

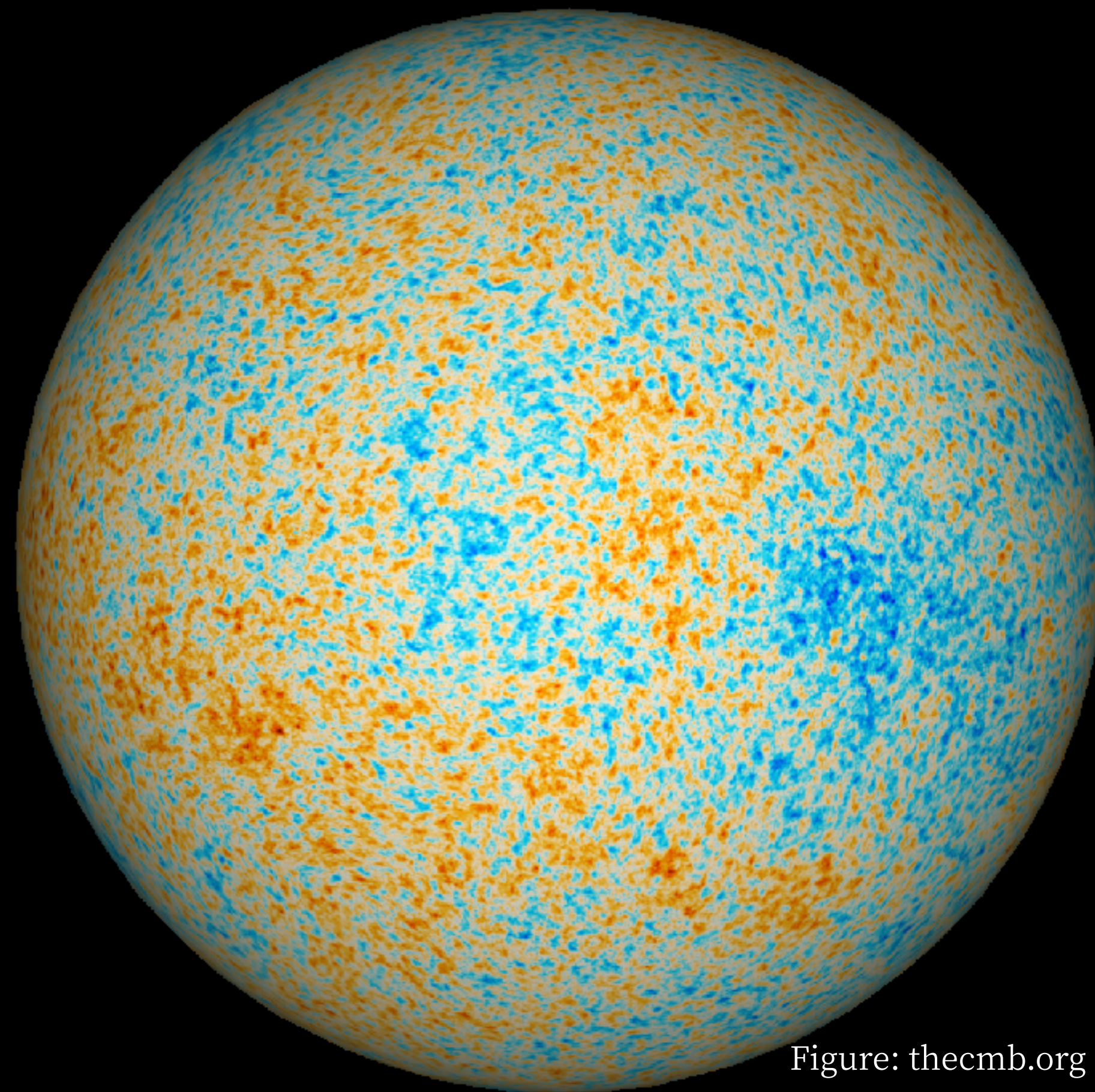


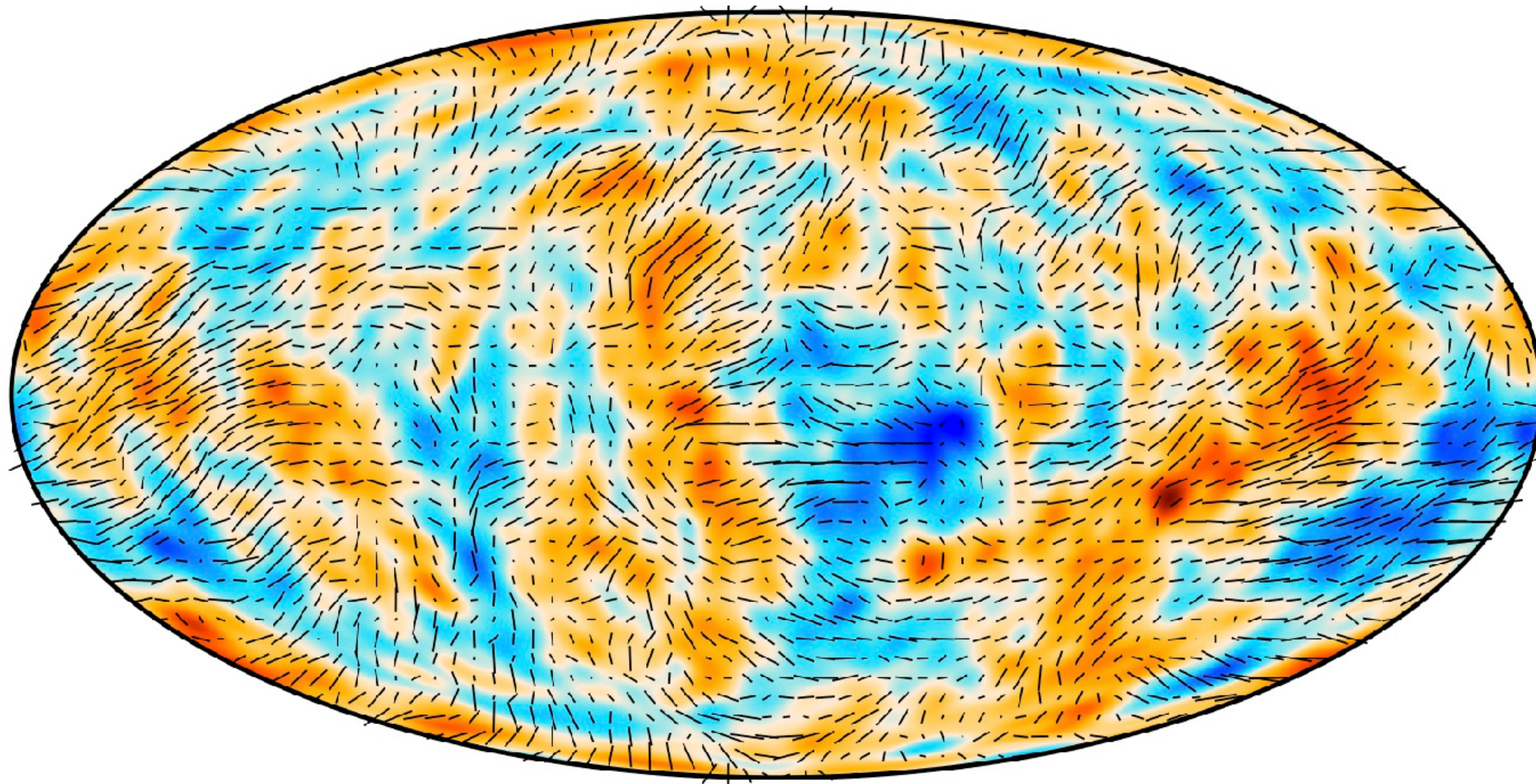
Figure: thecmb.org

MCMC와 우주론 모수 측정의 원리

정동희

(KIAS 고등과학원 / Penn State 펜실베니아 주립대학교)

Planck 우주배경복사 지도 (2018)



