldsymbol{x}$ and parameters $\boldsymbol{\theta}$ intractable in high-dimension
- Make marginal densities the target of inference
- Estimate marginal posterior probability density for pairs of parameters $\alpha, \beta \in \boldsymbol{\theta}$ by minimizing $-\sum_i \log q(\alpha_i, \beta_i|\boldsymbol{x}_i)$
- Result $q$ is an estimate (e.g. with normalizing flow) of marginal posterior $p(\alpha, \beta|\boldsymbol{x})$ if training parameters $\boldsymbol{\theta}_i$ drawn from prior $p(\boldsymbol{\theta})$ and used to simulate $\boldsymbol{x}_i$

***Markov chain Monte Carlo (MCMC) validation: see Figure A***

## Moment Networks

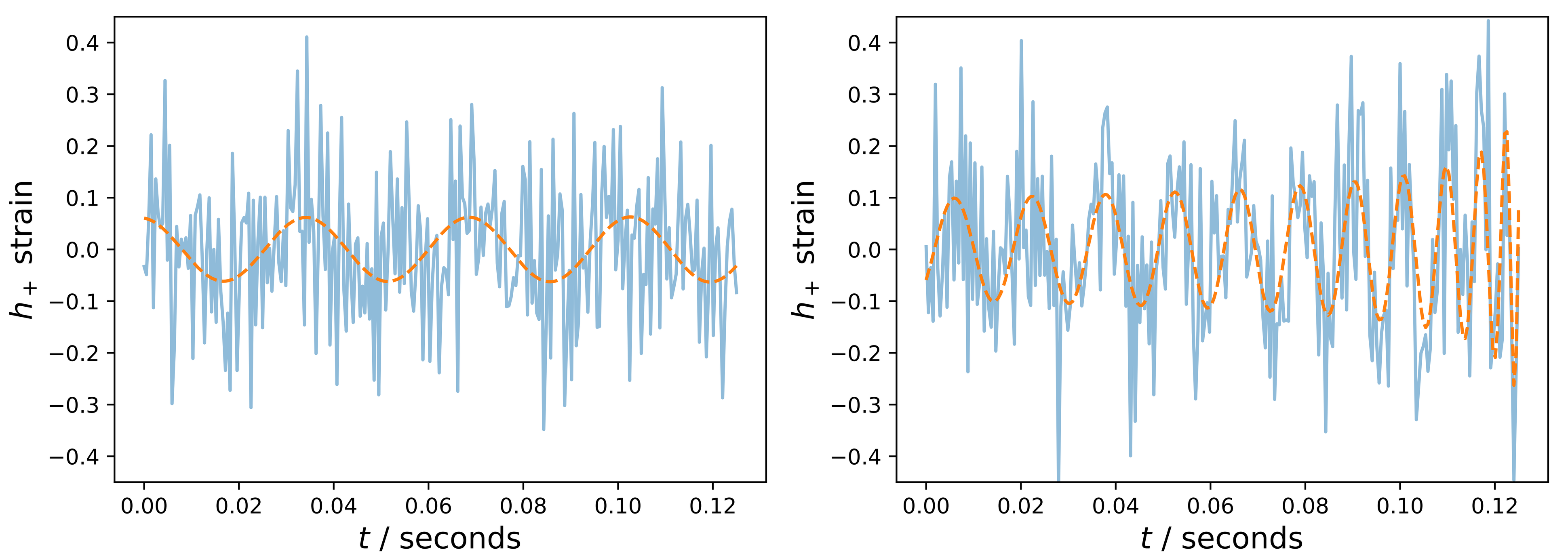
- Side-step the posterior density estimation problem
- Estimate location, scale, and covariance (and possibly higher-order moments) of marginal posteriors
- Network $F(\boldsymbol{x})$ minimizes $L_2$ loss over distribution of possible training examples $\{\boldsymbol{x}_i, \boldsymbol{\theta}_i\}$

$$J_0 = \int \|\boldsymbol{\theta} - F(\boldsymbol{x})\|^2 \, p(\boldsymbol{x}, \boldsymbol{\theta}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{\theta}$$

- For observed data it forms a posterior mean estimate $F(\boldsymbol{x}_{obs}) = \langle \boldsymbol{\theta} \rangle_{\theta|\boldsymbol{x}^{obs}}$
- Functions $F$ and $G$ can combine to output posterior means, variances, and covariances for subsets of the full set of parameters

$$J_1 = \int \left\| (\boldsymbol{\theta} - F_{\mathrm{fixed}}(\boldsymbol{x}))^2 - G(\boldsymbol{x}) \right\|^2 p(\boldsymbol{x}, \boldsymbol{\theta}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{\theta}$$

(B) Example simulated gravitational wave time series signals for the strain "+" polarization h+ with realistic LIGO-like noise.

