Title: Cosmology Observations 2

Date: Jul 12, 2018 09:00 AM

URL: http://pirsa.org/18070002

Abstract:

Part 3: forecasting and the Fisher matrix

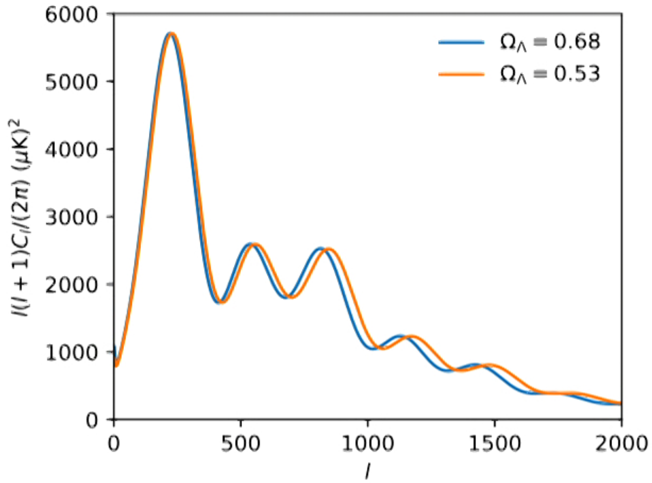
By my estimation, there are ~100 cosmology papers per year which are mainly concerned with computing a "Fisher matrix" F_{ab}.

What is the Fisher matrix and why is it so widespread?

By my estimation, there are ~100 cosmology papers per year which are mainly concerned with computing a “Fisher matrix” F_{ab}.

What is the Fisher matrix and why is it so widespread?

Motivating example: forecasting parameter sensitivity of the CMB



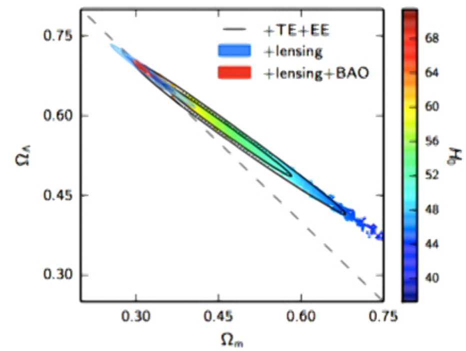


Fig. 26. Constraints in the Ω_m–Ω_Λ plane from the *Planck* TT+lowP data (samples; colour-coded by the value of H_0) and *Planck* TT,TE,EE+lowP (solid contours). The geometric degeneracy between Ω_m and Ω_Λ is partially broken because of the effect of lensing on the temperature and polarization power spectra. These limits are improved significantly by the inclusion of the *Planck* lensing reconstruction (blue contours) and BAO (solid red contours). The red contours tightly constrain the geometry of our Universe to be nearly flat.

Planck 2015

Abstract setup:

"model parameters" $\{\theta^a\}$
e.g. ρ_Λ, Ω_b, Ω_c, Δ_ζ^2, n_s, τ

$p(d|\theta)$ →

"data" d
e.g. CMB multipoles a_{lm}

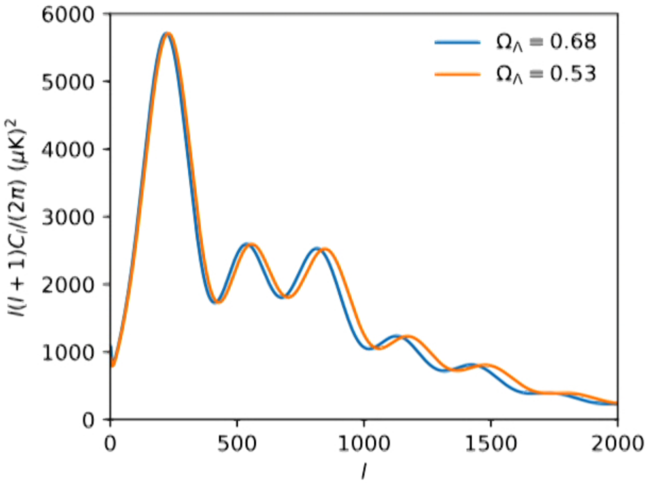
The data d is a random variable whose probability distribution depends on the model parameters θ^a.

$p(d|\theta)$ = conditional probability distribution of data d, given model parameters θ

By my estimation, there are ~100 cosmology papers per year which are mainly concerned with computing a “Fisher matrix” F_{ab}.

What is the Fisher matrix and why is it so widespread?

Motivating example: forecasting parameter sensitivity of the CMB



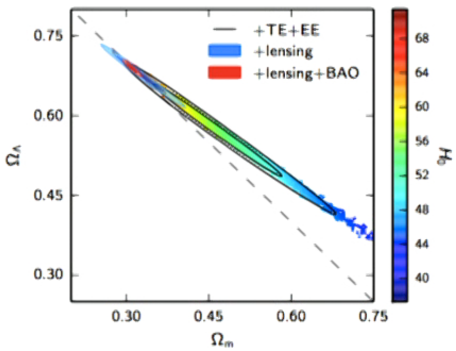


Fig. 26. Constraints in the Ω_m–Ω_Λ plane from the *Planck* TT+lowP data (samples; colour-coded by the value of H_0) and *Planck* TT,TE,EE+lowP (solid contours). The geometric degeneracy between Ω_m and Ω_Λ is partially broken because of the effect of lensing on the temperature and polarization power spectra. These limits are improved significantly by the inclusion of the *Planck* lensing reconstruction (blue contours) and BAO (solid red contours). The red contours tightly constrain the geometry of our Universe to be nearly flat.

Planck 2015

"model parameters" $\{\theta^a\}$

e.g. ρ_Λ, Ω_b, Ω_c, Δ_ζ^2, n_s, τ

$p(d|\theta)$ →

"data" d

e.g. CMB multipoles a_{lm}

We might be interested in:

1. Simulation. Given a model θ^a, how do we simulate a random data realization d? (i.e. sample the conditional PDF $p(d|\theta)$)

2. Analysis. Given a data realization d, what are the constraints (say at 95% CL) on the model space θ^a?

3. Forecasting. Given a rough fiducial guess θ_{fid} for the true model, what constraints on the model space do we expect to obtain, for a "typical" realization of the data?

The Fisher matrix is a tool for forecasting (#3).

"model parameters" $\{\theta^a\}$
e.g. ρ_Λ, Ω_b, Ω_c, Δ_ζ^2, n_s, τ

$p(d|\theta)$ →

"data" d
e.g. CMB multipoles a_{lm}

General definition of the Fisher matrix (to be motivated later!)

$$F_{ab} = - \left\langle \frac{\partial^2 \log p(d|\theta)}{\partial\theta^a \, \partial\theta^b} \right\rangle_d$$

where the expectation value is taken over random realizations of the data d, for a preferred fiducial choice of model parameters θ_{fid}

The Fisher matrix depends on the fiducial model θ_{fid}, but does not require a preferred realization of the data d (just the probability distribution $p(d|\theta)$).

