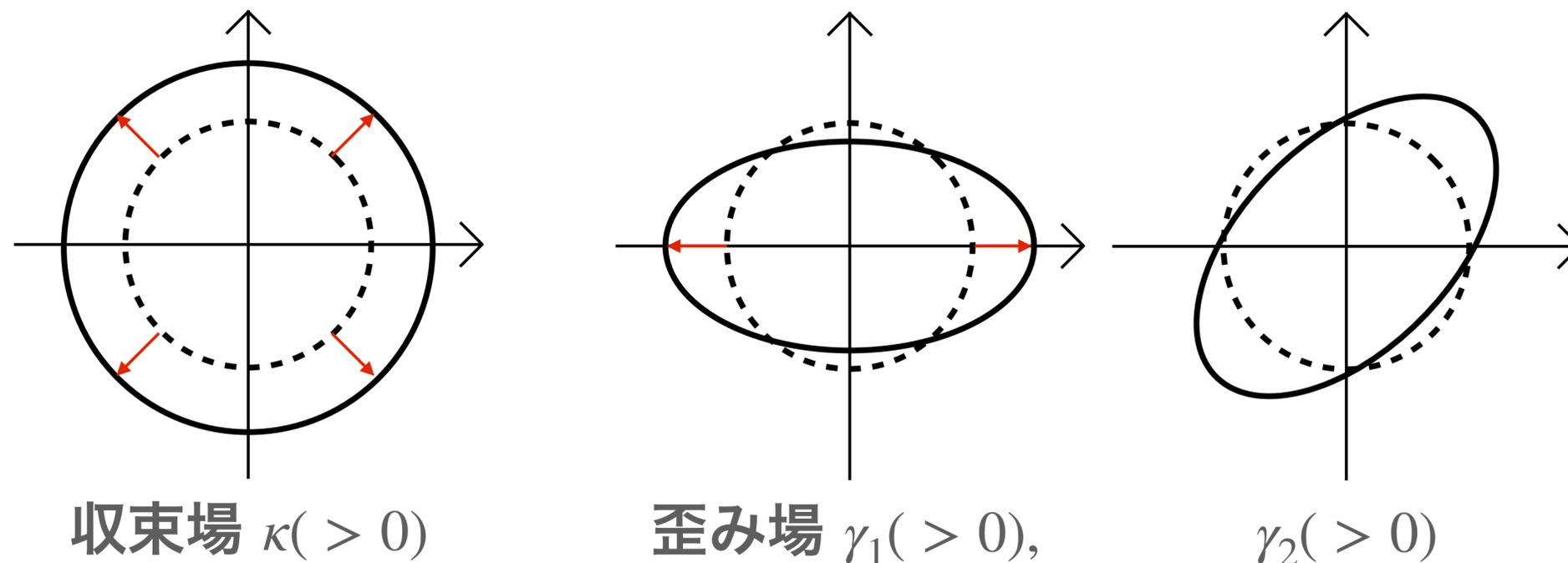


# 機械学習を用いた 重力レンズ収束場マップのノイズ除去

# 1. 重力レンズ

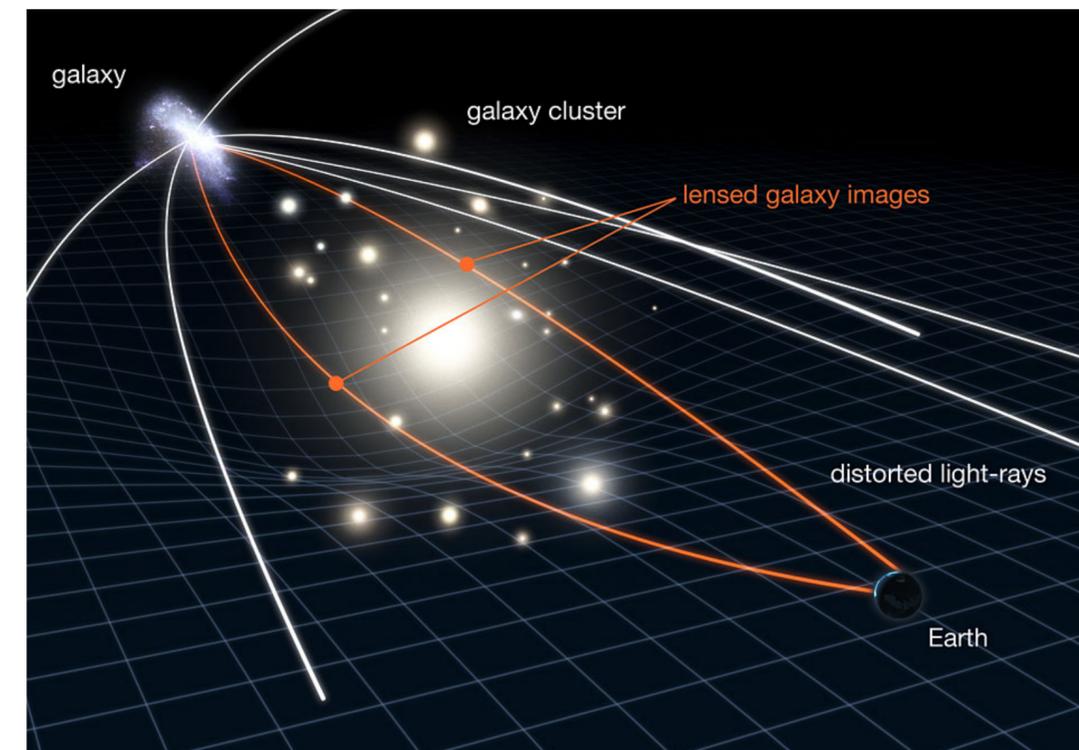
重力によって光の経路が曲げられて、見え方が変化



収束場は質量分布の情報を視線方向に投影したもの

$$\kappa(\theta) \propto \int_0^{x_s} dx \frac{S_K(x_s - x)S_K(x)}{S_K(x_s)} a[t(x)]^2 \bar{\rho}[t(x)] \delta[x, \theta, t(x)]$$

$\delta$  : 密度揺らぎ,  $\bar{\rho}$  : 平均密度,  $S_K(x)$  : 曲率を考慮した  $x$  までの距離,  $a$  : スケール因子



重力レンズが起こる様子

NASA <https://esahubble.org/images/heic1106c/>

# 1. 重力レンズ

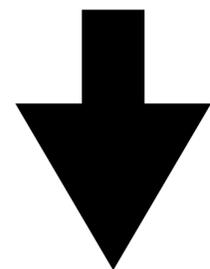
楕円率 $\epsilon_{\text{obs}}$ は測定できるが

光源自身の持つ固有の楕円率 $\epsilon_{\text{source}}$ も存在

$$\epsilon_{\text{obs}} = \epsilon_{\text{source}} + \gamma$$

$(\gamma = \gamma_1 + i\gamma_2)$

近くの光源は同じ $\gamma$ を持つ



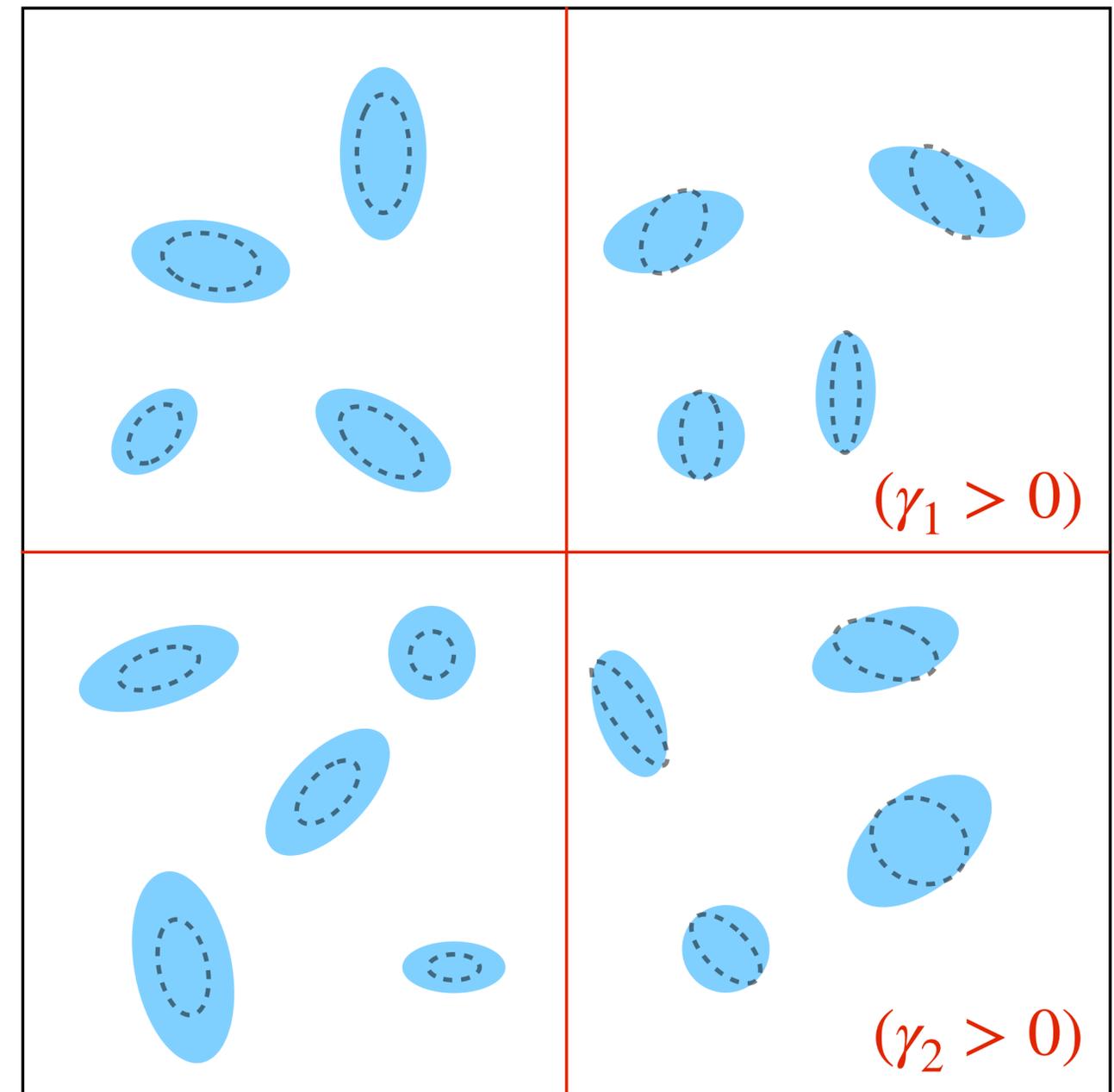
狭い範囲に銀河が多い

$$\langle \epsilon_{\text{source}} \rangle \rightarrow 0, \quad \langle \epsilon_{\text{obs}} \rangle \rightarrow \gamma$$

実際は銀河の数が不十分

$$\langle \epsilon_{\text{obs}} \rangle = \epsilon_N + \gamma (= \gamma_{\text{obs}})$$

$\epsilon_{\text{source}}$ によりシェイプノイズ( $\epsilon_N$ )が現れる



銀河の観測

# 1. 重力レンズ

$\gamma \rightarrow \kappa$  の変換は角度のフーリエ空間で簡単に表される

収束場にも同様にシェイプノイズが現れる

$$\tilde{\kappa}(\ell) = \frac{\ell^2}{(\ell_1 + i\ell_2)^2} \tilde{\gamma}(\ell)$$

$$\kappa_{\text{obs}} = \kappa_N + \kappa$$

このシェイプノイズ( $\kappa_N$ )を取り除きたい

ノイズを除去する手法

- $\kappa_{\text{obs}}$  にフィルター(Wiener filter など)を適用
- 機械学習を適用

**機械学習が有効** (Shirasaki+2019)

## 2. 研究目的

Masked Input



GAN

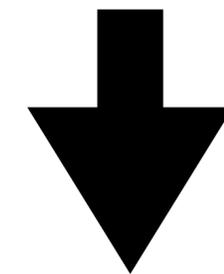


Palette (拡散モデル)



先行研究でのモデル  
GAN(敵対的生成ネットワーク)

GANよりも優れた画像生成モデル  
Palette Diffusion model  
(拡散モデル)



収束場画像に適用

Saharia+2022

シェイプノイズを精度よく除去できるか

Masked Input



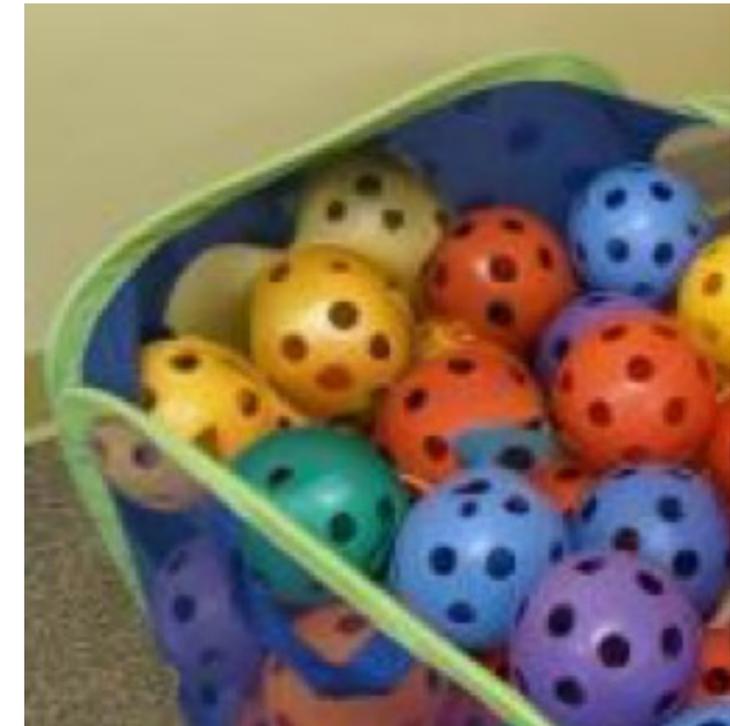
GAN



Palette (拡散モデル)



