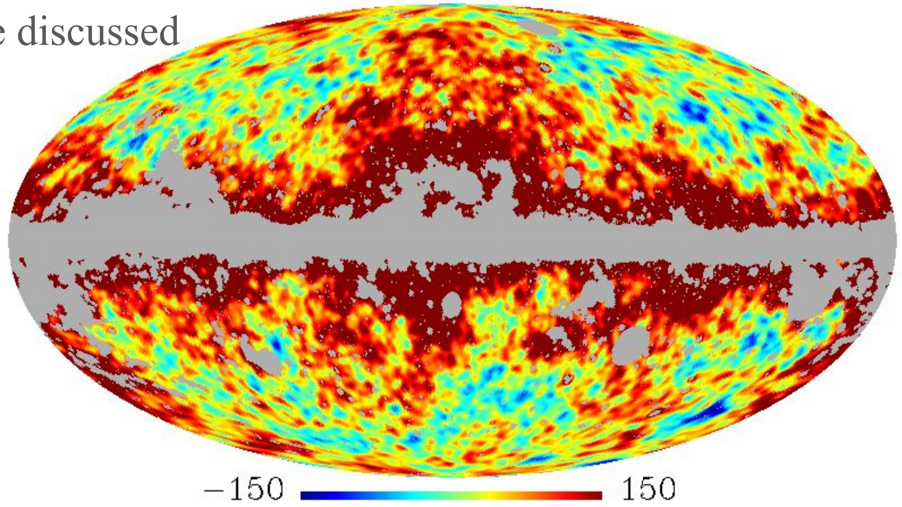
EB-leakage and minimum variance estimation of the fullsky signal & spectrum

Hao Liu

For the Ali-symposuim 2019

What is the question?

- The CMB signal is Gaussian and isotropic.
  - Some conversion might be necessary, will be discussed below.
- But part of the sky is missing.
- The available region is also contaminated.
  - E.g., by noise, systematics, various residuals
  - *The figure is exaggerated for illustration.*



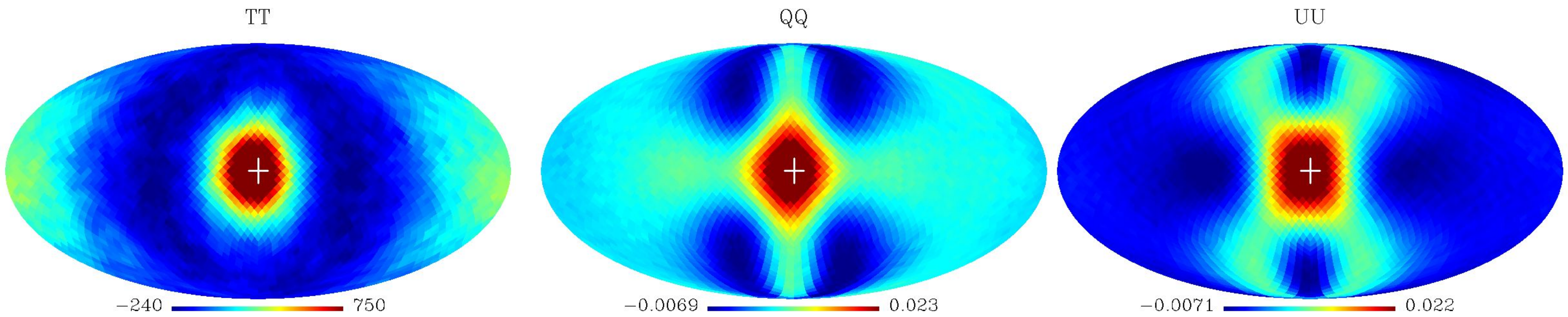
- How to get a minimum variance estimate of the fullsky CMB signal and angular power spectrum?

First consider the missing region: the symmetry of the sky

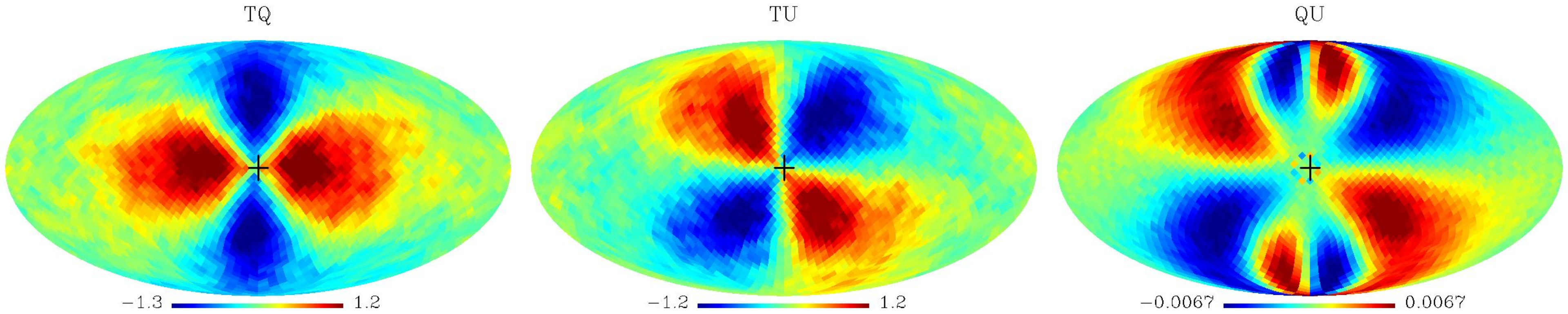
- **Three types of symmetries.**
  - Homogeneity Isotropy No apparent symmetry.
- **Homogeneity**: All points are the same (statistically).
  - Example: white noise.
  - It is very easy to reconstruct the missing part.
- **Isotropy**: Sitting at any position, the points at different distances are different, but points at the same distance are the same.
  - Example: CMB temperature.
  - It is possible to reconstruct the missing region, and there is a minimum variance solution.
- **No apparent symmetry**: Even the points from the same distance look different.
  - The reconstruction can be different for each case.

How about the CMB signal?

- T (temperature) is isotropic.
- E and B modes polarizations are also isotropic.
- But the Q and U stokes parameters are neighther homogeneous nor isotropic.
  - That is the main difference between the temperature & polarization reconstructions.
  - See next page for details.



One row of the pixel domain *TQU* covariance matrix (with one point fixed at the cross)



The full *TQU* covariance matrix in the pixel domain

It was given in Appendix A2 of Tegmark et al, 2001 (originally by Zaldarriaga 1998, *also seen in some later works like Chon et al, 2004, but the ones by Zaldarriaga & Tegmark are earlier* and *more clear*).