"model parameters" $\{\theta^a\}$

e.g. ρ_Λ, Ω_b, Ω_c, Δ_ζ^2, n_s, τ

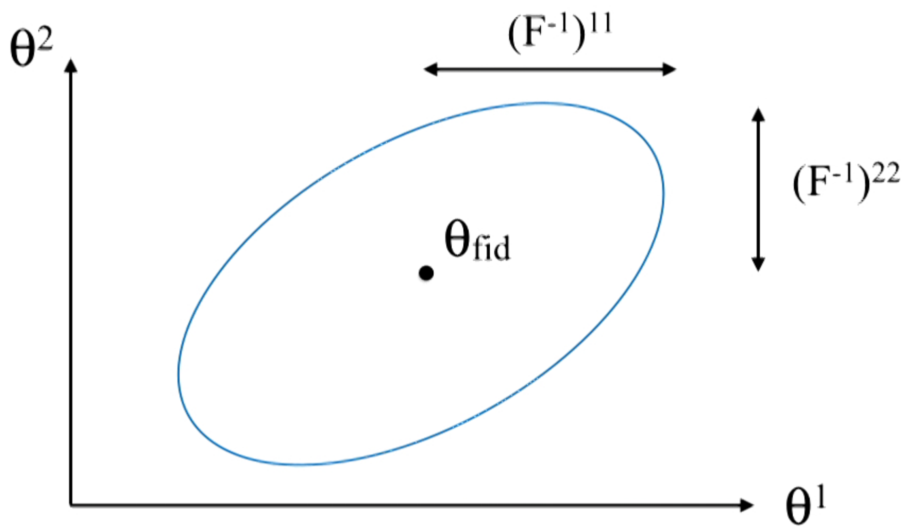
$p(d|\theta)$ →

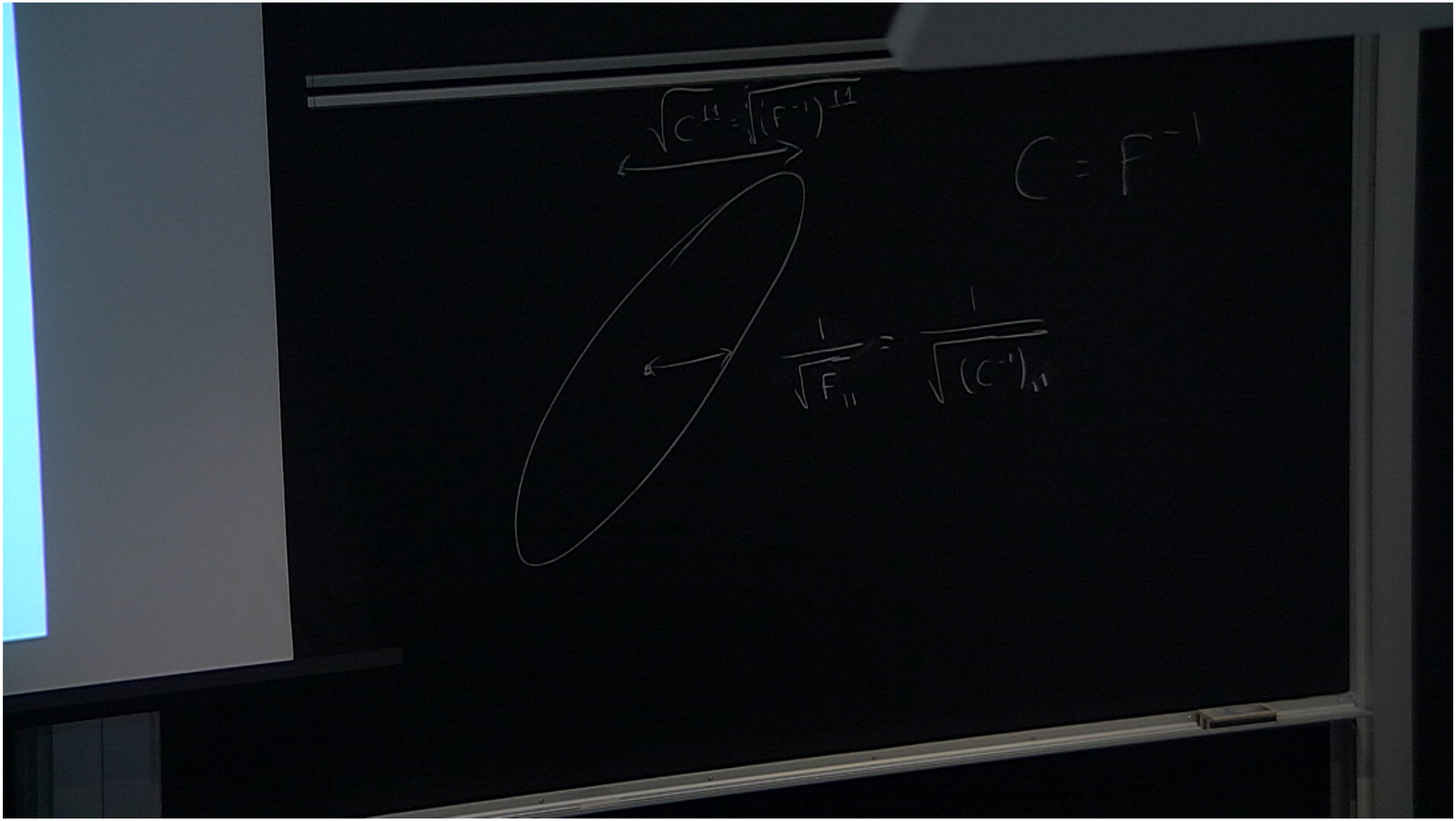
"data" d

e.g. CMB multipoles a_{lm}

Interpretation: the Fisher matrix F_{ab} is the forecasted *inverse* covariance matrix of the model constraints obtained from a "typical" realization of the data.

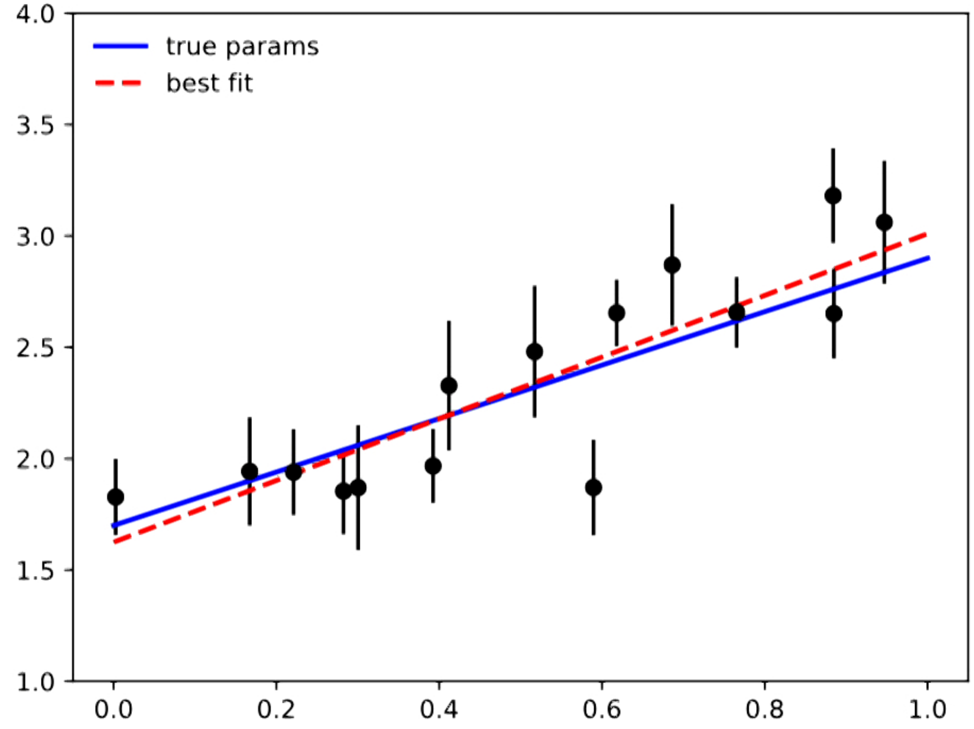
$$F_{ab} = -\left\langle \frac{\partial^2 \log p(d|\theta)}{\partial\theta^a\, \partial\theta^b} \right\rangle_d$$