Saharia+2022



# 3. 機械学習モデル

## GAN (敵対的生成ネットワーク)

損失関数 (Isola+2017)

$$L_{\text{GAN}} = \min_G \max_D \{L_1(G, D) + \lambda L_2(G)\}$$

G(生成子) : 偽画像を生成

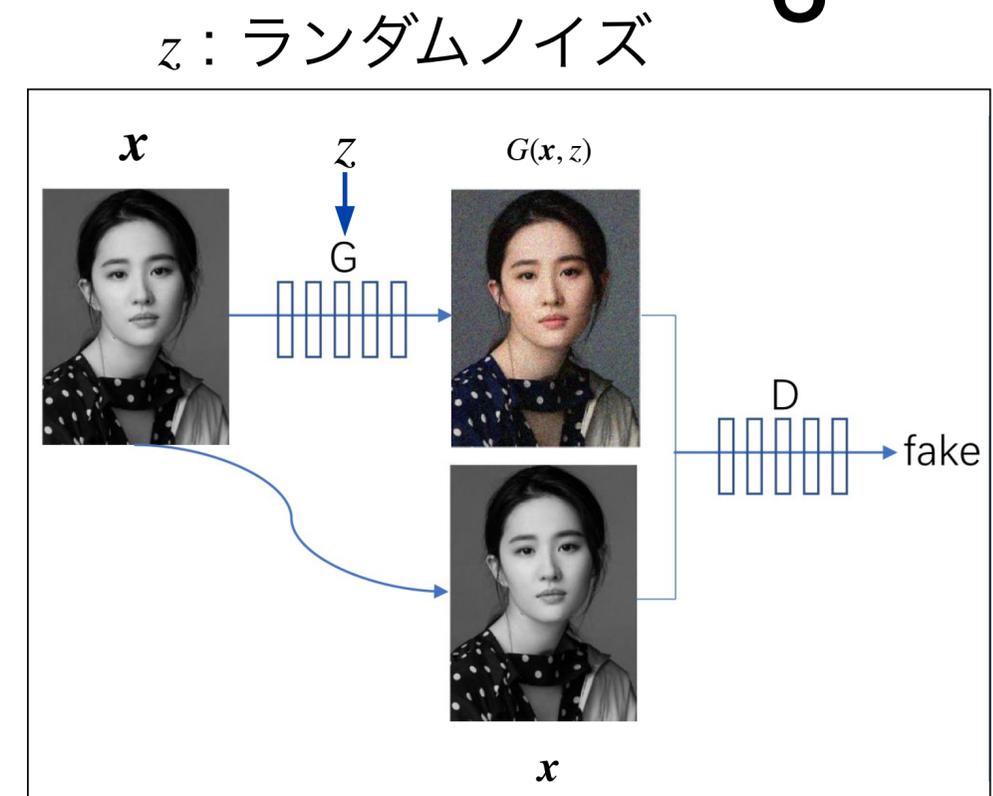
D(識別子) : 本(偽)物か判別 (1 ↔ 0)

損失関数に従って精度の良い生成子, 識別子を作成

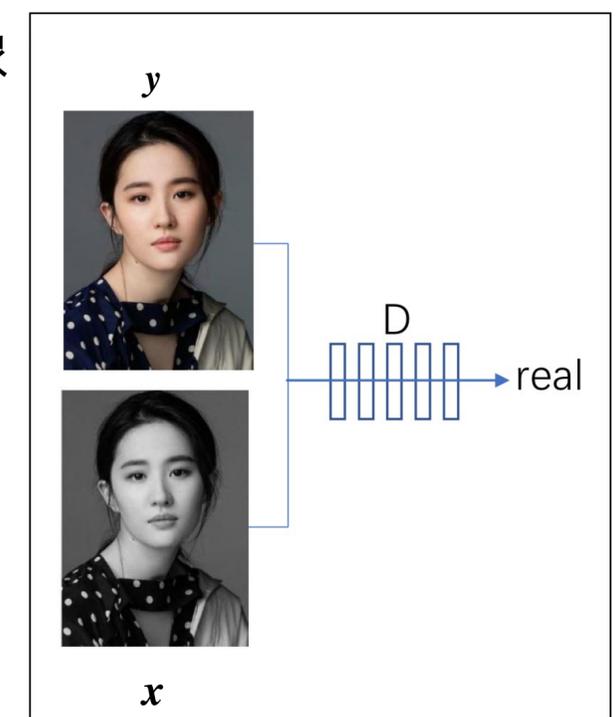
$$L_1(G, D) = \mathbb{E}_{x,y} \log D(x, y) + \mathbb{E}_{x,z} \log \{1 - D(x, G(x, z))\}$$

$$L_2(G) = \mathbb{E}_{x,y,z} \sum_{\text{pixels}} |y - G(x, z)|$$

$x$  : 入力画像



$y$  : 教師画像



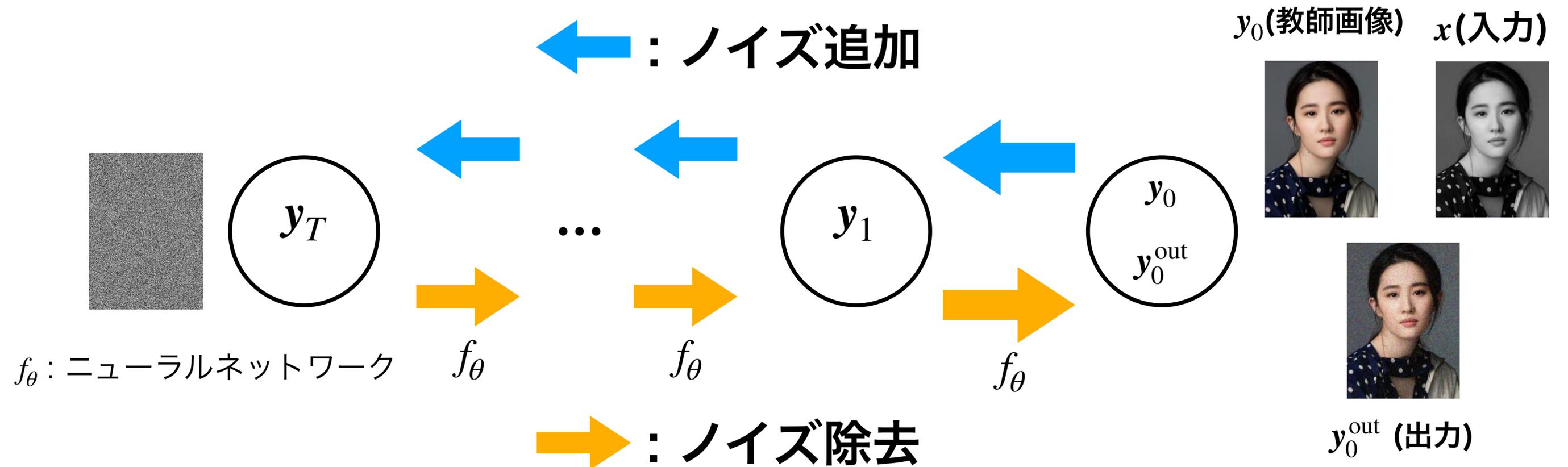
# 3. 機械学習モデル

## Diffusion model (拡散モデル)

損失関数 (Saharia+2022)

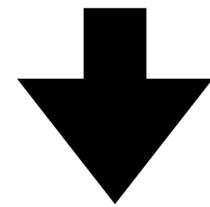
$$L_{\text{Diffusion}} = \min_{\theta} \mathbb{E}_{(x,y)} \mathbb{E}_{\epsilon,\gamma} \|f_{\theta}(x, \sqrt{\gamma}y_0 + \sqrt{1-\gamma}\epsilon, \gamma) - \epsilon\|$$

$f_{\theta}$ : ニューラルネットワーク,  $\epsilon$ : ノイズ,  $\gamma$ : ノイズの強さ



# 4. 手法 シミュレーション結果による収束場マップの作成 (Osato+2021)

IllustrisTNG流体シミュレーションの密度分布

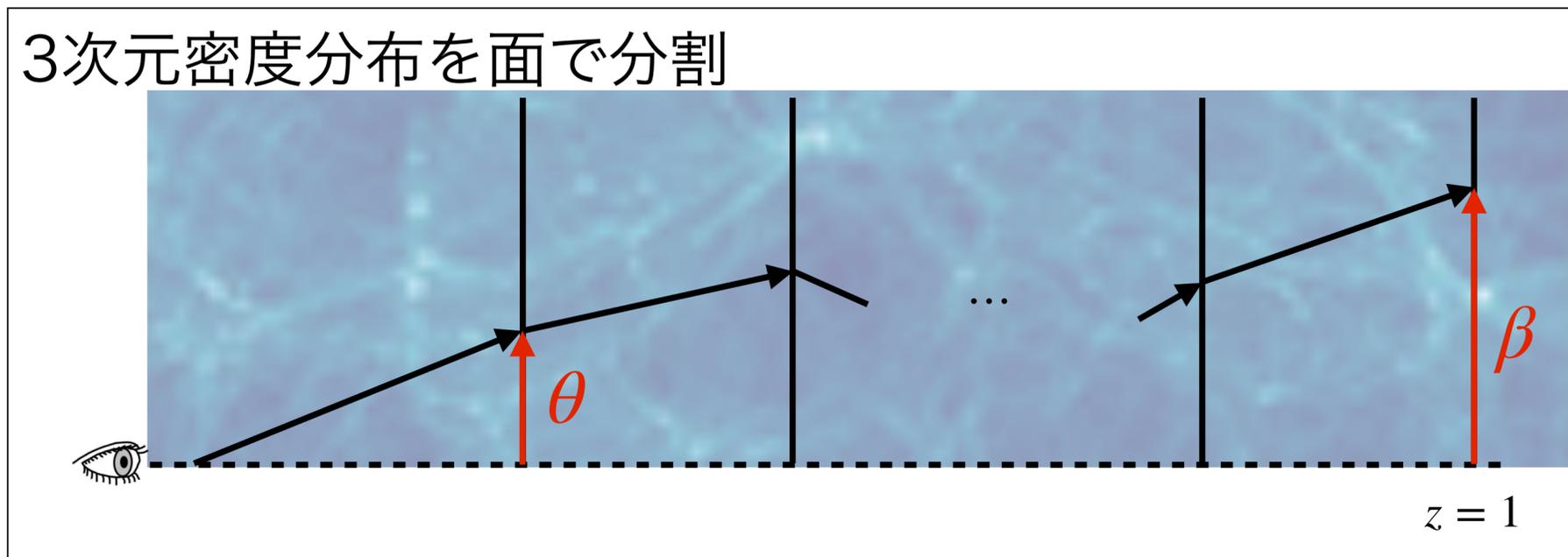


光の曲がりを計算

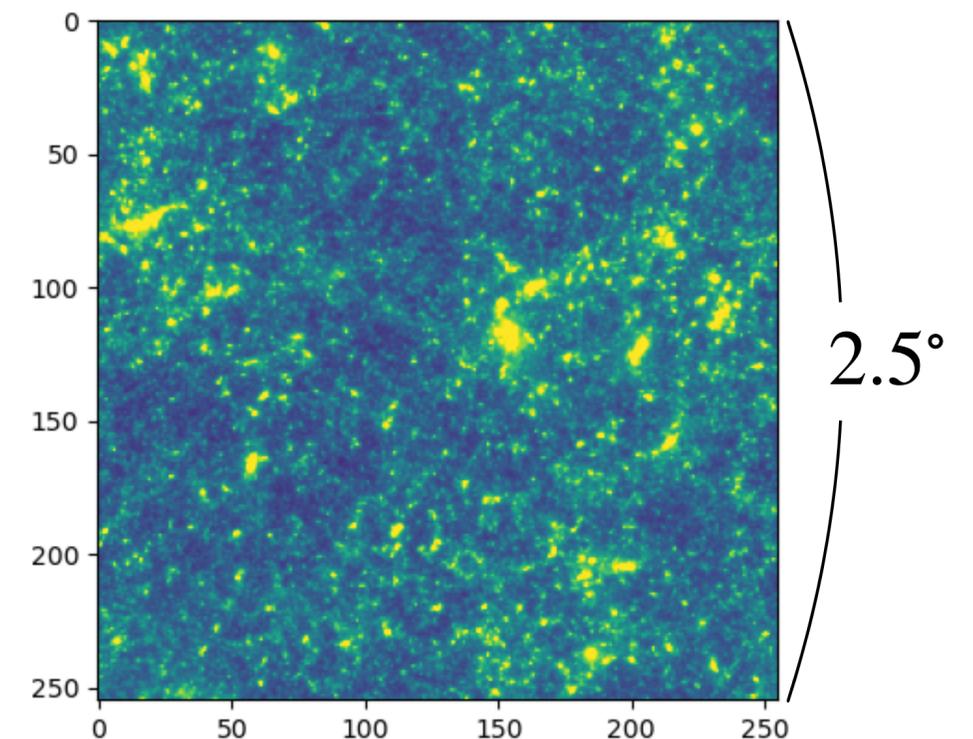
ヤコビ行列A (収束場 $\kappa$ )を計算

$$A = \left( \frac{\partial \beta_i}{\partial \theta_j} \right) = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}$$

