

Lecture 5

Outline:

- 1) Simplified data model
- 2) bandpowers & \hat{C}_e^{TT}
- 3) $\text{var}(\hat{C}_e^{\text{TT}})$, noise bias, cosmic variance
- 4) Posterior samples of C_e^{TT}

Start by studying estimation in a simplified case: $T(\hat{n})$ observed everywhere with white noise & a beam

Assume the observed field $d(\hat{n})$ is measured on all $\hat{n} \in S^2$ & has the form

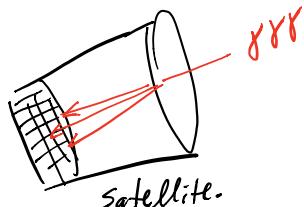
$$(*) \quad d(\hat{n}) = \Phi * T(\hat{n}) + \underbrace{\varepsilon(\hat{n})}_{\substack{\text{white noise s.t.} \\ \text{models a beam point spread function.}}} \quad E(\varepsilon_{\text{em}} \bar{\varepsilon}_{\text{em}}) = \sigma^2 \delta_{\text{em}} \delta_{\text{am}}$$

Note: the fact that $\varepsilon(\hat{n})$ is white noise we expect statistical properties of Model (*) to be similar to the case of finite pixel observations

①

The beam $\Phi(\hat{n})$

As the CMB pass through the detector lens, they get deflected and spread out



Satellite.

The effect on the observed $T(\hat{n})$ is modeled as a convolution with a point spread function $\Phi(\hat{n})$.

on S^2

$$\bullet \text{ Signal } T(\hat{n}), \hat{n} \in S^2 \quad | \quad \bullet \text{ Signal } T(x), x \in \mathbb{R}^2$$

$$\bullet \Phi * T(\hat{n}) = \int_{g \in S^2} T(g \cdot \hat{n}) \Phi(g) dg \quad | \quad \bullet \Phi * T(x) = \int_{\mathbb{R}^2} T(y) \Phi(x-y) dy$$

$$\begin{aligned} & \bullet (\Phi * T)_k = \iint_{\mathbb{R}^2} e^{-ik \cdot y} T(y) \Phi(x-y) dy dx \\ & \quad \text{models a beam point spread function.} \end{aligned}$$

How do we

model $\Phi(\hat{n})$
(equiv Φ_k) on S^2

$$\bullet \Phi(x) \text{ modeled as} \quad \frac{1}{(2\pi)^{d_x} b^{d_x}} \exp\left(-\frac{|x|^2}{2b^2}\right)$$

$$\therefore \Phi_k = \frac{1}{(2\pi)^{d_k} b^{d_k}} \exp\left(-\frac{|k|^2}{2b^2}\right)$$

$$\approx \frac{1}{(2\pi)^{d_k} b^{d_k}} \exp\left(-\frac{b^2}{8\log(2)} |k|^2\right)$$

where $b = \text{full width half max}$

$$\therefore (\Phi * T_k) = \exp\left(-\frac{b^2}{8\log(2)} |k|^2\right) T_k$$

For $x \in S^2$, $\varphi(x)$ has the property that the solution to

$$\Delta u^t(x) = \frac{d}{dt} u^t(x)$$

is expressed

$$u_k^t = \exp\left(-t \|k\|^2\right) u_k^0$$

$$= (\varphi * u^0)_k, \text{ for } t = \frac{b^2}{8 \log(2)}.$$

On the sphere Δ is defined and has the property that the solution to

$$\Delta u^t(\hat{n}) = \frac{d}{dt} u^t(\hat{n})$$

is expressed as

$$u_{em}^t = \exp\left(-t \ell(\ell+1)\right) u_{em}^0$$

$$= (\varphi * u^0)_{em}$$

$$\text{i.e. } \varphi_{em} = \frac{1}{2\pi} \sqrt{\frac{2\ell+1}{4\pi}} \exp(-t \ell(\ell+1))$$

is how we model the beam in the simplified case on S^2 .