$$\mathbf{M}(\hat{\mathbf{r}}_i \cdot \hat{\mathbf{r}}_j) \equiv \begin{pmatrix} \langle T_iT_j\rangle & \langle T_iQ_j\rangle & \langle T_iU_j\rangle \\ \langle T_iQ_j\rangle & \langle Q_iQ_j\rangle & \langle Q_iU_j\rangle \\ \langle T_iU_j\rangle & \langle U_iQ_j\rangle & \langle U_iU_j\rangle \end{pmatrix}, \quad \text{(A9)}$$

$$\langle T_iT_j\rangle \equiv \sum_\ell \left(\frac{2\ell+1}{4\pi}\right) P_\ell(z) C_\ell^T, \quad \text{(A10)}$$

$$\langle T_iQ_j\rangle \equiv -\sum_\ell \left(\frac{2\ell+1}{4\pi}\right) F_\ell^{10}(z) C_\ell^{TE}, \quad \text{(A11)}$$

$$\langle T_iU_j\rangle \equiv -\sum_\ell \left(\frac{2\ell+1}{4\pi}\right) F_\ell^{10}(z) C_\ell^{BT}, \quad \text{(A12)}$$

$$\langle Q_iQ_j\rangle \equiv \sum_\ell \left(\frac{2\ell+1}{4\pi}\right) \left[F_\ell^{12}(z) C_\ell^E - F_\ell^{22}(z) C_\ell^B\right], \quad \text{(A13)}$$

$$\langle U_iU_j\rangle \equiv \sum_\ell \left(\frac{2\ell+1}{4\pi}\right) \left[F_\ell^{12}(z) C_\ell^B - F_\ell^{22}(z) C_\ell^E\right], \quad \text{(A14)}$$

$$\langle Q_iU_j\rangle \equiv \sum_\ell \left(\frac{2\ell+1}{4\pi}\right) \left[F_\ell^{12}(z) + F_\ell^{22}(z)\right] C_\ell^{EB}, \quad \text{(A15)}$$

$$F^{10}(z) = 2\frac{\frac{\ell z}{(1-z^2)} P_{\ell-1}(z) - \left(\frac{\ell}{1-z^2} + \frac{\ell(\ell-1)}{2}\right) P_\ell(z)}{[(\ell-1)\ell(\ell+1)(\ell+2)]^{1/2}} \quad \text{(A16)}$$

$$F^{12}(z) = 2\frac{\frac{(\ell+2)z}{(1-z^2)} P_{\ell-1}^2(z) - \left(\frac{\ell-4}{1-z^2} + \frac{\ell(\ell-1)}{2}\right) P_\ell^2(z)}{(\ell-1)\ell(\ell+1)(\ell+2)} \quad \text{(A17)}$$

$$F^{22}(z) = 4\frac{(\ell+2)P_{\ell-1}^2(z) - (\ell-1)zP_\ell^2(z)}{(\ell-1)\ell(\ell+1)(\ell+2)(1-z^2)} \quad \text{(A18)}$$

However, this requires difference reference frames for each element of the covariance matrix.

$\mathbf{M}_{ij} = \mathbf{L}_{ij'} \cdot \mathbf{C}_{i'j'} \cdot \mathbf{R}_{i'j}$, or in tensor form:

$$C_{ij}^{pk} = M_{ij}^{p'k'} L_{k'}^{p} R_{p'}^{k}$$

After conversion, the real matrix look like...

That is the diffincult point:

- When some region is missing, we need reconstruction.
- For reconstruction, we need isotropy.
- To get isotropy, we need conversion.
- For the conversion, we need full sky.

This problem was also observed in Sec 6.2 of the BICEP2 paper arxiv:1403.3985

To break this circle, we need an explicit pixel-domain EB-leakage correction to restore isotropy of the B-mode (as much as possible). Then the reconstruction can be done just like the case of an isotropic signal.

Therefore, we need two steps for reconstruction

- First solve the EB-leakage.
  - Either by a blind estimate (Hao Liu et al., *PRD 100, 023538 & JCAP04(2019)046*).
  - Or by a minimal variance estimate with given prior information (Hao Liu *arXiv:1906.10381*).
  - Both are the best in the corresponding situation.
- Then choose the best method of sky-reconstruction, and get rid of the remaining problems.

The EB-leakage

- General solutions (also work for all integral transforms)
  - The blind case:

$$\boldsymbol{\mathcal{L}}_{ji}(\boldsymbol{p}) = \Psi_j(\boldsymbol{M}\boldsymbol{W}\Psi_i(\boldsymbol{M}\boldsymbol{f}))$$

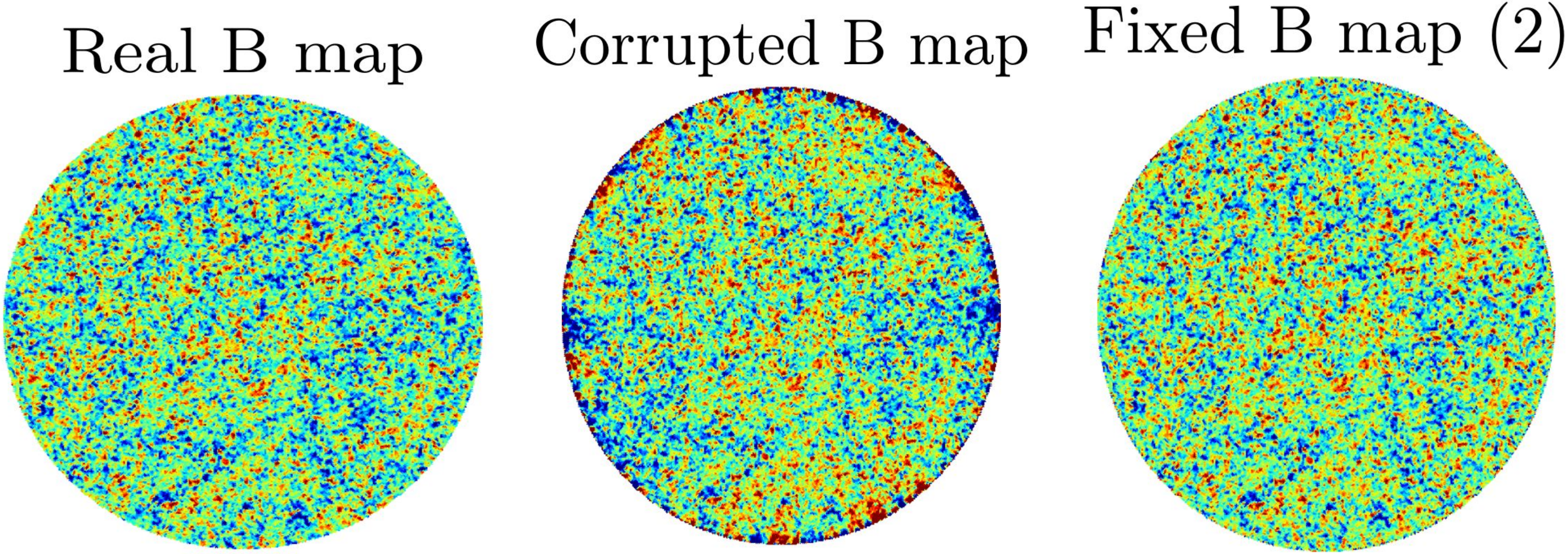
  - With prior information (unbiased estimate with minimal variance).

$$\boldsymbol{\mathcal{L}}^{\mathcal{I}}_{ji}(\boldsymbol{p}) = \boldsymbol{C}_1 \cdot \boldsymbol{C}_2^{-1} \cdot \boldsymbol{\mathcal{L}}_{ji}(\boldsymbol{p})$$

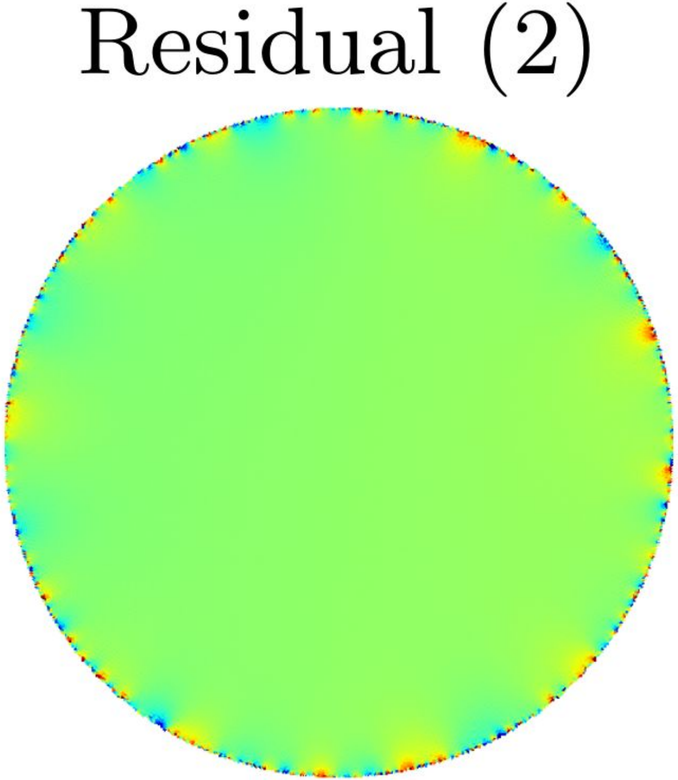
Minimum variance fitting

  - See arxiv:1906.10381 section 3 for details and proofs.