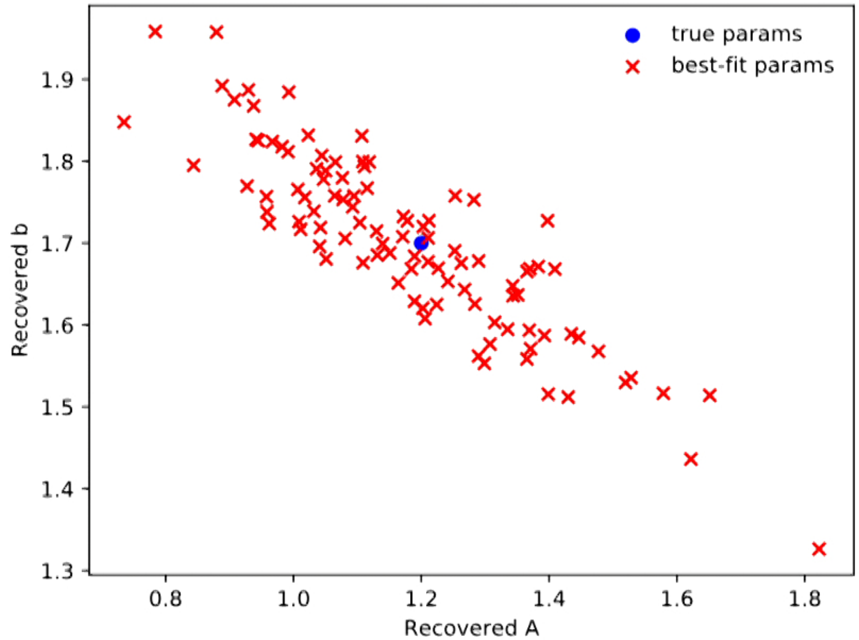

√C^11 = √(F^-1)^11
C = F^-1
1/√F_11 = 1/√(C^-1)_11

Toy example: linear regression

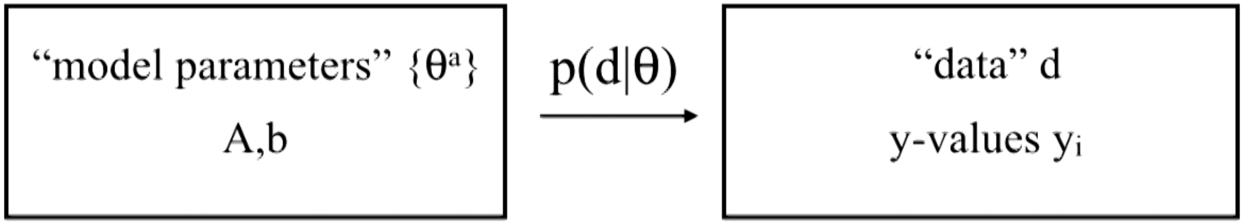
Fitting a line $y=Ax+b$ through points (x_i,y_i) with error bars σ_i.



The following scatterplot shows the result of repeating the linear regression 100 times. (Detail: the y-values were randomized, but x_i and σ_i were held fixed throughout.)



Let's compute the Fisher matrix for the model parameters $(\theta^1,\theta^2) = (A,b)$, and compare with this plot.



First we need to write down the conditional likelihood $p(d|\theta)$.
Given $(\theta^1,\theta^2) = (A,b)$, the PDF of an *individual* data value y_i is a Gaussian with mean (Ax_i+b) and variance σ_i^2:

$$\begin{aligned} p(y_i|\theta) &= \frac{1}{(2\pi\sigma_i^2)^{1/2}} \exp\left(-\frac{(y_i - Ax_i - b)^2}{2\sigma_i^2}\right) \\ &= \frac{1}{(2\pi\sigma_i^2)^{1/2}} \exp\left(-\frac{(y_i - \theta^a f_a(x_i))^2}{2\sigma_i^2}\right) \end{aligned}$$

introducing the two-vector notation $f_a(x) = \begin{pmatrix} x \\ 1 \end{pmatrix}$

"model parameters" $\{\theta^a\}$

A,b

$p(d|\theta)$ →

"data" d

y-values y_i

The conditional likelihood $p(d|\theta)$ is obtained by multiplying the PDF's for each individual data value y_i.

$$p(d|\theta) = \prod_i \frac{1}{(2\pi\sigma_i^2)^{1/2}} \exp\left(-\frac{(y_i - \theta^a f_a(x_i))^2}{2\sigma_i^2}\right)$$

$$f_a(x) = \begin{pmatrix} x \\ 1 \end{pmatrix}$$

Now we can compute the Fisher matrix $F_{ab} = -\left\langle \frac{\partial^2 \log p(d|\theta)}{\partial\theta^a\,\partial\theta^b} \right\rangle_d$

$$p(d|\theta) = \prod_i \frac{1}{(2\pi\sigma_i^2)^{1/2}} \exp\left(-\frac{(y_i - \theta^a f_a(x_i))^2}{2\sigma_i^2}\right)$$

$$\Rightarrow F_{ab} = -\left\langle \frac{\partial^2 \log p(d|\theta)}{\partial\theta^a\,\partial\theta^b} \right\rangle_d$$

$$= -\left\langle \frac{\partial^2}{\partial\theta^a\,\partial\theta^b} \sum_i \left(-\frac{1}{2}\log(2\pi\sigma_i^2) - \frac{1}{2\sigma_i^2}(y_i - \theta^c f_c(x_i))^2\right) \right\rangle_y$$

