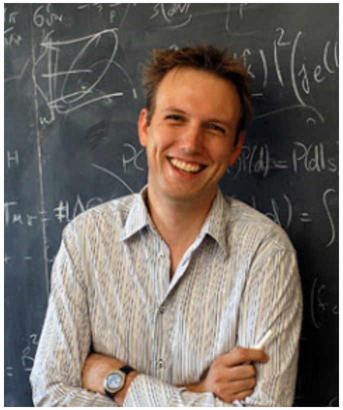
# Bayesian CMB delensing for optimal constraints on $r$

Based on *arxiv:1708.06753* (posted last week)

Marius Millea

with

Ethan Anderes



Ben Wandelt

COSMO 17 – Aug 28, 2017

How do we optimally delense future CMB data to obtain the best possible estimates of $r$?

CMB “fields”

$$f \equiv (T, Q, U)$$

Lensing potential

Cosmo params

Data

$$\mathcal{P}(f, \phi, r \,|\, d)$$


$$\mathcal{P}(\phi \,|\, r, d)$$

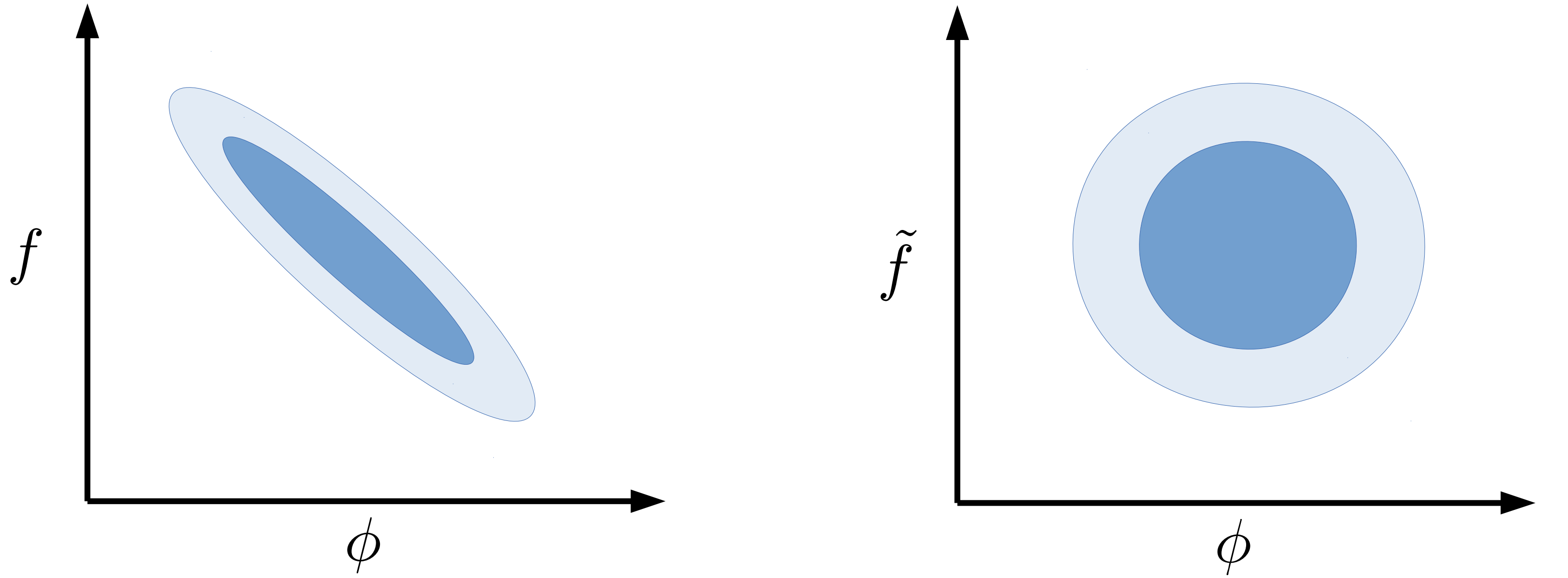
$$= \int df \, \mathcal{P}(f, \phi \,|\, r, d)$$

$$\hat{\phi}(\mathbf{L}) = \int d\mathbf{l_1} W(\mathbf{l_1}, \mathbf{l_2}) d(\mathbf{l_1})^* d(\mathbf{l_2})$$

Carron & Lewis (2017), Hirata & Seljak (2003) give algorithm to *maximize* this

All current analyses are based on this
Currently near-optimal but will be sub-optimal for next-gen noise levels