The effect in pixel domain



It is quite easy to get the covariance matrix of residual



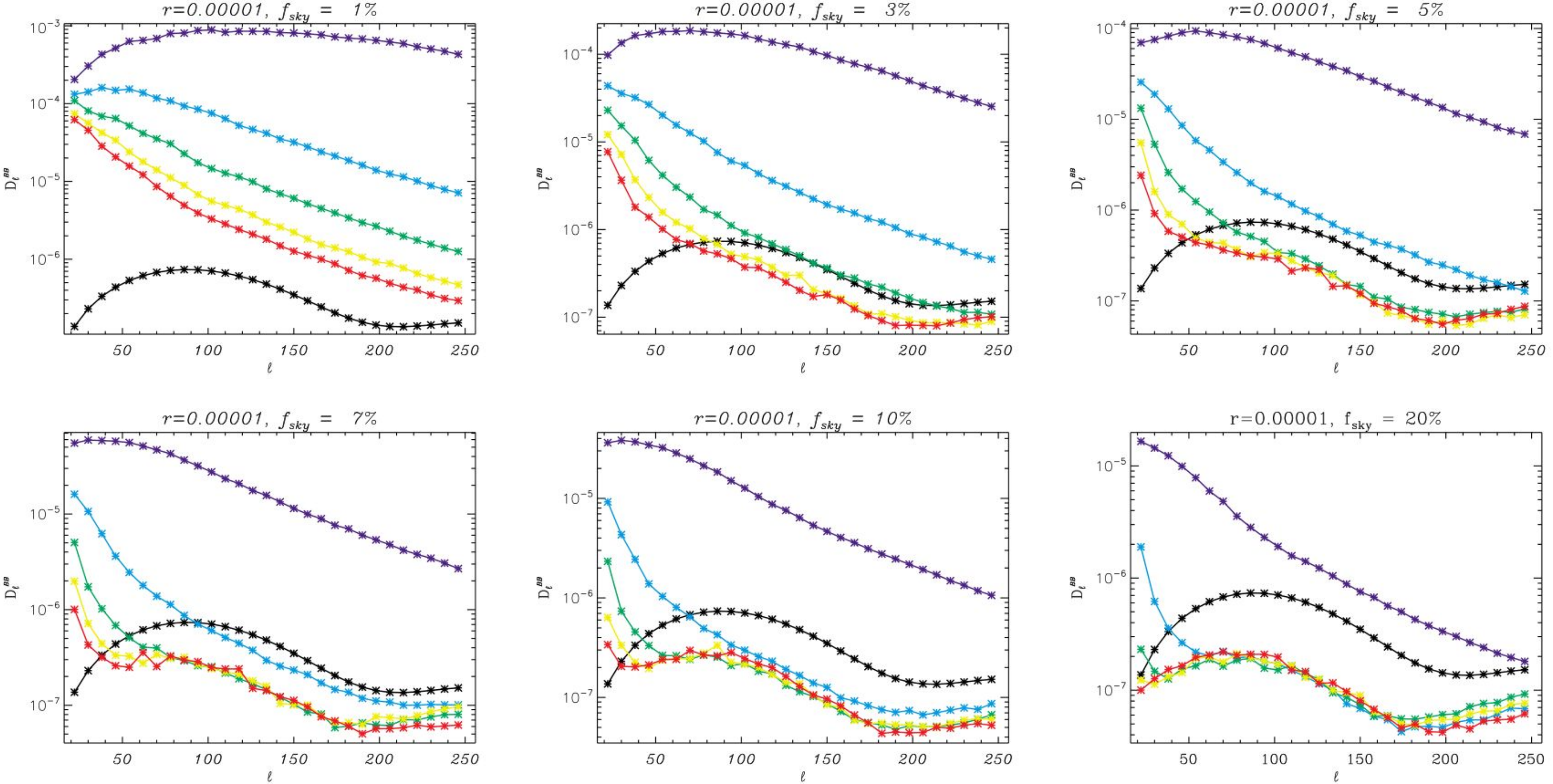
We know how to get the covariance matrix for them
and
Get covariance matrix = know how to fixe them
(details come later)

Estimate the maximum ability to detect primodial GW

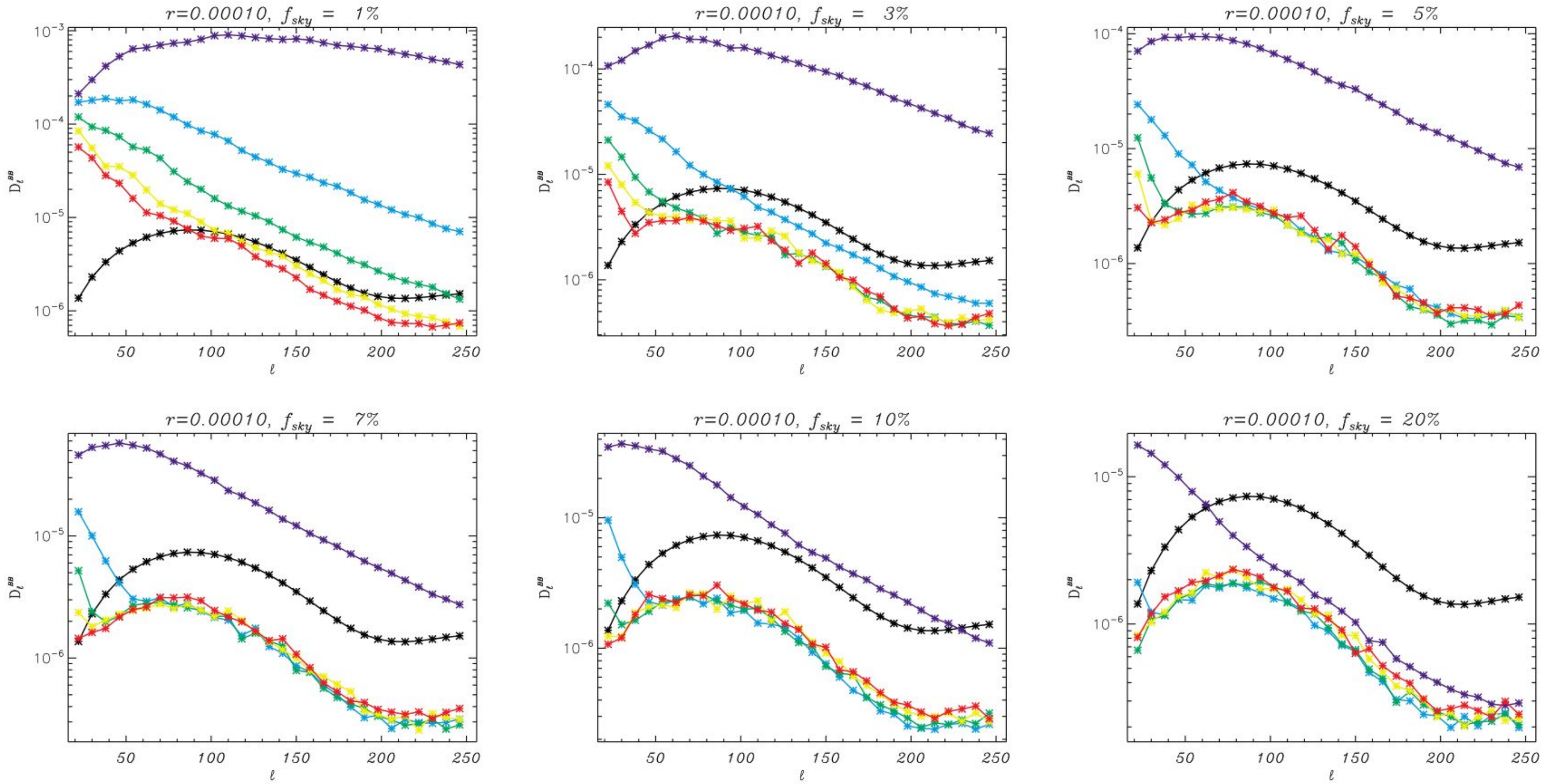
- Assume perfect conditions to compute the maximum detectability.
    - No noise, no systematics.
    - Perfect delensing,
    - No foreground.
    - Only one constraint: limited sky coverage.
- Also assume optimal EB-leakage correction and ideal estimate of the spectrum (fisher estimator, ignore time cost for now).
- The resulting error bar is amplified by 5 times for 5-sigma detectioin.
- The results (next slide)