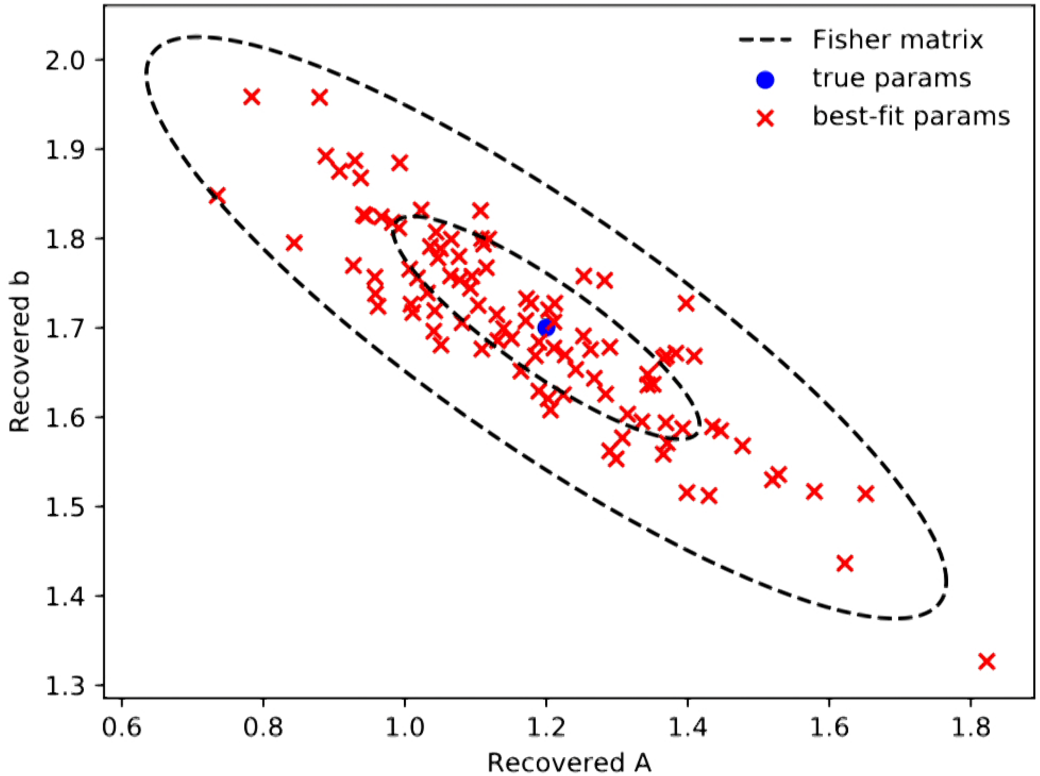
$$= -\left\langle \frac{\partial^2}{\partial\theta^a\,\partial\theta^b} \sum_i \left(-\frac{1}{2\sigma_i^2}(\theta^c f_c(x_i))(\theta^d f_d(x_i))\right) \right\rangle_y$$

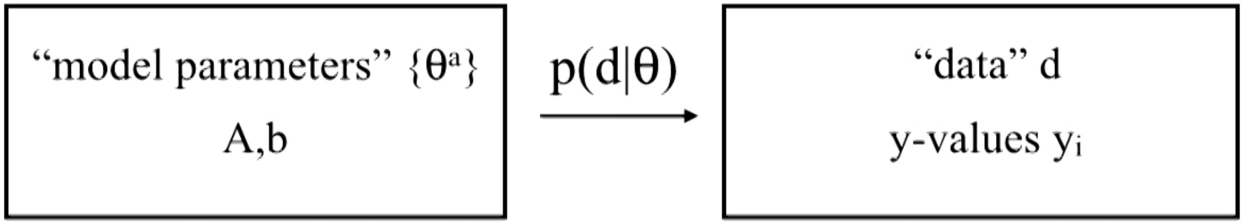
$$= \left\langle \sum_i \left(\frac{f_a(x_i) f_b(x_i)}{\sigma_i^2}\right) \right\rangle_y$$

$$= \sum_i \frac{f_a(x_i) f_b(x_i)}{\sigma_i^2}$$

In this example, the expectation value $\langle\,\rangle_y$ is trivial, and the Fisher matrix does not depend on a choice of fiducial model $\theta_{\text{fid}} = (A_{\text{fid}}, b_{\text{fid}})$.

Comparison between Fisher matrix and Monte Carlo scatterplot.



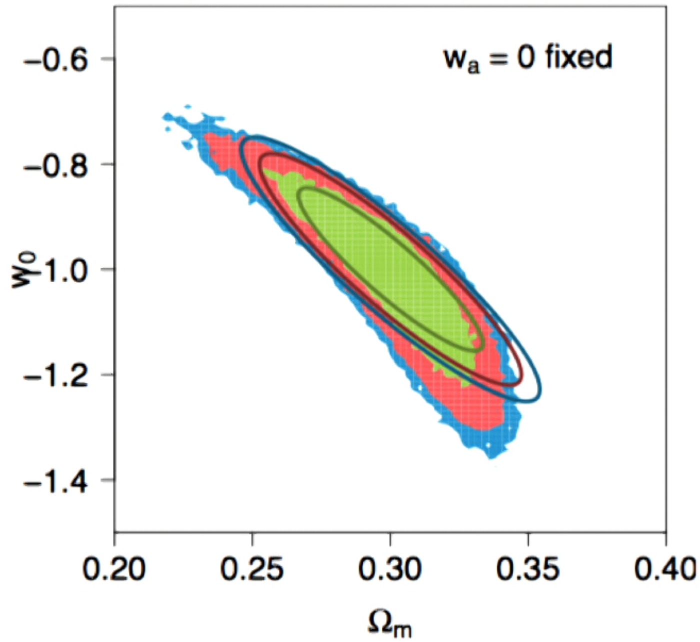


First we need to write down the conditional likelihood $p(d|\theta)$.
Given $(\theta^1,\theta^2) = (A,b)$, the PDF of an *individual* data value y_i is a Gaussian with mean (Ax_i+b) and variance σ_i^2:

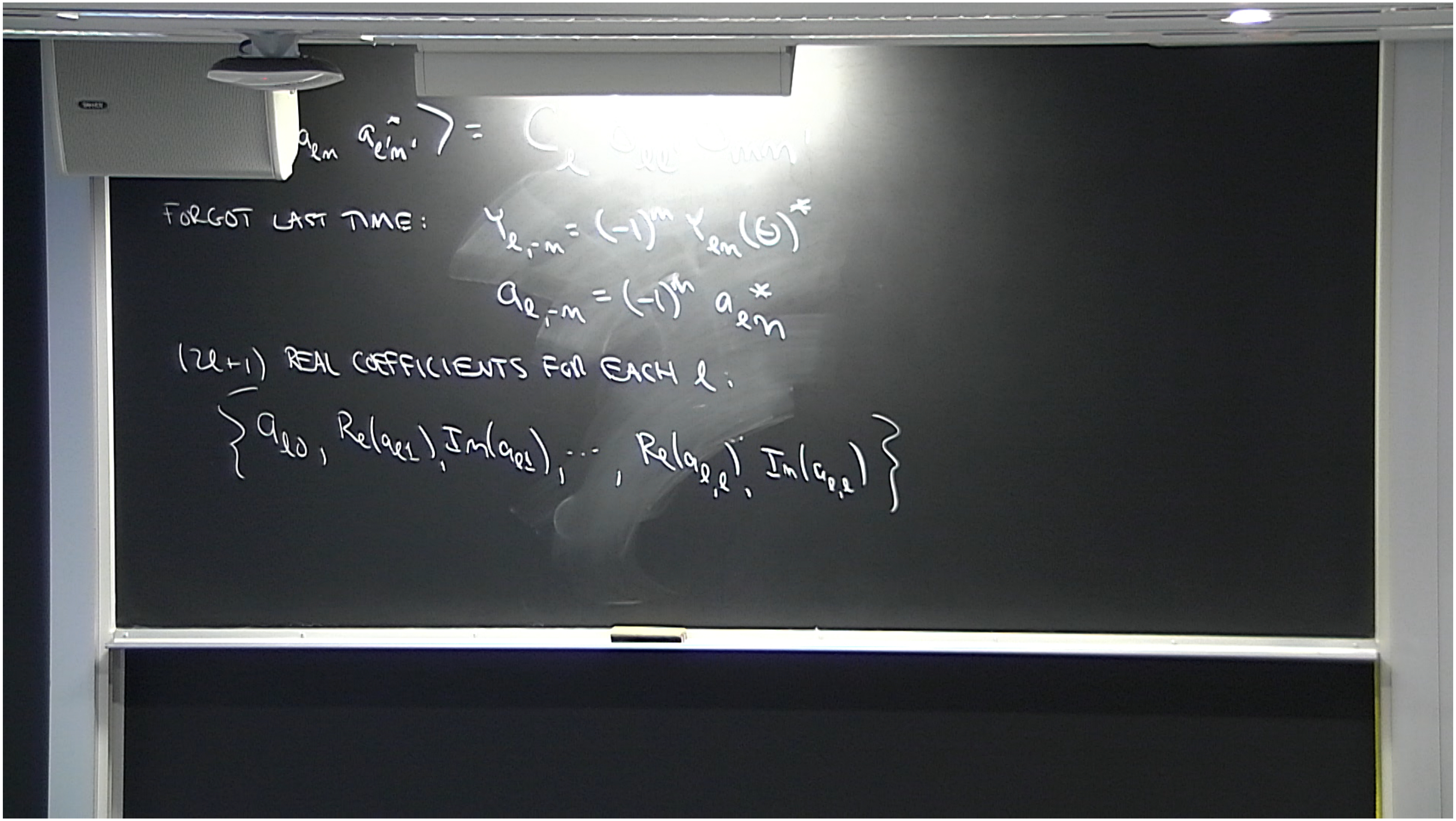
$$p(y_i|\theta) = \frac{1}{(2\pi\sigma_i^2)^{1/2}} \exp\left(-\frac{(y_i - Ax_i - b)^2}{2\sigma_i^2}\right)$$
$$= \frac{1}{(2\pi\sigma_i^2)^{1/2}} \exp\left(-\frac{(y_i - \theta^a f_a(x_i))^2}{2\sigma_i^2}\right)$$