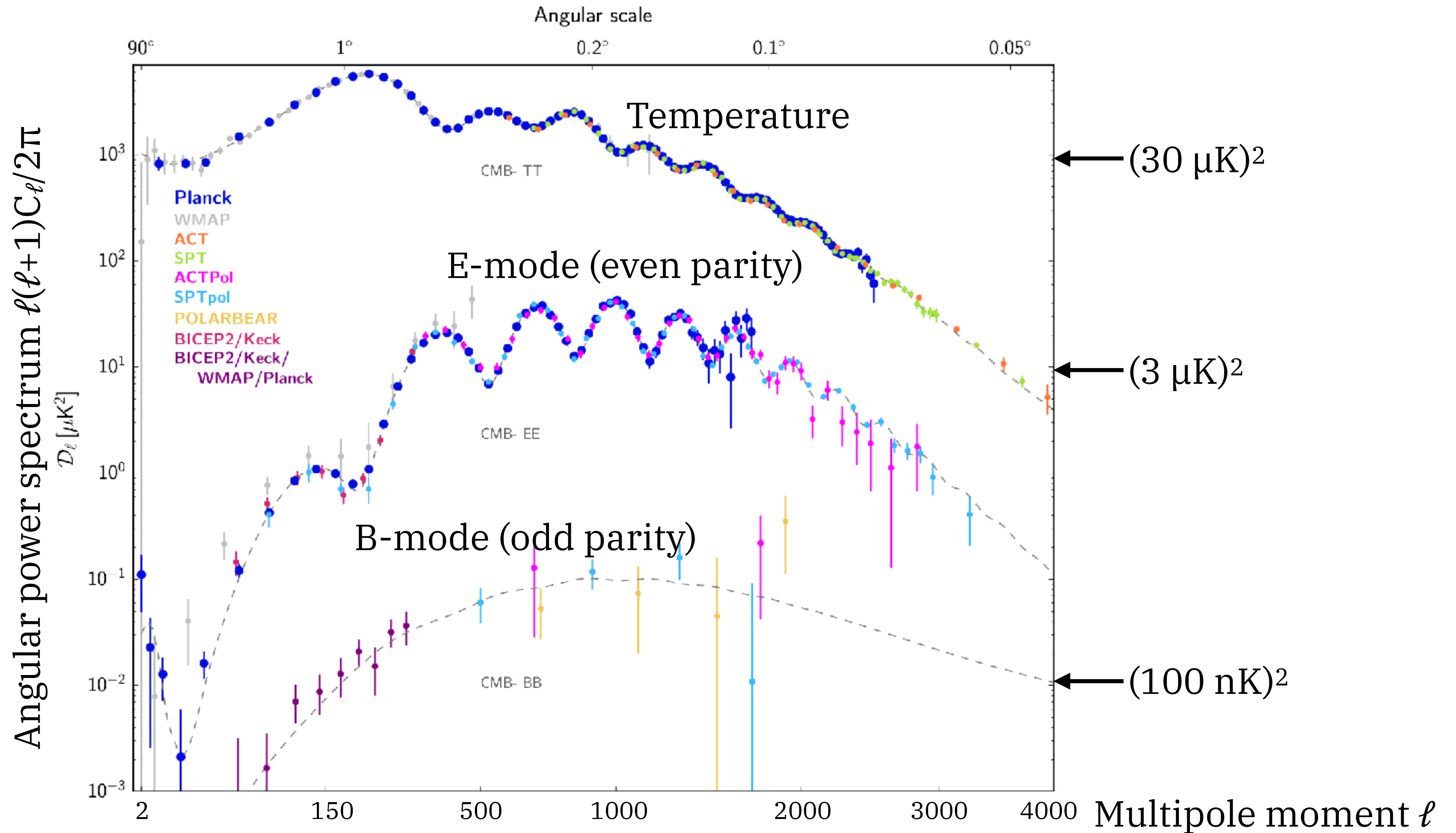
| 0.41 μK

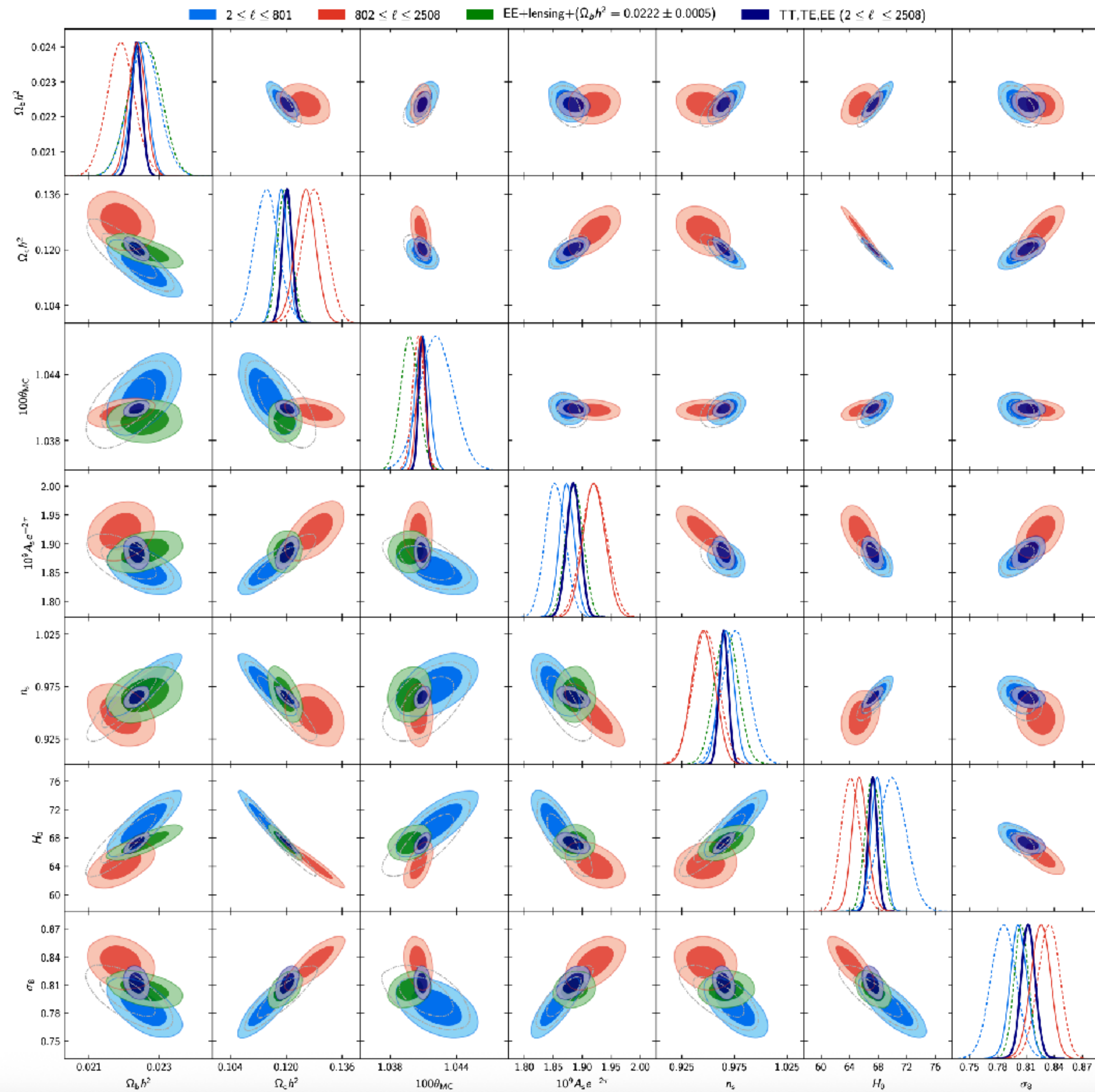
-160

160 μK

5° 필터로 스무딩한 Planck 온도 (색상) + 편광 (선) 지도

파워 스펙트럼 = 섭동의 크기

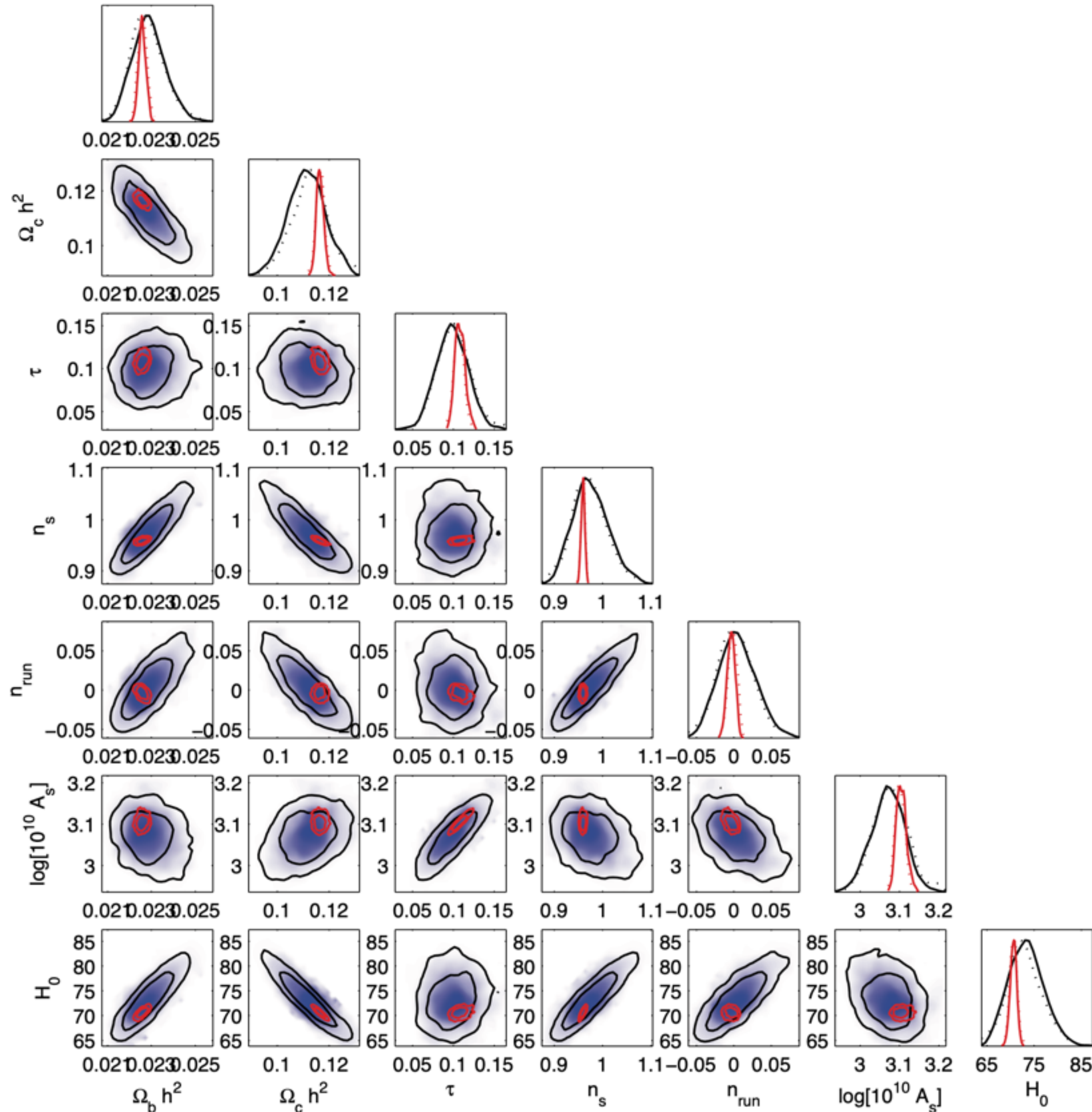




Planck 결과

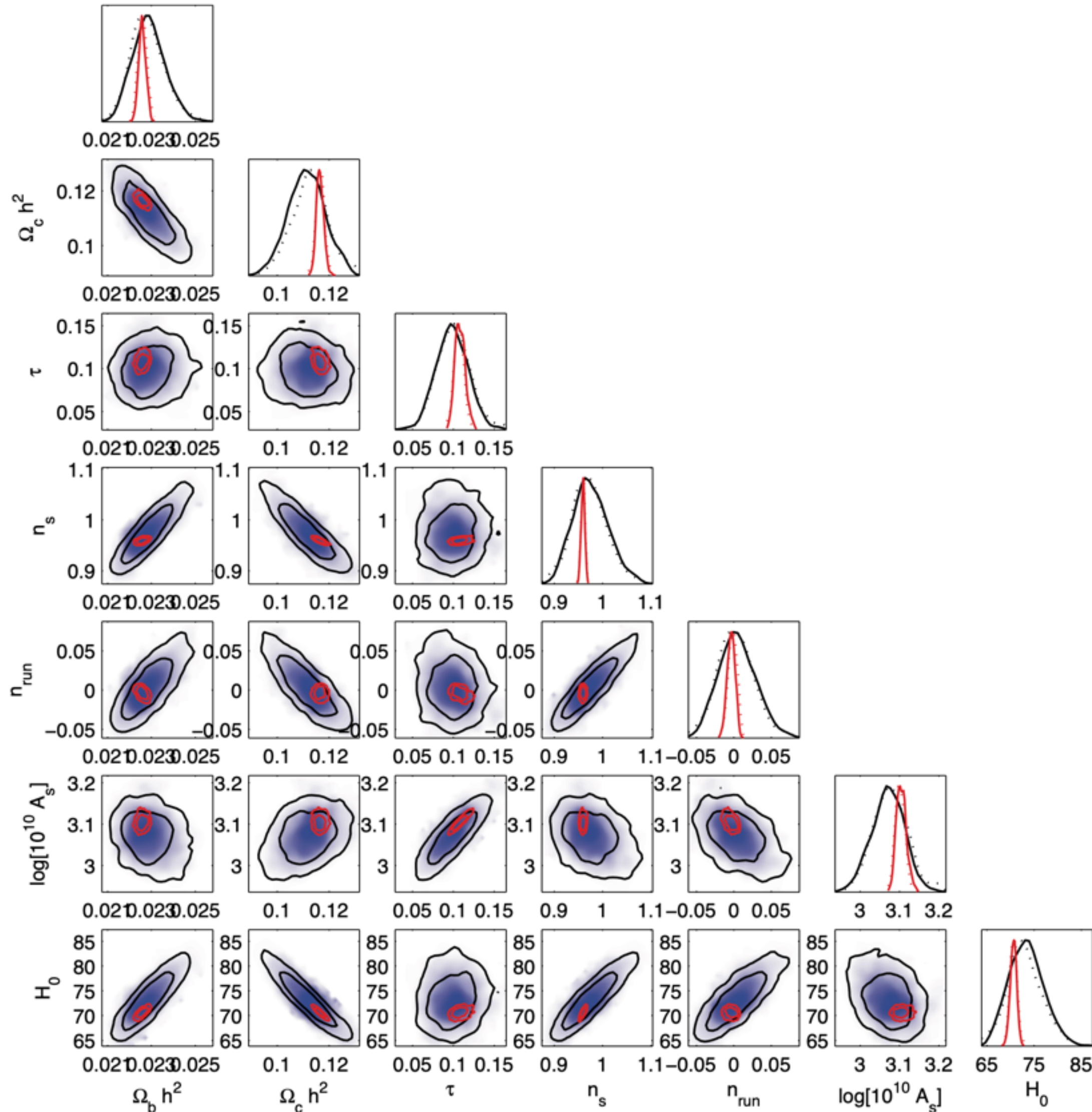
- Planck 2018: 우주론 파라미터 (1807.06209)
- Corner plot: $\{\Omega_b h^2, \Omega_c h^2, \theta_{MC}, \tau, n_s, H_0, \sigma_8\}$
- 좌하: T (대각: 끊김선)
- 우상: TT+TE+EE+lowE (대각: 실선)

Planck 청서 (blue book, 2006)



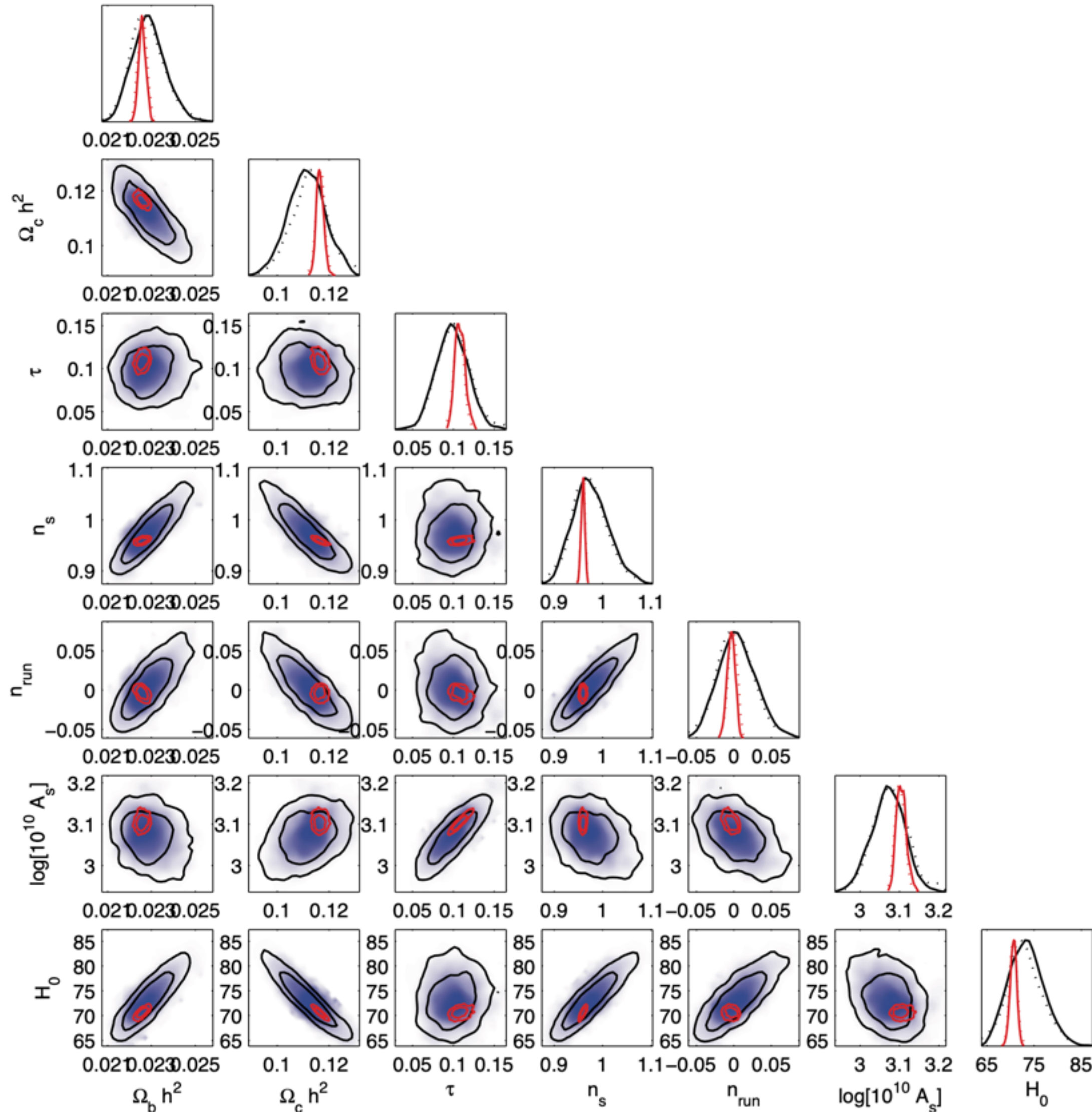
- 푸른색 + 검정 contour:
WMAP 4yr 예측치
- 붉은 contour:
Planck 1yr 예측치

Planck 청서 (blue book, 2006)



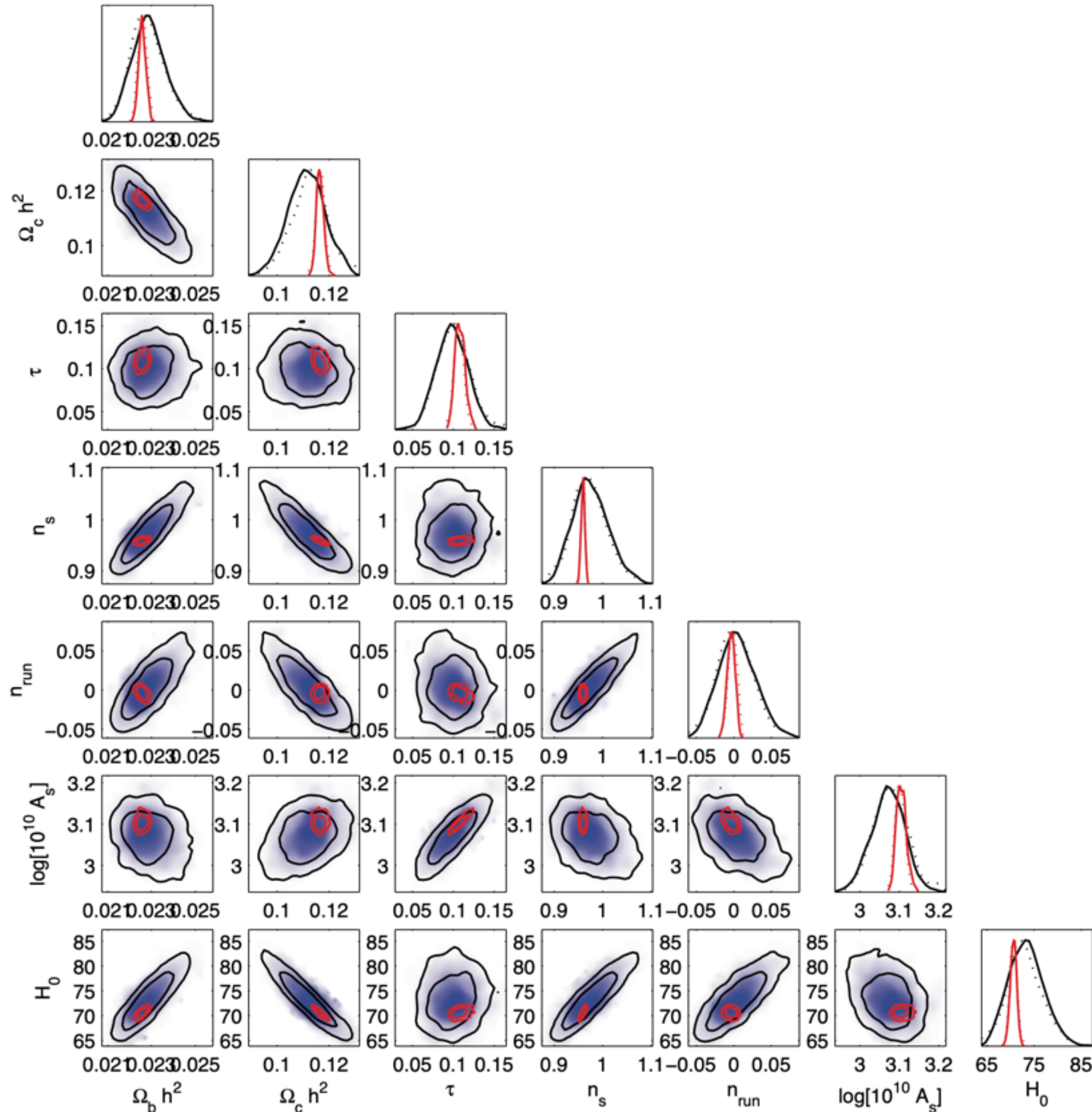
- 푸른색 + 검정 contour:
WMAP 4yr 예측치
- 붉은 contour:
Planck 1yr 예측치
- [질문 1] 어떻게 관측을 시작도 하지 않았는데, 파라미터 측정의 정확도를 예측 할 수 있을까?

Planck 청서 (blue book, 2006)



- 푸른색 + 검정 contour:
WMAP 4yr 예측치
- 붉은 contour:
Planck 1yr 예측치
- [질문 1] 어떻게 관측을 시작도 하지 않았는데, 파라미터 측정의 정확도를 예측 할 수 있을까?
- [질문 2] 저 contour들은 어떤 의미를 갖는가?

Planck 청서 (blue book, 2006)



- 푸른색 + 검정 contour:
WMAP 4yr 예측치
- 붉은 contour:
Planck 1yr 예측치
- [질문 1] 어떻게 관측을 시작도 하지 않았는데, 파라미터 측정의 정확도를 예측 할 수 있을까?
- [질문 2] 저 contour들은 어떤 의미를 갖는가?
- [질문 3] 2차원 contour와 1차원 그래프 사이에는 어떤 관계가 있는가?

천문 관측과 확률

- 왜 확률적인가?

천문 관측과 확률

- 왜 확률적인가?
 - 기본입자 (전자, 광자)를 이용한 관측기기의 확률적 특성
 - CCD/CMOS: 광전자 개수, MKID: 준입자 생성 (양자 효율이 1보다 작은 확률과정)
 - 입자 갯수 하나하나가 소중: X-선, 감마선, 중성미자

천문 관측과 확률

- 왜 확률적인가?
 - 기본입자 (전자, 광자)를 이용한 관측기기의 확률적 특성
 - CCD/CMOS: 광전자 개수, MKID: 준입자 생성 (양자 효율이 1보다 작은 확률과정)
 - 입자 갯수 하나하나가 소중: X-선, 감마선, 중성미자
 - 천문학의 많은 현상들은 확률과정 (Stochastic process):
 - 별/은하 생성: 난류, 자기장 등의 sub-grid physics, feedback 등
 - 빛의 전파 또한 경로상에서 방출, 흡수, 산란등을 반복하는 확률과정

우주론 관측과 확률

- 왜 확률적인가?

우주론 관측과 확률

- 왜 확률적인가?
 - 관측가능한 우주는 딱 하나.
 - 우리가 관측한 우주배경복사의 비등방성은 약 450억광년 반지름 천구에 딱 하나.
 - 각각의 적색이동 (또는 우주 시간)에 딱 정해진 하나의 은하분포.
 - 이 한계가 바로 Cosmic Variance.

우주론 관측과 확률

- 왜 확률적인가?
 - 관측가능한 우주는 딱 하나.
 - 우리가 관측한 우주배경복사의 비등방성은 약 450억광년 반지름 천구에 딱 하나.
 - 각각의 적색이동 (또는 우주 시간)에 딱 정해진 하나의 은하분포.
 - 이 한계가 바로 Cosmic Variance.
 - 우주배경복사의 관측결과, 우주거대구조를 만든 초기우주 밀도섭동의 분포는 정규분포로 매우 정확하게 근사할 수 있음
 - 중심극한정리로 가정한 것이 아니고, 3점함수를 통해 측정해서 나온 결론.

얼마나 잘 정규분포를 따르나?

- 국소(Local) 비정규성의 경우:

$$\Phi = \phi_G + f_{\text{NL}}^{\text{local}}(\phi_G^2 - \langle \phi_G^2 \rangle) + g_{\text{NL}}^{\text{local}} \phi_G^3 + \dots$$

- $\phi_G \simeq \Delta T/T \simeq 10^{-5}$ 이기 때문에, 95% C.L. 최대치를 잡을 경우에
 - 2차항은 1차항에 비해 0.01%,
 - 3차항은 1차항에 비해 10^{-10} 정도의 작은 차이만 줌.