# Why is sampling/minimizing $\mathcal{P}(f,\phi\,|\,d)$ such a hard problem?




So, as pointed out by Anderes et al. 2015, its *very* beneficial to reparametrize,

$$\mathcal{P}(\tilde{f},\phi\,|\,d) = \mathcal{P}(f(\tilde{f}),\phi\,|\,d)\left|\frac{df}{d\tilde{f}}\right|$$

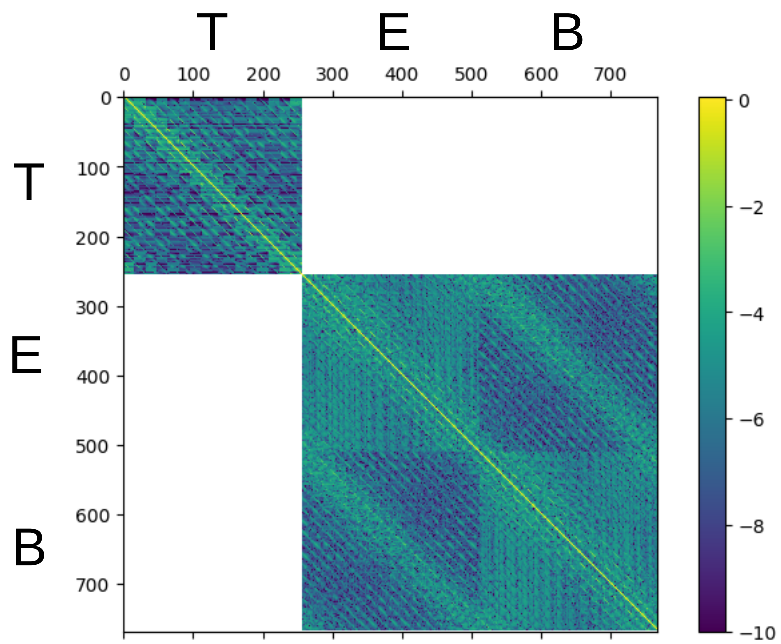
But now we introduce this determinant...

where $\tilde{f} = \mathcal{L}(\phi)f \implies \left|\frac{df}{d\tilde{f}}\right| = 1/\left|\mathcal{L}(\phi)\right|$

# What is the determinant of lensing?

- Infinite resolution: lensing is a remapping (i.e. permutation) so $\det|\mathcal{L}(\phi)| = 1$
- This is not the case when we have *pixelization*. Consider the Taylor series approx:

$$\tilde{f}(x) = f(x + \nabla\phi(x)) \approx \underbrace{\left[1 + \nabla\phi(x)\cdot\nabla + ...\right]}_{\mathcal{L}(\phi)} f(x)$$




$\log(\text{abs}(\mathcal{L}(\phi)_{ij}))$

Matrix representation of $\mathcal{L}(\phi)$ for 16x16 1' pixel TEB maps for 7$^{\text{th}}$ order Taylor series approximation

not close to 1!

$$\det|\mathcal{L}(\phi)| = 1.9 \times 10^{-9}$$

Additionally, the variation of the determinant with $\phi$ is significant.