introducing the two-vector notation $f_a(x) = \begin{pmatrix} x \\ 1 \end{pmatrix}$

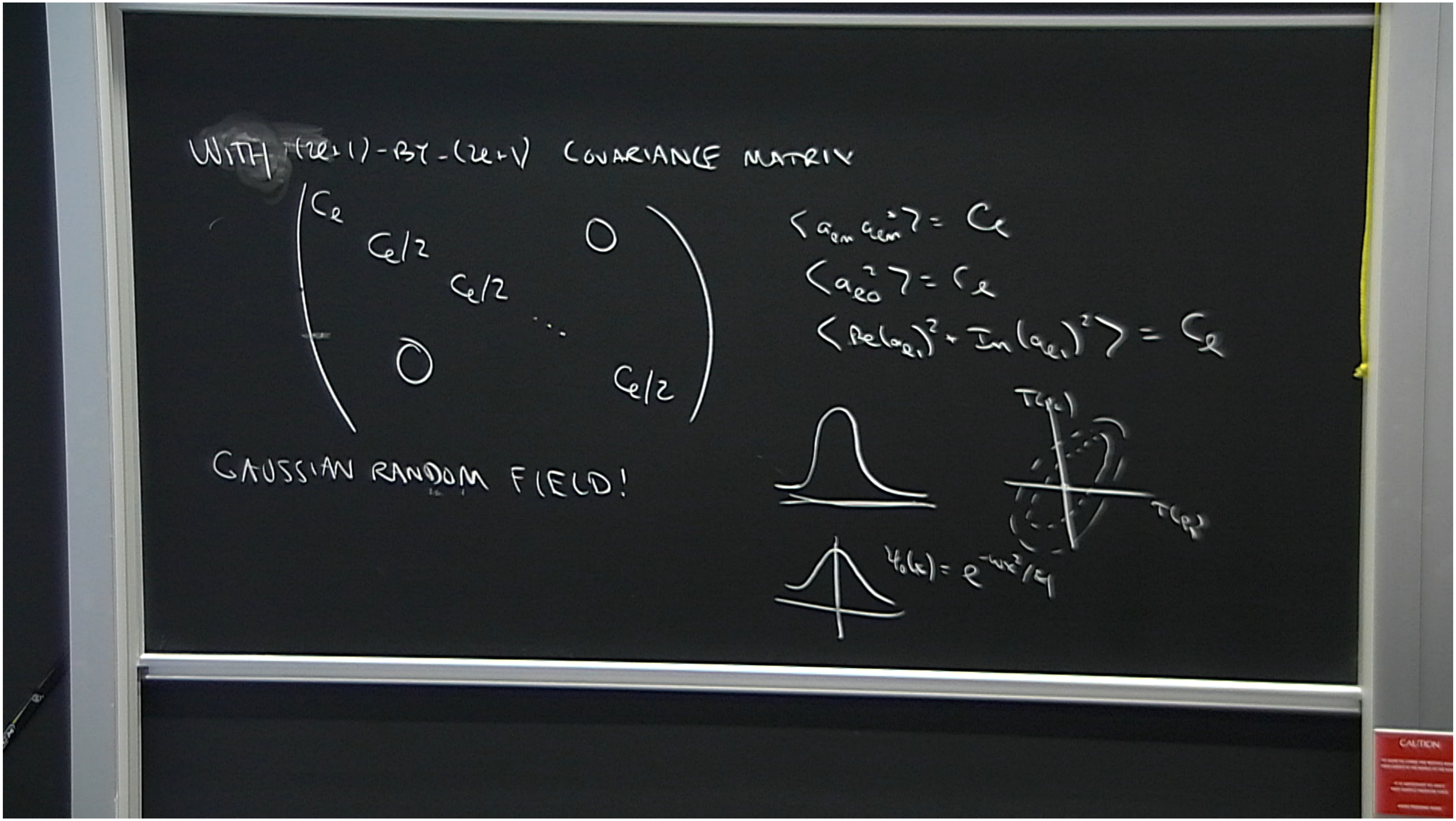
A cosmological example. In general, the Fisher matrix is not expected to agree precisely with the Monte Carlo scatterplot. In particular, the contours of the scatterplot need not be ellipses. However, the Fisher matrix is usually easy to compute, and is usually a good approximation.