(3)

Since we assume observations of (*) everywhere on S^2 we can take the spherical transform to obtain:

$$d_{em} = \exp\left(-\frac{b^2}{8 \log(2)} \ell(\ell+1)\right) T_{em} + \varepsilon_{em}$$

Dividing out the beam (& renaming d_{em}, ε_{em}) we can assume w.l.o.g. the data is in the form of an infinite sequence of spherical coefficients:

$$d_{em} = T_{em} + \varepsilon_{em}$$

for $\ell = 0, 1, 2, \dots$, $m = -\ell, -\ell+1, \dots, \ell$ s.t.

$$E(T_{em} T_{e'm'}^*) = \delta_{ee} \delta_{mm'} C_e^{TT}$$

$$E(\varepsilon_{em} \varepsilon_{e'm'}^*) = \delta_{ee} \delta_{mm'} \sigma^2 \exp\left(\frac{b^2}{8 \log(2)} \ell(\ell+1)\right)$$

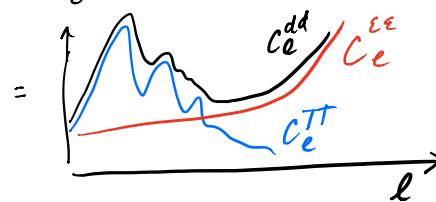
Notice

$$C_e^{dd} = E|d_{em}|^2$$

$$= E|T_{em}|^2 + 2 \underbrace{E(T_{em} \bar{\varepsilon}_{em})}_{=0} + E|\varepsilon_{em}|^2$$

$$= C_e^{TT} + \sigma^2 \exp\left(\frac{b^2}{8 \log(2)} \ell(\ell+1)\right)$$

$$= C_e^{dd}$$



(4)

Nonparametric estimates of C_e^{TT}

(5)

Let $\sigma_e = \frac{1}{2l+1} \sum_{m=-l}^l |d_{em}|^2$ "band power"

Now $E(\sigma_e) = \frac{1}{2l+1} \sum_{m=-l}^l E|d_{em}|^2$

$$= C_e^{dd}$$

$$= C_e^{TT} + C_e^{EE}$$

\curvearrowleft Noise bias...
but known
and can be
subtracted.

Now

$$\begin{aligned} \text{var}(\sigma_e) &= E[(\sigma_e)^2] - (C_e^{dd})^2 \\ &= \frac{1}{(2l+1)^2} \sum_{m=-l}^l \sum_{m'= -l}^l E(d_{em}^* d_{em'}^* d_{em'} d_{em}) \\ &\quad - (C_e^{dd})^2 \end{aligned}$$

where

$$\begin{aligned} E(d_{em}^* d_{em'}^* d_{em'} d_{em}) &= E(d_{em}^* d_{em}^* d_{em'} d_{em'}) \\ &\quad + E(d_{em}^* d_{em'}^* d_{em'} d_{em}) \\ &\quad + E(d_{em}^* d_{em'} d_{em'} d_{em}) \\ &\quad + E(d_{em}^* d_{em'} d_{em} d_{em'}) \\ &= E|d_{em}|^2 E|d_{em'}|^2 \\ &\quad + E(d_{em} d_{em'}) E(d_{em}^* d_{em'}^*) \\ &\quad + E(d_{em} d_{em'}^*) E(d_{em}^* d_{em'}) \\ &= (C_e^{dd})^2 \\ &\quad + (-1)^m \int_{M,-m} f_{M,-m} C_e^{dd} (-1)^m \int_{-M,m} f_{-M,m} C_e^{dd} \\ &\quad + \int_{m,m} f_m C_e^{dd} \int_{-m,-m} f_{-m,-m} C_e^{dd} \end{aligned}$$