A solution: **LenseFlow**

Define $f_t(x) \equiv f(x + t\nabla\phi(x))$ s.t. $f_{t=0}(x) = f(x)$

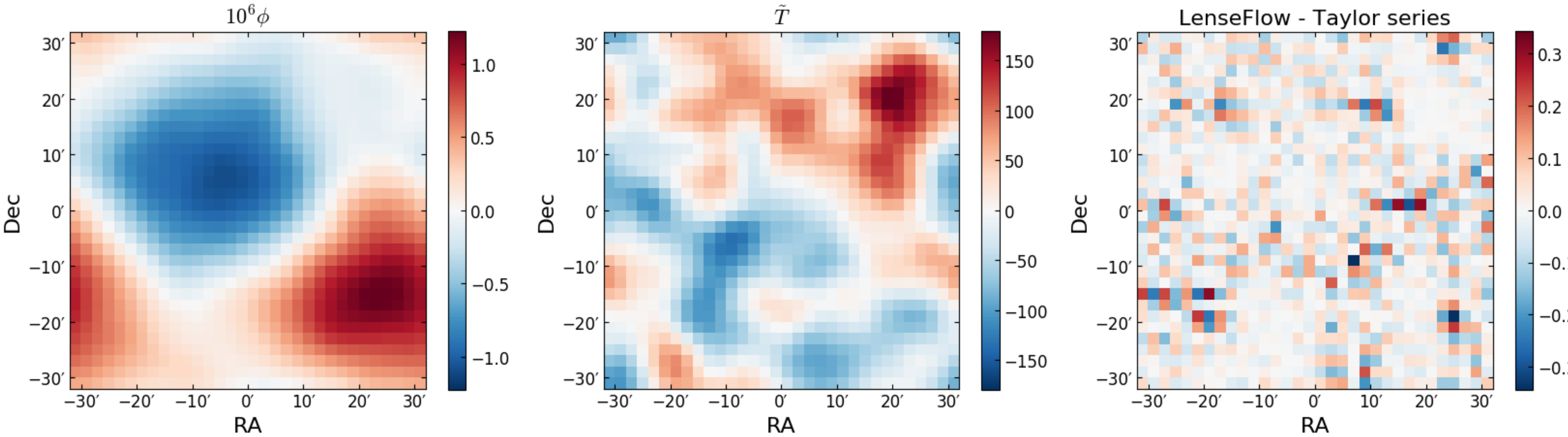
$f_{t=1}(x) = \tilde{f}(x)$

One can show $f_t$ obeys an ODE "flow" equation

$$\frac{df_t(x)}{dt} = \nabla\phi(x) \cdot \left[\mathbb{1} + t\nabla\nabla\phi(x)\right]^{-1} \cdot \nabla f_t(x)$$

- To *lense* a map, just run the ODE from t=0 to t=1
- To *delense* a map, just run it backwards from t=1 to t=0
- This operation provably has determinant = 1