収束場マップを40,000枚作成  
そのうち4,000枚を使用



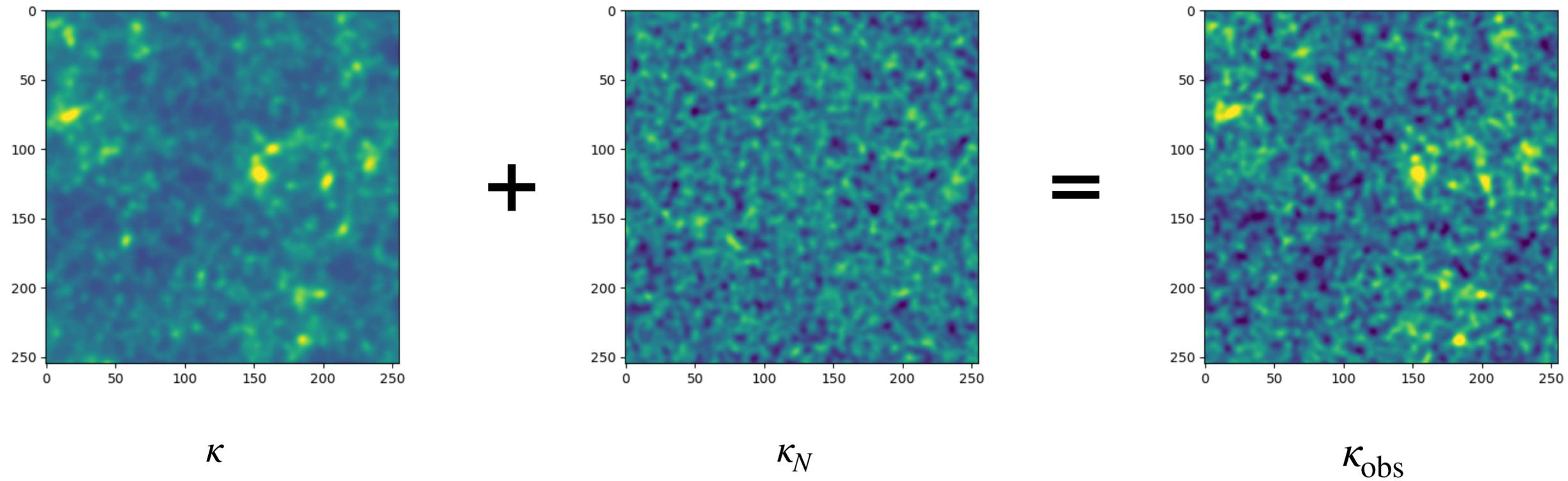
$\theta$  : 見かけの角度,  $\beta$  : 実際の角度



収束場マップ



# 4. 手法 スムージング



## フィルター関数

小スケールのノイズを抑えるフィルター関数

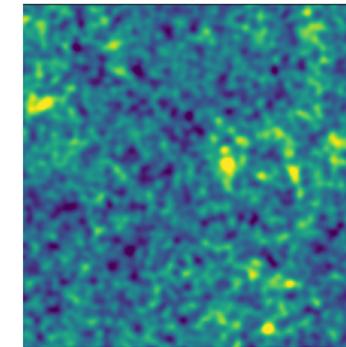
$$W(\theta) = \frac{1}{\pi\theta_G^2} \exp\left(-\frac{\theta^2}{\theta_G^2}\right)$$

$\theta_G = 1.5$  [arcmin] : スムージングスケール  
銀河団 ( $z \sim 1$ ) が観測できる値

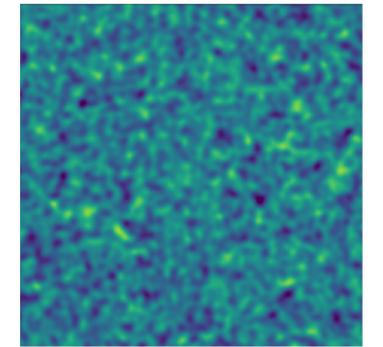
# 4. 手法

先行研究 (GAN)	スムージング	入出力
<b>smooth</b>	○	$\kappa_{\text{obs}} \rightarrow \hat{\kappa}_N$
direct	○	$\kappa_{\text{obs}} \rightarrow \hat{\kappa}$
No smooth	×	$\kappa_{\text{obs}} \rightarrow \hat{\kappa}_N$

先行研究 (GAN) から **smooth** での機械学習が**精度良い**



$$\kappa_{\text{obs}} \rightarrow \hat{\kappa}_N$$



$\kappa_{\text{obs}} \rightarrow \hat{\kappa}_N$  の場合は  $\hat{\kappa}$  に変換

$$\hat{\kappa} = \underbrace{(\kappa + \kappa_N)}_{\kappa_{\text{obs}}} - \hat{\kappa}_N$$

訓練データ 3600枚  
 テストデータ 400枚 } 4000枚

本研究では

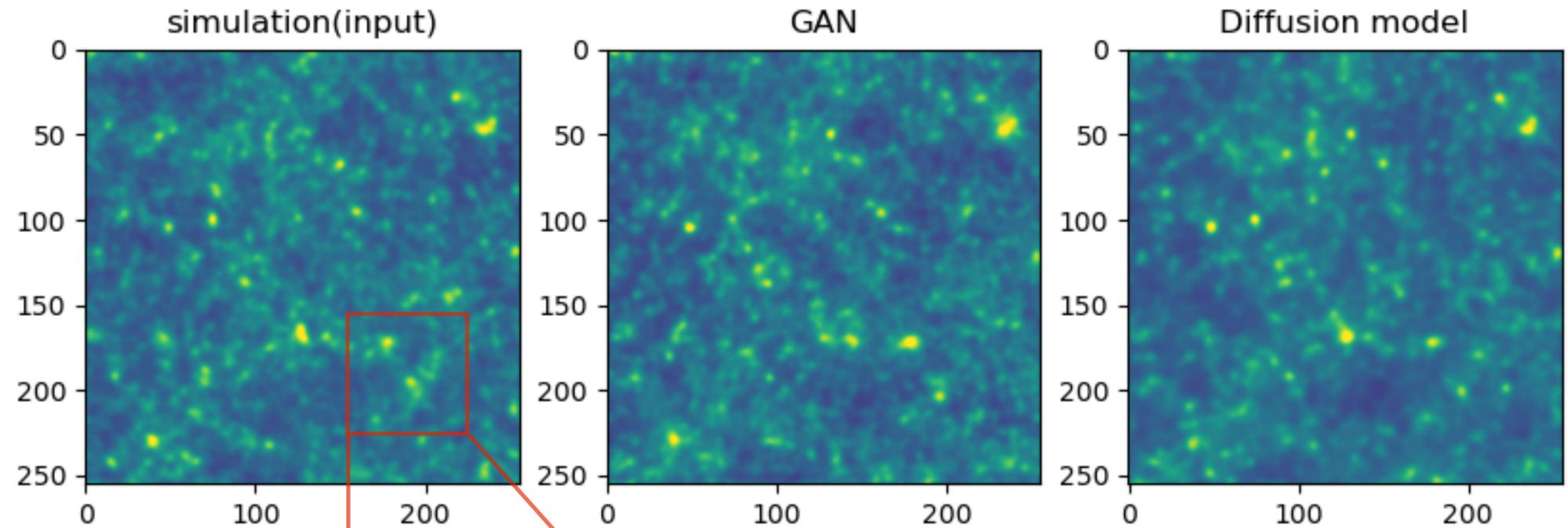
- **smooth** での Palette(拡散モデル) と GAN の比較
- スムージングの有無, 出力の違いによる結果の変化

# 5. 結果

## 出力画像の比較

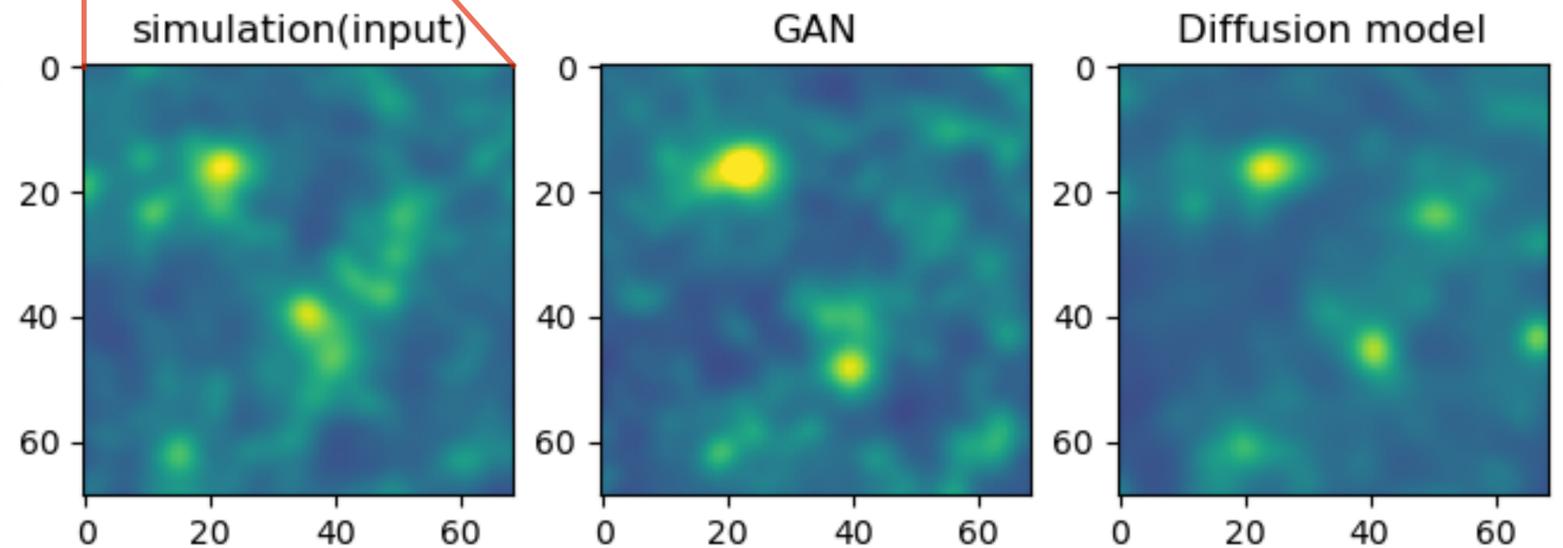
全体の構造は似ている

細かい違いが多い



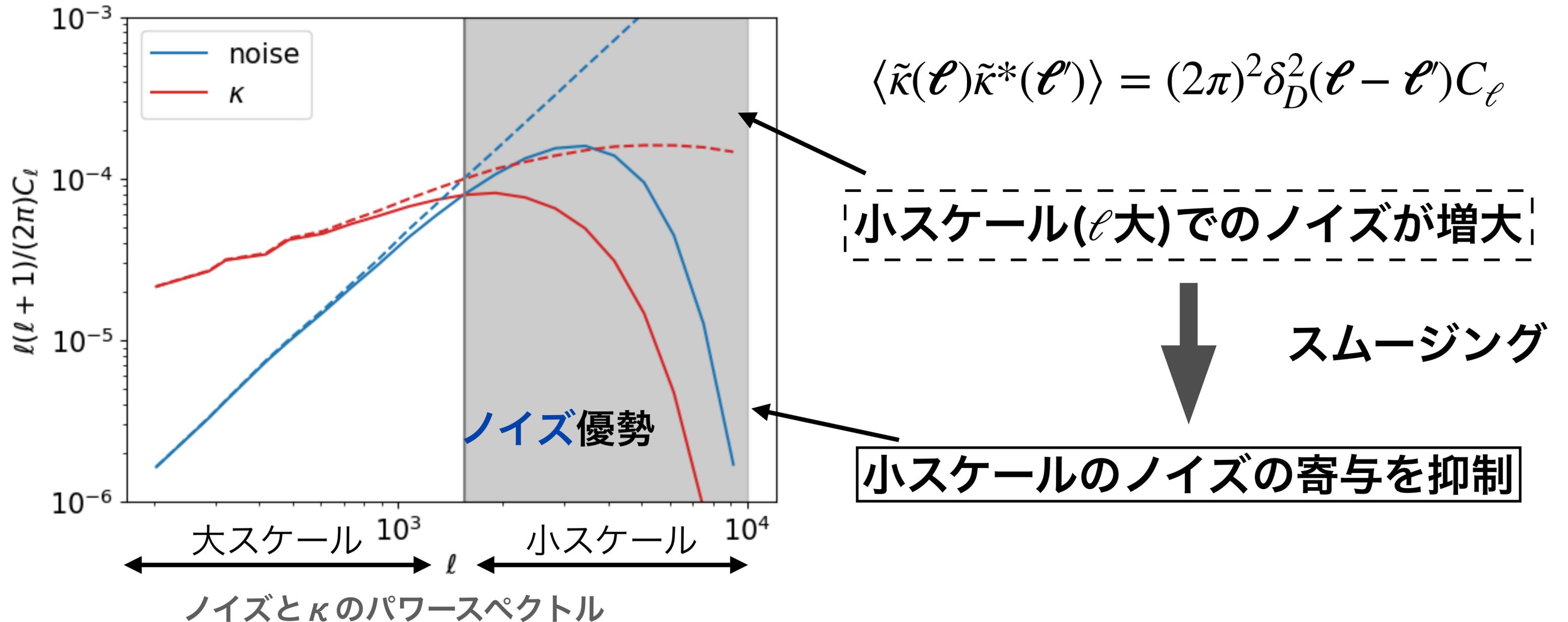
以下の統計量を計算し、評価を行う

- パワースペクトル
- PDF(probability density function)
- peak, minima



# 5. 結果 パワースペクトルについて

パワースペクトル  $C_\ell$  はフーリエ空間の二点相関関数  $\langle \tilde{\kappa}(\ell) \tilde{\kappa}'(\ell') \rangle$  により表される