		SMICA <i>T+E</i>				$g_{\text{NL}}^{\text{local}} = (-5.8 \pm 6.5) \times 10^4$
f_{NL}^X	Local	4.1	$\pm$ 5.1	-0.9	$\pm$ 5.1	
	Equilateral	-25	$\pm$ 47	-26	$\pm$ 47	
	Orthogonal	-47	$\pm$ 24	-38	$\pm$ 24	

초기섭동이 정규분포일 때

- 현재 우주론에서 CMB의 온도 비등방성과 편광은 모두 우주 초기의 인플레이션 중에 양자요동이 지평선 밖으로 나가면서 생겨났다는게 정설.
- 만약 인플레이션 동안 우리 우주가 단 하나의 잘 맞춰진 시계에 의해 기술될 수 있고, 양자요동 또한 그 시계의 미세한 차이만을 주는 역할을 한다면 (single-clock inflation), 그리고 양자요동이 진공상태에서 생겨났다면, 초기섭동은 정규분포를 따라야 한다.

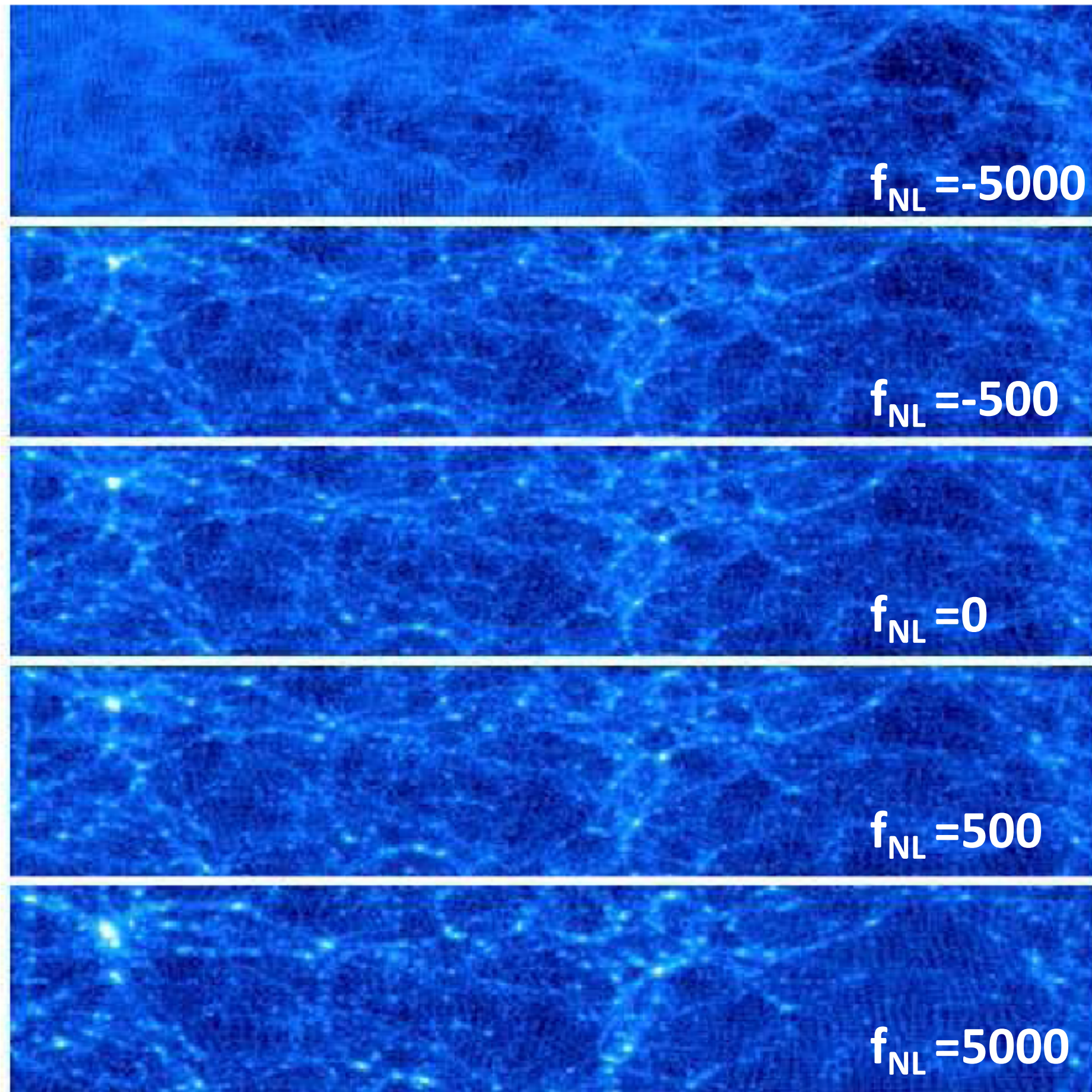
비정규성(non-Gaussianity) 관측의 중요성

- 현재 우주론에서 CMB의 온도 비등방성과 편광은 모두 우주 초기의 인플레이션 중에 양자요동이 지평선 밖으로 나가면서 생겨났다는게 정설.
- 만약 인플레이션 동안 우리 우주가 단 하나의 잘 맞춰진 시계에 의해 기술될 수 있고, 양자요동 또한 그 시계의 미세한 차이만을 주는 역할을 한다면 (single-clock inflation), 그리고 양자요동이 진공상태에서 생겨났다면, 초기섭동은 정규분포를 따라야 한다.
- 그러니까, 우주초기의 비정규성이 관측된다면 single-clock inflation모형이 틀렸다는 관측적 증거가 된다!
 - 이 명제는 인플라톤(inflaton)의 운동항 (kinetic term), 포텐셜 (potential)과 상관 없이 참임.

은하에서의 비정규성 (non-Gaussianity)

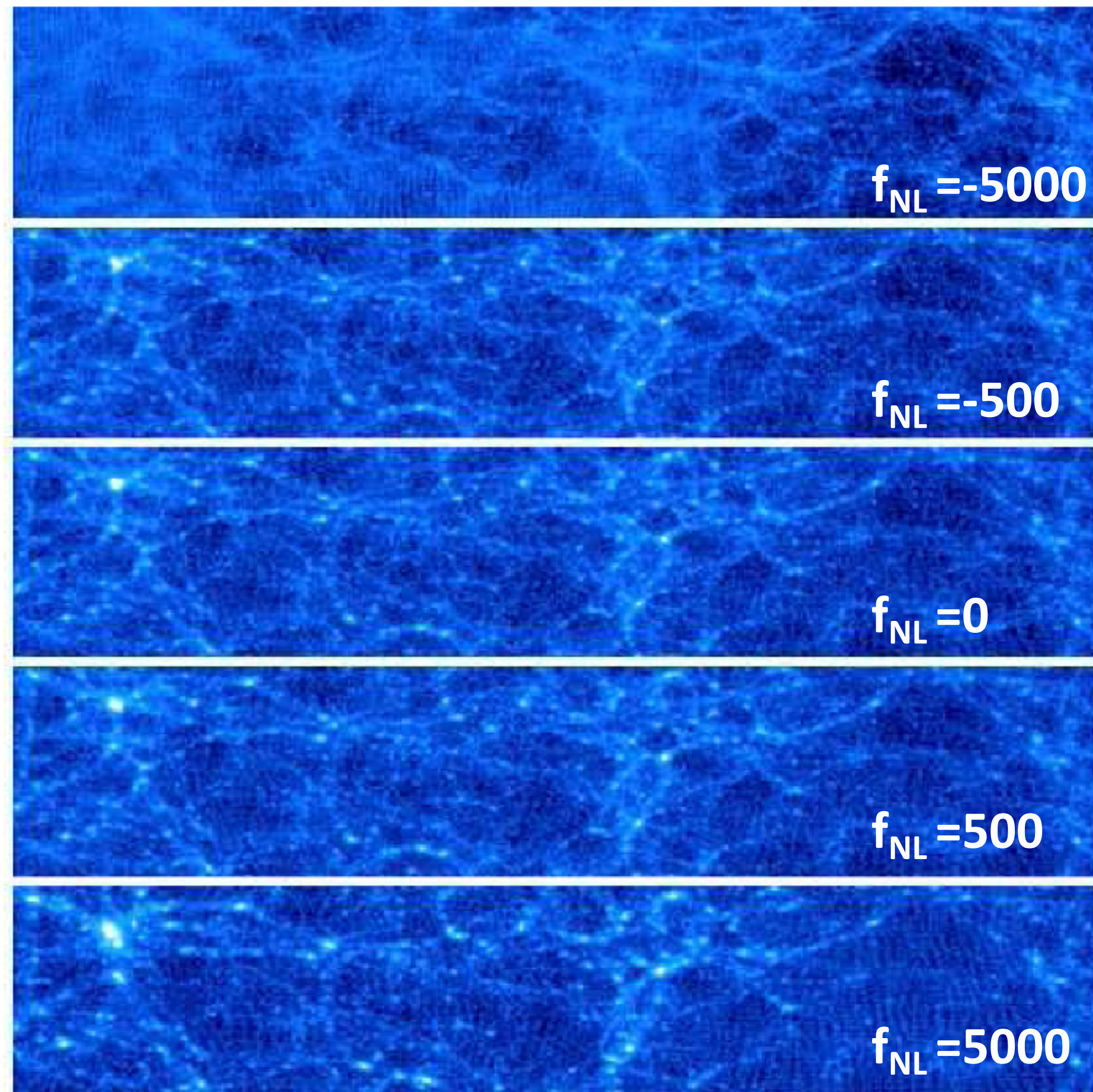
- 큰 스케일로 갈수록 커지는 (scale-dependent) 은하 바이어스 (galaxy bias)를 통해서 CMB보다 잘 측정할 수 있음.
 - 초기에 국소 비정규성 (local primordial non-Gaussianity)이 있으면:
중력 포텐셜이 컸던 (작았던) 곳의 밀도 파워 스펙트럼이 크기 (작기) 때문에, 은하가 더 많이 (적게) 분포해야한다.

비정규성(non-Gaussianity)의 효과

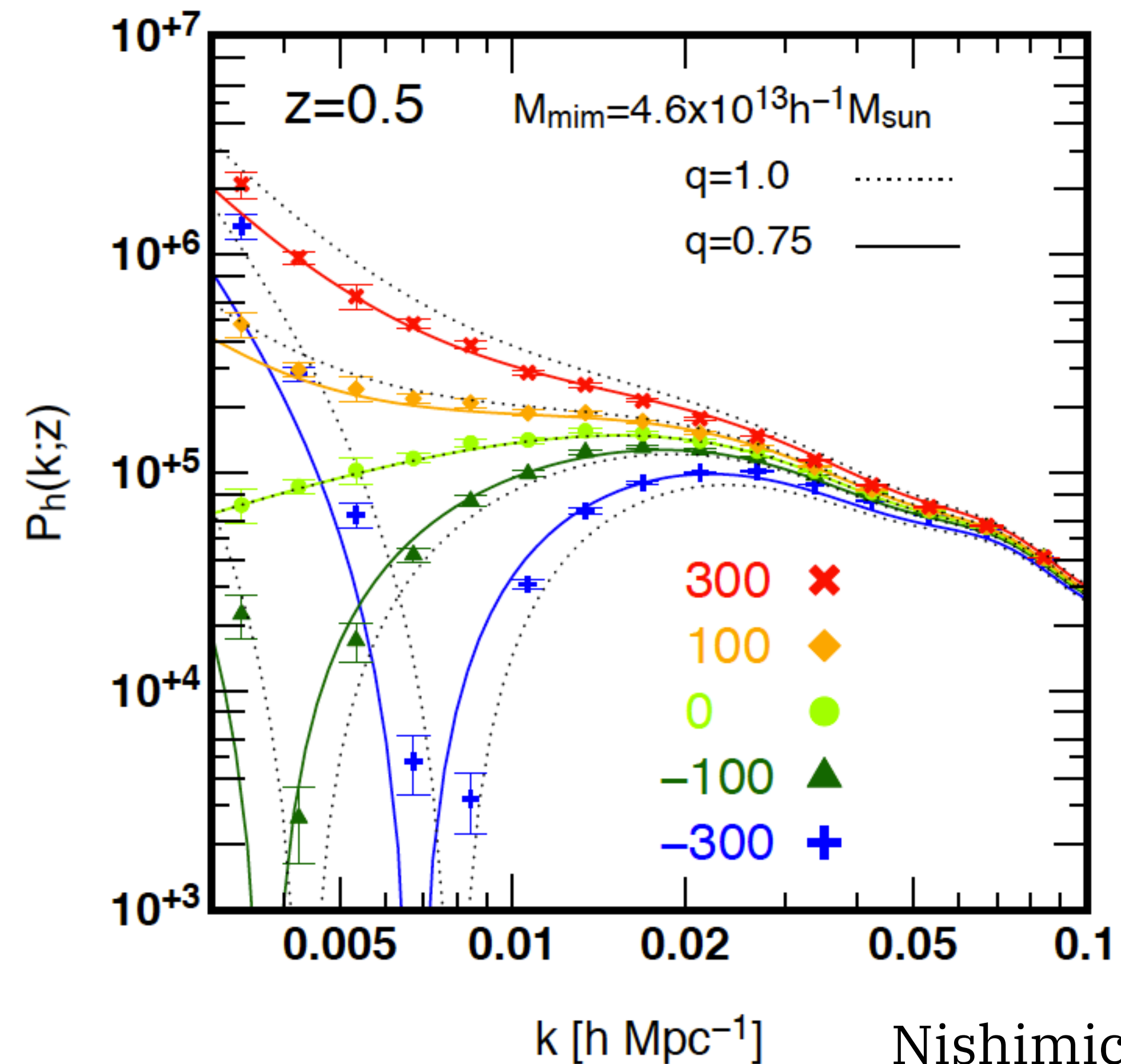


Dalal et al. (2008)

비정규성(non-Gaussianity)의 효과



Dalal et al. (2008)



Nishimichi et al. (2010)

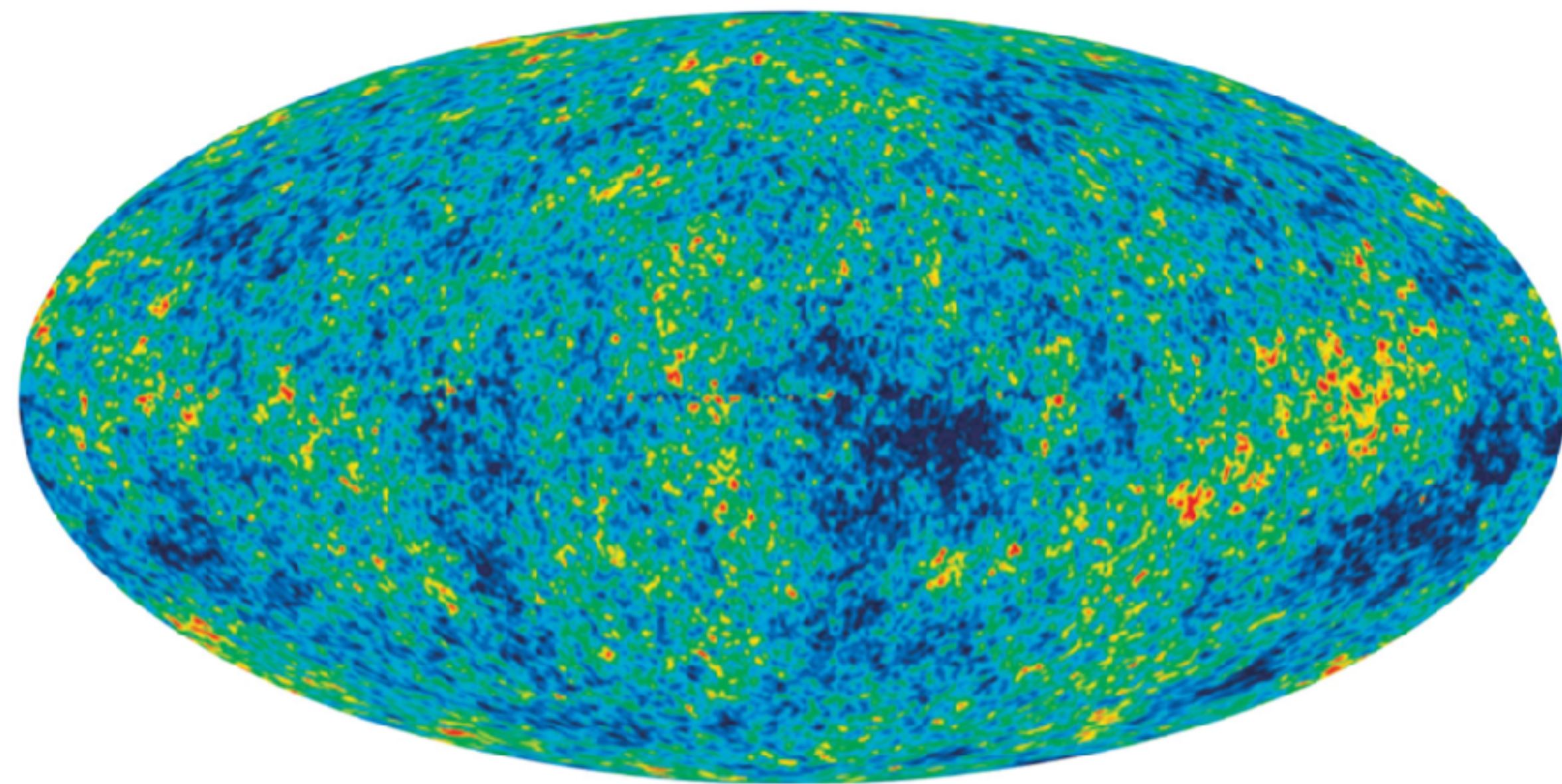
은하에서의 비정규도 (non-Gaussianity)

- 큰 스케일로 갈수록 커지는 (scale-dependent) 은하 바이어스 (galaxy bias)를 통해서 CMB보다 잘 측정할 수 있음.
 - 초기에 국소 비정규성 (local primordial non-Gaussianity)이 있으면:
중력 포텐셜이 컸던 (작았던) 곳의 밀도 파워 스펙트럼이 크기 (작기) 때문에, 은하가 더 많이 (적게) 분포해야한다.
- SPHEREx (NASA, 천문연구원 등)의 목표 정확도는 $\sigma_{f_{\text{NL}}} = 0.5!$
(Heinrich, Dore, Krause, 2311.13082)
- 물론, 이 정도 정확도에서는 일반상대성이론 효과 등 다른 여러가지 효과도 함께 고려해 줘야 함.