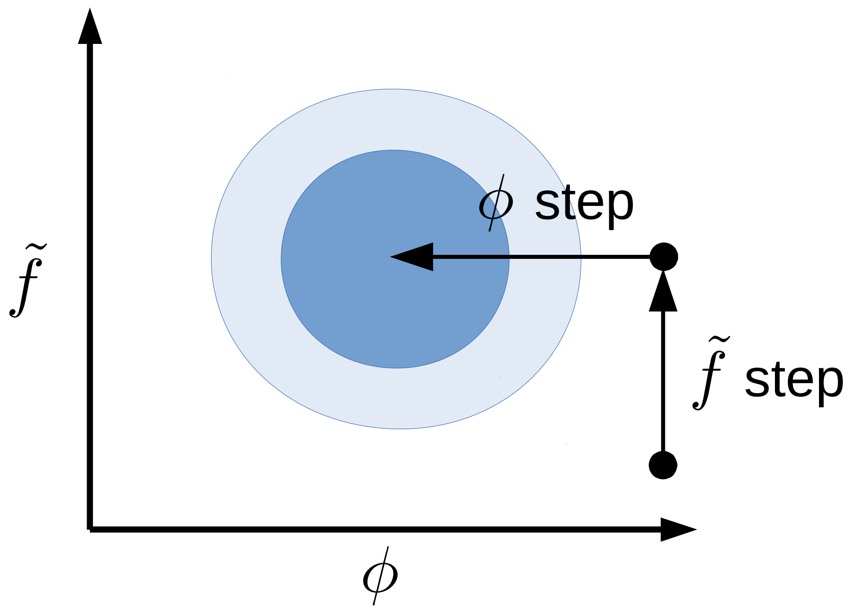
# LenseFlow vs. Taylor series




Differences between the two which lead to different determinants

# Ok, let's maximize & sample!

The algorithm we devise is a *coordinate descent*




$$-2 \ln \mathcal{P}(\tilde{f}, \phi \,|\, d) =$$

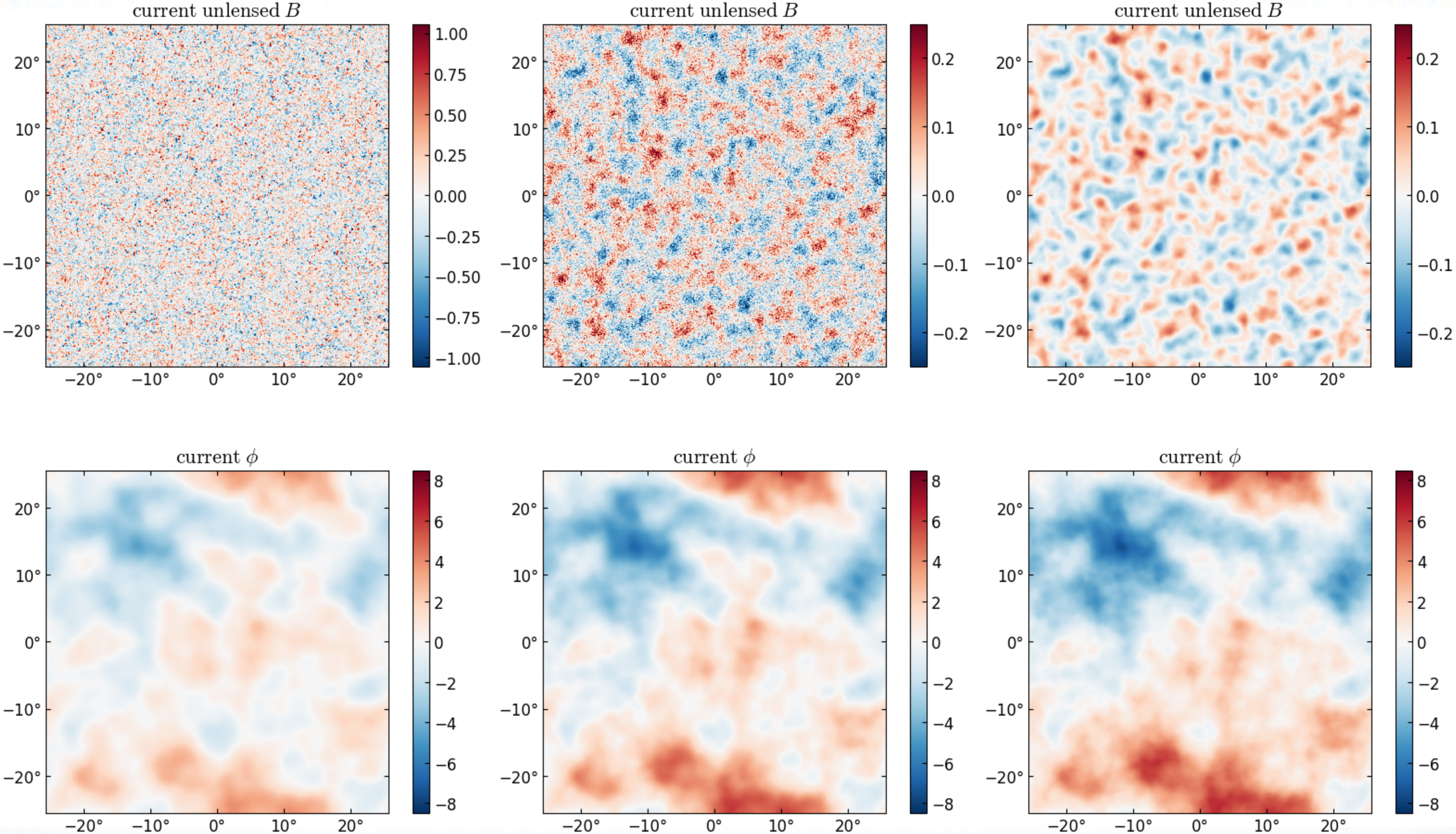
$\tilde{f}$ step : a Wiener filter

$$= \underbrace{(d - \tilde{f})^\dagger \mathcal{C}_n^{-1} (d - \tilde{f})}_{\text{likelihood}} + \underbrace{\tilde{f}^\dagger \mathcal{L}(\phi)^{-\dagger} \mathcal{C}_f^{-1} \mathcal{L}(\phi)^{-1} \tilde{f}}_{\text{prior on } f} + \underbrace{\phi^\dagger {\mathcal{C}_\phi}^{-1} \phi}_{\text{prior on } \phi}$$