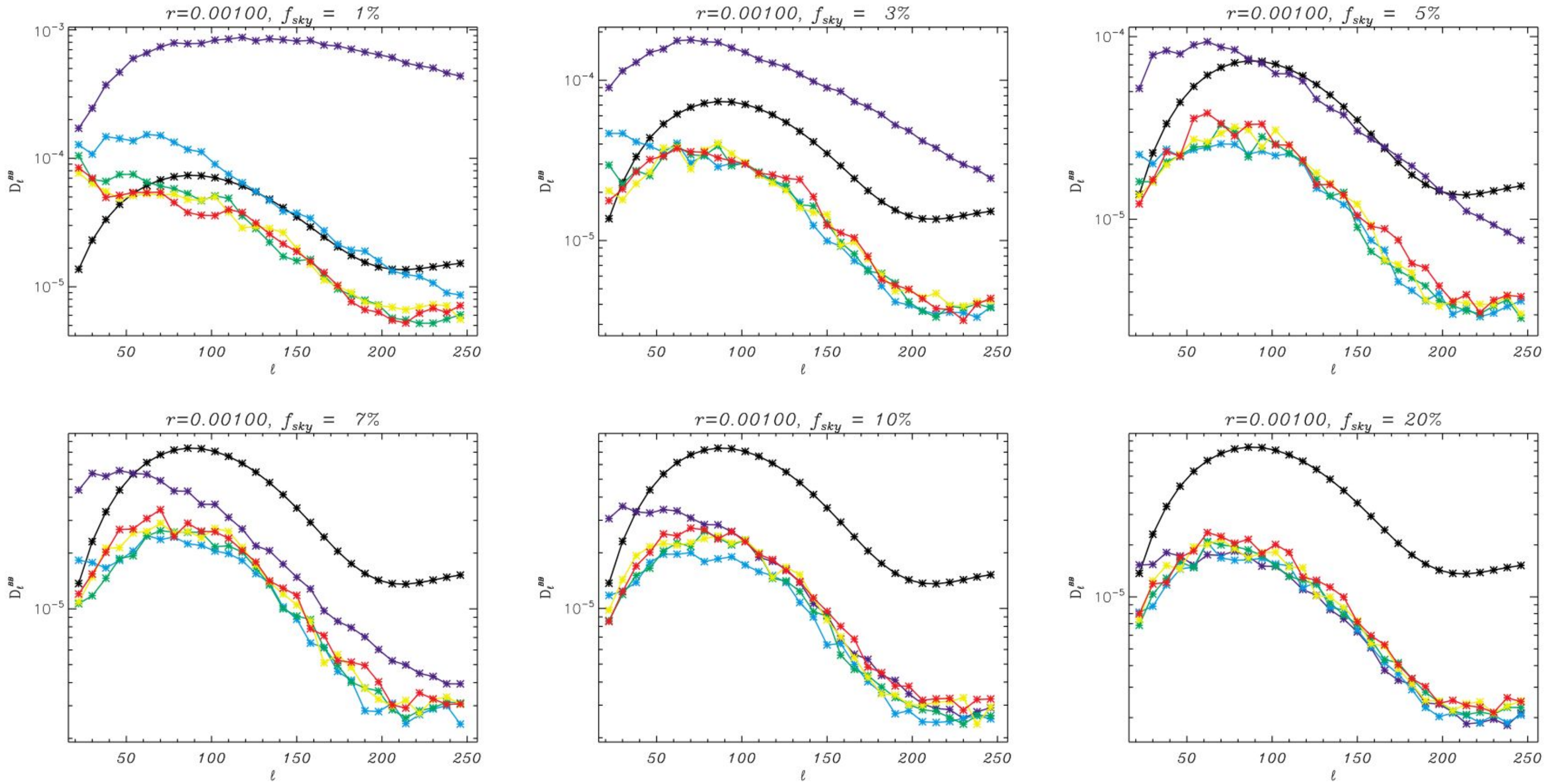
Summary of the detectability

	$r = 10^{-5}$	$r = 10^{-4}$	$r = 10^{-3}$	$r = 10^{-2}$
$f_{sky} = 0.01$	Impossible	Impossible	Barely	Possible
$f_{sky} = 0.03$	Barely	Barely	Possible	Hopeful
$f_{sky} = 0.05$	Barely	Possible	Hopeful	Hopeful
$f_{sky} = 0.07$	Barely	Hopeful	Hopeful	Hopeful
$f_{sky} = 0.10$	Possible	Hopeful	Hopeful	Hopeful
$f_{sky} = 0.20$	Hopeful	Hopeful	Hopeful	Hopeful



r=0.00001, f_sky = 1%
r=0.00001, f_sky = 3%
r=0.00001, f_sky = 5%
r=0.00001, f_sky = 7%
r=0.00001, f_sky = 10%
r=0.00001, f_sky = 20%
D_ℓ^BB
ℓ
50
100
150
200
250


r=0.00010, f_sky = 1%
r=0.00010, f_sky = 3%
r=0.00010, f_sky = 5%
r=0.00010, f_sky = 7%
r=0.00010, f_sky = 10%
r=0.00010, f_sky = 20%
D_ℓ^BB
ℓ
50
100
150
200
250



r=0.00100, f_sky = 1%
r=0.00100, f_sky = 3%
r=0.00100, f_sky = 5%
r=0.00100, f_sky = 7%
r=0.00100, f_sky = 10%
r=0.00100, f_sky = 20%
D_ℓ^BB
ℓ