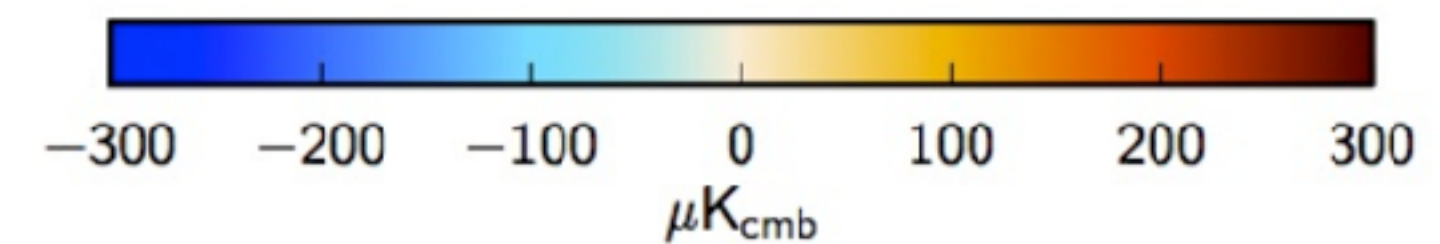
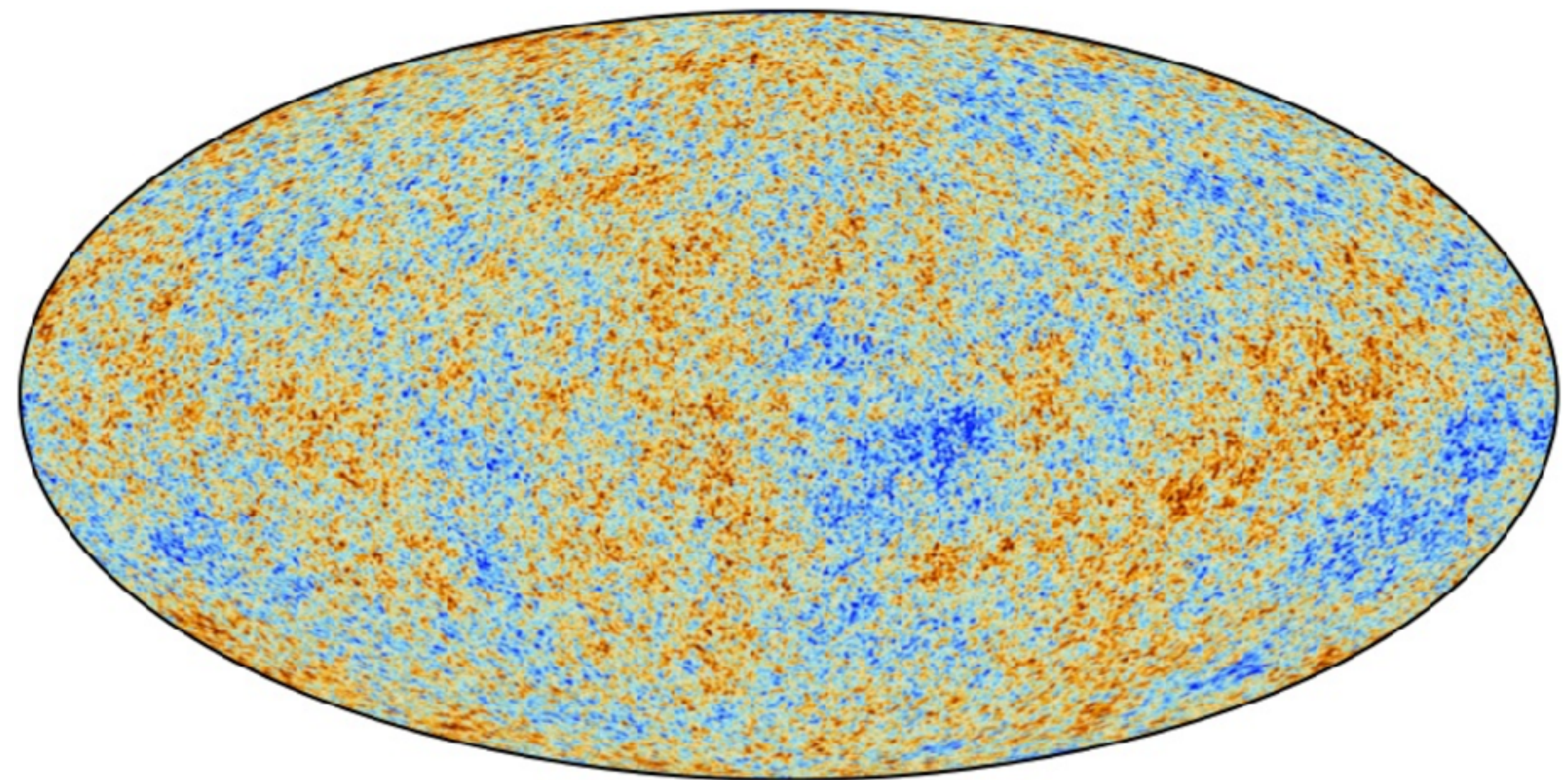
우주론 관측과 확률

- 왜 확률적인가?
 - 관측가능한 우주는 딱 하나.
 - 우리가 관측한 우주배경복사의 비등방성은 약 450억광년 반지름 천구에 딱 하나.
 - 각각의 적색이동 (또는 우주 시간)에 딱 정해진 하나의 은하분포.
 - 이 한계가 바로 Cosmic Variance.
 - 우주배경복사의 관측결과, 우주거대구조를 만든 초기우주 밀도섭동의 분포는 정규분포로 매우 정확하게 근사할 수 있음
 - 원리 설명을 위해, 간단한 온도 비등방성을 예시로 사용 (E/B도 근본은 같음)

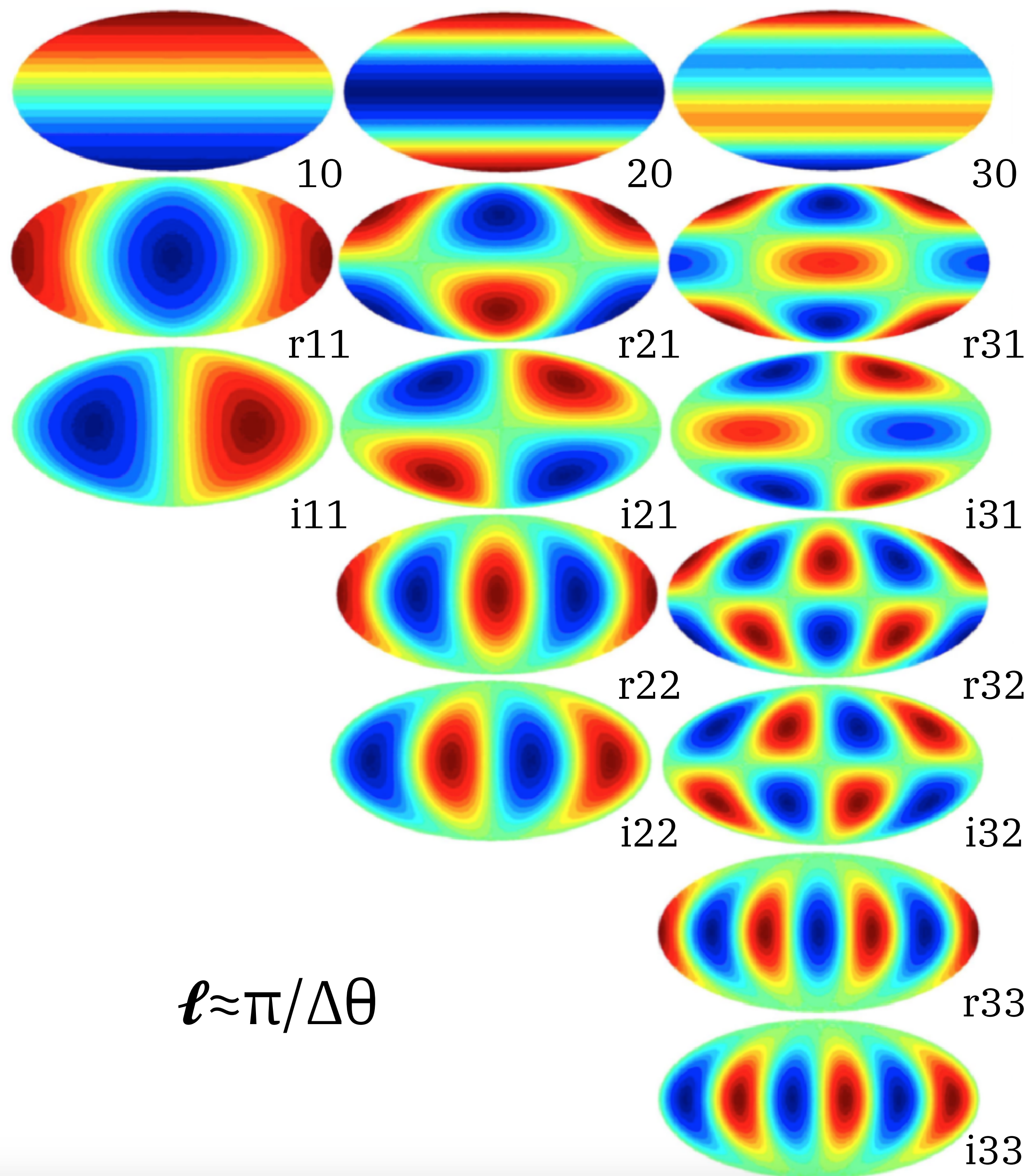
우주의 아기사진: 우주배경복사 지도 $T(\hat{n})$



9년차 WMAP ILC (Internal-Linear-Combination) 지도



Planck 2018 최우도 (Maximum Likelihood) 지도



구면조화함수 $Y_{\ell m}(\hat{n})$

$$\hat{L}^2 Y_{\ell m}(\hat{n}) = \ell(\ell+1) Y_{\ell m}(\hat{n}) = -\nabla_{\Omega}^2 Y_{\ell m}(\hat{n})$$

$$\hat{L}_z Y_{\ell m}(\hat{n}) = m Y_{\ell m}(\hat{n})$$

$$\int d^2 \hat{n} Y_{\ell m}(\hat{n}) Y_{\ell' m'}^*(\hat{n}) = \delta_{\ell \ell'} \delta_{m m'}$$

$$\sum_{\ell, m} Y_{\ell m}(\hat{n}) Y_{\ell m}^*(\hat{n}') = \delta(\hat{n} - \hat{n}')$$

$$Y_{\ell m}^* = (-1)^m Y_{\ell -m}$$

온도섭동의 파워 스펙트럼

- 온도섭동은 구면조화함수 ($Y_{\ell m}(\hat{n})$; spherical harmonics)로 전개해서 분석:

$$T(\hat{n}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m}^T Y_{\ell m}(\hat{n})$$

- 온도 파워 스펙트럼: 통계적으로 등방한 우주에서는 파워 스펙트럼이 대각화 되어있음.

$$\left\langle a_{\ell m}^T a_{\ell' m'}^{T*} \right\rangle = C_{\ell}^{TT} \delta_{\ell \ell'} \delta_{m m'}$$

- 초기섭동이 정규분포를 따르면, 초기섭동의 선형변환인 온도섭동 또한 정규분포를 따르게 되고, 2점함수가 (또는, 파워 스펙트럼이) 그 분포의 모든 것을 결정함.

$a_{\ell m}$: 조화공간의 온도섭동

- 조화공간에서 온도섭동은 조화전개 계수 $\{a_{\ell m}\}$ 로 표현됨:

$$a_{\ell m} = \int d^2 \hat{n} T(\hat{n}) Y_{\ell m}^*(\hat{n})$$

- $a_{\ell m}$ 의 특성:

- ensemble 평균이 $\langle T \rangle \delta_{\ell 0} \delta_{m 0}$: $\langle T(\hat{n}) \rangle = \bar{T}$.

- $a_{\ell m} = (-1)^m a_{\ell -m}^*$. 왜냐하면,

$$a_{\ell m}^* = \int d^2 \hat{n} T(\hat{n}) Y_{\ell m}(\hat{n}) = (-1)^m \int d^2 \hat{n} T(\hat{n}) Y_{\ell -m}^*(\hat{n}) = (-1)^m a_{\ell -m}$$

- 각각의 $a_{\ell m}$ 은 독립적으로 분산이 C_ℓ 인 정규분포를 따름.

$a_{\ell m}$ 의 확률밀도함수

- 각각의 $a_{\ell m}$ 은 독립적으로 분산이 C_ℓ 인 정규분포를 따른다:

$$P(a_{\ell m}) = \frac{1}{\sqrt{2\pi C_\ell}} e^{-\frac{|a_{\ell m}|^2}{2C_\ell}}$$

$\{a_{\ell m}\}$ 의 확률밀도함수

- 각각의 $a_{\ell m}$ 은 독립적으로 분산이 C_ℓ 인 정규분포를 따른다:

$$P(a_{\ell m}) = \frac{1}{\sqrt{2\pi C_\ell}} e^{-\frac{|a_{\ell m}|^2}{2C_\ell}}$$

- 각각의 $a_{\ell m}$ 이 독립사건이므로, 전체 온도비등방성 지도의 확률밀도함수는

$$P(\{a_{\ell m}\}) = \prod_{\ell, m} \frac{1}{\sqrt{2\pi C_\ell}} e^{-\frac{|a_{\ell m}|^2}{2C_\ell}}$$

우주론 모수들은 어떻게 측정하나?

- 우주배경복사의 확률밀도함수를 알게 되었다.

$$P(\{a_{\ell m}\}) = \prod_{\ell, m} \frac{1}{\sqrt{2\pi C_\ell}} e^{-\frac{|a_{\ell m}|^2}{2C_\ell}}$$

- 이로부터 우리가 정말로 알고 싶은것은 우주론의 모수(parameter)들 $\vec{\theta} = \{\Omega_b h^2, \Omega_c h^2, 100\theta_{\text{MC}}, \tau, A_s, n_s, \dots\}$ 에 대한 통계정보.
- 지금 측정한 우주배경복사 지도에 부합하는 우주론 모수를 알고 싶은 것. 즉, 우리가 원하는 것은 $p(\vec{\theta} | \vec{a})$ 이다.

베이즈 정리 (Bayes' theorem)

- 조건부 확률의 정의로부터: $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- 결합 확률은 조건부 확률로 적을 수 있고, $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$
- 위의 두 식을 잘 변형하면, 베이즈 정리를 얻는다: $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

베이즈 정리 (Bayes' theorem)

- 베이즈 정리:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

- $P(B)$ 는 확률의 총합이 1인 것을 이용해 정규화할 수 있으므로, 별로 중요하지 않고, 보통은 이렇게 쓴다:

$$P(A|B) \propto \mathcal{L}(B|A)P(A)$$

사후확률 (posterior distribution) = 우도 (likelihood) x 사전확률 (prior distribution)

베이즈 정리 (Bayes' theorem)

- 우주배경복사 데이터에 적용하면,

$$P(\vec{\theta}|\vec{a}) \propto \mathcal{L}(\vec{a}|\vec{\theta})P(\vec{\theta})$$

- 우도함수 (likelihood function)는 앞서 구한 CMB 온도섭동의 확률밀도함수로 부터,

$$\mathcal{L}(\{a_{\ell m}\}|\vec{\theta}) = \prod_{\ell, m} \frac{1}{\sqrt{2\pi C_{\ell}(\vec{\theta})}} \exp \left[-\frac{|a_{\ell m}|^2}{2C_{\ell}(\vec{\theta})} \right]$$

- 사전확률은 그 파라미터에 대해 우리가 알고 있는 사전 지식을 반영하고 있음.