$\phi$ step

Starting point: $\phi = 0$

Simulated data with: 1uK-arcmin noise, r=0.05



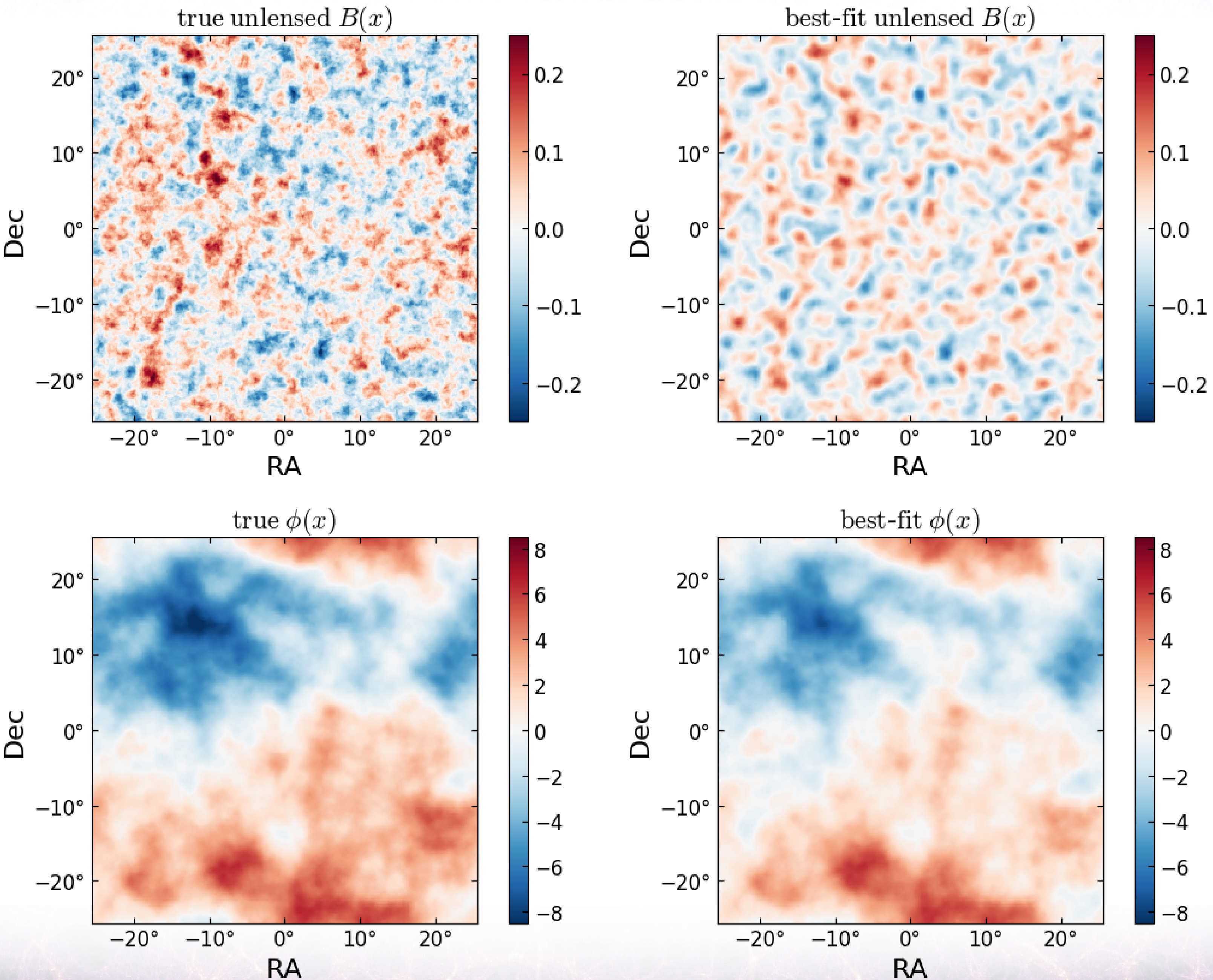

Step 1

Step 3

Step 30

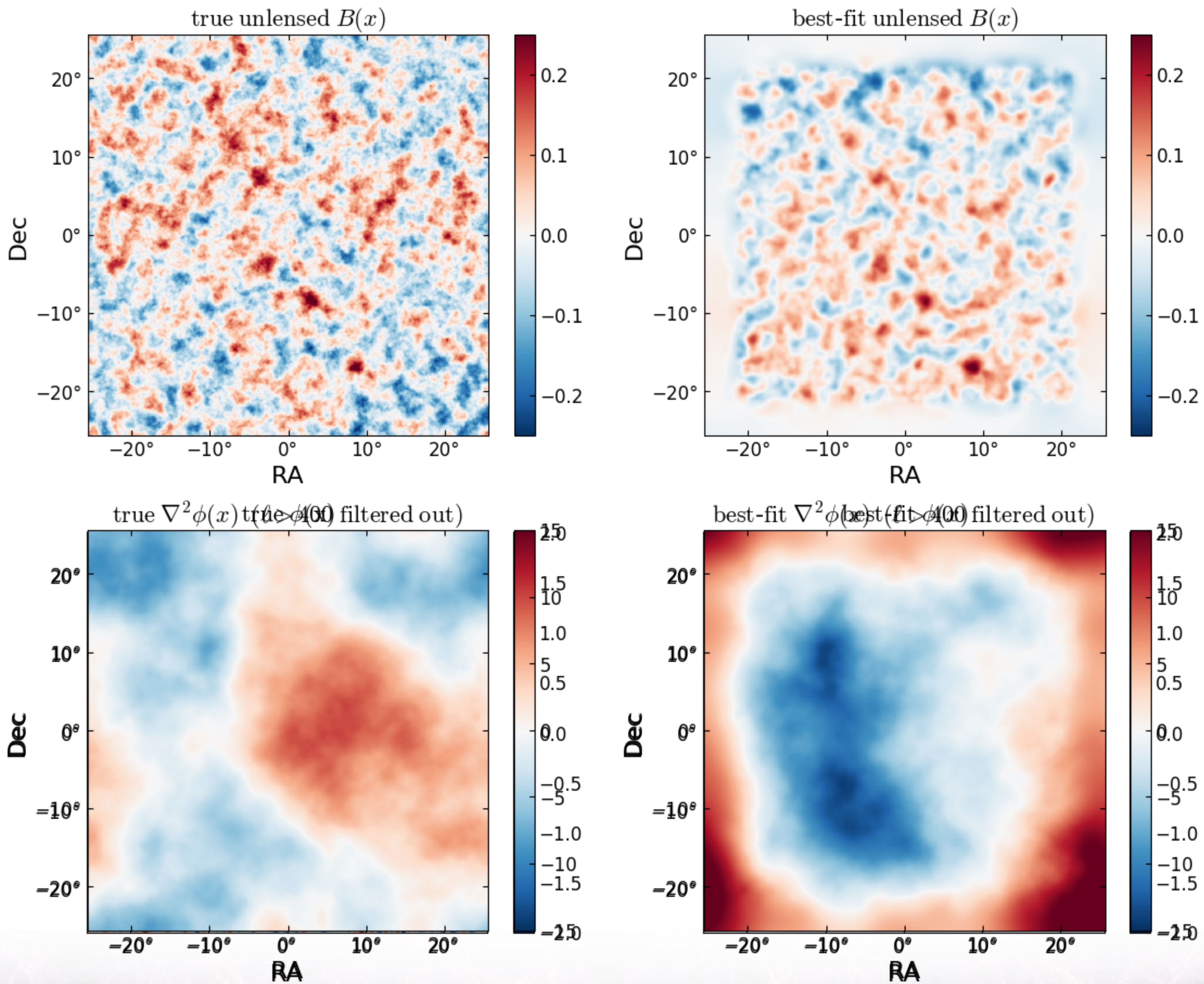
# 30min on 1 single multi-core CPU for these 2500deg$^2$