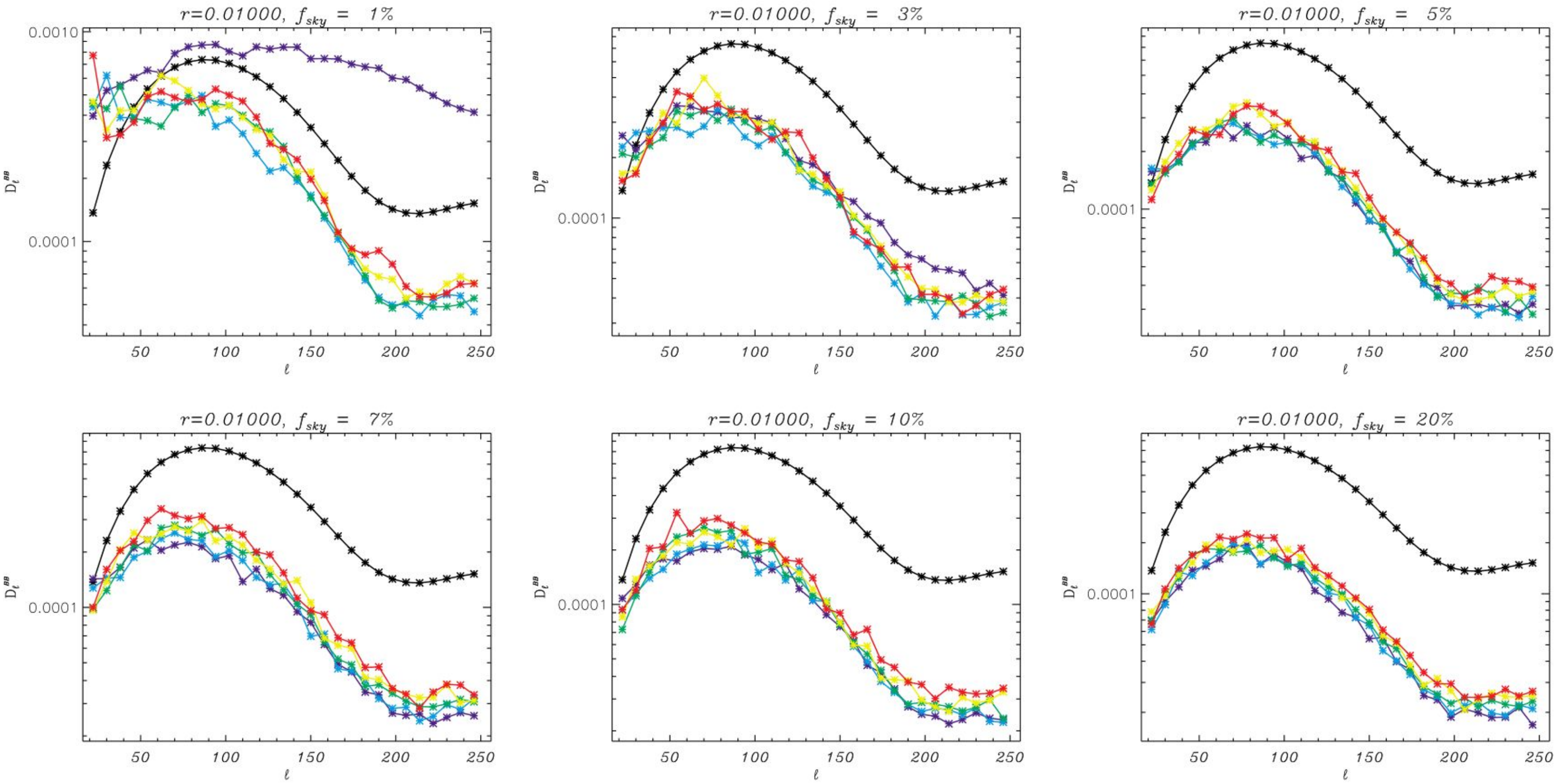


Figure 6. Similar to Figure 3 but for $r = 10^{-2}$.

After fixing the EB-leakage in pixel domain:

- We can consider reconstructing the fullsky map/spectrum
  - **This is not a new problem (next page for a brief review). For example, we already know that the minimum variance estimate can be given by the fisher estimator . But...**
    - The time cost is huge.
    - High-l is impossible.
  - What we want to do:
    - Keep minimum variance.
    - Make it greatly faster (Maybe other min-var methods).
    - Make it work for high-l .
- Let’s start with a very brief review

A very brief review: methods for getting fullsky signal/APS

- All foreground removal methods *(maybe one hundred papers here, so I will skip the citation)*
    - The idea: remove the foreground to get fullsky CMB.
    - Requirements: 1) fullsky map for each band; 2) as many bands as possible.
    - Problem: some region cannot be cleaned, especially for polarization and for the Gal-center and Gal-plane.
- Reconstructing the missing part *(again maybe one hundred papers here, so I will skip the citation)*
    - Non-minimum variance estimator
        - All pseudo-Cl methods (like MASTER).
        - The method by two point correlation functions (like PolSpice).
        - Non-minimum variance inpaintings (like diffusive inpaint).
    - Minimum variance estimator
        - Fisher estimator (Tegmark 1997, and many following works)
        - Other ways of finding the point of maximum likelihood.
            - Most of them contain slight simplification (that can be a good thing).
        - Minimal variance inpainting (Jaiseung Kim et al., 2012).
    - Connection between minimum variance and maximum likelihood?

Likelihood

$$L(\boldsymbol{X}|\boldsymbol{\Theta}) \propto \frac{e^{-\frac{1}{2}\boldsymbol{X}^T C^{-1}\boldsymbol{X}}}{\sqrt{|C|}},$$

Arxiv:1906.10381, appendix A

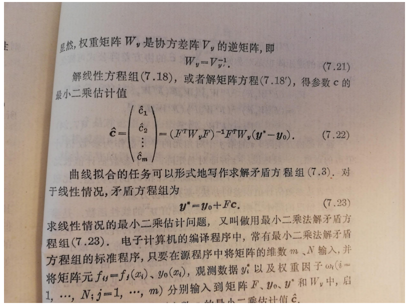
Fisher estimator

$$F_{\ell\ell'} = 2\mathbf{Tr}[CE^{\ell}CE^{\ell'}]$$
$$\widetilde{C}_{\ell} = F^{-1}_{\ell\ell'} y_{\ell'}$$
$$E^{\ell} = \frac{1}{2}C^{-1}\frac{\partial C}{\partial C_{\ell}}C^{-1},\ y_{\ell} = \boldsymbol{X}^T E^{\ell}\boldsymbol{X}$$

Minimum variance fitting

In many textbooks, like:

“实验的数学处理”, 李惕碚, P271, equ. (7.22)



显然, 权重矩阵 W_y 是协方差阵 V_y 的逆矩阵, 即

$$W_y = V_{y'}^{-1}. \qquad (7.21)$$

解线性方程组(7.18), 或者解矩阵方程(7.18′), 得参数 $\boldsymbol{c}$ 的最小二乘估计值

$$\hat{\boldsymbol{c}} = \begin{pmatrix}\hat{c}_1\\ \hat{c}_2\\ \vdots\\ \hat{c}_m\end{pmatrix} = (F^T W_y F)^{-1} F^T W_y (\boldsymbol{y}^* - \boldsymbol{y}_0). \qquad (7.22)$$