(6)

$$\begin{aligned} \therefore \text{var}(\sigma_e) &= (C_e^{dd})^2 + \frac{2}{(2l+1)^2} \sum_{m=-l}^l (C_e^{dd})^2 \\ &\quad - (C_e^{dd})^2 \\ &= \frac{2}{2l+1} (C_e^{dd})^2 \end{aligned}$$

$$\therefore \text{var}(\sigma_e) = \frac{2}{2l+1} (C_e^{TT} + C_e^{EE})^2$$

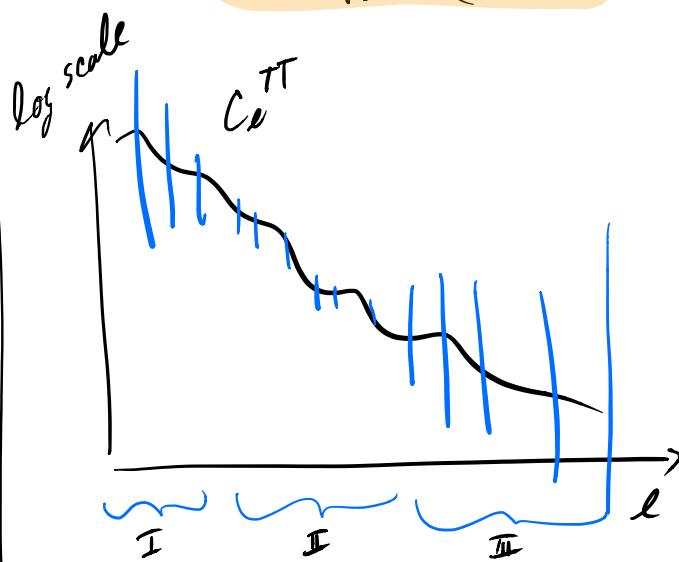
Now a nonparametric estimate of C_e^{TT} can be constructed as:

$$\hat{C}_e^{TT} := \sigma_e - C_e^{EE}$$

where

$$E(\hat{C}_e^{TT}) = C_e^{TT}$$

$$\text{sd}(\hat{C}_e^{TT}) = \sqrt{\frac{2}{2l+1}} (C_e^{TT} + C_e^{EE})$$



$$\text{Relative S.d.} = \frac{\sqrt{\frac{2}{2l+1} (C_e^{TT} + C_e^{EE})}}{C_e^{TT}}$$

$$= \sqrt{\frac{2}{2l+1} \left(1 + \frac{C_e^{EE}}{C_e^{TT}} \right)}$$

$$= \begin{cases} \text{big } (1 + \frac{\text{small}}{\text{big}}) & \text{in I} \\ \text{small } (1 + 1) & \text{in II} \\ \text{small } (1 + \frac{\text{big}}{\text{small}}) & \text{in III} \end{cases}$$

Cosmic Variance

Notice that even in the noise free case

$$d(\hat{n}) = T(\hat{n}) \Leftrightarrow d_{\text{em}} = T_{\text{em}}$$

The estimate $\hat{C}_e^{\text{TT}} = C_e$ one still has uncertainty @ each ℓ :

$$\text{sd}(\hat{C}_e^{\text{TT}}) = \sqrt{\frac{2}{2\ell+1}} C_e^{\text{TT}}$$

This is called cosmic variance and comes from the fact we only have one realization of the universe i.e. just one map $T(\hat{n})$.

\therefore if a parameter θ_i only effects C_e^{TT} for $0 \leq \ell \leq l_{\max}$ then θ_i can not be consistently estimated from $d(\hat{x})$ as noise $\rightarrow 0$.

However since $\text{sd}(\hat{C}_e^{\text{TT}}) \rightarrow 0$ as $\ell \rightarrow \infty$ (in the noise free case) consistent estimation is possible when observing only one realization of $T(\hat{n})$ on $\hat{n} \in S^2$.

(7)

$$P(\vec{d} | C_e^{\text{TT}}, C_e^{\text{EE}})$$

(8)