1024x1024, 3 arcmin pixels



# Masking works too

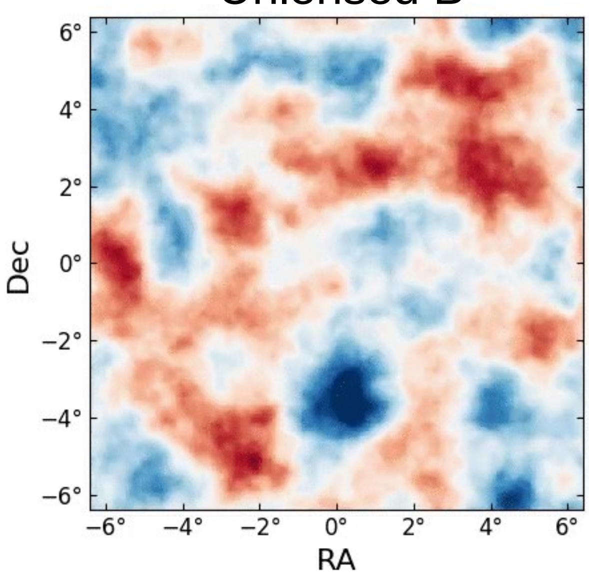
(Only affects the Wiener filter step which needs more conjugate gradient steps => 4 hours)