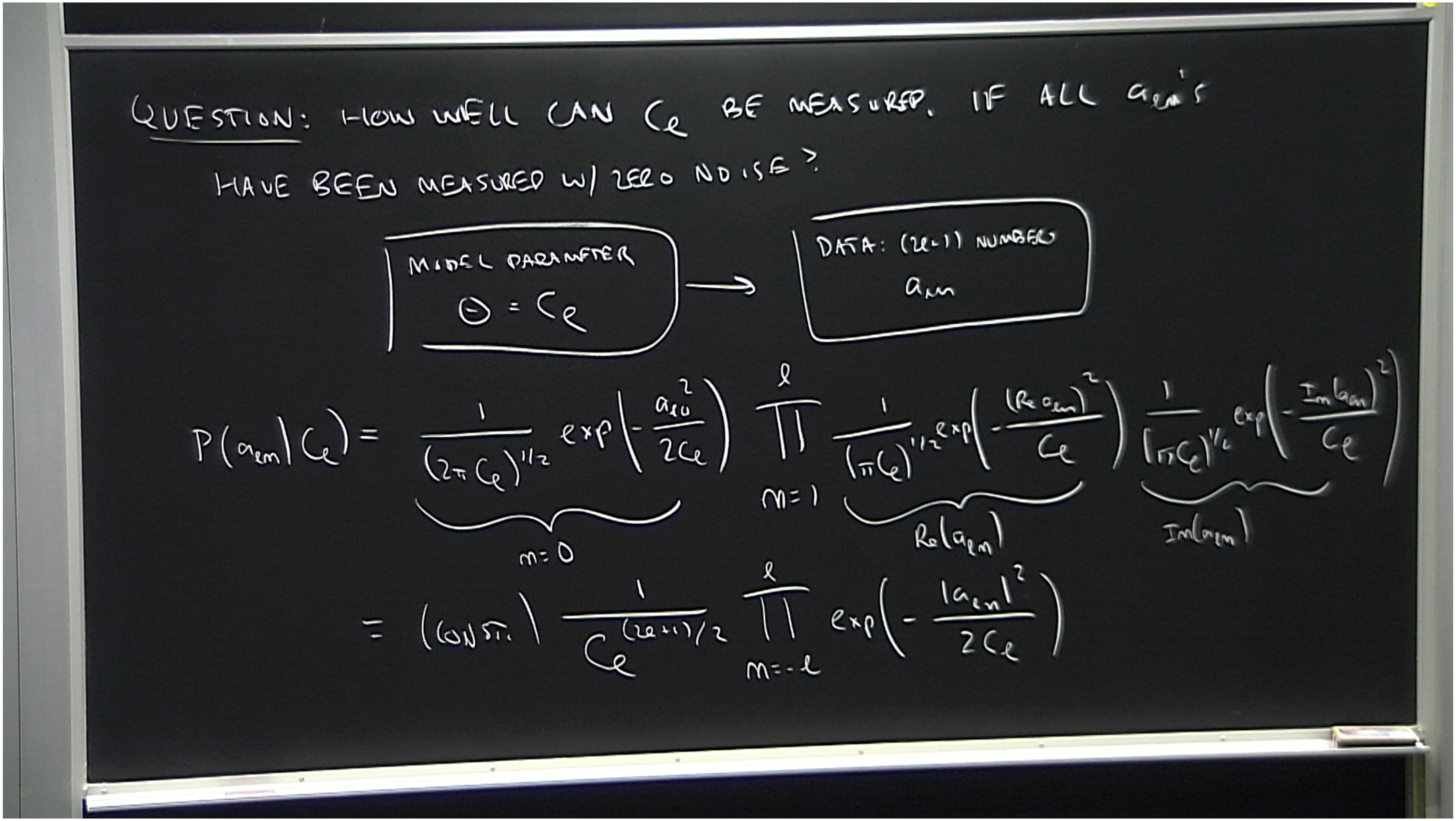


Example: supernova (not CMB) Fisher matrix from Wolz et al 1205.3984



$\langle a_{\ell m} a^*_{\ell' m'} \rangle = C_\ell \delta_{\ell\ell'} \delta_{mm'}$
FORGOT LAST TIME: $Y_{\ell,-m} = (-1)^m Y_{\ell m}(\theta)^*$
$a_{\ell,-m} = (-1)^m a^*_{\ell m}$
$(2\ell+1)$ REAL COEFFICIENTS FOR EACH ℓ:
$\{ a_{\ell 0}, \mathrm{Re}(a_{\ell 1}), \mathrm{Im}(a_{\ell 1}), \dots, \mathrm{Re}(a_{\ell,\ell}), \mathrm{Im}(a_{\ell,\ell}) \}$


WITH (2ℓ+1)-BY-(2ℓ+1) COVARIANCE MATRIX
C_ℓ $C_\ell/2$ $C_\ell/2$ ... $C_\ell/2$ 0 0
$\langle a_{\ell m} a_{\ell m}^* \rangle = C_\ell$
$\langle a_{\ell 0}^2 \rangle = C_\ell$
$\langle \mathrm{Re}(a_{\ell 1})^2 + \mathrm{Im}(a_{\ell 1})^2 \rangle = C_\ell$
GAUSSIAN RANDOM FIELD!
$T(\hat{n}_1)$
$T(\hat{n}_2)$
$\psi_0(x) = e^{-\omega x^2/\hbar}$


QUESTION: HOW WELL CAN C_ℓ BE MEASURED, IF ALL $a_{\ell m}$'s HAVE BEEN MEASURED W/ ZERO NOISE?
MODEL PARAMETER
$\Theta = C_\ell$
DATA: $(2\ell+1)$ NUMBERS
$a_{\ell m}$
$P(a_{\ell m} | C_\ell) = \frac{1}{(2\pi C_\ell)^{1/2}} \exp\left(-\frac{a_{\ell 0}^2}{2C_\ell}\right) \prod_{m=1}^{\ell} \frac{1}{(\pi C_\ell)^{1/2}} \exp\left(-\frac{(\mathrm{Re}\, a_{\ell m})^2}{C_\ell}\right) \frac{1}{(\pi C_\ell)^{1/2}} \exp\left(-\frac{\mathrm{Im}(a_{\ell m})^2}{C_\ell}\right)$
$m=0$
$\mathrm{Re}(a_{\ell m})$
$\mathrm{Im}(a_{\ell m})$
$= (\mathrm{const.}) \frac{1}{C_\ell^{(2\ell+1)/2}} \prod_{m=-\ell}^{\ell} \exp\left(-\frac{|a_{\ell m}|^2}{2C_\ell}\right)$

1-BY-1 FISHER "MATRIX"