曲线拟合的任务可以形式地写作求解矛盾方程组(7.3). 对于线性情况, 矛盾方程组为

$$\boldsymbol{y}^* = \boldsymbol{y}_0 + F\boldsymbol{c}. \qquad (7.23)$$

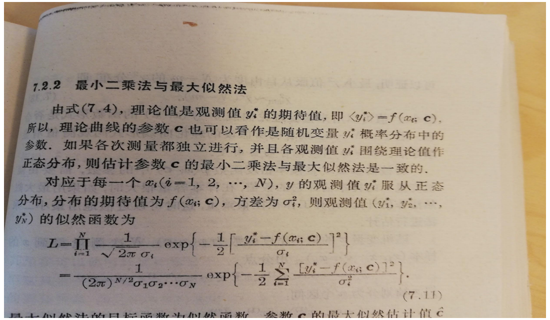
求线性情况的最小二乘估计问题, 又叫做用最小二乘法解矛盾方程组(7.23). 电子计算机的编译程序中, 常有最小二乘法解矛盾方程组的标准程序, 只要在源程序中将矩阵的维数 m、N 输入, 并将矩阵元 $f_{ij} = f_j(x_i)$、$y_0(x_i)$, 观测数据 y_i^* 以及权重因子 $\omega_i(i=1, \cdots, N; j=1, \cdots, m)$ 分别输入到矩阵 F、$\boldsymbol{y}_0$、$\boldsymbol{y}^*$ 和 W_y 中, 启

$$\langle \boldsymbol{Y}\rangle = \boldsymbol{P}\boldsymbol{Q}^{-1}\boldsymbol{X}$$
$$\boldsymbol{P} = \langle \boldsymbol{Y}\boldsymbol{X}^T\rangle$$
$$\boldsymbol{Q} = \langle \boldsymbol{X}\boldsymbol{X}^T\rangle$$

Multi-variant Normal distribution

$$\boxed{f_{\mathbf{X}}(x_1,\ldots,x_k) = \frac{\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)}{\sqrt{(2\pi)^k|\boldsymbol{\Sigma}|}}}$$

For CMB, “minimum variance” and “maximum likelihood” are equivallent. See e.g., Tegmark 1997, or some textbooks, like



7.2.2 最小二乘法与最大似然法

由式(7.4), 理论值是观测值 y_i^* 的期待值, 即 $\langle y_i^*\rangle = f(x_i; \boldsymbol{c})$, 所以, 理论曲线的参数 $\boldsymbol{c}$ 也可以看作是随机变量 y_i^* 概率分布中的参数. 如果各次测量都独立进行, 并且各观测值 y_i^* 围绕理论值作正态分布, 则估计参数 $\boldsymbol{c}$ 的最小二乘法与最大似然法是一致的.

对应于每一个 $x_i(i=1, 2, \cdots, N)$, y 的观测值 y_i^* 服从正态分布, 分布的期待值为 $f(x_i; \boldsymbol{c})$, 方差为 σ_i^2, 则观测值 $(y_1^*, y_2^*, \cdots, y_N^*)$ 的似然函数为

$$L = \prod_{i=1}^{N}\frac{1}{\sqrt{2\pi}\,\sigma_i}\exp\left\{-\frac{1}{2}\left[\frac{y_i^* - f(x_i; \boldsymbol{c})}{\sigma_i}\right]^2\right\}$$
$$= \frac{1}{(2\pi)^{N/2}\sigma_1\sigma_2\cdots\sigma_N}\exp\left\{-\frac{1}{2}\sum_{i=1}^{N}\frac{[y_i^* - f(x_i; \boldsymbol{c})]^2}{\sigma_i^2}\right\}. \qquad (7.11)$$

Wiener filter

Causal solution [edit]

$$G(s) = \frac{H(s)}{S_x^+(s)},$$

How to choose a best estimator?

- Some requirements:
  - Minimum variance (Do not waste the work of others).
  - As simple as possible in mathematics; Robust and sufficiently fast.
- What we observe:
  - No current method satisfies all of them and is ready-for-use.
  - The minimal variance inpainting (Jaiseung Kim et al., 2012) is the closest candidate.
  - Improve it → **Multi-Resolution Minimal Variance inPaint (Mr.MVP)**
- For understanding: the main ideas of Mr.MVP and Fisher estimator:
  - Both based on the same constraints:
    - signals in the available region + Gaussianity, isotropy + noise covariane matrix
  - Fisher estimator: Given all those constrains, what is the most probable C_l ?
  - Mr.MVP: use a three-step approach:
    - How to find realizations of the missing part that satisfy all constraints (coherent realizations)?
    - With the realizations, what is the expectation of the map?
    - With the realizations, what is the expectation of C_l ?

Mathematical basis of Mr.MVP

Original ideas: Hoffman and Ribak (1991), Jaiseung Kim et al., 2012 (in harmonic domain)

Two point covariance

$$C_{ij} = \frac{1}{4\pi}\sum_{\ell}(2l+1)W_\ell^2 C_\ell P_\ell[\cos(\theta_{ij})],$$

Solution of the map expectation

missing region

$$\langle \boldsymbol{Y} \rangle = \boldsymbol{P}\boldsymbol{Q}^{-1}\boldsymbol{X}$$

Available region

$$\boldsymbol{P} = \langle \boldsymbol{Y}\boldsymbol{X}^T \rangle$$

$$\boldsymbol{Q} = \langle \boldsymbol{X}\boldsymbol{X}^T \rangle$$

Q is square but P can be non-square

This is exactly a minimum variance solution

To get a coherent realization of the missing region:

$$\boldsymbol{Y}_i^{rlz} = \langle \boldsymbol{Y} \rangle + \boldsymbol{Y}_i - \boldsymbol{P}\boldsymbol{Q}^{-1}\boldsymbol{X}_i$$

One coherent realization of the missing part

available/missing parts of one simulation

Why this is the best way?

- Minimal variance.
- Extremely simple in mathematics.
  - No Wigner 3j or 6j symbols.
  - Even no spherical harmonics.
- Can be **greatly** accelerated by multi-resolution with neglegible loss of accuracy (2%).

Multi-resolution?

The time performance (compared with fisher estimator)

- Fisher estimator:
  - *Nside=64, 1000 simulations, f_{sky}~80%:* roughly one day on 16384 CPUs (Diego et al., 2014)
  - **Time cost scales as ~$Nside^6$**
- Mr.MVP:
  - *Nside=64, 1000 simulations, f_{sky}~80%:* 2 minutes on my laptop.
  - **With the help of multi-resolution, the time cost scales as ~$Nside^3$.**
- In summary: Mr.MVP is about 5-6 orders of magnitudes faster than QML, with merely 2% larger error bar.
- For high-*l*, this is probably the only way to get a minimum variance estimate.

Noise, residual, beam transfer function, extra apodization

$$C_{ij} \rightarrow C_{ij} + N_{ij}$$

For noise, sys, residual, just add the corresponding covariance matrix into C_{ij}

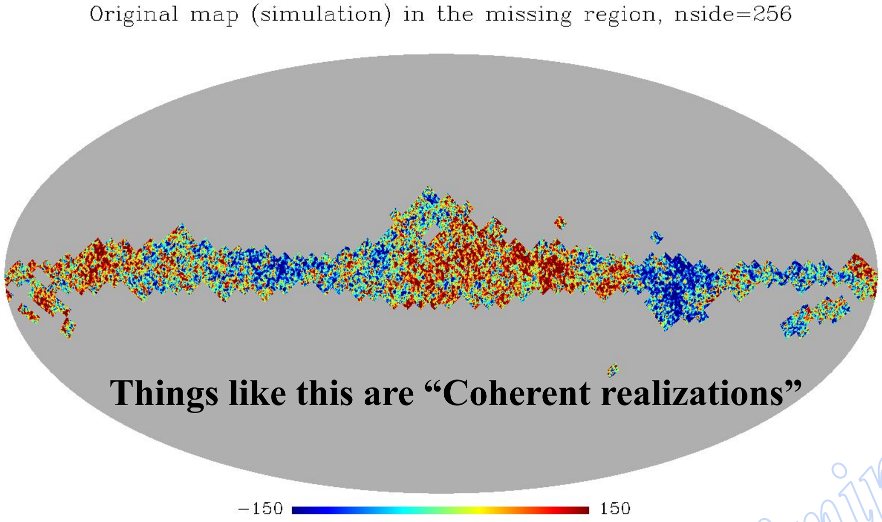
$$\boldsymbol{P} \rightarrow \boldsymbol{P}\boldsymbol{m}^T$$
$$\boldsymbol{Q} \rightarrow \boldsymbol{m}\boldsymbol{Q}\boldsymbol{m}^T$$