# 5. 結果 peak, minimaについて

1	1	2	4	2
1	7	2	1	6
5	3	6	3	4
4	0	8	2	1
1	1	5	-1	1
3	-1	2	4	2

## Peak

→ 周りの8個のセルよりも値が**大きい**  $\kappa$  の値をカウント

## minima

→ 周りの8個のセルよりも値が**小さい**  $\kappa$  の値をカウント

Peakの  $\kappa$  の方向に質量が多く存在

→  $(S/N \equiv ) \kappa / \sigma_{noise} > 4$  の数は

ダークマターハローの数に関係

minimaはその方向に質量が少ない  
(ボイド)

統計量 peak, minimaは宇宙論パラメータや宇宙の構造決定に関与

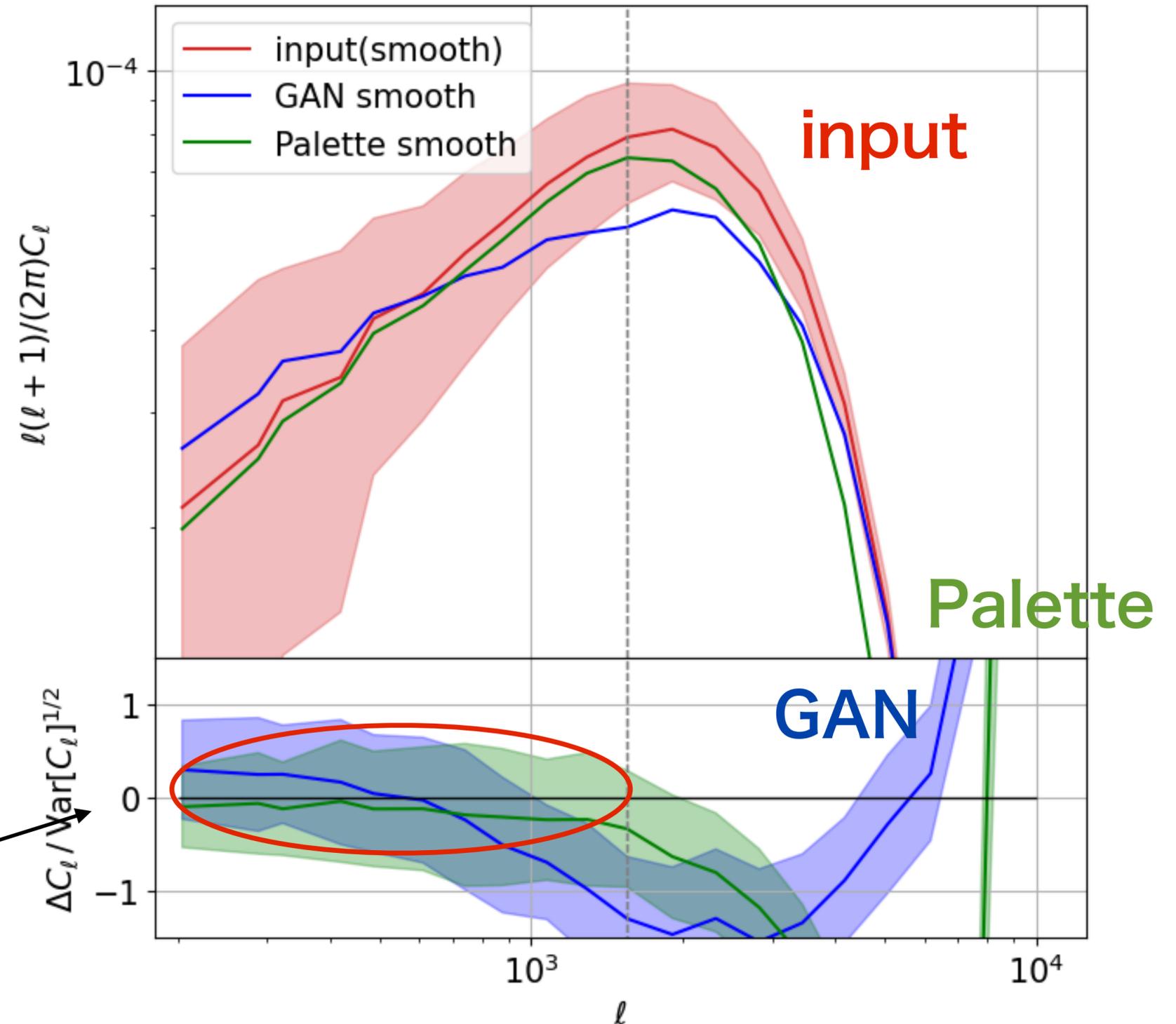
# 5. 結果 smoothでのモデル比較 パワースペクトル

- ノイズの小さい領域では Paletteの方が精度が良い
- ノイズ優勢領域では どちらのモデルも精度が悪い

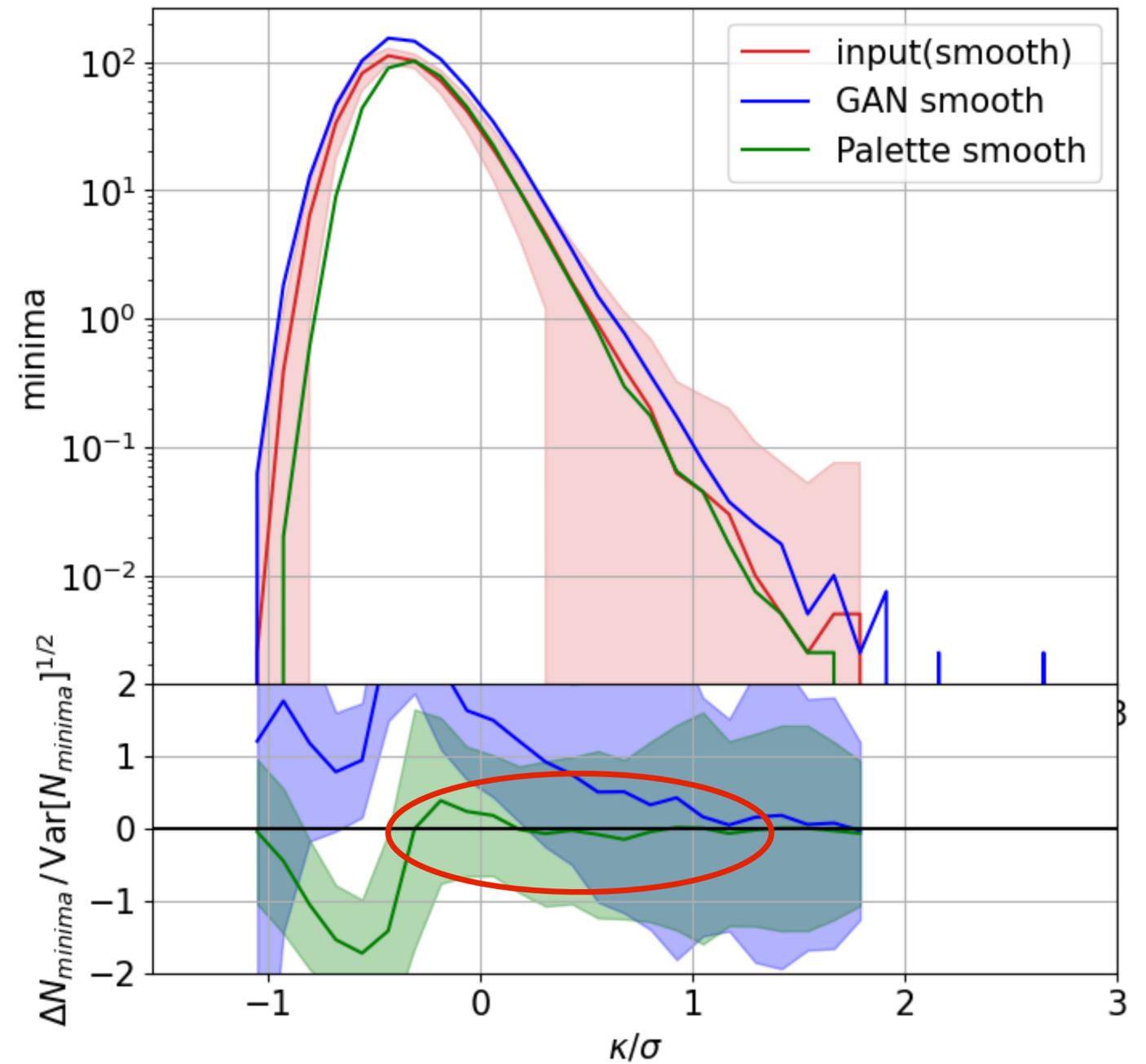
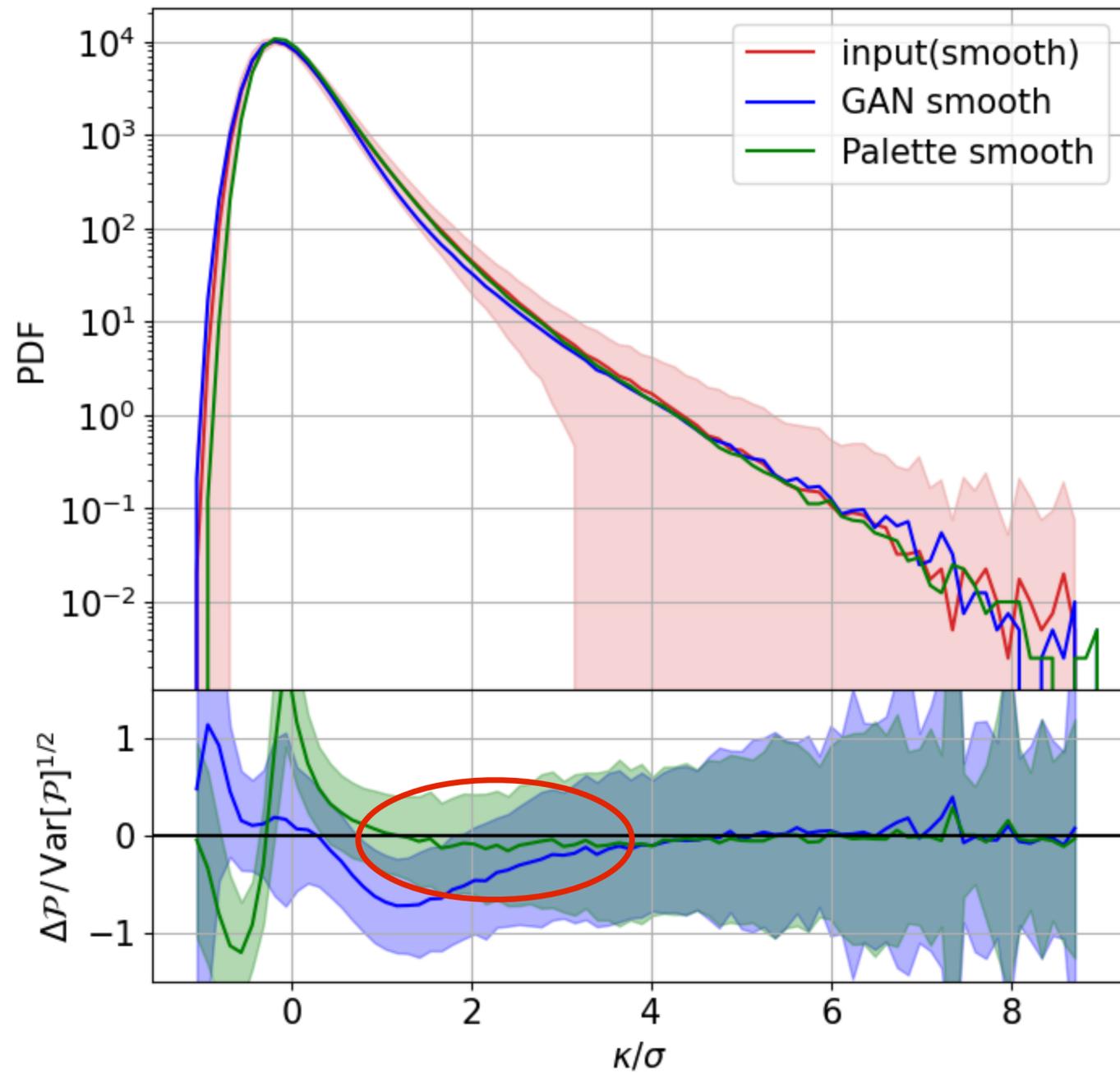
画像ごとの統計量の差

$$\frac{C_{\text{model}}^i - C_{\text{input}}^i}{\text{std}[C_{\text{input}}]}, i = 1, 2, \dots, 400$$