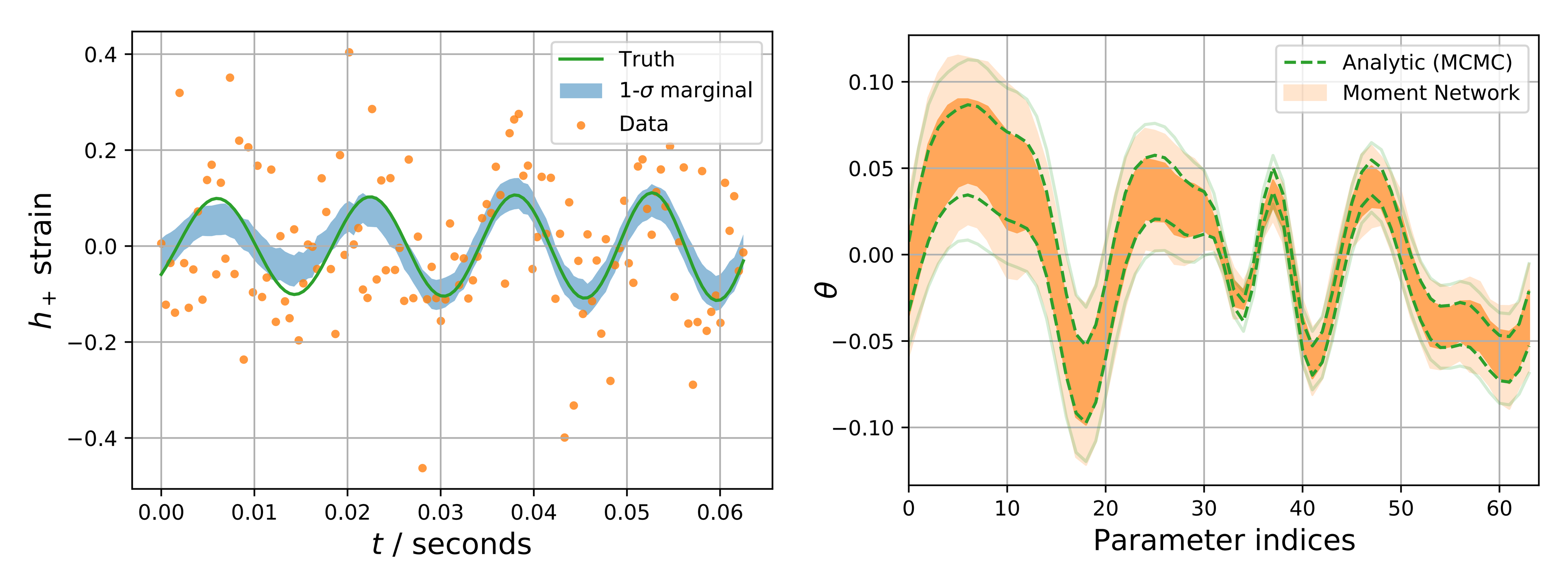



## LIGO-like gravitational wave time series

- Fig. B: two example simulated gravitational wave time series. The simplified signals (dashed orange) are ~0.12s intervals from the 1 second before a binary black hole merger
- Fig. C *left panel:* Moment Network estimate of the marginal mean and variance for each time step parameter
- Fig. C *right panel:* validation case of similar complexity, but with known likelihood
- Trained Moment Network accurately matches long-run MCMC chain, which validates our approach

(C) *Left panel:* Moment Network (MN) estimate of the 1-σ per strain parameter.

*Right panel:* For each of 64 parameters (timestep), contours are marginal posterior σ from MN (shaded orange) and MCMC (dashed green).




## Cosmological applications

- **Mapping the Universe**: tractable simulation-based inference of high-dimensional cosmological fields & dark matter maps
- **Robustness**: Moment Networks use simpler architectures, reducing training failure risk and boosting inference speed
- **Cross-validation**: moments of estimated marginal posteriors should match those from Moment Networks
- **Evading MCMC:** Even when the likelihood can be sampled, marginal flows and Moment Networks still provide advantages, as many of the drawbacks of high-dimensional MCMC can be simply avoided