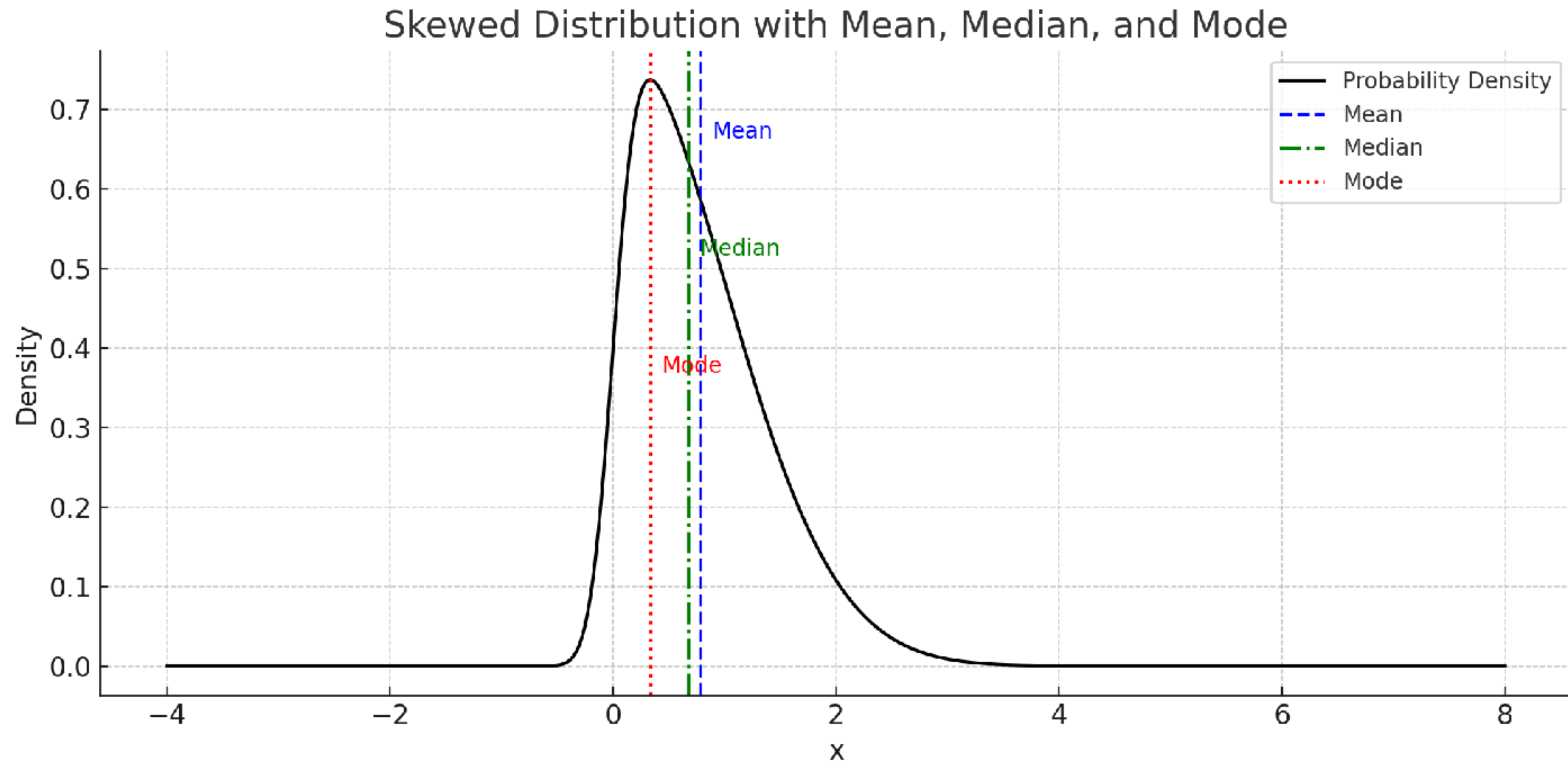
우주배경복사에서의 우주론 모수 추정

- 모수벡터 $\vec{\theta}$ 에 대한 사후확률

$$P(\vec{\theta}|\{a_{\ell m}\}) \propto \mathcal{L}(\{a_{\ell m}\}|\vec{\theta})P(\vec{\theta}) = \prod_{\ell, m} \frac{1}{\sqrt{2\pi C_{\ell}(\vec{\theta})}} \exp \left[-\frac{|a_{\ell m}|^2}{2C_{\ell}(\vec{\theta})} \right]$$

- 이 식이 우주론 모수추정의 기본식이며, 사후확률 계산을 위해 필요한 것은
 - 사전확률 (전에 행해진 우주론적 관측결과, 우리의 믿음: $m_{\nu}>0$, 등)
 - 우도함수, 이를 위해 $C_{\ell}(\vec{\theta})$ 를 알아야 함 [CAMB, CLASS 등을 이용]

다양한 모수 추정방법: 평균, 최우도



최우도 추정의 예시: C_ℓ

- 모수 $\theta = C_\ell$ 에 대한 1차원 사후확률을 생각하자. 사전확률 $P(C_\ell)$ 은 flat

$$P(C_\ell | \{a_{\ell m}\}) \propto \mathcal{L}(\{a_{\ell m}\} | C_\ell) P(C_\ell) = \prod_m \frac{1}{\sqrt{2\pi C_\ell}} \exp \left[-\frac{|a_{\ell m}|^2}{2C_\ell} \right]$$

- 다른 ℓ 은 C_ℓ 추정에 기여하지 않음으로 무시. 최우도 추정(MLE)을 해보자:

$$\log \mathcal{L}(C_\ell | \{a_{\ell m}\}) = -\frac{2\ell + 1}{2} \log C_\ell - \frac{1}{2C_\ell} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2 + \text{const}$$

최우도 추정의 예시: C_ℓ

- 모수 $\theta = C_\ell$ 에 대한 1차원 사후확률을 생각하자. 사전확률 $P(C_\ell)$ 은 flat

$$P(C_\ell | \{a_{\ell m}\}) \propto \mathcal{L}(\{a_{\ell m}\} | C_\ell) P(C_\ell) = \prod_m \frac{1}{\sqrt{2\pi C_\ell}} \exp \left[-\frac{|a_{\ell m}|^2}{2C_\ell} \right]$$

- 다른 ℓ 은 C_ℓ 추정에 기여하지 않음으로 무시. 최우도 추정(MLE)을 해보자:

$$\log \mathcal{L}(C_\ell | \{a_{\ell m}\}) = -\frac{2\ell + 1}{2} \log C_\ell - \frac{1}{2C_\ell} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2 + \text{const}$$

$$\frac{d}{dC_\ell} \log \mathcal{L}(C_\ell) = -\frac{2\ell + 1}{2C_\ell} + \frac{1}{2C_\ell^2} \sum_m |a_{\ell m}|^2 = 0$$

최우도 추정의 예시: C_ℓ

- 모수 $\theta = C_\ell$ 에 대한 1차원 사후확률을 생각하자. 사전확률 $P(C_\ell)$ 은 flat

$$P(C_\ell | \{a_{\ell m}\}) \propto \mathcal{L}(\{a_{\ell m}\} | C_\ell) P(C_\ell) = \prod_m \frac{1}{\sqrt{2\pi C_\ell}} \exp \left[-\frac{|a_{\ell m}|^2}{2C_\ell} \right]$$

- 다른 ℓ 은 C_ℓ 추정에 기여하지 않음으로 무시. 최우도 추정(MLE)을 해보자:

$$\log \mathcal{L}(C_\ell | \{a_{\ell m}\}) = -\frac{2\ell + 1}{2} \log C_\ell - \frac{1}{2C_\ell} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2 + \text{const}$$

$$\frac{d}{dC_\ell} \log \mathcal{L}(C_\ell) = -\frac{2\ell + 1}{2C_\ell} + \frac{1}{2C_\ell^2} \sum_m |a_{\ell m}|^2 = 0$$

$$\widehat{C}_\ell = \frac{1}{2\ell + 1} \sum_m |a_{\ell m}|^2$$

최우도 추정된 C_ℓ 의 분산

- 모수 $\theta = C_\ell$ 에 대한 1차원 사후확률을 최우도 추정치로 다시 쓰면,

$$P(C_\ell | \{a_{\ell m}\}) \propto \prod_m \frac{1}{\sqrt{2\pi C_\ell}} \exp\left[-\frac{|a_{\ell m}|^2}{2C_\ell}\right] \propto \left(\frac{1}{C_\ell}\right)^{(2\ell+1)/2} \exp\left[-\frac{(2\ell+1)\widehat{C}_\ell}{2C_\ell}\right]$$

- 이 pdf는 chi-square분포이고 (당연히!!), 특히 $\widehat{C}_\ell \sim \frac{C_\ell}{2\ell+1} \cdot \chi_{2\ell+1}^2$.

- chi-square분포 χ_k^2 의 분산은 $2k$.

- 최우도 추정치의 분산은:

$$\text{Var}[\widehat{C}_\ell] = \left(\frac{C_\ell}{2\ell+1}\right)^2 2(2\ell+1) = \frac{2}{2\ell+1} C_\ell^2$$

Cosmic Variance

- 통계적으로, 우리의 CMB에서 추정된 C_ℓ 값과 실제 C_ℓ 값의 차이는

$$\sqrt{\left\langle \left(\Delta \widehat{C}_\ell \right)^2 \right\rangle} = \sqrt{\frac{2}{2\ell + 1}} C_\ell$$

- 왜냐하면, C_ℓ 을 측정하기 위해, 각 ℓ -모드에 대해 $(2\ell + 1)/2$ 개의 $|a_{\ell m}|^2$ 를 이용하였으므로, 그 정확도는 이용한 ℓ -모드의 개수에 따라 늘어난다.
- 우리가 관측할 수 있는 우주는 딱 하나의 전체 우주배경복사이기 때문에, 이 값이 우리가 할 수 있는 최선. 그래서, 이 분산 값이 바로 앞에서 본 **Cosmic Variance!**

다차원 모수공간에서 최우도 추정

- 예시는 1차원 모수공간이었지만, 다차원 모수공간에서 최우도를 추정하기 위해서는 다양한 최적화 방법을 이용:

$$\nabla_{\vec{\theta}} \log \mathcal{L} = \frac{\partial \log \mathcal{L}}{\partial \theta_i} \hat{\theta}_i = 0$$

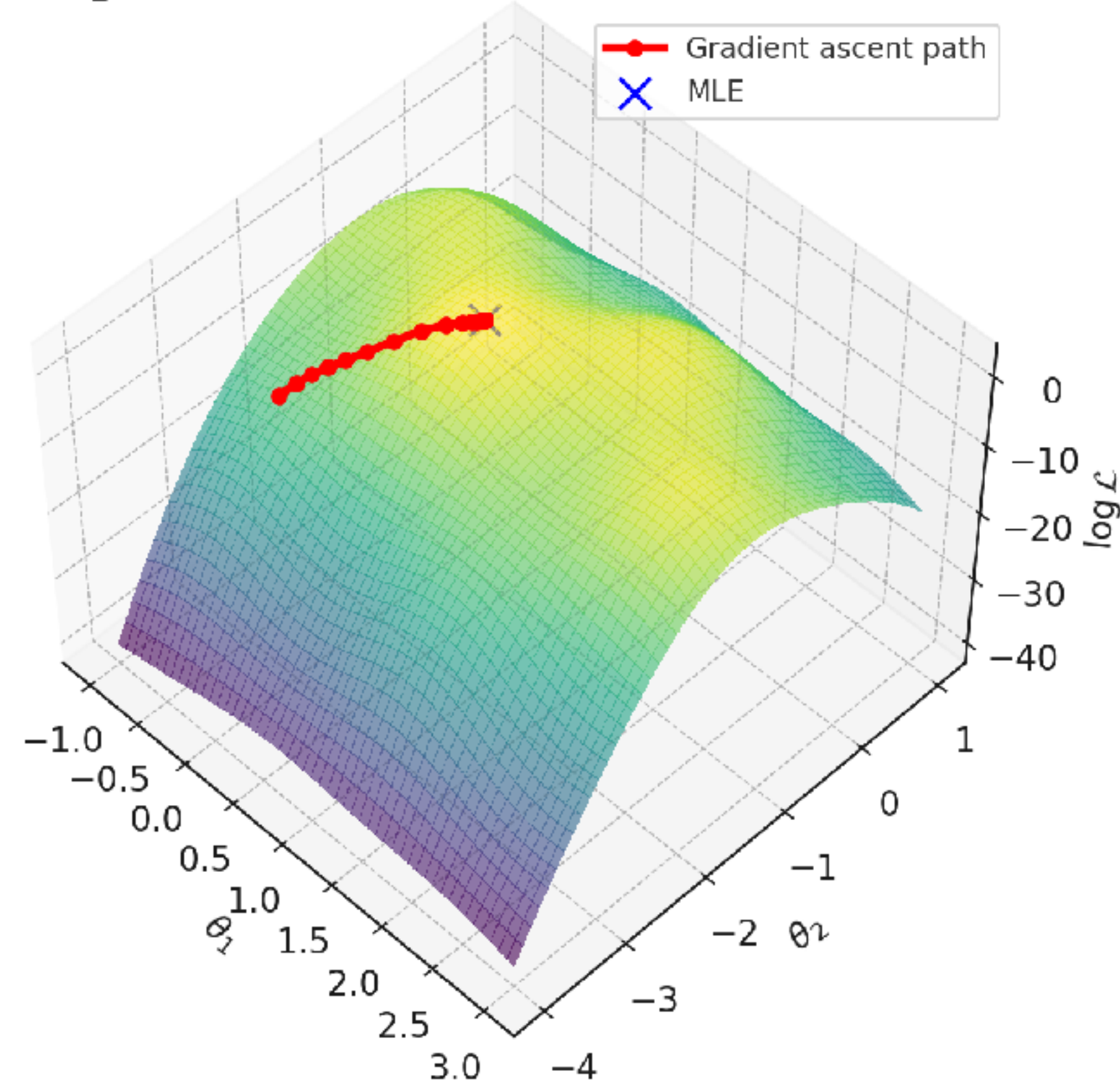
- 최근의 AI에서는 아주 고(백만!)차원에서 경사하강법(Gradient Decent)을 응용함:

$$\theta^{(t+1)} = \theta^{(t)} - \eta \nabla (-\log \mathcal{L}(\theta^{(t)}))$$

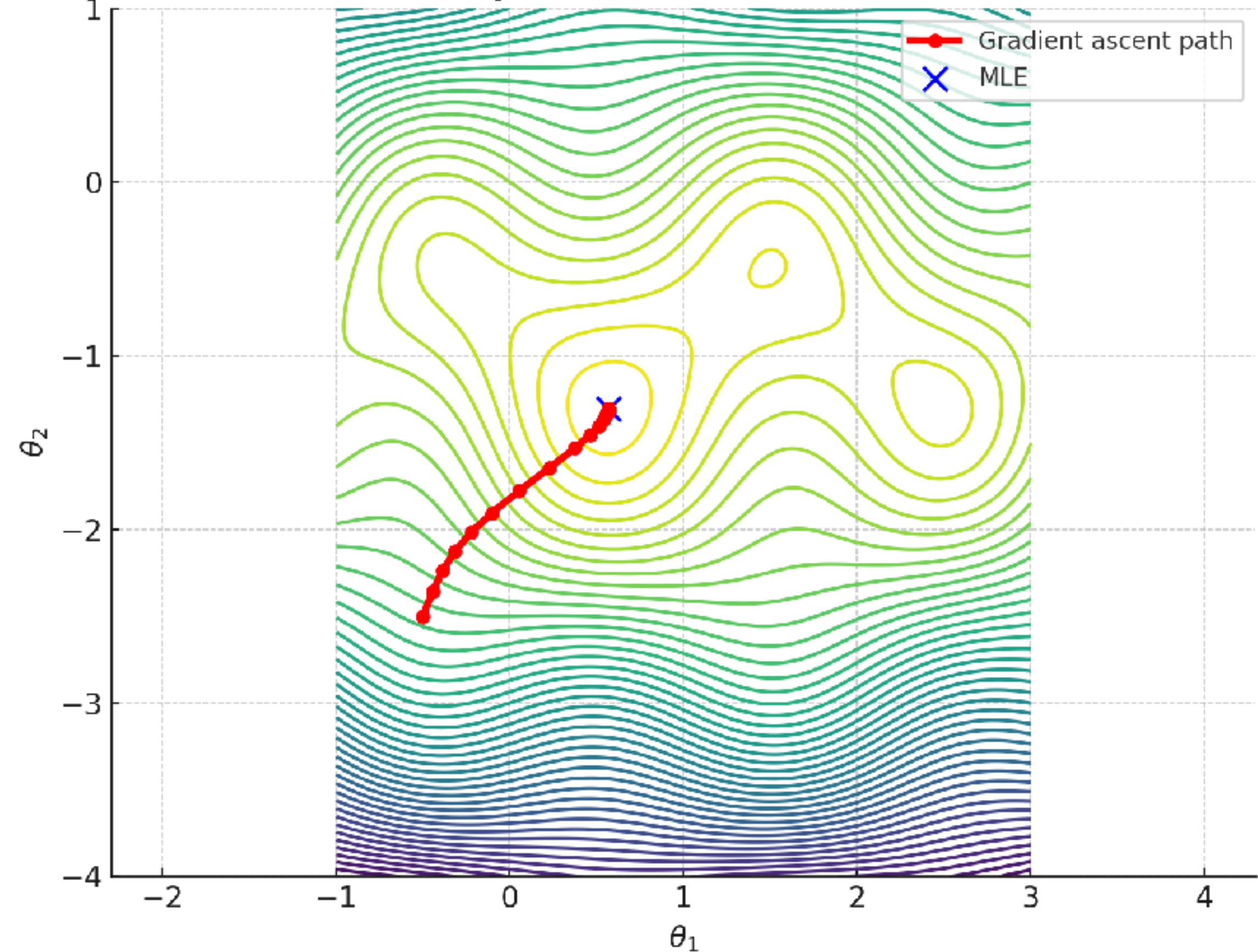
- 많이 쓰는 알고리즘: (주의: 미적분이 아님! $\epsilon \rightarrow 0$ 극한은 아주 안 좋은 경우가 많음.)
Stochastic Gradient Decent, Adaptive Moment Estimate (ADAM)