0に近いほどinputに近い



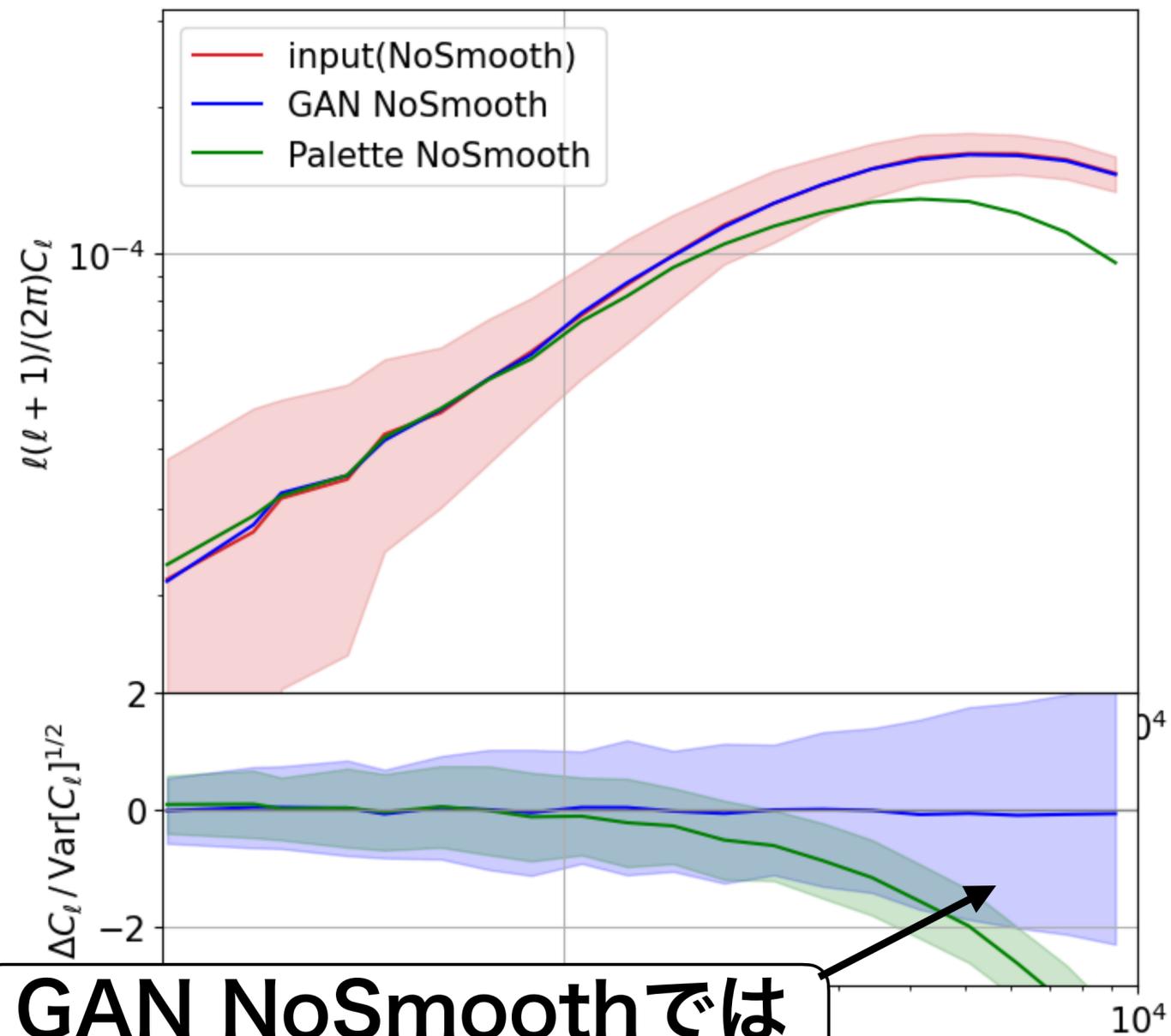
# 5. 結果 smoothでのPDF, minimaのモデル比較



PDF, minima とともにPaletteの方がすぐにinputに近づいている(peakも同様)

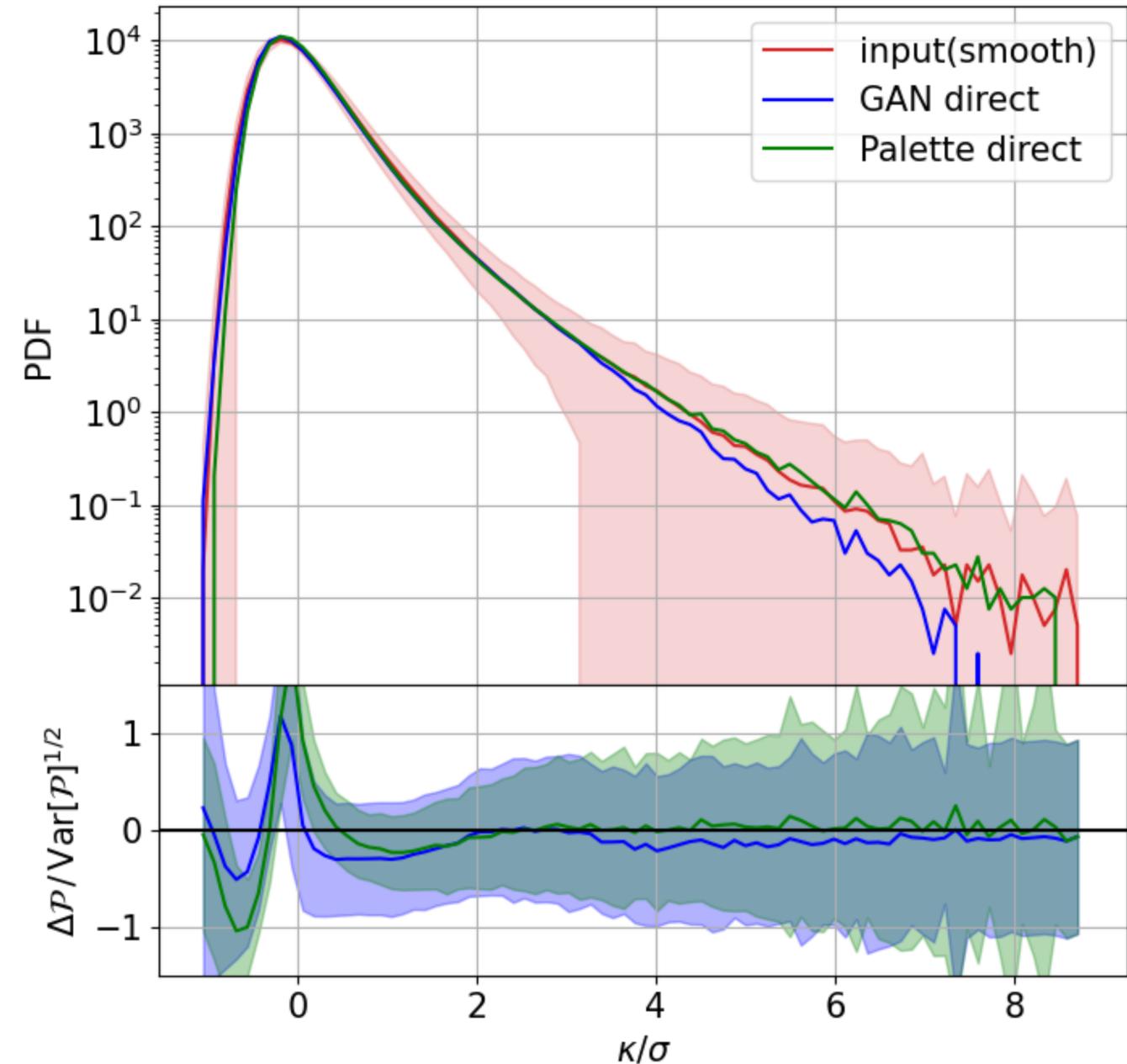
# 5. 結果 出力の違い, スムージングの有無による変化

NoSmooth



GAN NoSmoothでは  
標準偏差が大きい

direct



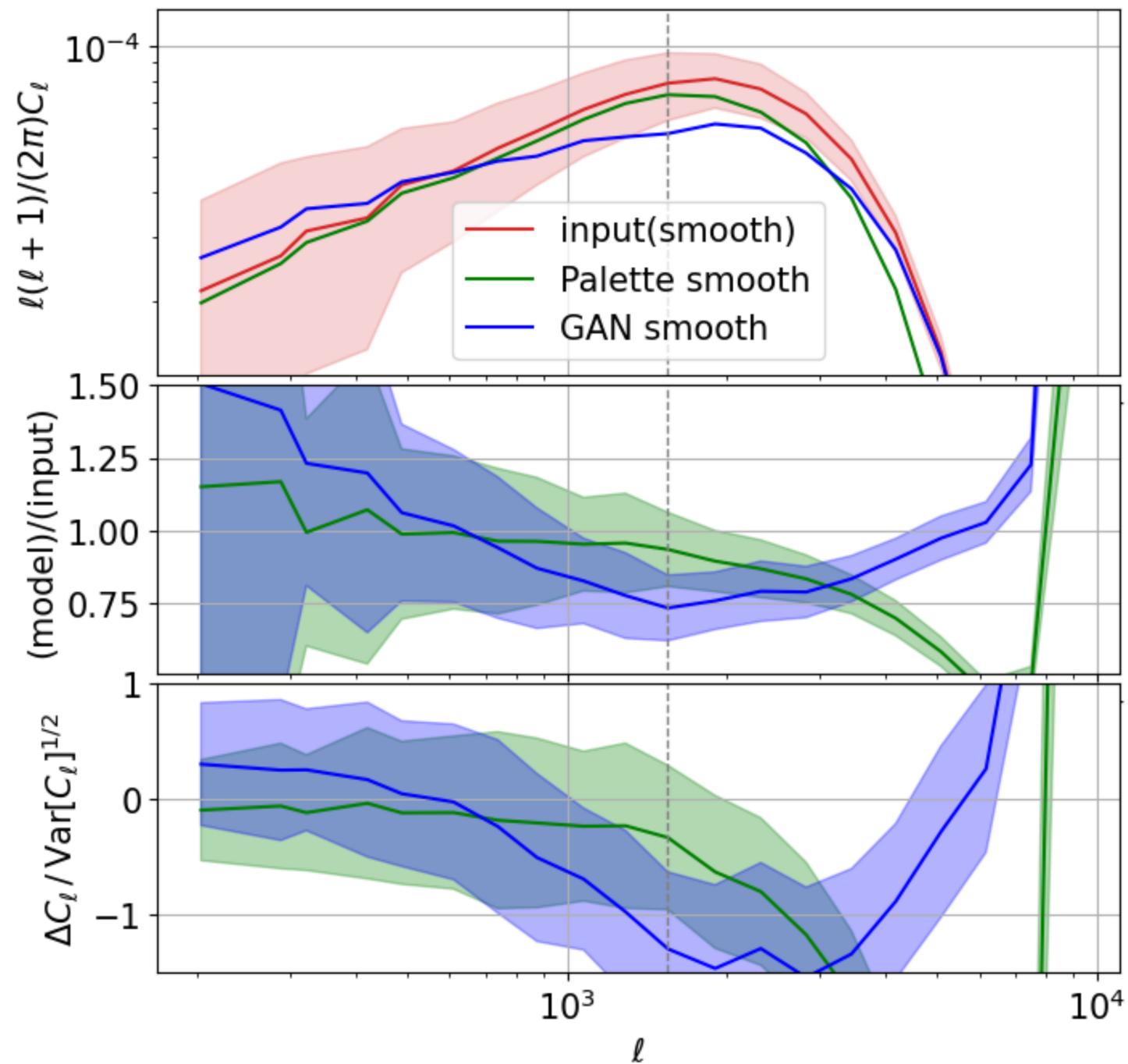
Palette directは基準の場合と大きな差は見られない

## 6. まとめ

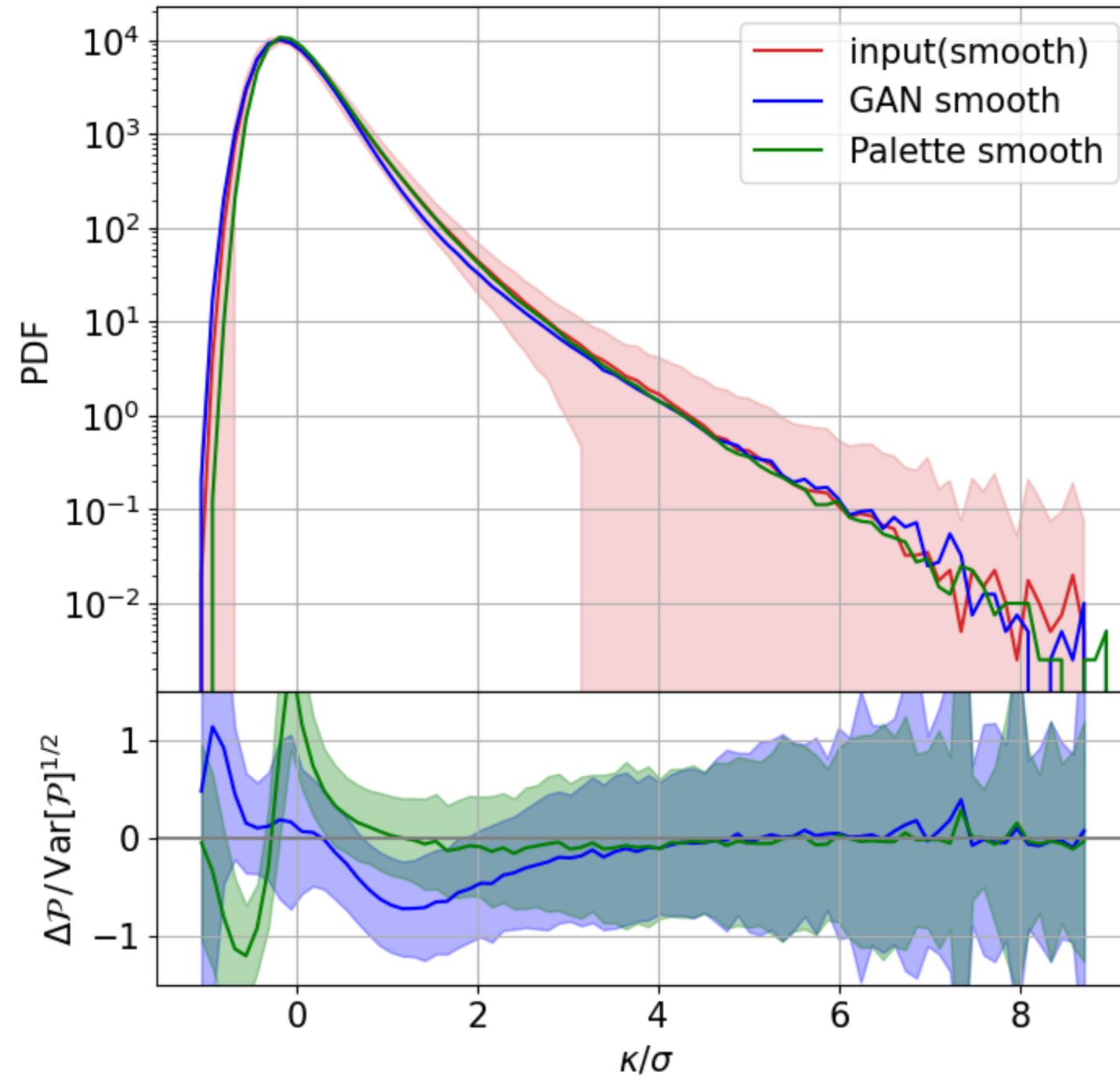
- 観測した収束場画像にはシェイプノイズが含まれる
- シェイプノイズを取り除く手法として機械学習が有効
- 学習モデルとして 先行研究では敵対的生成ネットワーク (GAN),  
**本研究では拡散モデル(diffusion model) を使用**
- 先行研究(スムージング有,  $\kappa_{obs} \rightarrow \hat{\kappa}_N$ )と同じ入出力での比較では,  
**拡散モデルの方が精度が良い結果になった**
- **GAN NoSmoothが平均では精度が良いが, 画像ごとのズレが大きい**  
これから
  - 学習モデルの調整や、学習データを増やすなど進めていく
  - 他の統計量で検証 (3点相関関数など)