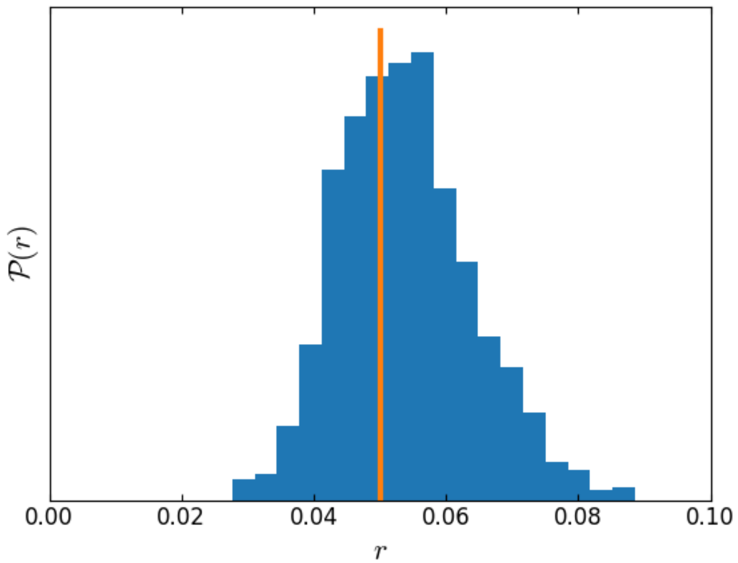


# What about sampling ?

This is fixed at its best-fit

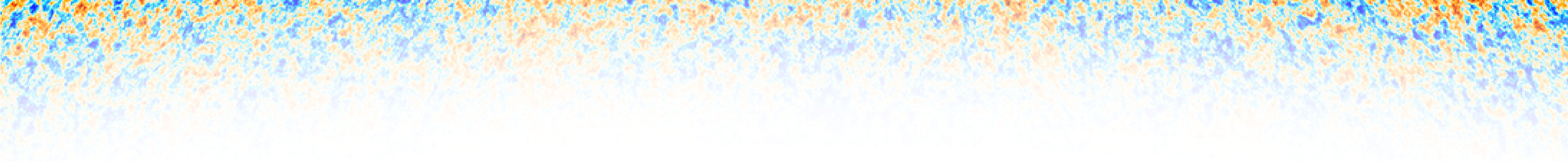
For now, a slightly simplified preview: $\mathcal{P}(f, \hat{\phi}, r \,|\, d)$

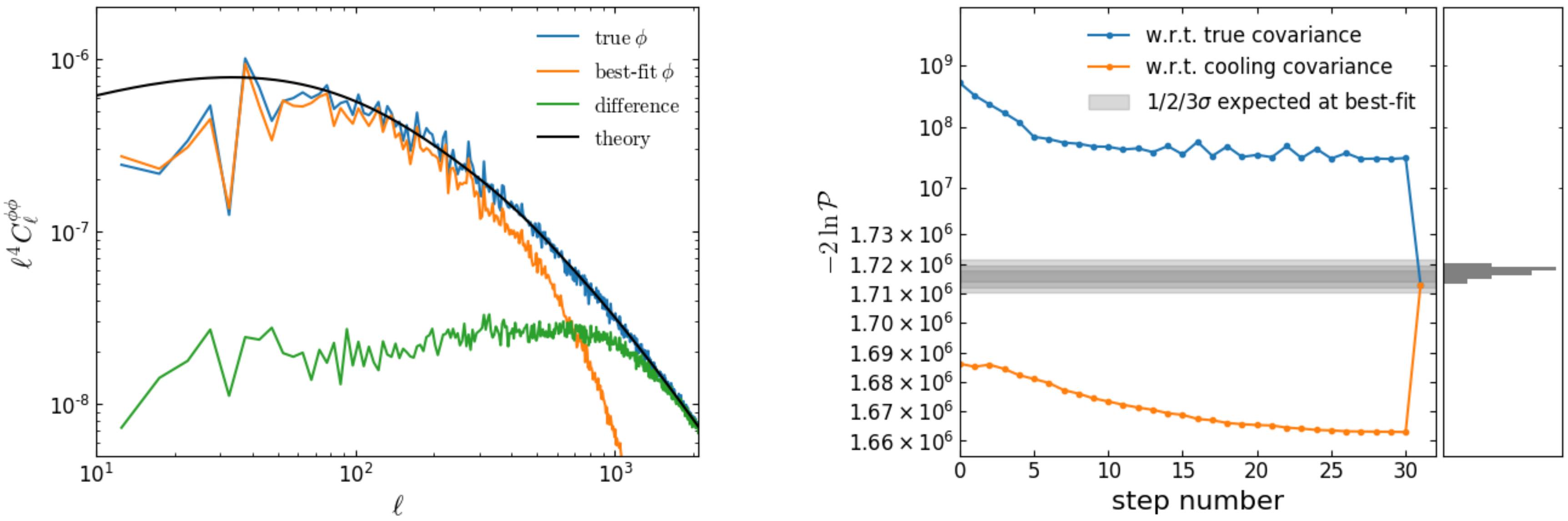


# Conclusions

- We can maximize $\mathcal{P}(f, \phi, r \,|\, d)$
- Sampling is coming up and I've given you a preview of it
- Looking forward to more improvement, application to data, and feedback from the community (see *arxiv:1708.06753*)





true $\phi$
best-fit $\phi$
difference
theory
$\ell^4 C_\ell^{\phi\phi}$
$\ell$
w.r.t. true covariance
w.r.t. cooling covariance
1/2/3$\sigma$ expected at best-fit
$-2\ln\mathcal{P}$
step number