다차원 모수공간에서 최우도 추정

3D Log-Likelihood Surface with Gradient Ascent



2D Projection of Gradient Ascent



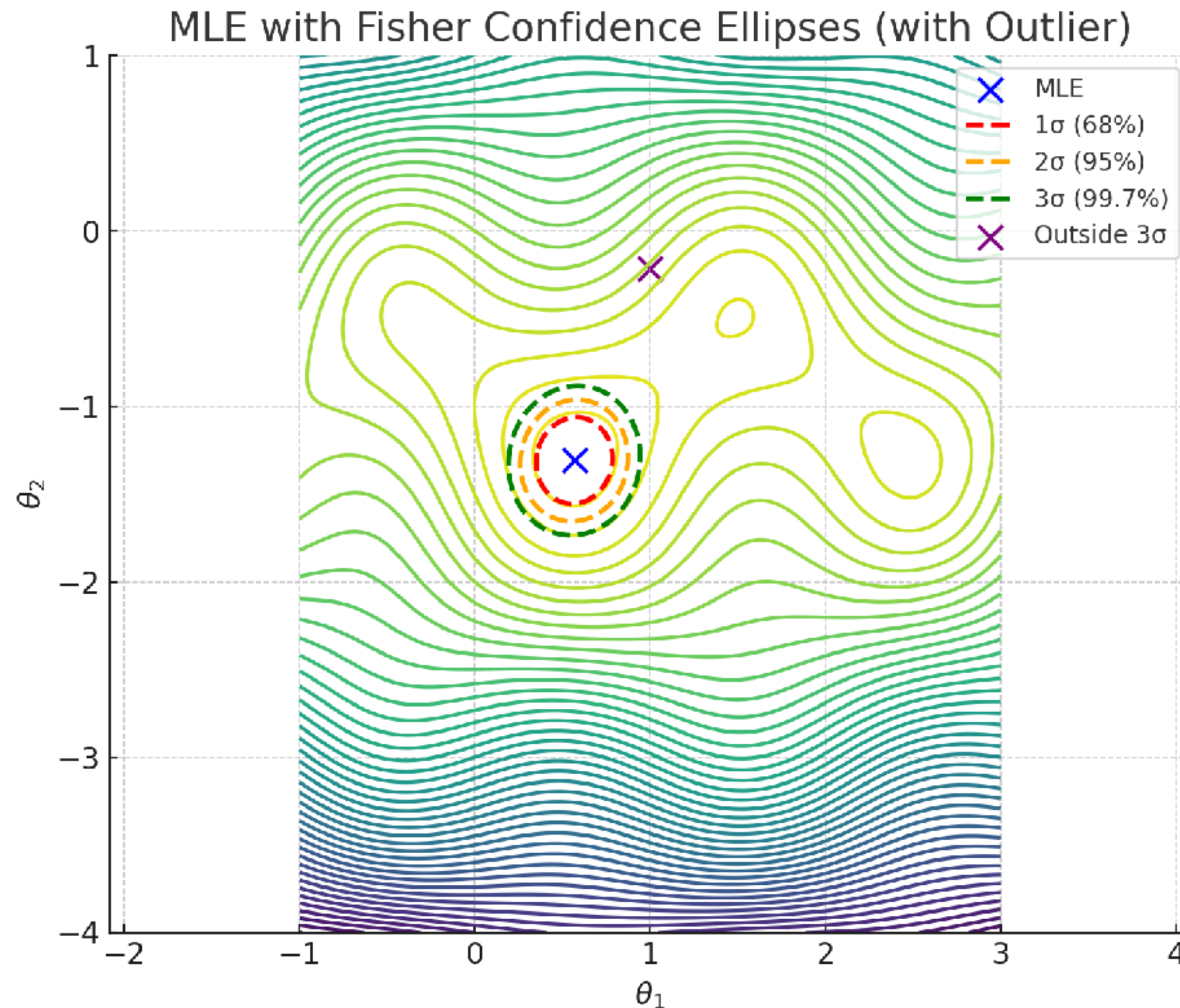
피셔 행렬 (Fisher matrix)

- 피셔 정보 행렬 (Fisher Information Matrix)의 정의

$$F_{ij} = - \left\langle \frac{\partial^2 \log \mathcal{L}}{\partial \theta_i \partial \theta_j} \right\rangle$$

- 피셔 행렬은 모수공간의 MLE점에 서서 본 Hessian (곡률행렬)을 나타냄.
즉, MLE에 서서 각각 다른 파라미터 방향으로 얼마나 빠르게 우도함수가 변하는가를 나타내는 척도.
- 그 방향으로 Fisher matrix가 크면 우도함수의 곡률이 가파르고 (추정에러가 작고), 작으면 완만함 (추정에러가 큼).

다차원 모수공간에서 피셔 행렬



피셔 행렬 (Fisher matrix) 계산

- 피셔 정보 행렬 (Fisher Information Matrix)의 정의에 따라,

$$\log \mathcal{L}(\vec{\theta}|\{a_{\ell m}\}) = -\sum_{\ell} \frac{2\ell+1}{2} \log C_{\ell} - \frac{1}{2C_{\ell}} \sum_{\ell m} |a_{\ell m}|^2 + \text{const}$$

$$\frac{\partial \log \mathcal{L}(\vec{\theta})}{\partial \theta_i} = -\sum_{\ell} \frac{2\ell+1}{2} \frac{C_{\ell,i}}{C_{\ell}} + \sum_{\ell} \frac{C_{\ell,i}}{2C_{\ell}^2} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2$$

$$\frac{\partial \log^2 \mathcal{L}(\vec{\theta})}{\partial \theta_i \partial \theta_j} = -\sum_{\ell} \left(\frac{2\ell+1}{2} \right) \left[\frac{C_{\ell,ij}}{C_{\ell}} - \frac{C_{\ell,i}C_{\ell,j}}{C_{\ell}^2} \right] + \sum_{\ell} \left(\frac{C_{\ell,ij}}{2C_{\ell}^2} - \frac{C_{\ell,i}C_{\ell,j}}{C_{\ell}^3} \right) \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2$$

피셔 행렬 (Fisher matrix) 계산

- 피셔 정보 행렬 (Fisher Information Matrix)의 정의에 따라,

$$\log \mathcal{L}(\vec{\theta} | \{a_{\ell m}\}) = - \sum_{\ell} \frac{2\ell + 1}{2} \log C_{\ell} - \frac{1}{2C_{\ell}} \sum_{\ell m} |a_{\ell m}|^2 + \text{const}$$

$$\frac{\partial \log \mathcal{L}(\vec{\theta})}{\partial \theta_i} = - \sum_{\ell} \frac{2\ell + 1}{2} \frac{C_{\ell,i}}{C_{\ell}} + \sum_{\ell} \frac{C_{\ell,i}}{2C_{\ell}^2} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2$$

$$\frac{\partial \log^2 \mathcal{L}(\vec{\theta})}{\partial \theta_i \partial \theta_j} = - \sum_{\ell} \left(\frac{2\ell + 1}{2} \right) \left[\frac{C_{\ell,ij}}{C_{\ell}} - \frac{C_{\ell,i} C_{\ell,j}}{C_{\ell}^2} \right] + \sum_{\ell} \left(\frac{C_{\ell,ij}}{2C_{\ell}^2} - \frac{C_{\ell,i} C_{\ell,j}}{C_{\ell}^3} \right) \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2$$

$$F_{ij} = - \left\langle \frac{\partial \log^2 \mathcal{L}}{\partial \theta_i \partial \theta_j} \right\rangle = \sum_{\ell} \left(\frac{2\ell + 1}{2} \right) \left[\frac{C_{\ell,i} C_{\ell,j}}{C_{\ell}^2} \right]$$

피셔 행렬 (Fisher matrix)

- CMB에서의 피셔 정보 행렬 (Fisher Information Matrix)

$$F_{ij} = - \left\langle \frac{\partial \log^2 \mathcal{L}}{\partial \theta_i \partial \theta_j} \right\rangle = \sum_{\ell} \left(\frac{2\ell + 1}{2} \right) \frac{\partial \log C_{\ell}}{\partial \theta_i} \frac{\partial \log C_{\ell}}{\partial \theta_j}$$

- MLE지점 근처에서 우도함수의 2차근사를 줌.
- 심지어 데이터가 없어도, $C_{\ell}(\theta)$ 만 알면 계산할 수 있음!
- 전천이 CMB관측이 아닌 경우에는 $(2\ell+1)$ 를 $(2\ell+1)f_{\text{sky}}$ 로 바꿔서 동일한 식 사용 가능.
- Cramer-Rao: $\text{Var}(\theta) \geq (F^{-1})_{\theta\theta}$

C_ℓ 의 피셔 행렬 (Fisher matrix)

- 원하는 모수가 $\theta = C_\ell$ 일 경우, 피셔 정보 행렬 (Fisher Information Matrix)은

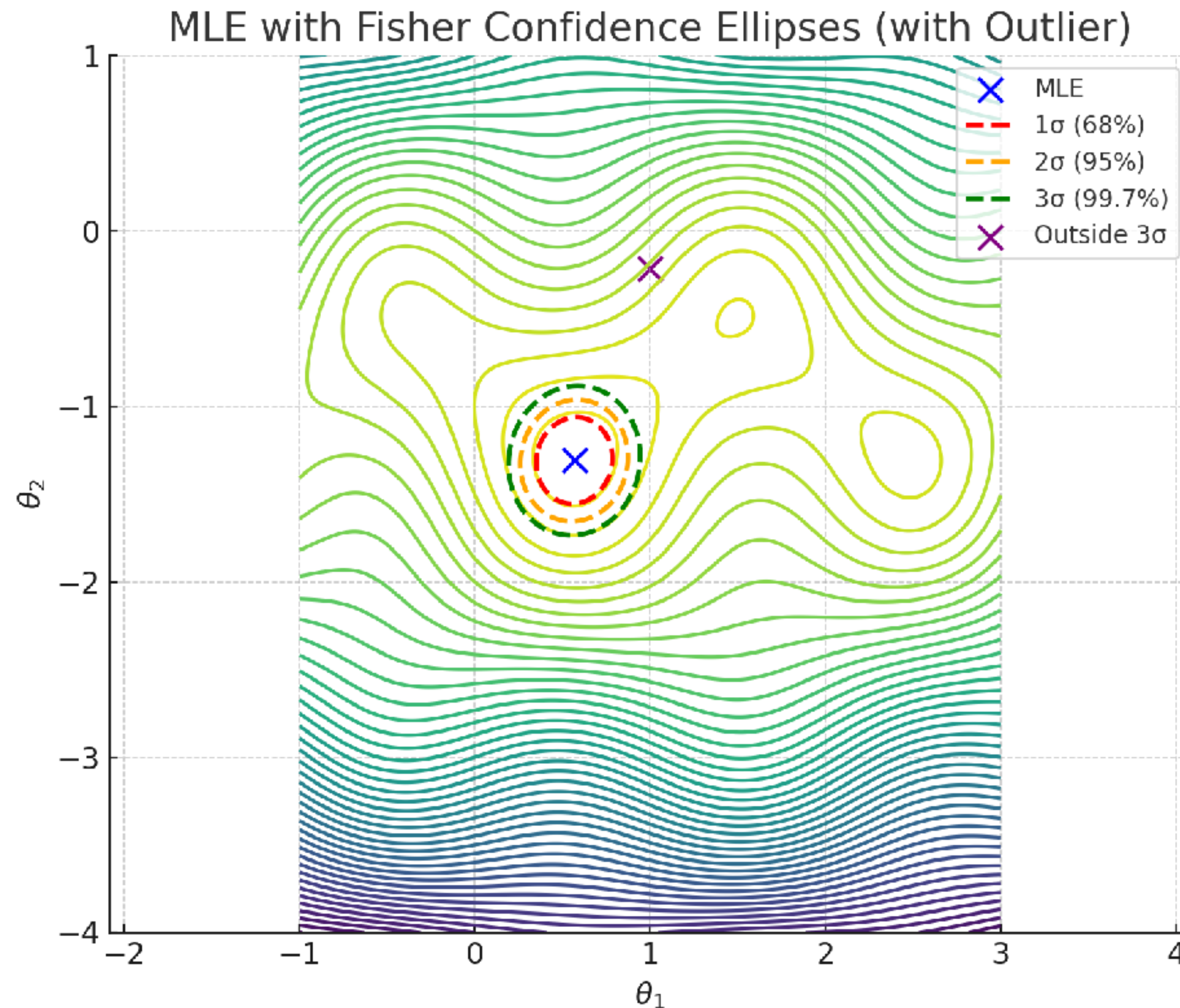
$$F_{C_\ell C_\ell} = \sum_{\ell'} \left(\frac{2\ell' + 1}{2} \right) \frac{\partial \log C_{\ell'}}{\partial C_\ell} \frac{\partial \log C_{\ell'}}{\partial C_\ell} = \frac{2\ell + 1}{2C_\ell^2}$$

- 앞에서 구한 1차원 추정의 분산은 Cramar-Rao 부등식을 만족:

$$\text{Var}[\widehat{C}_\ell] = \frac{2}{2\ell + 1} C_\ell^2 = \frac{1}{F_{C_\ell C_\ell}}$$

- 이 경우, 최소 분산 불편 추정량 (minimum-variance-unbiased estimator) 라고 함!

피셔 행렬은 국소적: 외삽할 수 없음!



다른 모수의 주변화 (marginalization)

- 흔히 고차원의 모수공간에서 최적화를 해야함.
- 예: 플랑크 위성 자료분석의 경우: 94개의 모수

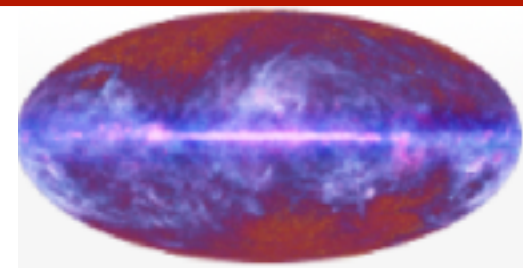
```
>> wc -l base_plikHM_TTTEEE_lowl_lowE_lensing.paramnames  
94 base_plikHM_TTTEEE_lowl_lowE_lensing.paramnames
```

- 두 개의 모수 θ_1, θ_2 만 남기고 나머지는 주변화: 92차원 적분이 필요!!!!

$$P(\theta_1, \theta_2) = \int d\theta_3 \int d\theta_4 \cdots \int d\theta_{94} P(\theta_1, \theta_2, \cdots, \theta_{94})$$

94차원 pdf를 정확히 반영한 샘플이 있다면?

https://wiki.cosmos.esa.int/planck-legacy-archive/index.php/Cosmological_Parameters



Main Index

▼ Introduction

The Planck mission

The spacecraft

Ground segment and
Early operations

Survey scanning and
Routine operations

Q&A

► The Instruments

► Data processing

► PLA Mission
Products

► Planck Added Value
Tools

► Operational data

► Appendix

► Categories

► Tools

Page

Discussion

Read

[View source](#)

[View history](#)



2018 Cosmological parameters and MC chains

[CMB spectrum & Likelihood Code](#) > [Cosmological Parameters](#)

Contents [\[hide\]](#)

- 1 [Description](#)
- 2 [Production process](#)
- 3 [Caveats and known issues](#)
- 4 [Related products](#)
- 5 [Parameter tables](#)
- 6 [Parameter chains](#)
- 7 [File formats](#)
- 8 [Previous Releases: \(2015\) and \(2013\) Cosmological Parameters and MC Chains](#)
- 9 [References](#)

Description [\[edit\]](#)

The 2018 cosmological parameter results explore a variety of cosmological models with combinations of Planck and other data. We provide results from MCMC exploration chains, as well as best fits, and sets of parameter tables. Definitions, conventions, and references are contained in [Planck-2013-XVI](#)^[1], [Planck-2015-A15](#)^[2], and the 2018 parameter paper [Planck-2018-L06](#)^[3].

Production process [\[edit\]](#)

Parameter chains are produced using CosmoMC, a sampling package available [here](#). This includes the sample analysis package (and GUI) GetDist, and the scripts for managing, analysing, and plotting results from the full grid of runs. The Python [GetDist](#) sample analysis package is also available separately.

Chain products provided here have had burn-in removed. Some results with additional data are produced by importance sampling. Note that the baseline model includes one massive (0.06eV) neutrino.

94차원 pdf를 정확히 반영한 샘플이 있다면?

- 92차원 적분 대신 나머지 92차원의 숫자를 그냥 무시!
- 예: 플랑크 위성의 MCMC체인의 경우, $\Omega_{\text{b}} h^2$ [3열]- $\Omega_{\text{c}} h^2$ [4열]
1: 1.000000E+00 1.397755E+03 2.223634E-02 1.190220E-01 1.040860E+00 5.773151E-02
3.044784E+00 9.728228E-01 9.994213E-01 4.526123E+01 2.205585E-01 7.089804E+00
2.557510E+02 4.243939E+01 3.342349E+01 ...
2: 1.000000E+00 1.401902E+03 2.221187E-02 1.190419E-01 1.040930E+00 6.001946E-02
3.050158E+00 9.710518E-01 9.994618E-01 3.366526E+01 6.638596E-01 6.944407E+00
2.341132E+02 4.339938E+01 4.416310E+01 1.296760E+02 1.251694E+00 9.926168E+00
1.383219E+01 2.743546E+01 1.121982E+02 1.348152E-01 8.585275E-02 6.160616E-01
2.030023E-01 6.915661E-01 2.504823E+00 9.993913E-01 9.972368E-01 6....
3: 2.000000E+00 1.394904E+03 2.218835E-02 1.190339E-01 1.040894E+00 6.037516E-02
3.050733E+00 9.712151E-01 9.994621E-01 3.838740E+01 6.305131E-01 6.452556E+00
2.647166E+02 5.021033E+01 4.927893E+01 1.307783E+02 7.840395E-01 8.674568E+00
1.068322E+01 2.088690E+01 9.925467E+01 1.523667E-01 1.183208E-01 ...