# smooth

## パワースペクトル

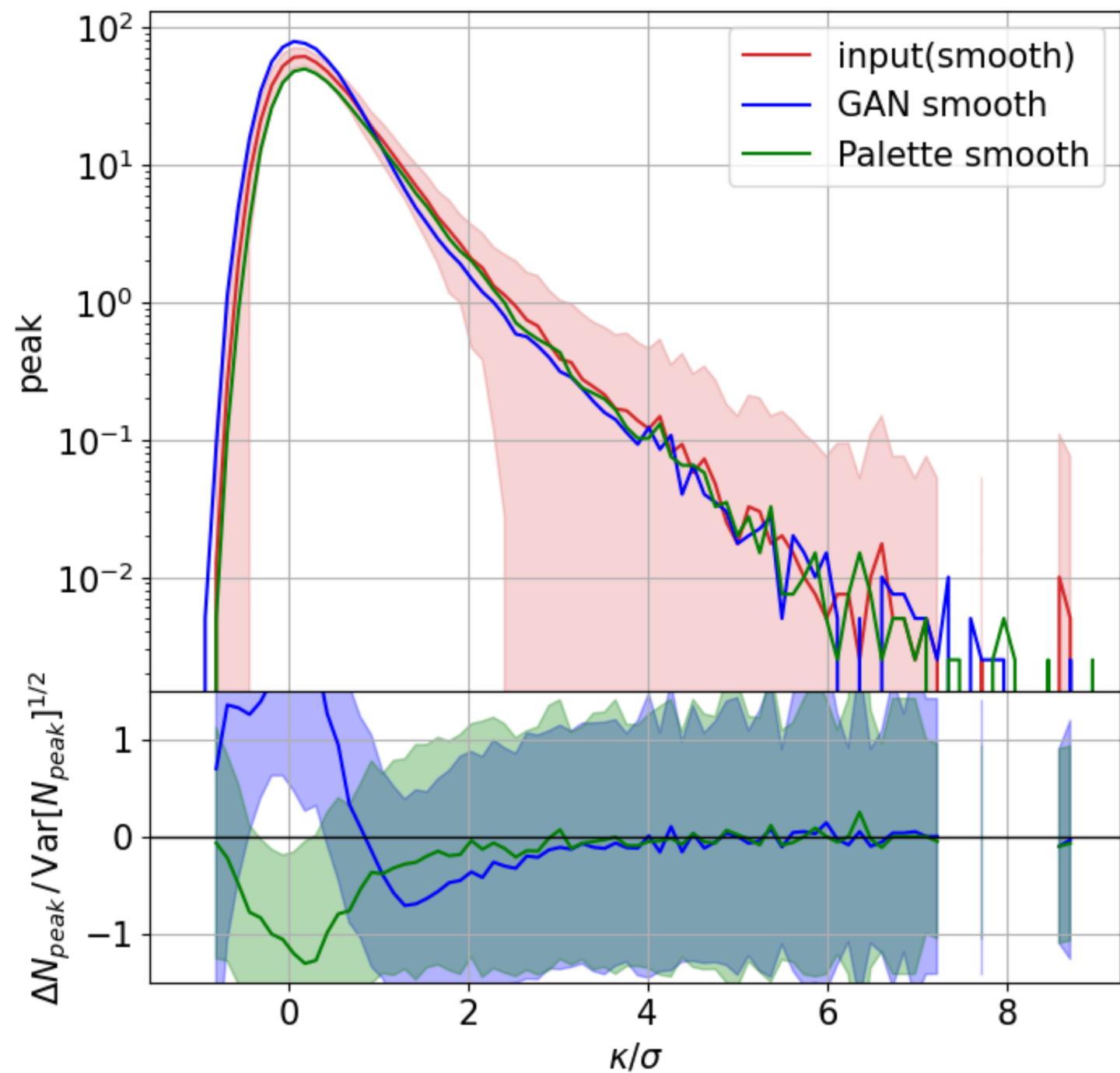


## PDF

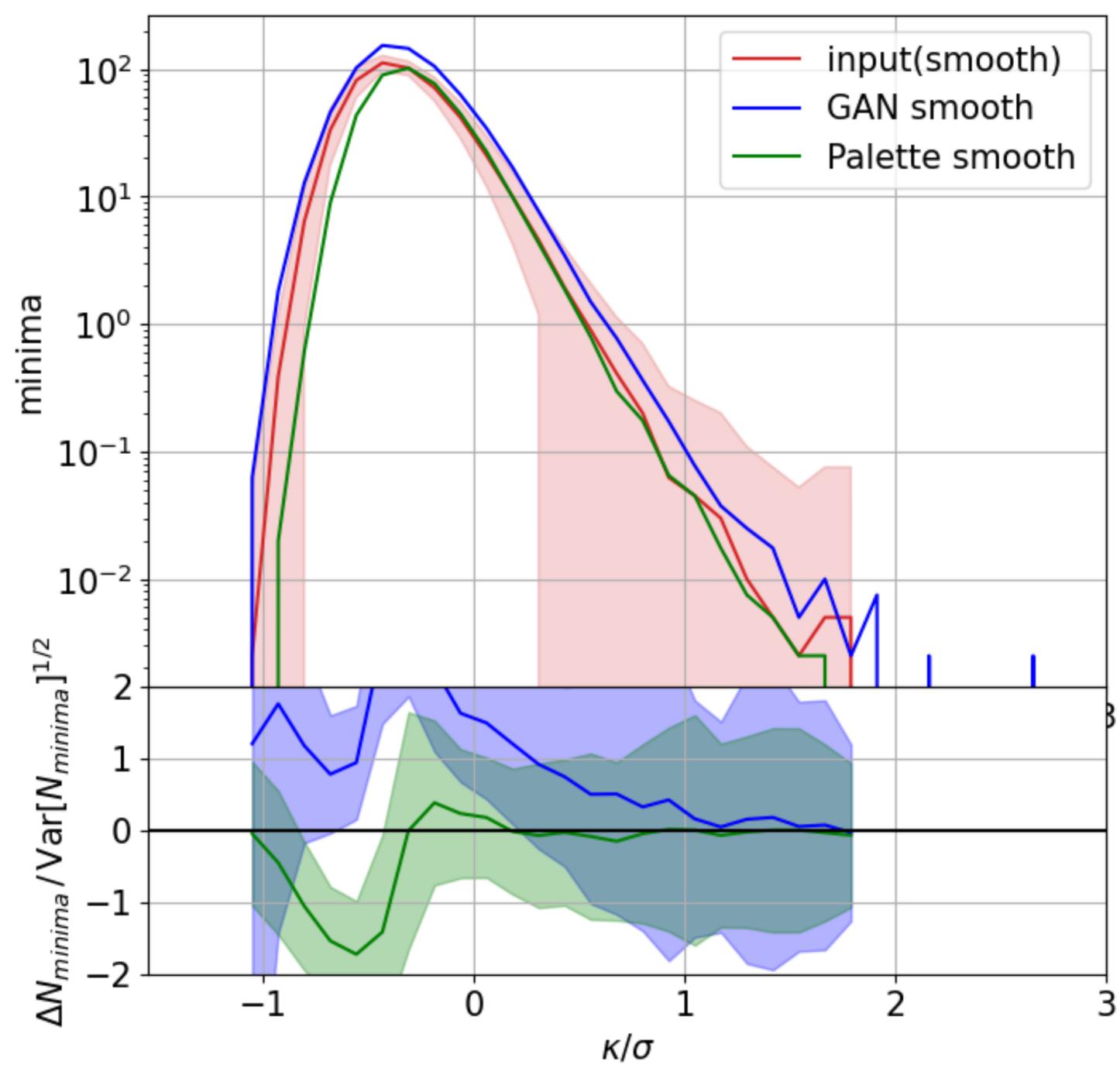


# smooth

## peak

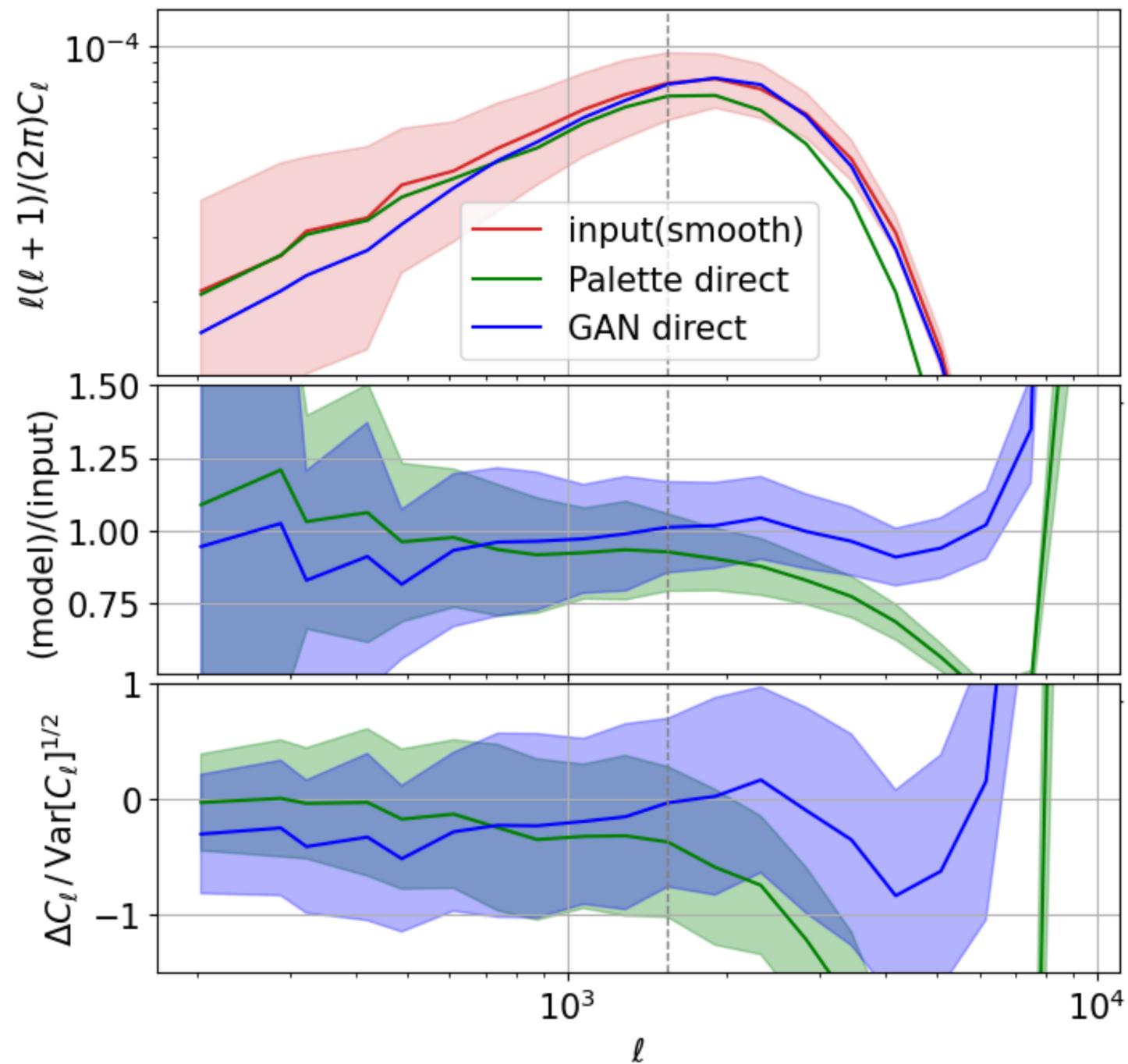


## minima

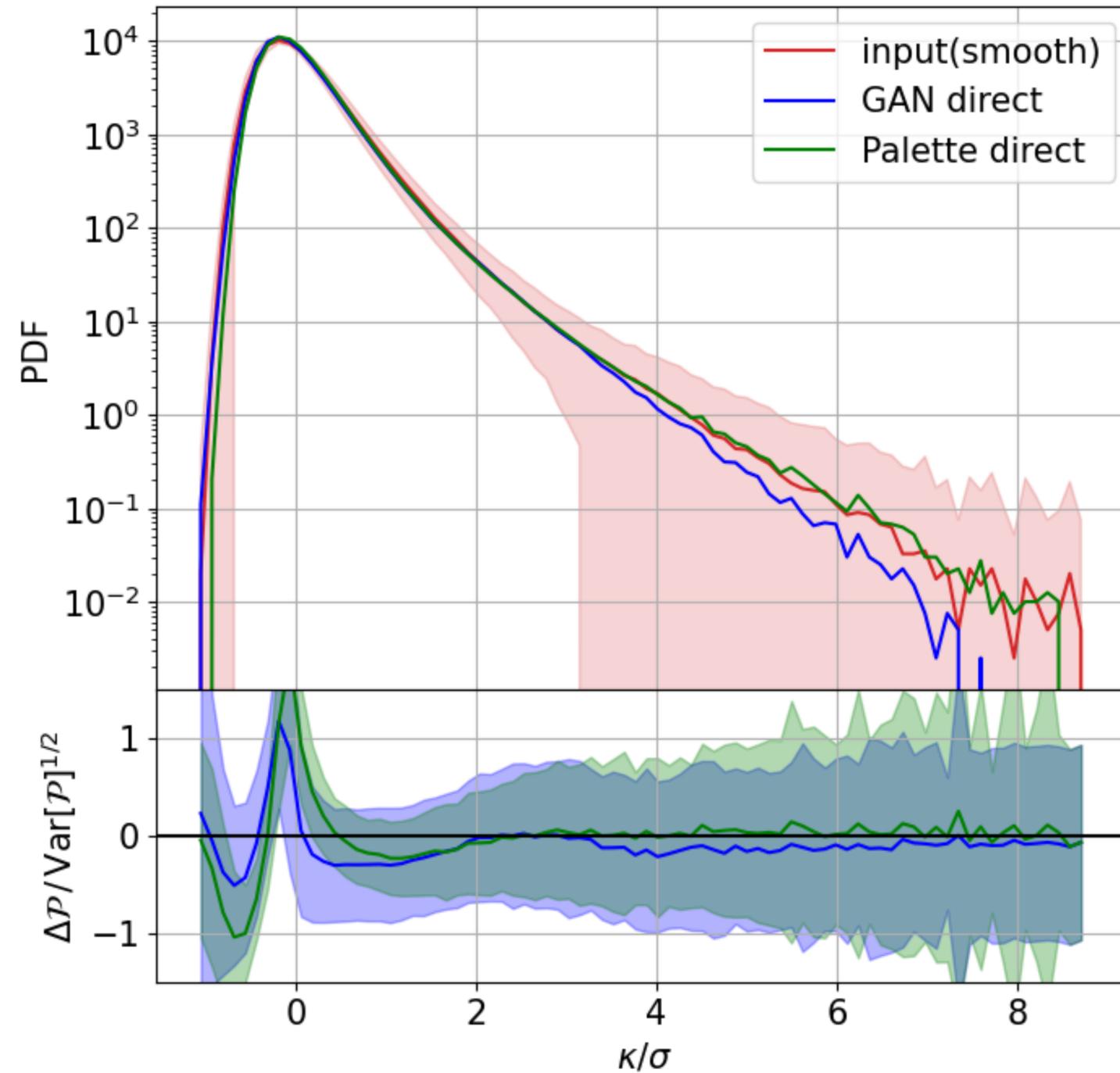


# direct

## パワースペクトル

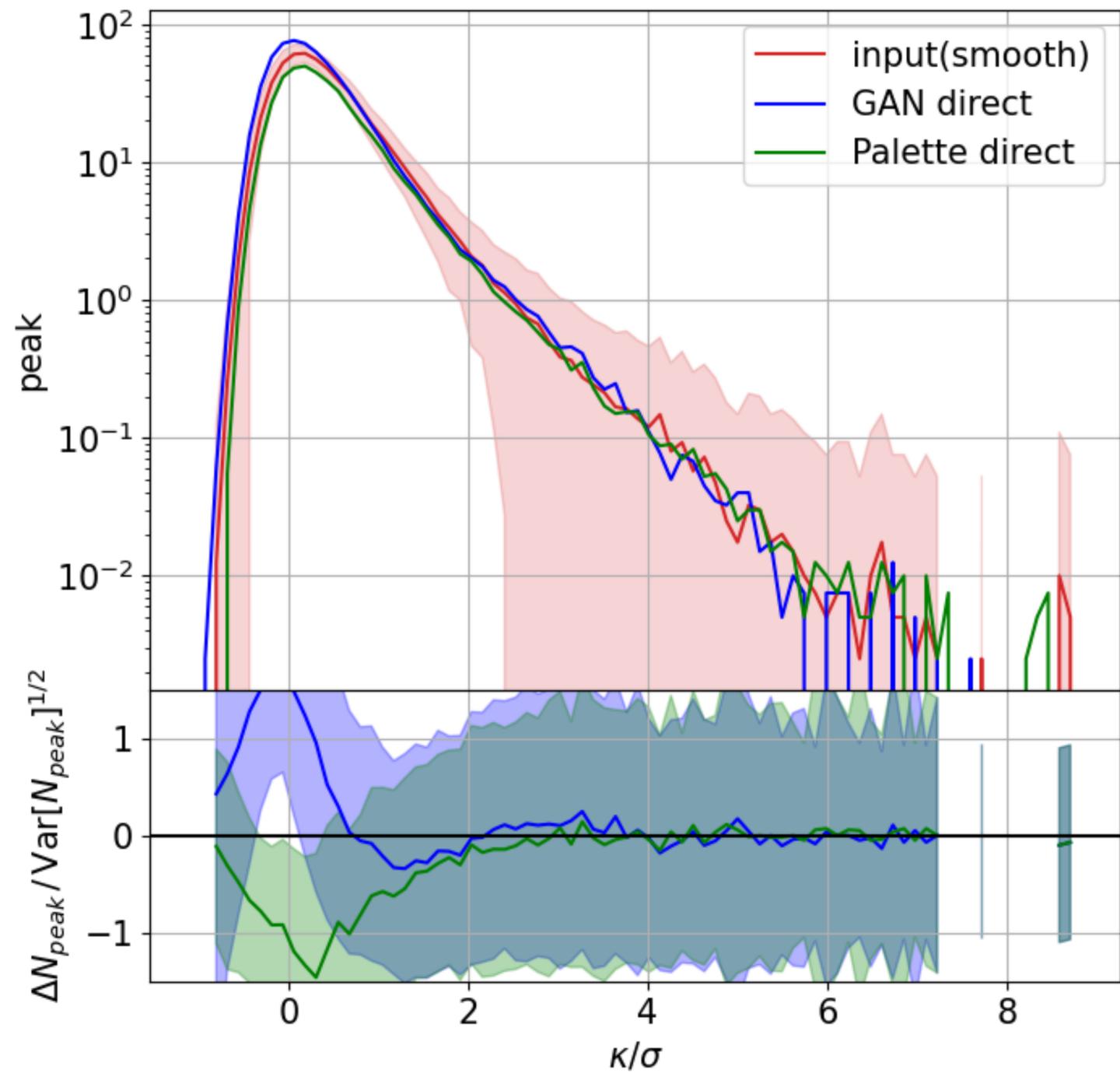


## PDF

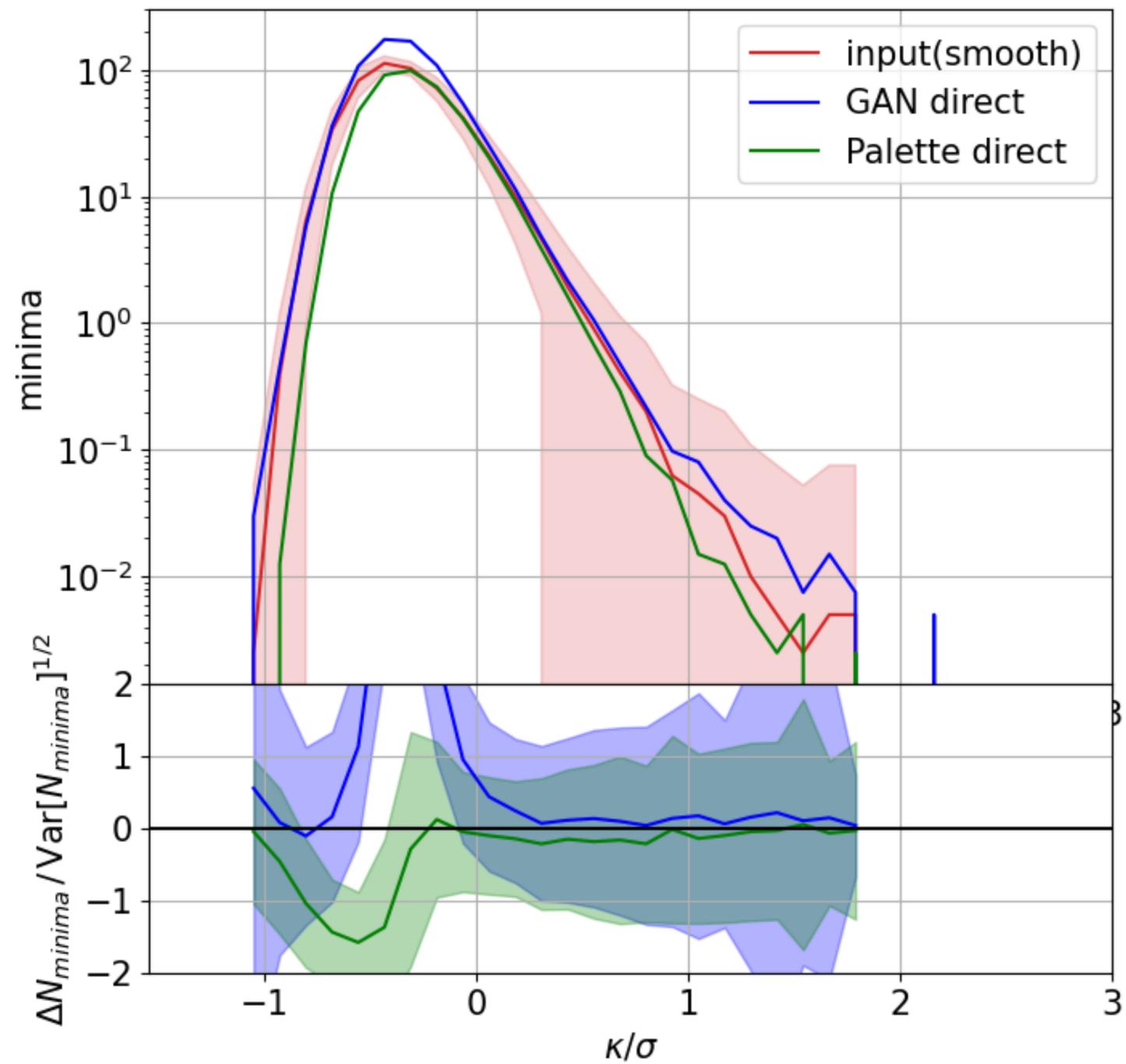