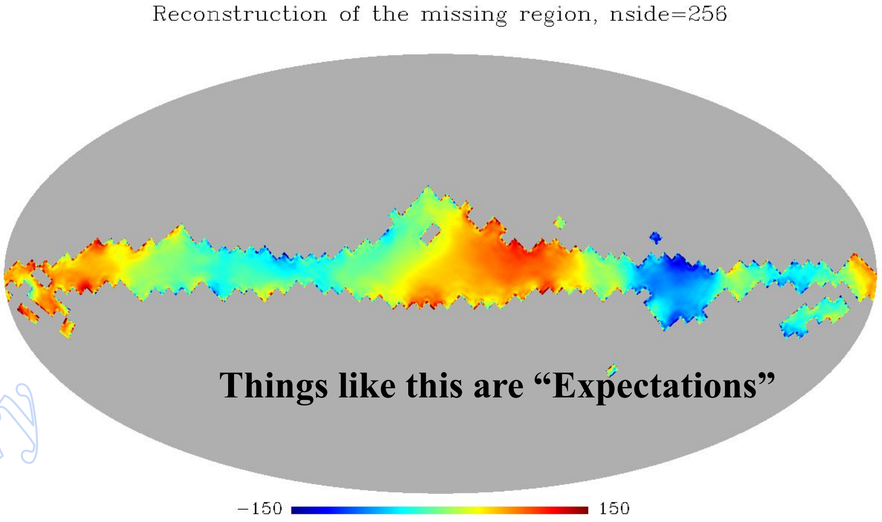
For additional apodization $\boldsymbol{m}$, put it into $\boldsymbol{P}$ and $\boldsymbol{Q}$ as shown in the left.

The beam transfer function W_l is already in C_{ij}

Because the estimate is already optimal, there is **no need for apodization if the noise covariance matrix is good enough**. If not, the apodization serves like a trade between accuracy and safety.

Pixel domain performance of Mr.MVP (expectation, one simulation for example, nside=256)





Note:

1) The difference between expectation and coherent realization

2) The difference between map expectation and power spectrum expectation

Note: the pixel domain expectation is NOT the expectation of the angular power spectrum!

$$\alpha_k^{\ell m} = a^{\ell m} + b_k^{\ell m}$$

Full-sky

index of realization

Available region

Missing region

The expectation of the map (not spectrum!)

$$\langle \alpha_k^{\ell m} \rangle = a^{\ell m} + \langle b_k^{\ell m} \rangle = a^{\ell m} + e^{\ell m}$$

The expectation of the spectrum

$$\langle |\alpha_k^{\ell m}|^2 \rangle = \langle |a^{\ell m} + b_k^{\ell m}|^2 \rangle = |a^{\ell m}|^2 + 2\mathrm{Re}(a^{\ell m *} e^{\ell m}) + \langle |b_k^{\ell m}|^2 \rangle$$

The spectrum of expectation

$$|\langle \alpha_k^{\ell m} \rangle|^2 = |a^{\ell m} + e^{\ell m}|^2 = |a^{\ell m}|^2 + 2\mathrm{Re}(a^{\ell m *} e^{\ell m}) + |e^{\ell m}|^2$$

Wanted: Expectation of spectrum

$$\langle |\alpha_k^{\ell m}|^2 \rangle = |\langle \alpha_k^{\ell m} \rangle|^2 + \langle |b_k^{\ell m}|^2 \rangle - |e^{\ell m}|^2$$

Spectrum of the missing region expectation

Spectrum of expectation

Expectation of the missing region spectrum

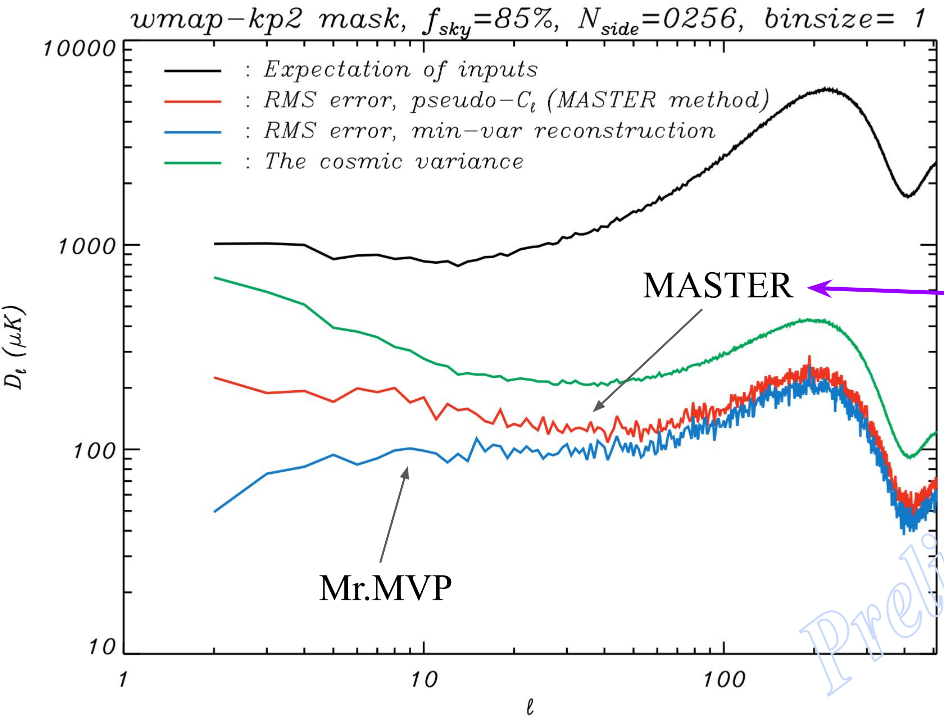
A real Tongue twister!

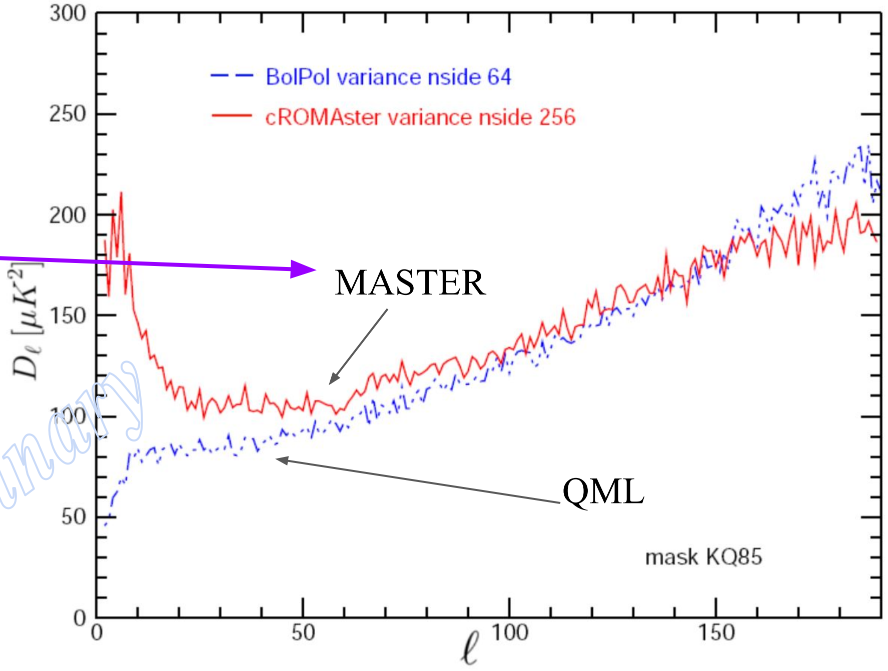
The performance of angular power spectrum reconstruction