94차원 pdf를 정확히 반영한 샘플이 있다면?

- 92차원 적분 대신 나머지 92차원의 숫자를 그냥 무시!
- 예: 플랑크 위성 자료분석의 경우, $\Omega_{\text{b}} h^2$ [3열]- $\Omega_{\text{c}} h^2$ [4열]

1: 1.000000E+00 1.397755E+03 2.223634E-02 1.190220E-01 1.040860E+00 5.773151E-02
3.044784E+00 9.728228E-01 9.994213E-01 4.526123E+01 2.205585E-01 7.089804E+00
2.557510E+02 4.243939E+01 3.342349E+01 ...

2: 1.000000E+00 1.401902E+03 2.221187E-02 1.190419E-01 1.040930E+00 6.001946E-02
3.050158E+00 9.710518E-01 9.994618E-01 3.366526E+01 6.638596E-01 6.944407E+00
2.341132E+02 4.339938E+01 4.416310E+01 1.296760E+02 1.251694E+00 9.926168E+00
1.383219E+01 2.743546E+01 1.121982E+02 1.348152E-01 8.585275E-02 6.160616E-01
2.030023E-01 6.915661E-01 2.504823E+00 9.993913E-01 9.972368E-01 6....

3: 2.000000E+00 1.394904E+03 2.218835E-02 1.190339E-01 1.040894E+00 6.037516E-02
3.050733E+00 9.712151E-01 9.994621E-01 3.838740E+01 6.305131E-01 6.452556E+00
2.647166E+02 5.021033E+01 4.927893E+01 1.307783E+02 7.840395E-01 8.674568E+00
1.068322E+01 2.088690E+01 9.925467E+01 1.523667E-01 1.183208E-01 ...

94차

- 92차
- 예: 플

1: 1.000

3.044784

2.557510

2: 1.000

3.050158

2.341132

1.383219

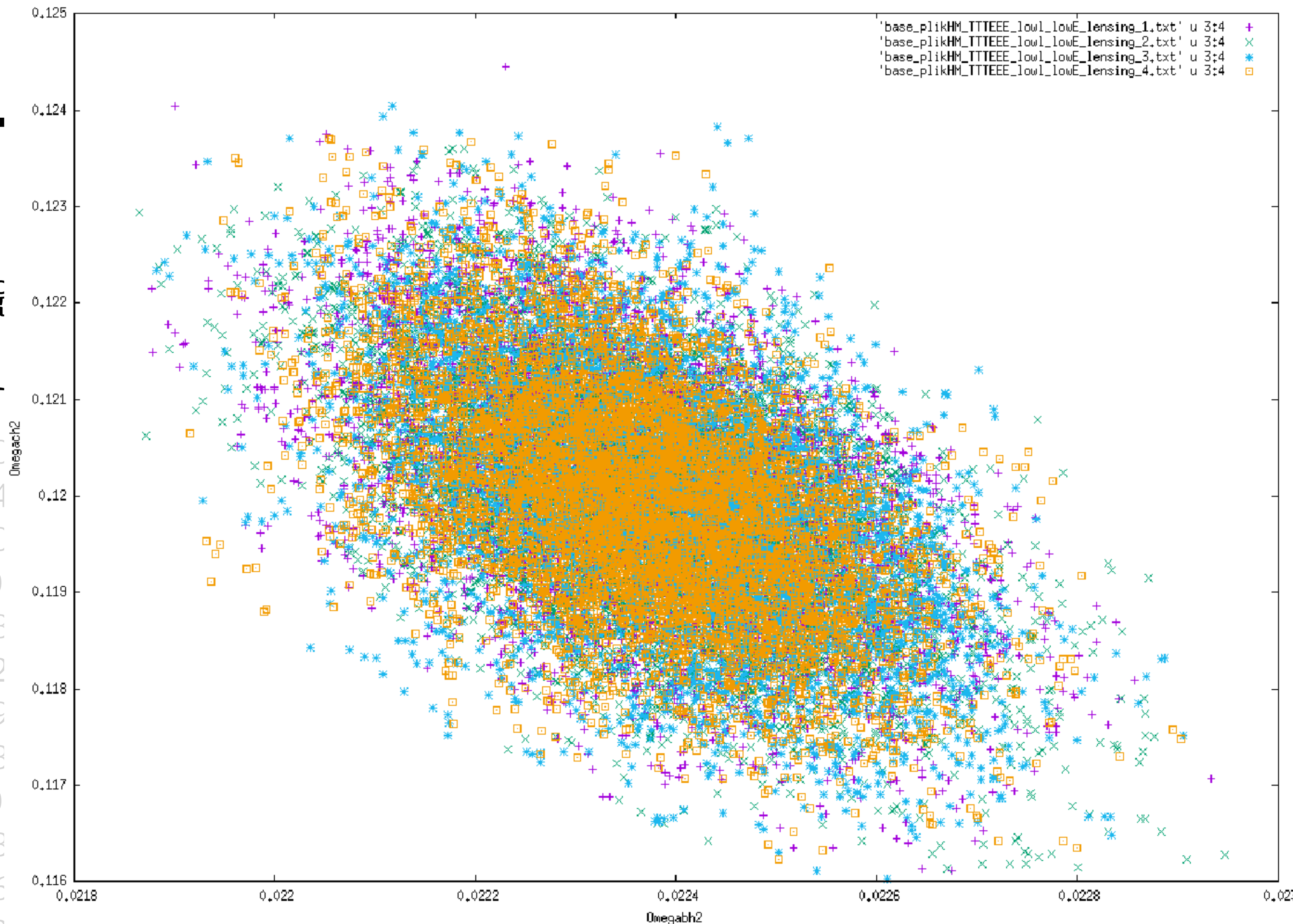
2.030023

3: 2.000

3.050733

2.647166

1.068322



다면?

3151E-02

E+00

946E-02

E+00

8E+00

5E-01

516E-02

E+00

8E+00

1.068322E+01 2.000070E+01 7.725407E+01 1.525007E-01 1.105200E-01 ...

2차원 등고선 그림 그리기

- Python package: getdist
- 직접하고 싶으면: 다른 파라미터를 주변화한 후, 2D 우도함수에서 χ^2 정의

$$\chi^2 \equiv C - 2 \log \mathcal{L}$$

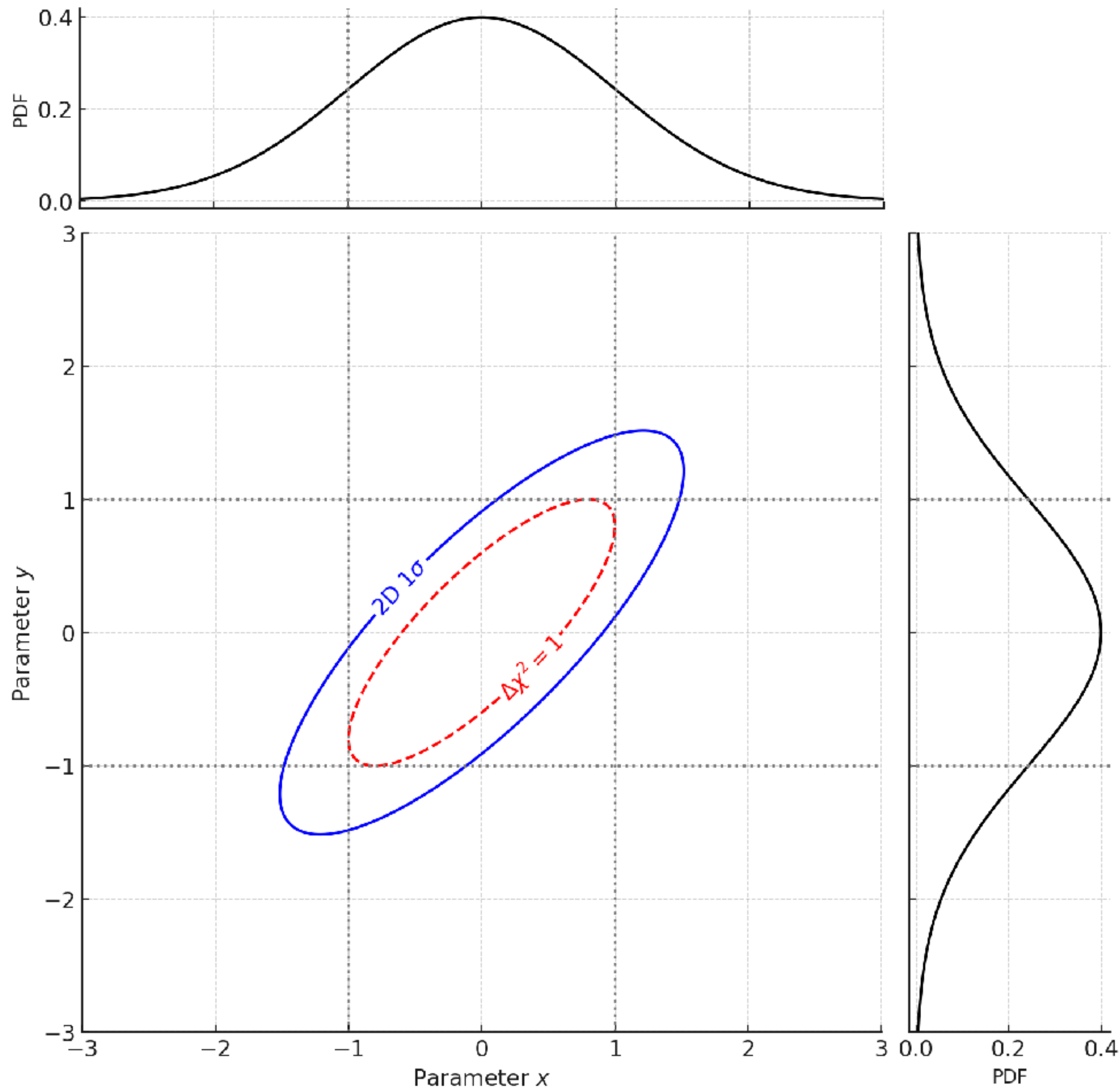
χ^2 값이 가장 작은 점이 2D의 MLE(별 의미는 없음), 그리고 Gaussian인 경우

$$\Delta\chi^2 = \chi^2(\theta_1, \theta_2) - \chi_{\min}^2 = -2 \ln [1 - 0.01p] = 2.30, 6.18, 11.8$$

가 되는 점들을 잇는 등고선이 1-sigma (68%), 2-sigma (95%), 3-sigma (99.9%) 영역.

- 아니면, 직접 χ^2 등고선을 낮춰가면서 68%, 95%, 99.9% 영역 찾기.

Corner-style plot: 1D vs 2D 1σ comparison



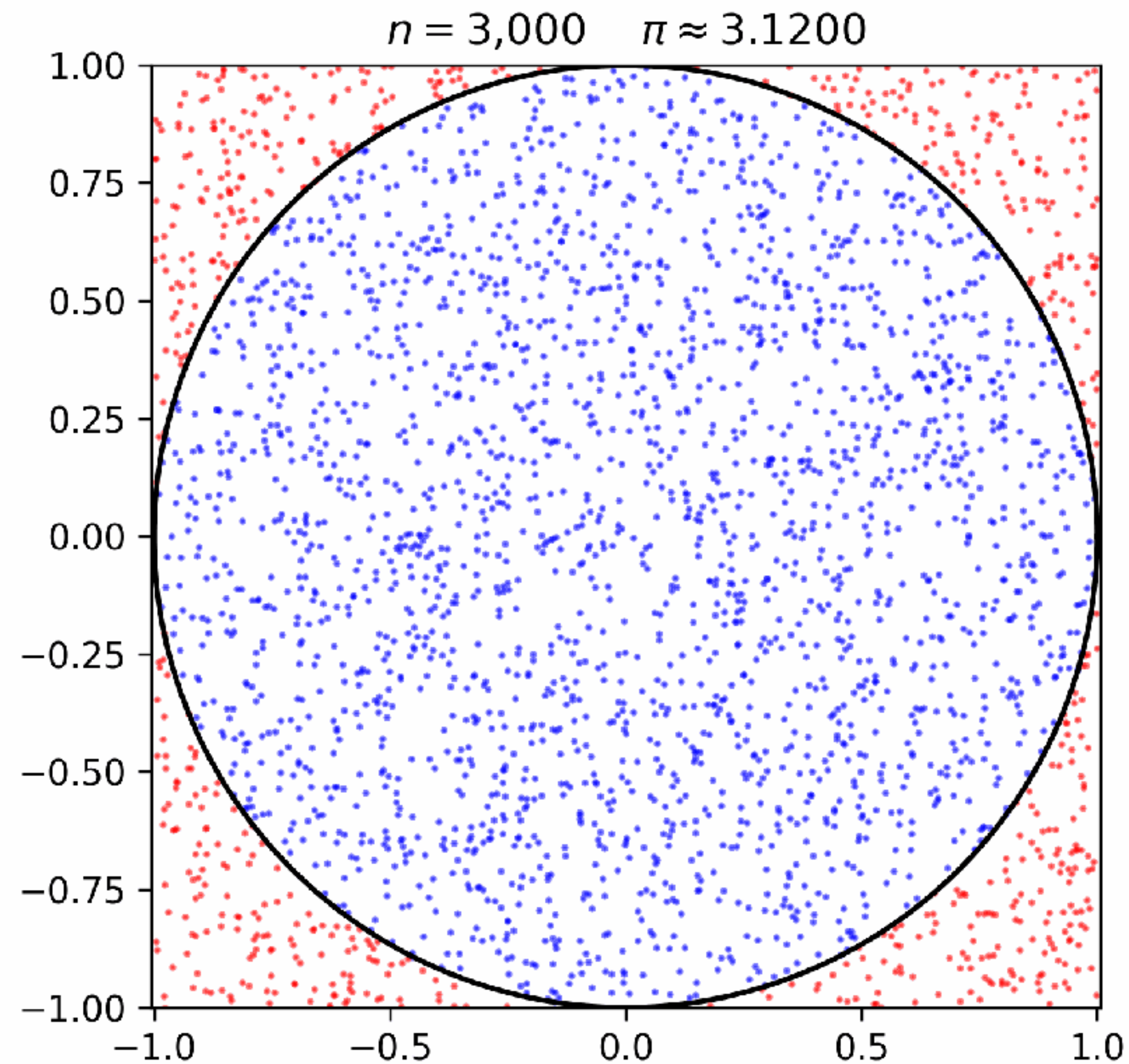
2D vs. 1D pdf

- 2차원 등고선에서 $\Delta\chi^2=1$ 인 타원의 최대 점이 주변화된 1차원 pdf의 1-sigma 영역과 같다.
- 2차원의 1-sigma 영역은 $\Delta\chi^2=2.3$ 으로, 1차원에 비해서 더 넓음.

Markov chain Monte Carlo

- n차원 확률밀도함수를 효율적으로 sampling하는 기법
- Markov 체인: $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_i, \mathbf{x}_{i+1}, \dots\}$
 - 연속된 n차원 벡터의 연결 (체인)인데, $\mathbf{x}_{i+1}$ 의 통계량이 오직 $\mathbf{x}_i$ 에 의해서만 결정됨.
 - 즉, 그 전의 모든 기억을 지우고, 그 전단계에만 의존하는 체인
- Monte-Carlo
 - Random number를 이용해 확률과정을 모사하는 기법.
 - 응용 예) 고차원 적분, 확률밀도함수 sampling

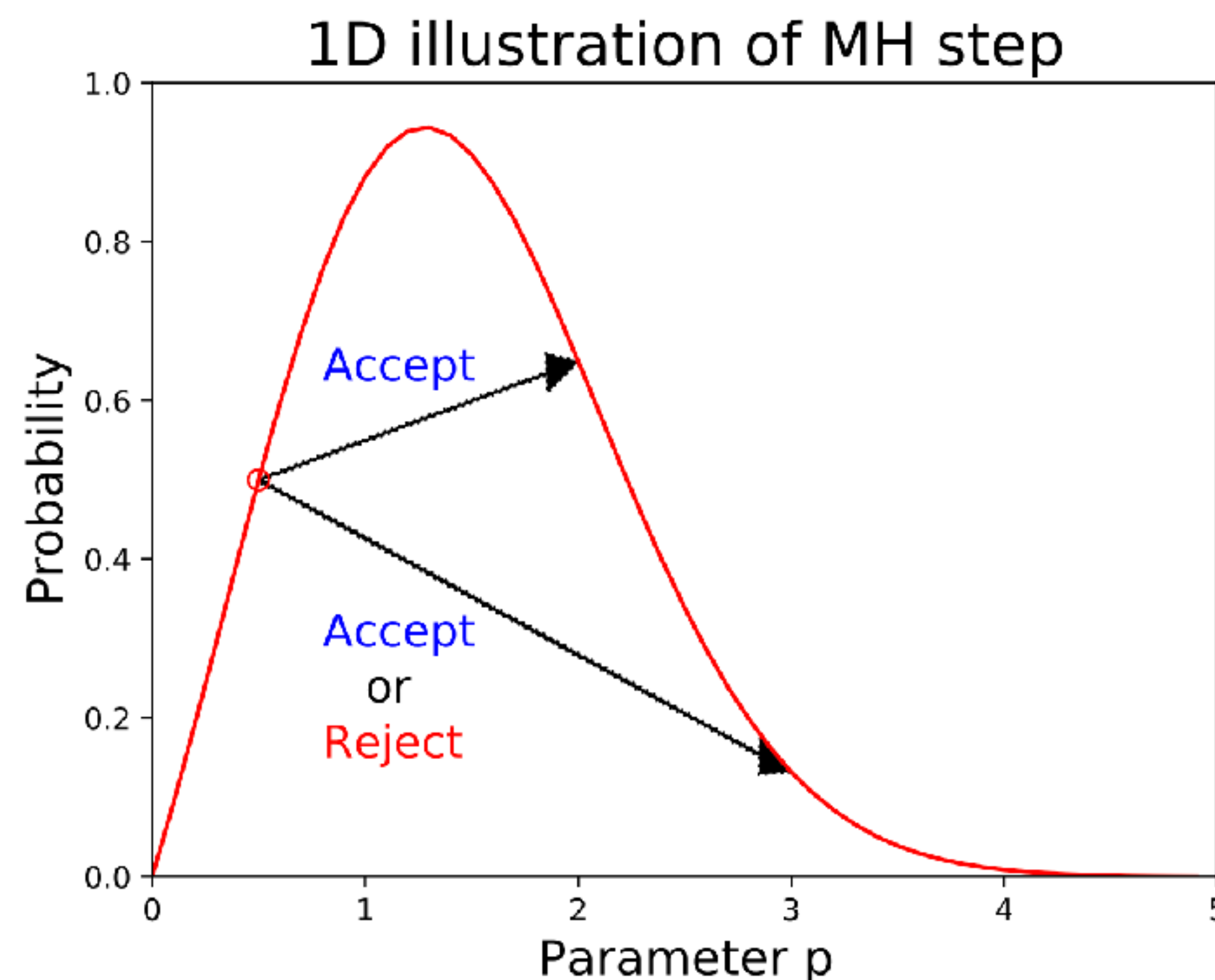
Monte Carlo Integration example



Markov chain Monte Carlo

MCMC: Metropolis-Hastings algorithm

- ▶ choose proposal function $Q(p'|p_t)$
- ▶ at step t , at some parameters p_t
- ▶ propose move to $p' = p_t + \Delta p_t$ (draw Δp_t from Q)
- ▶ evaluate $r = P(p')/P(p_t)$
- ▶ MH step:
 - ▶ if $r > 1$ **accept move**
 - ▶ if $r < 1$ generate random number $\alpha \in [0, 1]$
 - ▶ if $\alpha < r$, **accept move**
 - ▶ if $\alpha > r$, **reject move**
- ▶ $t = t + 1$
- ▶ end when MCMC has converged