# direct

## peak

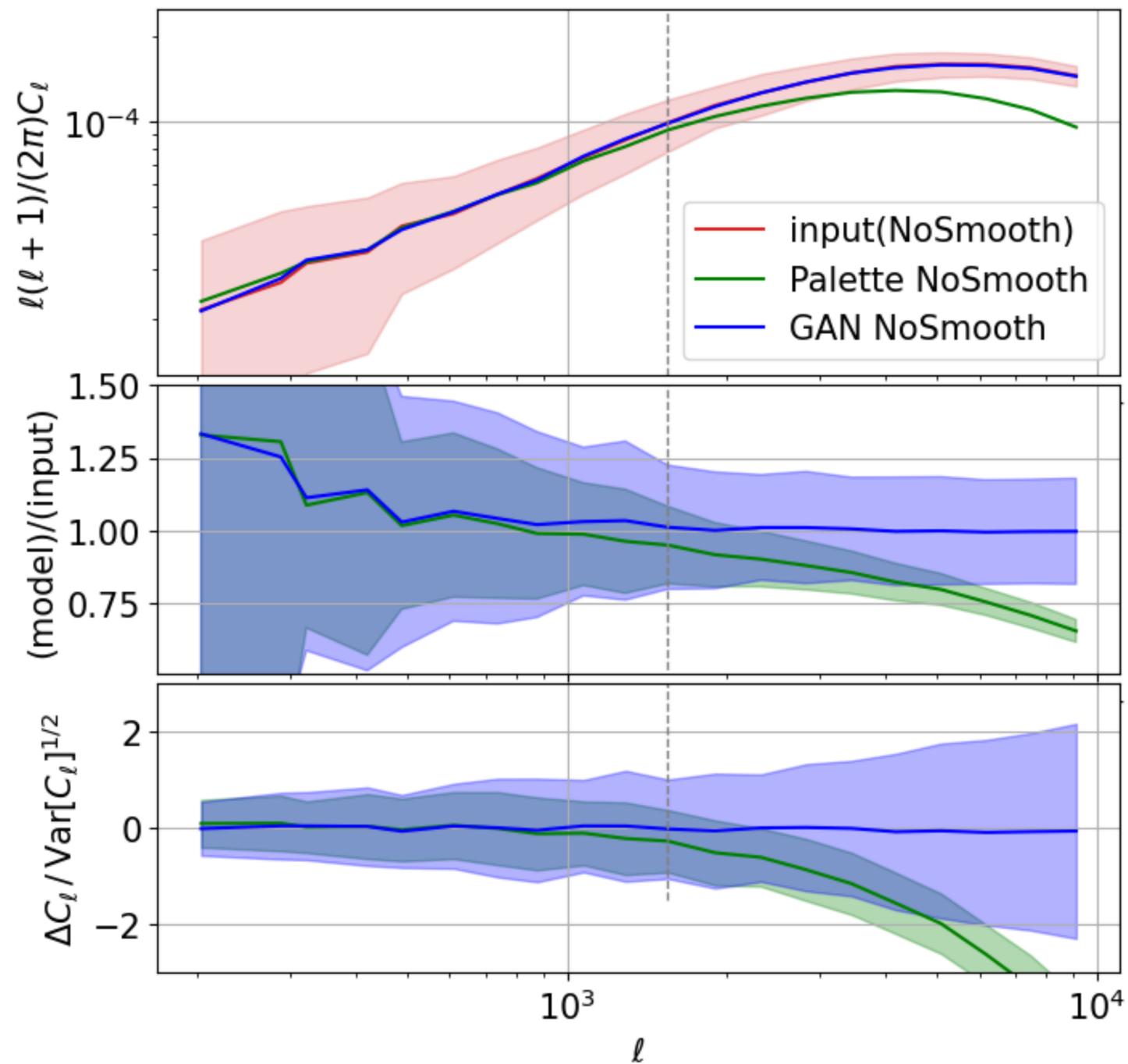


## minima

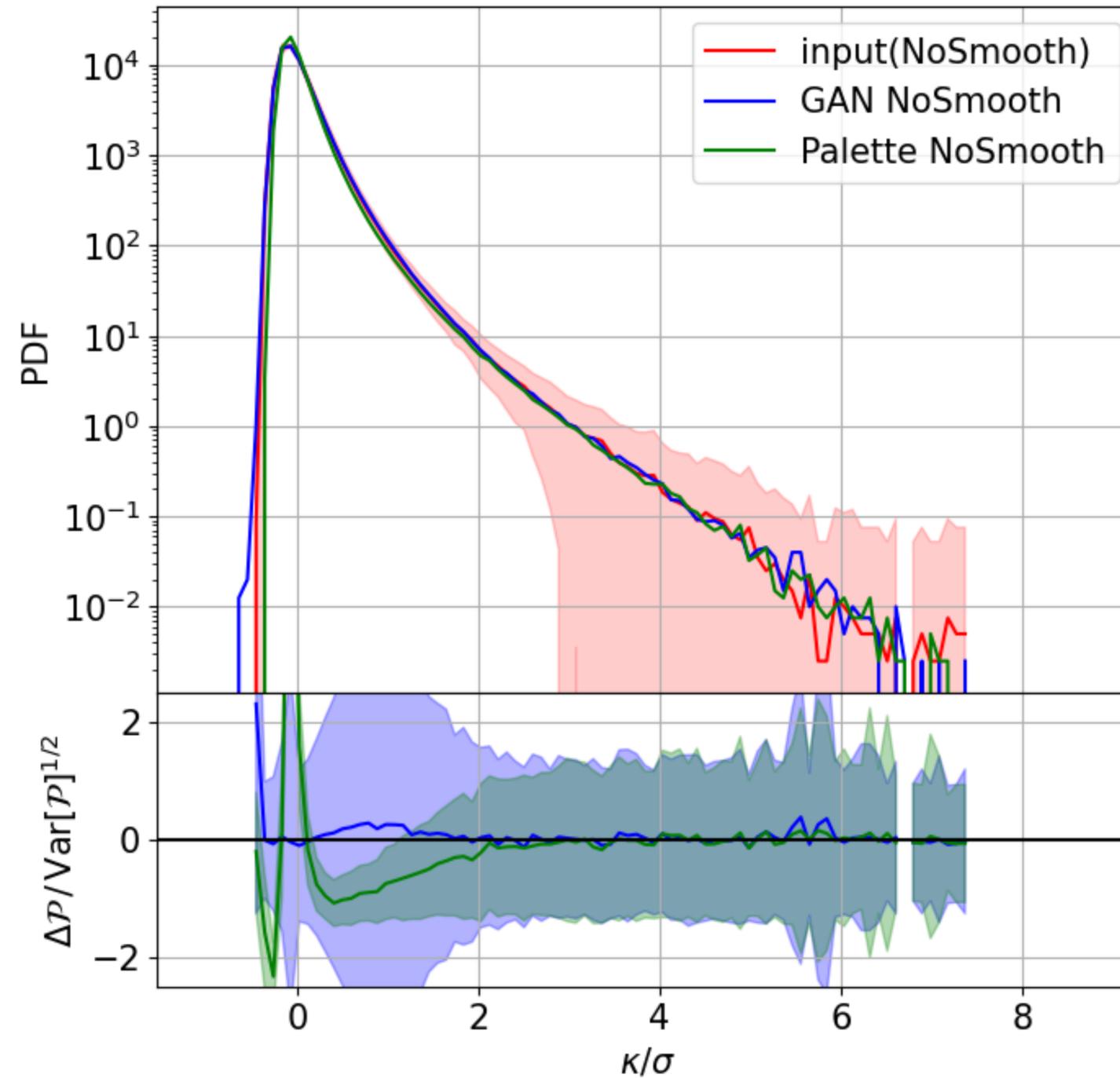


# NoSmooth

## パワースペクトル

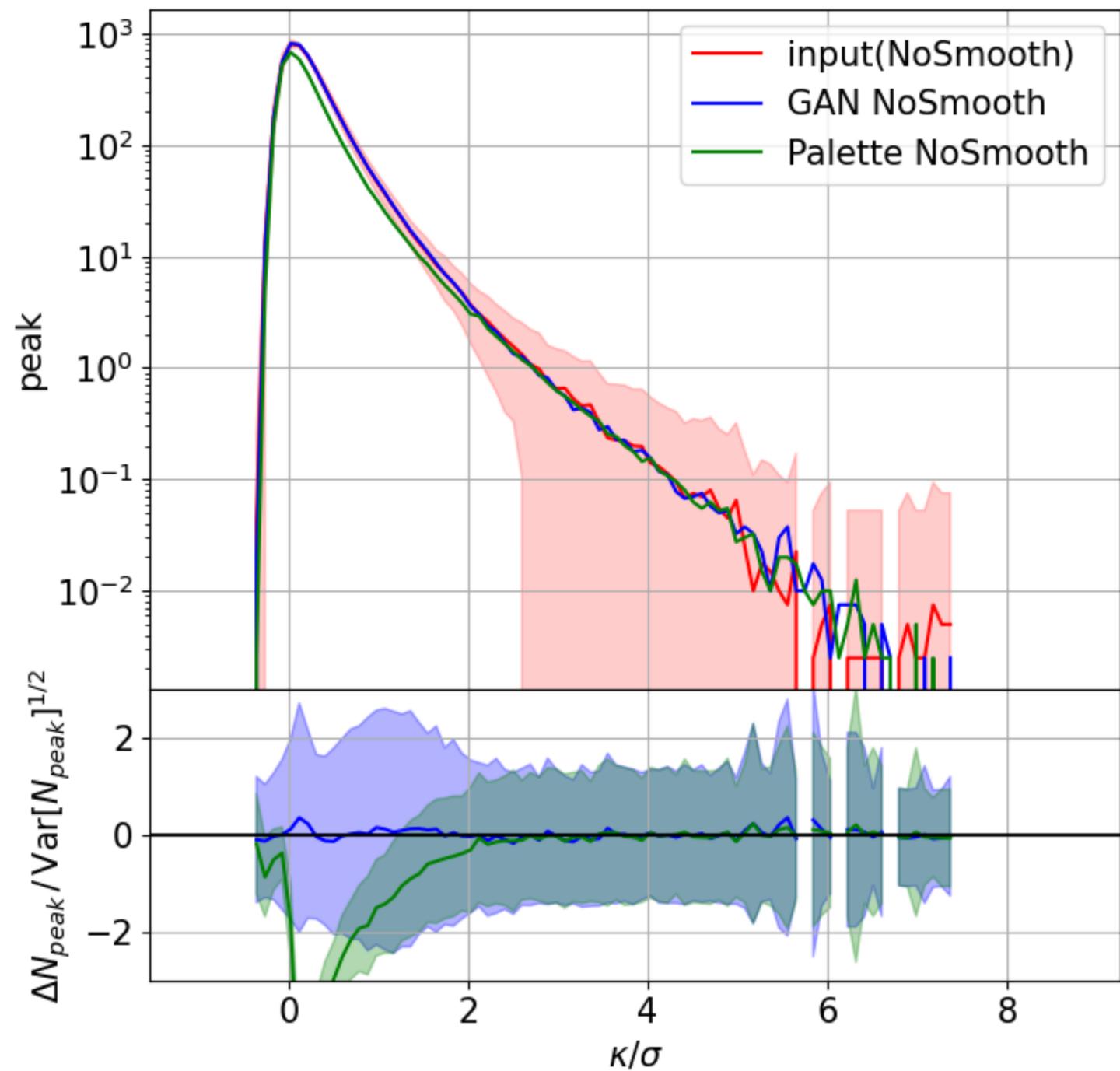


## PDF

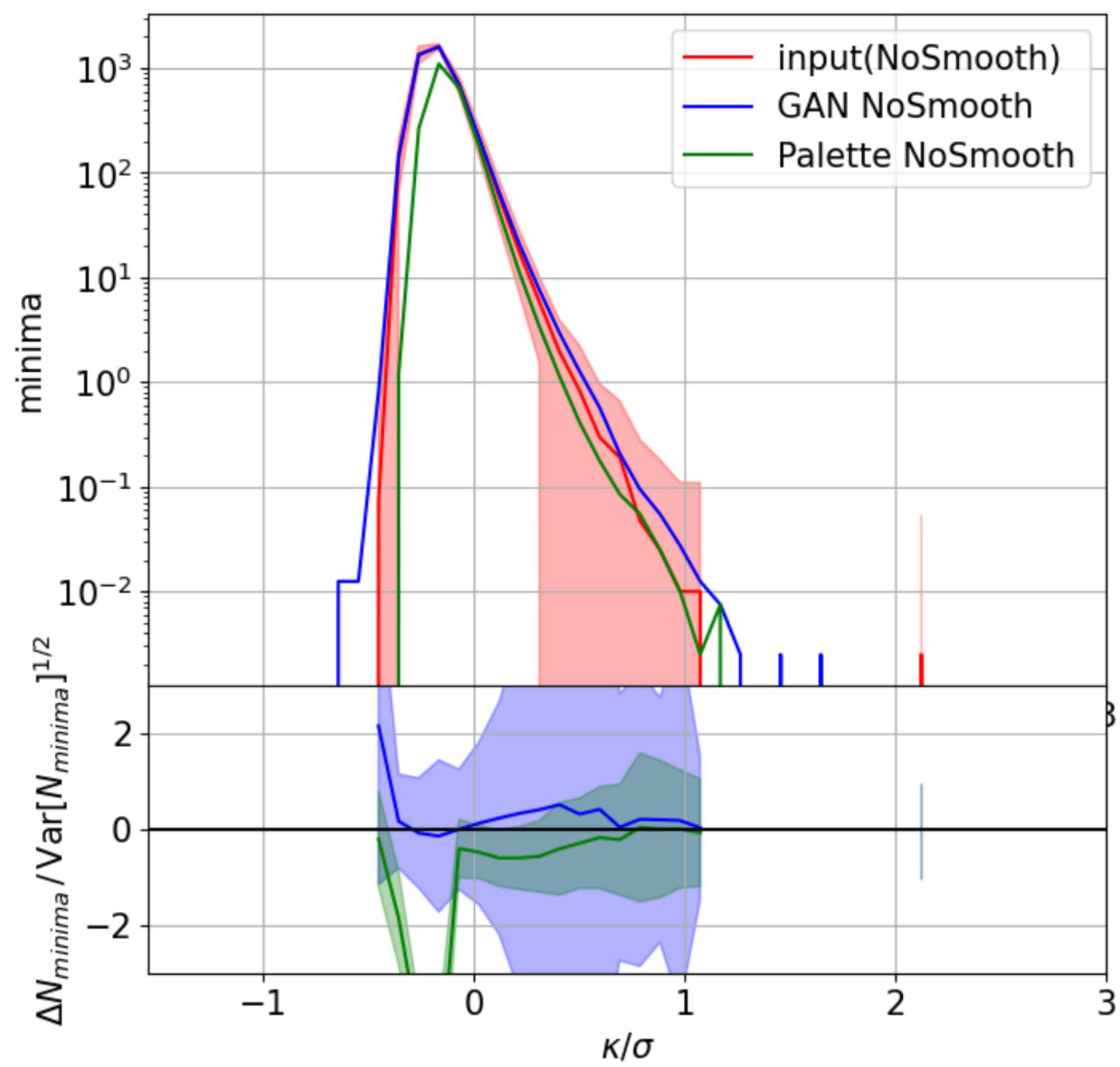


# NoSmooth

## peak



## minima



## フーリエ空間において収束場と歪み場の関係

$$\tilde{\kappa} = \frac{1}{2}(\ell_1^2 + \ell_2^2)\varphi, \quad \tilde{\gamma}_1 = \frac{1}{2}(\ell_1^2 - \ell_2^2)\varphi, \quad \tilde{\gamma}_2 = \ell_1\ell_2\varphi,$$

$$\text{レンズポテンシャル: } \varphi(\boldsymbol{\theta}) = \frac{2}{c^2 S_K(x_s)} \int_0^{x_s} dx \frac{S_K(x_s - x)}{S_K(x)} \Phi(\boldsymbol{\theta}, x, t)$$

## レンズポテンシャルを消去して

$$\tilde{\kappa}(\boldsymbol{\ell}) = \frac{\ell^2}{(\ell_1 + i\ell_2)^2} \tilde{\gamma}(\boldsymbol{\ell})$$

$\gamma \rightarrow \kappa$  に変換可能