Going back to the Model

$$d_{\text{em}} = T_{\text{em}} + E_{\text{em}}$$

where E_{em} 's are mean zero complex Gaussian s.t. $E_{\ell,-m} = (-1)^m \bar{E}_{\ell,m}$ &

$$E(E_{\ell,m} \bar{E}_{\ell',m'}) = \delta_{\ell\ell'} \delta_{mm'} \sigma^2 \exp\left(\frac{b^2}{8\log(2)} \ell(\ell+1)\right).$$

claim: If m is restricted to be ≥ 0 then the E_{em} 's are independent &

$$m > 0 \Rightarrow \begin{pmatrix} \text{Re}(d_{\text{em}}) \\ \text{Im}(d_{\text{em}}) \end{pmatrix} \stackrel{(1)}{\sim} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{C_e^{\text{TT}} + C_e^{\text{EE}}}{2} & 0 \\ 0 & \frac{C_e^{\text{TT}} + C_e^{\text{EE}}}{2} \end{pmatrix}\right)$$

$$m = 0 \Rightarrow \begin{pmatrix} \text{Re}(d_{\text{em}}) \\ \text{Im}(d_{\text{em}}) \end{pmatrix} \stackrel{(2)}{\sim} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} C_e^{\text{TT}} & 0 \\ 0 & 0 \end{pmatrix}\right)$$

Proof: First Note $E|d_{\text{em}}|^2 = C_e^{\text{TT}} + C_e^{\text{EE}}$ implies $E(\text{Re}(d_{\text{em}})^2) + E(\text{Im}(d_{\text{em}})^2) = C_e^{\text{TT}} + C_e^{\text{EE}}$. Since d_{em} is real we immediately get (2).

Also $E(d_{\text{em}} d_{\text{em}}) = 0$ implies

$$E(\text{Re}(d_{\text{em}})^2) - E(\text{Im}(d_{\text{em}})^2) = 0 \quad (4)$$

$$2E(\text{Re}(d_{\text{em}}) \text{Im}(d_{\text{em}})) = 0 \quad (5)$$

$\therefore (3), (4) \& (5)$ implies (1)

$$\begin{aligned} (z_1 + z_2)(\bar{z}_1 + \bar{z}_2) &= z_1^2 - z_2^2 \\ &\quad + i z_1 z_2 + i \bar{z}_1 \bar{z}_2 \end{aligned}$$

To see why the d_{em} 's are indep $\forall m \geq 0$, let $\ell \neq \ell'$ & $m \neq m'$ both ≥ 0 .

write $d_{\text{em}} = R + iI$ &

$$e'_{\ell'm'} = R' + iI'$$

Now

$$E(RR' + II') + iE(II' - RI')$$

$$= E(d_{\text{em}} \bar{d}_{\ell'm'}) = 0 \quad (6)$$

$$(R+iI)(R'-iI')$$

Also

$$\begin{aligned} E(RR' - II') + iE(IR' + RI') \\ = E(\text{dem dem}') \\ = E(\text{dem}^{(1)} \overline{\text{dem}}^{(1)}) \\ \stackrel{(7)}{=} 0 \quad \text{since } m \neq -m' \text{ because } m, m' \geq 0 \end{aligned}$$

Adding (6) & (7) gives

$$E(RR') = 0 \quad \& \quad E(II') = 0$$

Subtracting (6) & (7) gives

$$E(II') = 0 \quad \& \quad E(RI') = 0$$

$\therefore R, R', I, I'$ are indep.

If we stack the dem's into a vector \vec{d} for all $l=0, 1, \dots, l_{\max}$ & $m=0, 1, \dots, l$ we get a Gaussian likelihood:

$$\begin{aligned} P(\vec{d} | C_e^{\text{dd}}) &= \prod_{l=0}^{l_{\max}} \prod_{m=0}^l P(\text{dem}_l | C_e^{\text{dd}}) \\ &= \prod_{l=0}^{l_{\max}} P(d_{e,0} | C_e^{\text{dd}}) \prod_{m=1}^l P(\text{dem}_l | C_e^{\text{dd}}) \\ &\propto \prod_{l=0}^{l_{\max}} \frac{1}{\sqrt{C_e^{\text{dd}}}} e^{-\frac{1}{2} \frac{|d_{e,0}|^2}{C_e^{\text{dd}}}} \prod_{m=1}^l \frac{1}{\sqrt{C_e^{\text{dd}}}} e^{-\frac{1}{2} \frac{|d_{em}|^2}{C_e^{\text{dd}}}} \\ &\propto \prod_{l=0}^{l_{\max}} \left(C_e^{\text{dd}} \right)^{-\frac{2l+1}{2}} \exp \left(-\frac{1}{2} \frac{|d_{e,0}|^2 + \sum_{m=1}^l |d_{em}|^2}{C_e^{\text{dd}}} \right) \\ &= \prod_{l=0}^{l_{\max}} \left(C_e^{\text{dd}} \right)^{-\frac{2l+1}{2}} \exp \left(-\frac{2l+1}{2} \frac{\sigma_e}{C_e^{\text{dd}}} \right) \end{aligned}$$