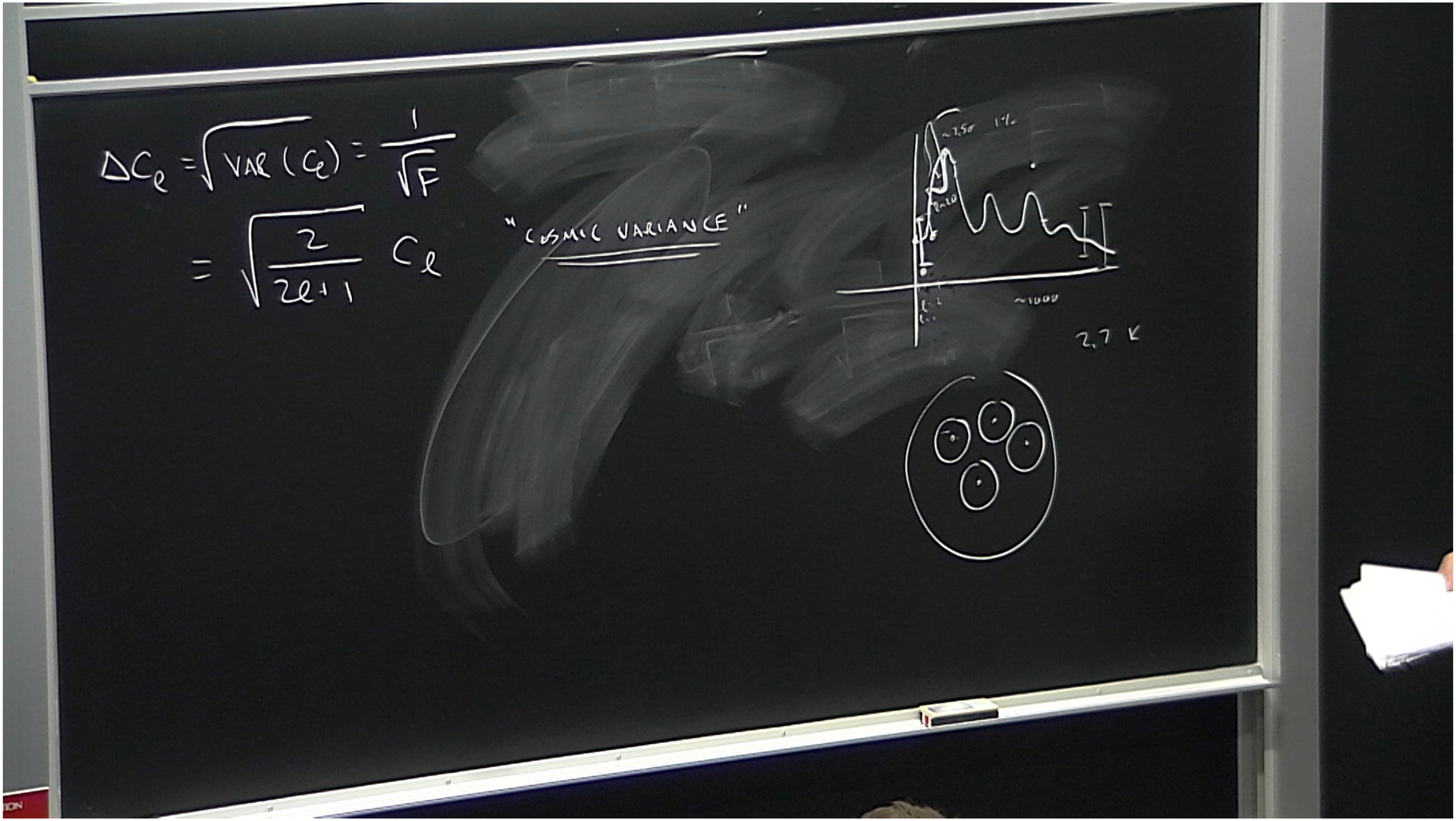
$$F = -\left\langle \frac{\partial^2 \log P(a_{\ell m} | C_\ell)}{\partial C_\ell \, \partial C_\ell} \right\rangle_{a_{\ell m}}$$

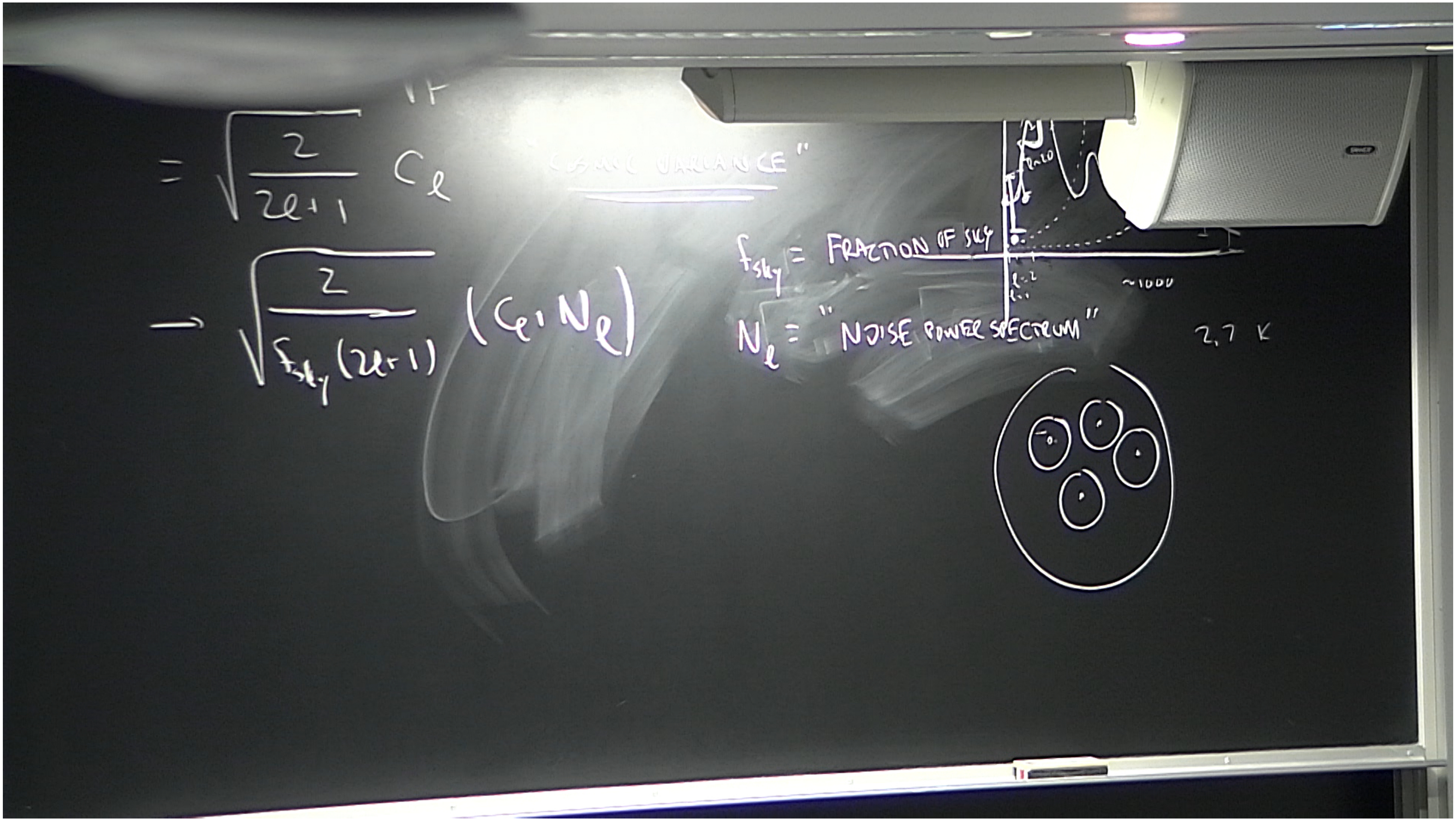
$$= -\left\langle \frac{\partial^2}{\partial C_\ell \, \partial C_\ell} \left(-\frac{2\ell+1}{2} \log C_\ell - \sum_{m=-\ell}^{\ell} \frac{|a_{\ell m}|^2}{2 C_\ell} \right) \right\rangle_{a_{\ell m}}$$

$$= -\left\langle \frac{2\ell+1}{2} \frac{1}{C_\ell^2} - \sum_{m=-\ell}^{\ell} \frac{|a_{\ell m}|^2}{C_\ell^3} \right\rangle_{a_{\ell m}}$$