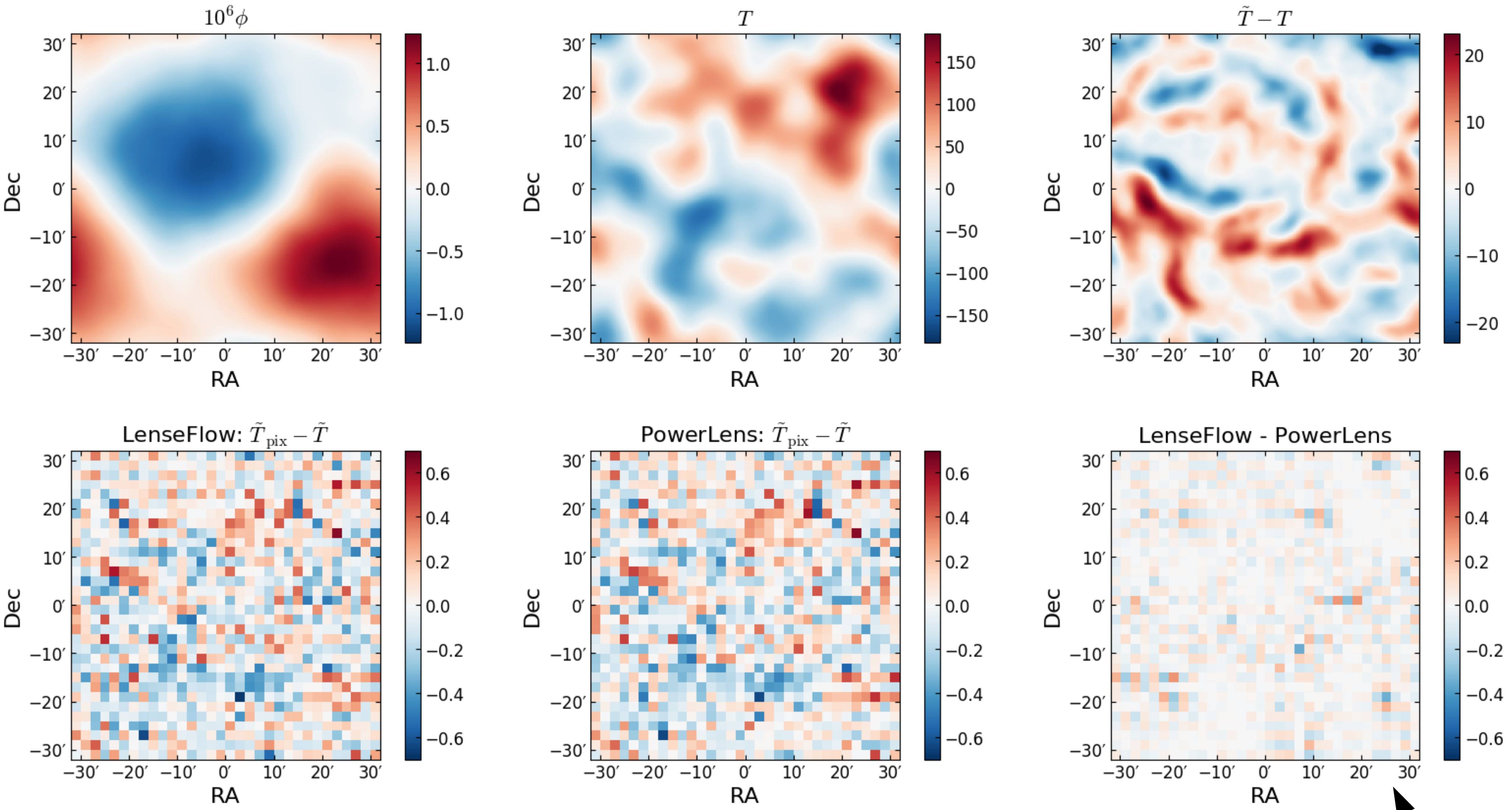
# LenseFlow determinant

$$\frac{df_t(x)}{dt} = \underbrace{\nabla\phi(x)\cdot\left[\mathbb{1} + t\nabla\nabla\phi(x)\right]^{-1}}_{p_t} \cdot \nabla f(x)$$

$$\mathcal{L}(\phi) = \left[\mathbb{1} + \varepsilon\; p_{t_n}\cdot\nabla\right]\cdots\left[\mathbb{1} + \varepsilon\; p_{t_0}\cdot\nabla\right]$$

$$\mathrm{logdet}\left[\mathbb{1} + \varepsilon\; p_t\cdot\nabla\,\right] = \varepsilon\,\mathrm{Tr}\left[\,p_t\cdot\nabla\,\right] + \mathcal{O}(\epsilon^2)$$

So for LenseFlow $\det|\mathcal{L}(\phi)| = 1$ so we can ignore it!




LenseFlow

Taylor series

Differences between two which lead to different determinants