(9)

so

(10)

$$P(\vec{d} | C_e^{\text{dd}}) \propto \prod_{l=0}^{l_{\max}} \left(C_e^{\text{dd}} \right)^{-\frac{2l+1}{2}} \exp \left(-\frac{2l+1}{2} \frac{\sigma_e}{C_e^{\text{dd}}} \right)$$

$$\text{where } C_e^{\text{dd}} = C_e^{\text{TT}} + C_e^{\text{EE}}$$

* This also shows that the bandpower σ_e is a sufficient statistic for estimating C_e^{TT} at any fixed l .
 (Note: this derivation was done under the assumption that $\epsilon(n)$ is isotropic generalized noise.)

Bayesian Posterior $P(C_e^{\text{TT}} | \sigma_e)$

Since the likelihood factors over l one can construct Bayesian posteriors for C_e^{TT} given σ_e at each l . In particular one can model the uncertainty in C_e^{TT} with a prior probability distribution $P(C_e^{\text{TT}})$. Following Bayes Rule for swapping the events in a conditional probability calculation we have

$$\begin{aligned} P(C_e^{\text{TT}} | \sigma_e) &= \frac{P(C_e^{\text{TT}}, \sigma_e)}{P(\sigma_e)} \\ &\stackrel{\substack{\text{denotes density} \\ \text{treating } C_e^{\text{TT}} \text{ as random}}}{\longrightarrow} \frac{P(\sigma_e | C_e^{\text{TT}}) P(C_e^{\text{TT}})}{\int P(\sigma_e | C_e^{\text{TT}}) P(C_e^{\text{TT}}) dC_e^{\text{TT}}} \\ &\stackrel{\substack{\text{prior} \\ \text{a normalizing constant w.r.t. } C_e^{\text{TT}}}}{\longrightarrow} \end{aligned}$$

(12)

Jeffreys prior for C_e^{TT} is given by

$$P(C_e^{TT}) \propto \sqrt{E\left(-\frac{d^2}{dC_e^{TT}} \log P(\delta_e/C_e^{TT})\right)}$$

(11)

$$\left. -\frac{2\ell+1}{2} \log (C_e^{TT} + C_e^{EE}) - \frac{2\ell+1}{2} \frac{\sigma_e}{C_e^{TT} + C_e^{EE}} \right)$$

$$\frac{d}{dC_e^{TT}} (\dots) = -\frac{2\ell+1}{2} \frac{1}{C_e^{TT} + C_e^{EE}} + \frac{(2\ell+1)}{(C_e^{TT} + C_e^{EE})^2} \frac{\sigma_e}{(2\ell+1)}$$

$$\frac{d^2}{dC_e^{TT}^2} (\dots) = \frac{(2\ell+1)}{(C_e^{TT} + C_e^{EE})^2} - \frac{2(2\ell+1)}{(C_e^{TT} + C_e^{EE})^3} \frac{\sigma_e}{(2\ell+1)}$$

$$\therefore P(C_e^{TT}) \propto \sqrt{\frac{(2\ell+1)}{(C_e^{TT} + C_e^{EE})^2} + \frac{2(2\ell+1)}{(C_e^{TT} + C_e^{EE})^3} \frac{\sigma_e}{(2\ell+1)}}$$

$$\propto \frac{1}{C_e^{TT} + C_e^{EE}}$$

Resulting a posterior

$$P(C_e^{TT}/\delta_e) \propto (C_e^{TT} + C_e^{EE})^{-\frac{2\ell+1-1}{2}} \exp\left(-\frac{2\ell+1}{2} \frac{\sigma_e}{C_e^{TT} + C_e^{EE}}\right)$$