MCMC기법들

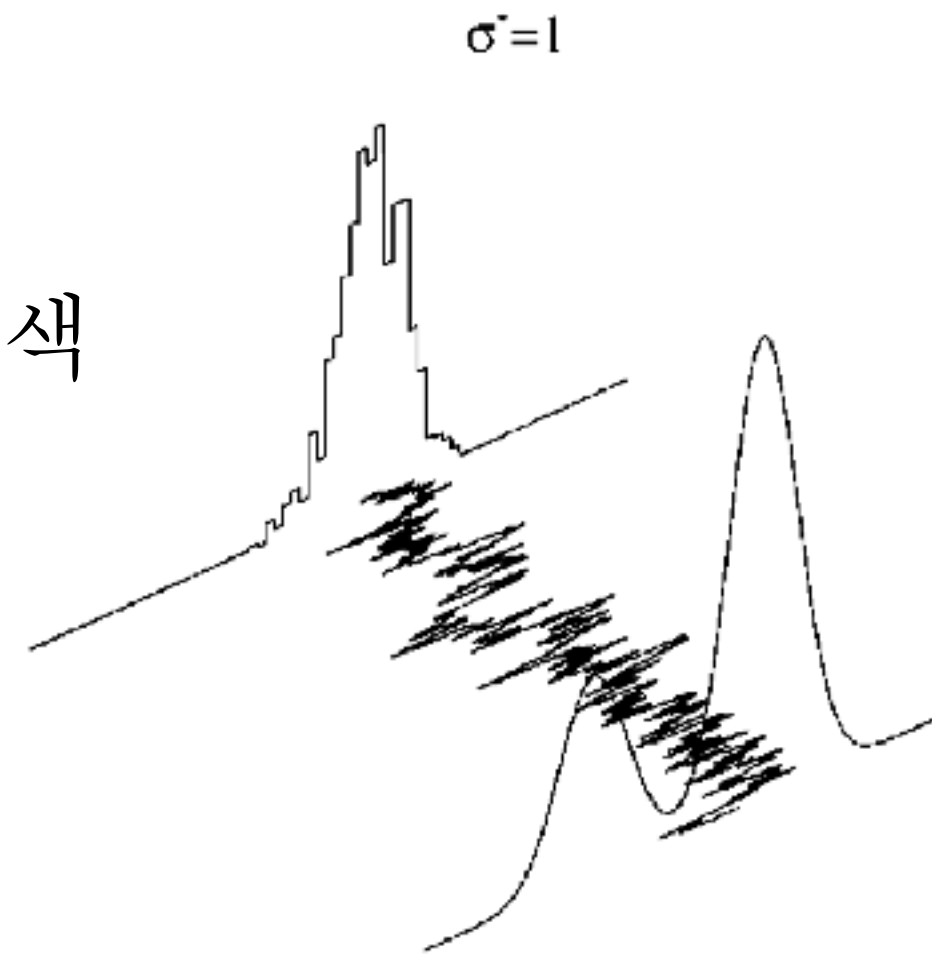
- CosmoMC, mecee, PyMC, Stan (cmdstanpy, Stan.jl)
 - Metropolis-Hastings
 - Gibbs sampling: $P(x_i \mid x_1, x_2, x_3, x_{i-1}, x_{i+1}, x_n)$ 로 매번 1차원씩
 - Hamiltonian Monte Carlo: 해밀턴 역학계에서 $p \sim e^{-E}$ 을 이용
- 원하는 확률밀도함수를 sampling하는 것은 정상상태에서만 가능
- 초기에 계산된 체인들은 수락률이 0.234 (MH확산의 경우), 0.65 - 0.8 (HMC의 경우) 정도가 되도록 MCMC의 계수들을 정하고 태워없앰 (burning phase).

수락률과 MCMC의 성능

너무 작은 step:

수락률~1

꼭대기 주변만 탐색

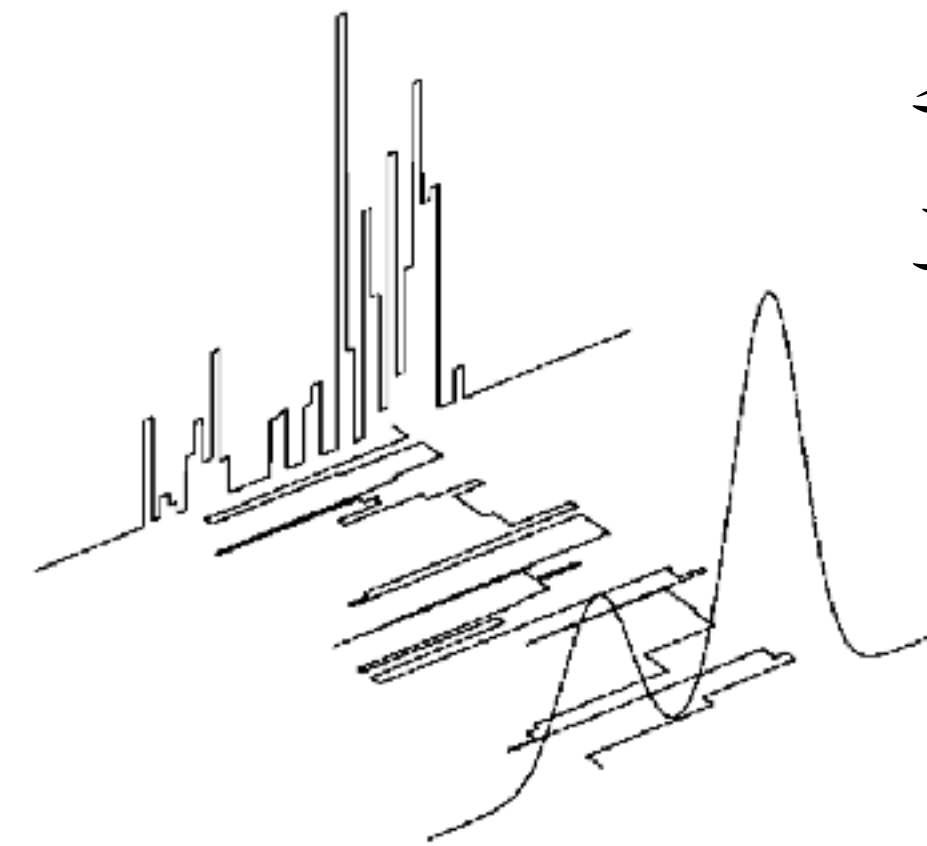


$\sigma^*=100$

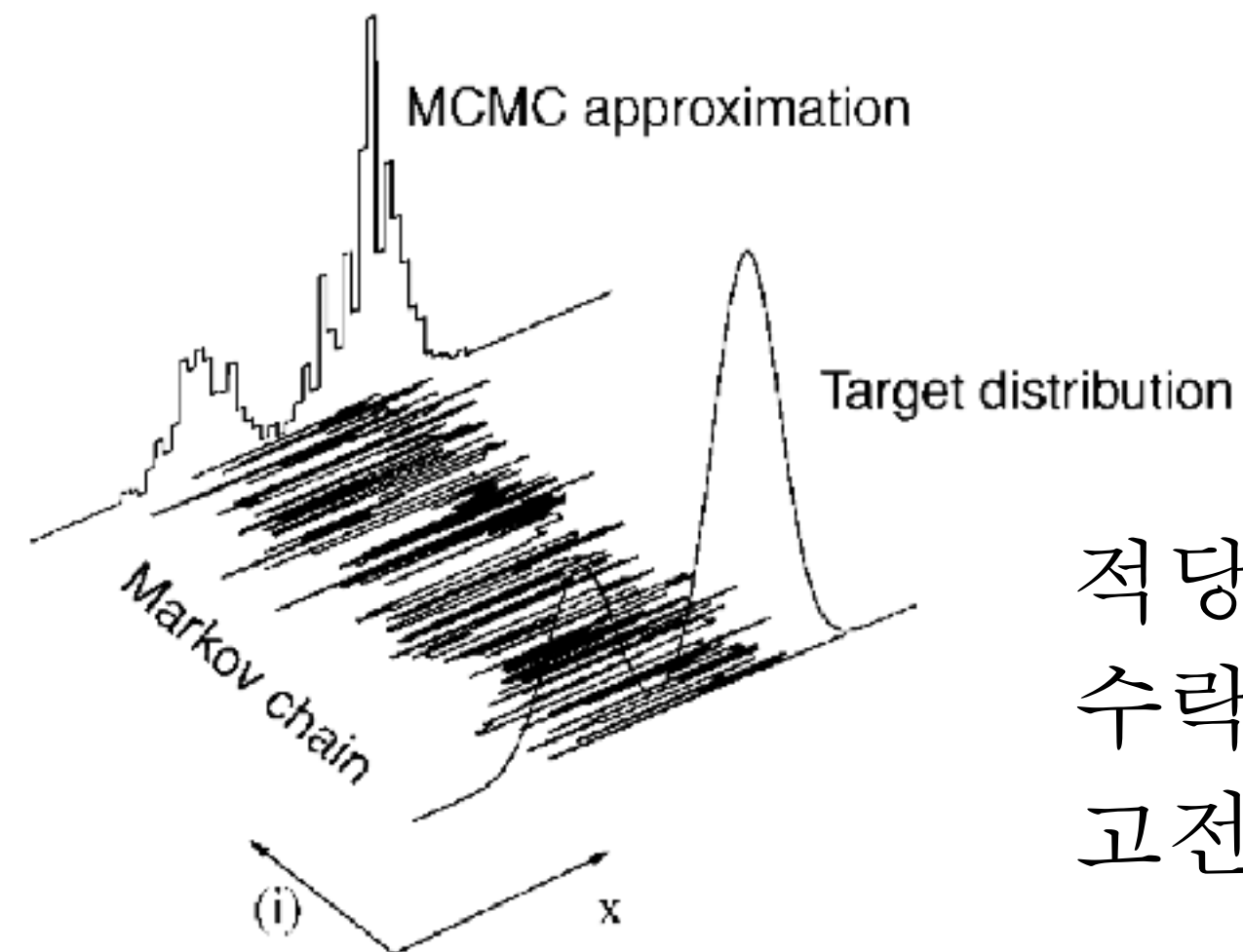
너무 큰 step:

수락률~0

효용성이 떨어짐



$\sigma^*=10$

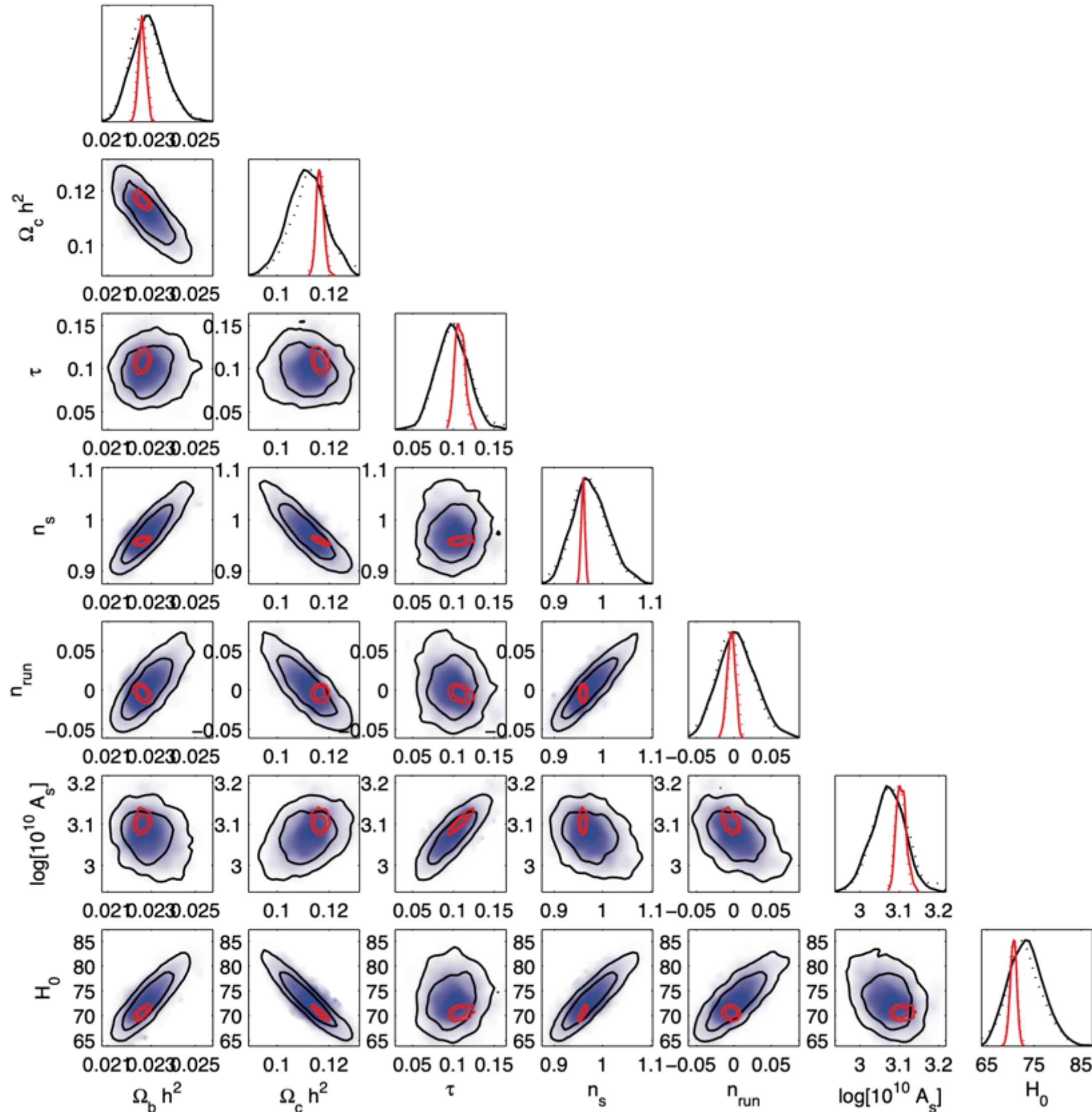


적당한 step:

수락률~0.234

고전적 확산의 경우에 추천되는 값

Planck 청서 (blue book, 2006)



- 푸른색 + 검정 contour:
WMAP 4yr 예측치
- 붉은 contour:
Planck 1yr 예측치
- [질문 1] 어떻게 관측을 시작도 하지 않았는데, 파라미터 측정의 정확도를 예측 할 수 있을까?
- [질문 2] 저 contour들은 어떤 의미를 갖는가?
- [질문 3] 2차원 contour와 1차원 그래프 사이에는 어떤 관계가 있는가?

생각해 볼 문제들

- 왜 이렇게 $\log(\mathcal{L})$ 이 자주 등장할까?
- 우주배경복사는 정규분포로 잘 근사할 수 있어서 저런 likelihood를 쓸 수 있는데, 은하 분포에 대해서는 비슷한 일을 할 수 있나? (Tegmark 1997)
- 여기서는 cosmic variance만 넣었는데, 관측 noise는 어떻게 넣어줄까?
- 우주론 모수추정에서 prior는 어떻게 정해야하나?
- 왜 지금까지 한 모든 우주론적 관측들을 더한 “**CMB (TT+TE+EE+lensing) + BAO + SN 초신성 + 약중력렌즈 + 강중력렌즈 + 암흑물질탐색실험 + 은하단 + 표준크레파스 + 표준 나이**” 이런 모수추정은 없을까?