compare with Fisher estimator (QML method), nearly same mask

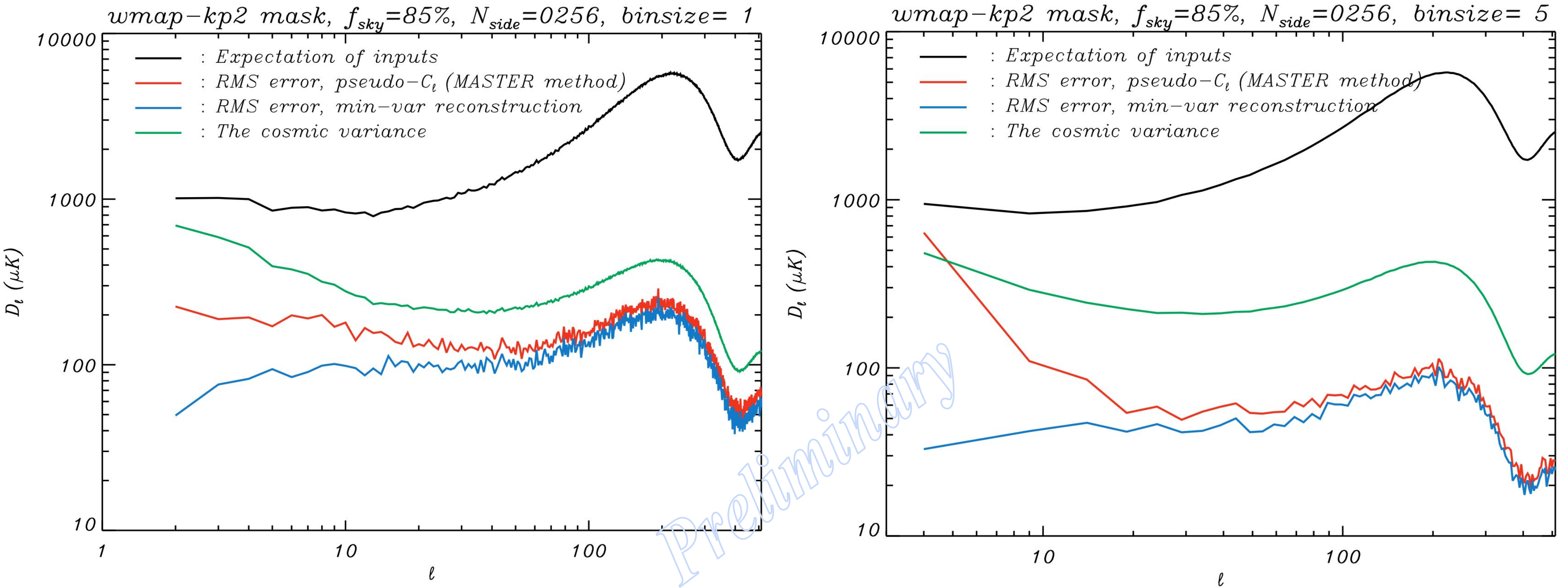
Mr.MVP
nside=256, and nside=512 for MASTER

Fisher estimator
(Diego et al., arXiv:1403.1089)



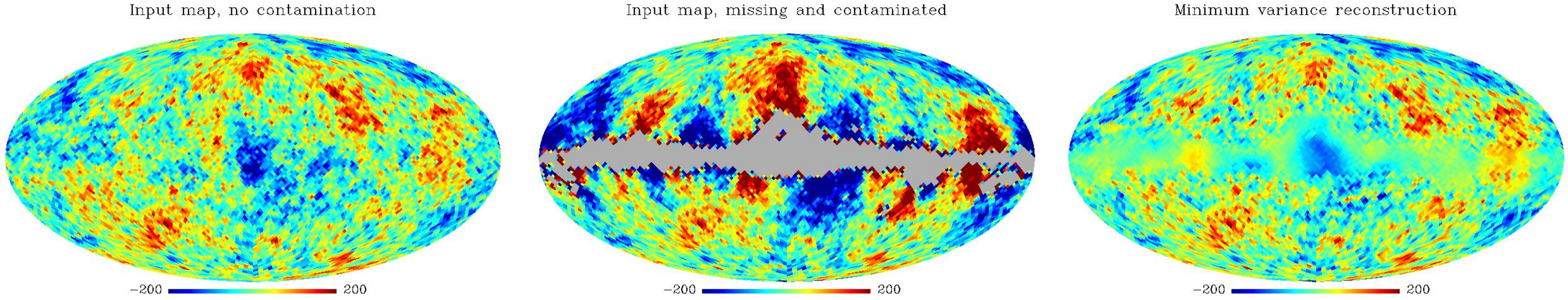


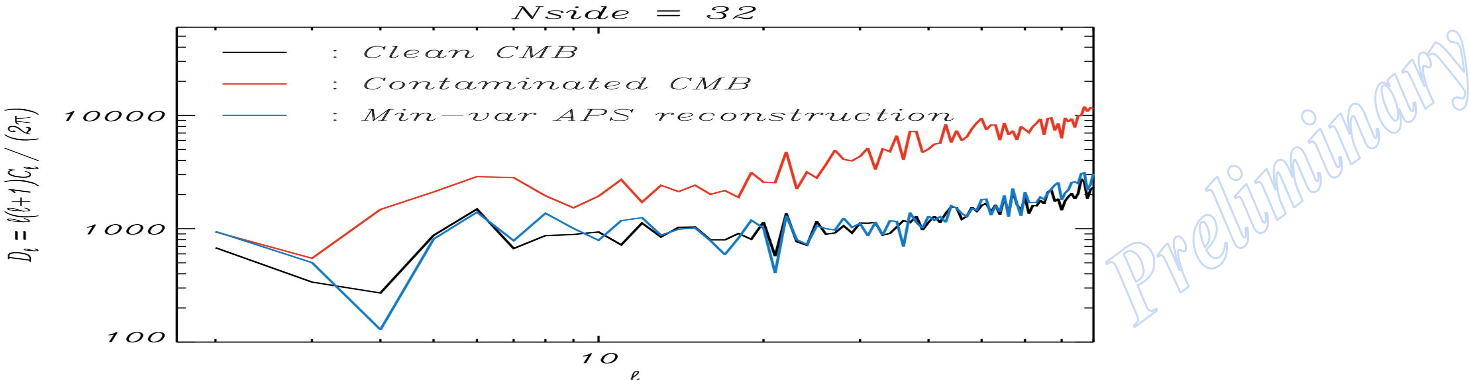
The performance of angular power spectrum reconstruction
Compare different binsizes



What can we do for the next step

- **2nd part of Mr. MVP: Multi-Resolution Minimal Variance outPaint.**
  - Out paint = paint the available region
  - This is necessary when the available region is also contaminated...





Can we remove the foreground like this?

- Yes, but not the first choice.
  - More band = more information = better results.
  - Multi-band foreground removal is always the first choice.
  - Similar for noise, systematics, EB-leakage...
- The minimal variance estimate is the last step V.S. the “cannot” things.
  - Regions that “cannot” be used.
  - Residuals that “cannot” be removed directly.
  - If covariance matrix is availabe, use it.
  - If covariance matrix is unavailable, use empirical weights.

Thank you for your attention!