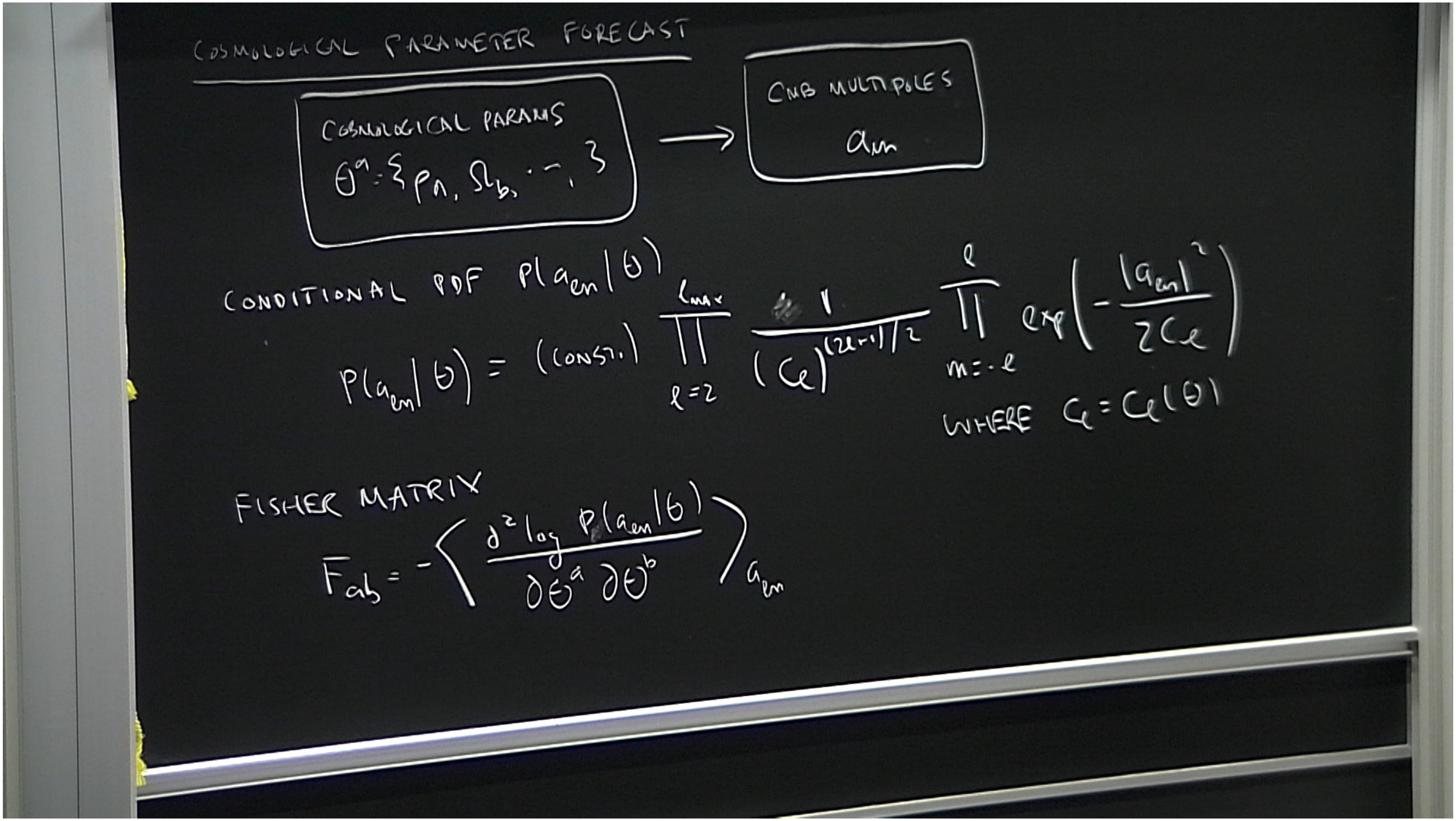
$$= -\left(\frac{2\ell+1}{2} \frac{1}{C_\ell^2} - (2\ell+1) \frac{1}{C_\ell^2} \right)$$

$$= \frac{2\ell+1}{2} \frac{1}{C_\ell^2}$$


$\Delta C_\ell = \sqrt{VAR(C_\ell)} = \frac{1}{\sqrt{F}}$
$= \sqrt{\frac{2}{2\ell+1}}\, C_\ell$
"COSMIC VARIANCE"
~250
1%
~1000
2.7 K



$= \sqrt{\frac{2}{2\ell+1}} \, C_\ell$
"COSMIC VARIANCE"
$\rightarrow \sqrt{\frac{2}{f_{sky}(2\ell+1)}} \, (C_\ell + N_\ell)$
f_{sky} = FRACTION OF SKY
N_ℓ = "NOISE POWER SPECTRUM"
~1000
2.7 K


COSMOLOGICAL PARAMETER FORECAST
COSMOLOGICAL PARAMS
$\Theta^a = \{\Omega_\Lambda, \Omega_b, \ldots, \}$
CMB MULTIPOLES
a_{lm}
CONDITIONAL PDF $P(a_{lm}|\Theta)$
$P(a_{lm}|\Theta) = (\text{CONST.}) \prod_{l=2}^{l_{max}} \frac{1}{(C_l)^{(2l+1)/2}} \prod_{m=-l}^{l} \exp\left(-\frac{|a_{lm}|^2}{2C_l}\right)$
WHERE $C_l = C_l(\Theta)$
FISHER MATRIX
$F_{ab} = -\left\langle \frac{\partial^2 \log P(a_{lm}|\Theta)}{\partial\Theta^a \partial\Theta^b} \right\rangle_{a_{lm}}$

$$F_{ab} = \sum_{\ell=2}^{\ell_{max}} \frac{2\ell+1}{2} \frac{1}{C_\ell^2} \frac{\partial C_\ell}{\partial \theta^a} \frac{\partial C_\ell}{\partial \theta^b}$$

$$\rightarrow f_{sky} \sum_{\ell} \frac{2\ell+1}{2} \frac{1}{(C_\ell + N_\ell)^2} \frac{\partial C_\ell}{\partial \theta^a} \frac{\partial C_\ell}{\partial \theta^b}$$

IN PRACTICE, C_ℓ AND $\frac{\partial C_\ell}{\partial \theta^a}$ ARE COMPUTED NUMERICALLY

$\theta \rightarrow$ [CAMB / CLASS] $\longrightarrow C_\ell$

$$\frac{\partial C_\ell}{\partial \theta^a} \approx \frac{C_\ell(\theta + \Delta\theta^a) - C_\ell(\theta - \Delta\theta^a)}{2(\Delta\theta^a)}$$

COSMOLOGICAL PARAMETER FORECAST

COSMOLOGICAL PARAMS $\Theta^a = \{ \rho_\Lambda, \Omega_b, \cdots, f_{NL} \}$ → C_ℓ → $a_{\ell m}$

$\langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3} \rangle \sim$

CONDITIONAL PDF $p(a_{\ell m} | \Theta)$

$$p(a_{\ell m} | \Theta) = (\text{CONST.}) \prod_{\ell=2}^{\ell_{max}} \frac{1}{(C_\ell)^{(2\ell+1)/2}} \prod_{m=-\ell}^{\ell} \exp\left(- \frac{|a_{\ell m}|^2}{2 C_\ell} \right)$$

WHERE $C_\ell = C_\ell(\Theta)$

FISHER MATRIX

$$F_{ab} = - \left\langle \frac{\partial^2 \log p(a_{\ell m} | \Theta)}{\partial \Theta^a \, \partial \Theta^b} \right\rangle_{a_{\ell m}}$$