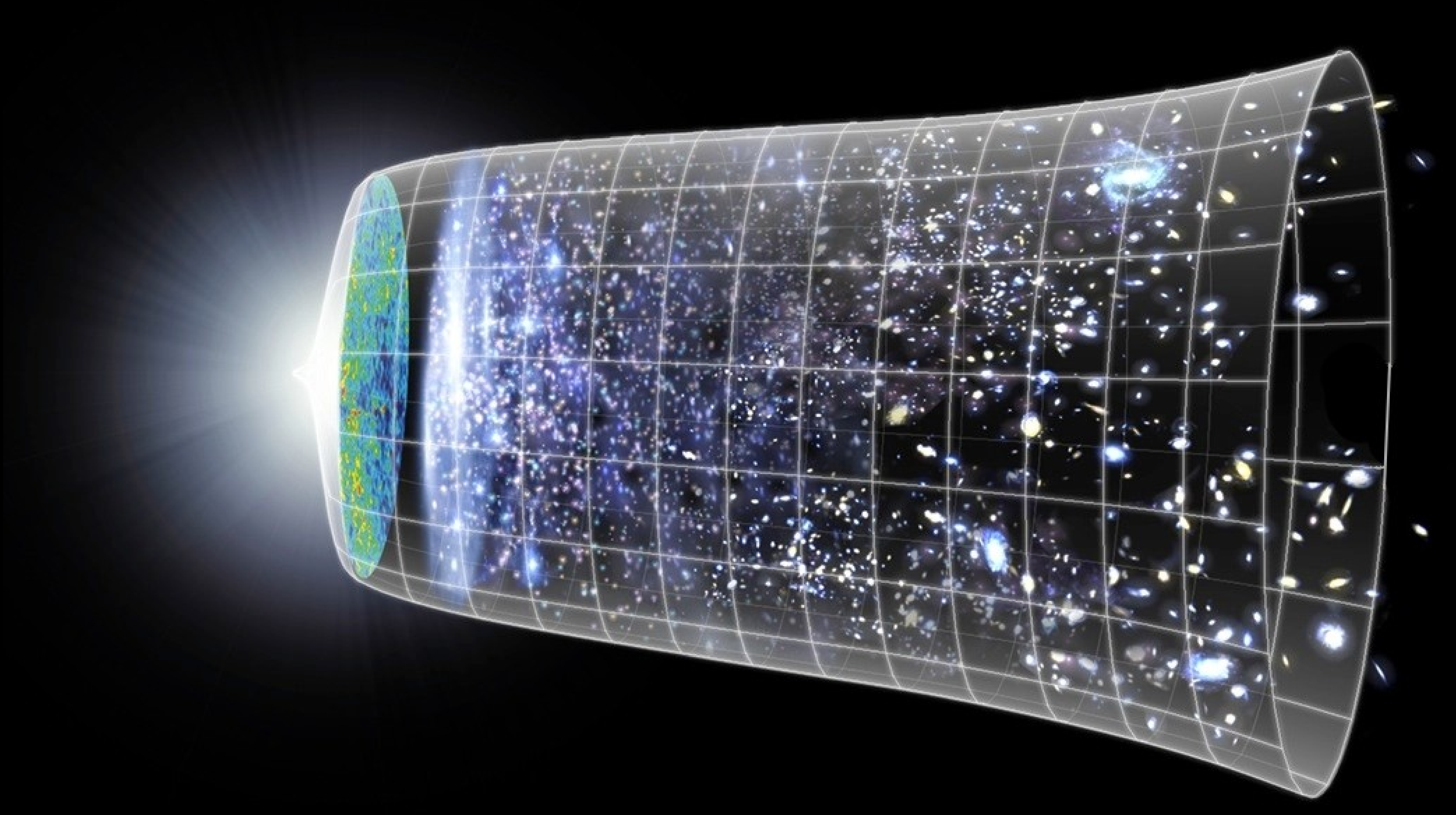
Why we need to care about unobservable modes

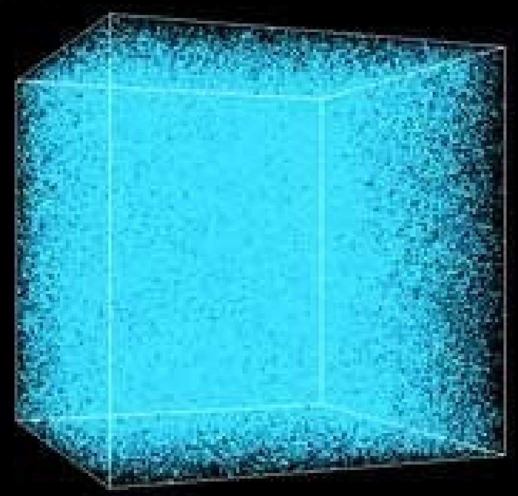
Based on
Lacasa & Rosenfeld 2016, Lacasa et al. 2018,
Lacasa & Kunz 2017, Lacasa & Grain 2019,
Gouyou Beauchamps et al. 2022,
Lacasa et al. arXiv:2209.14421

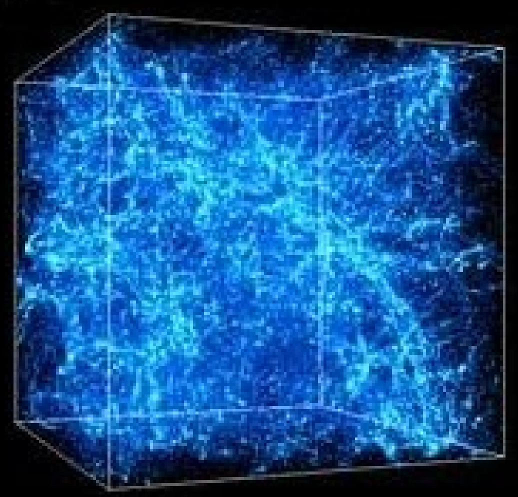
Context

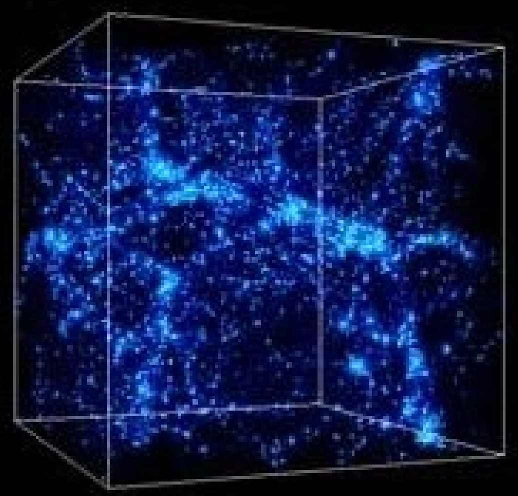


NASA/WMAP Science Team

Time and (non-) linearity



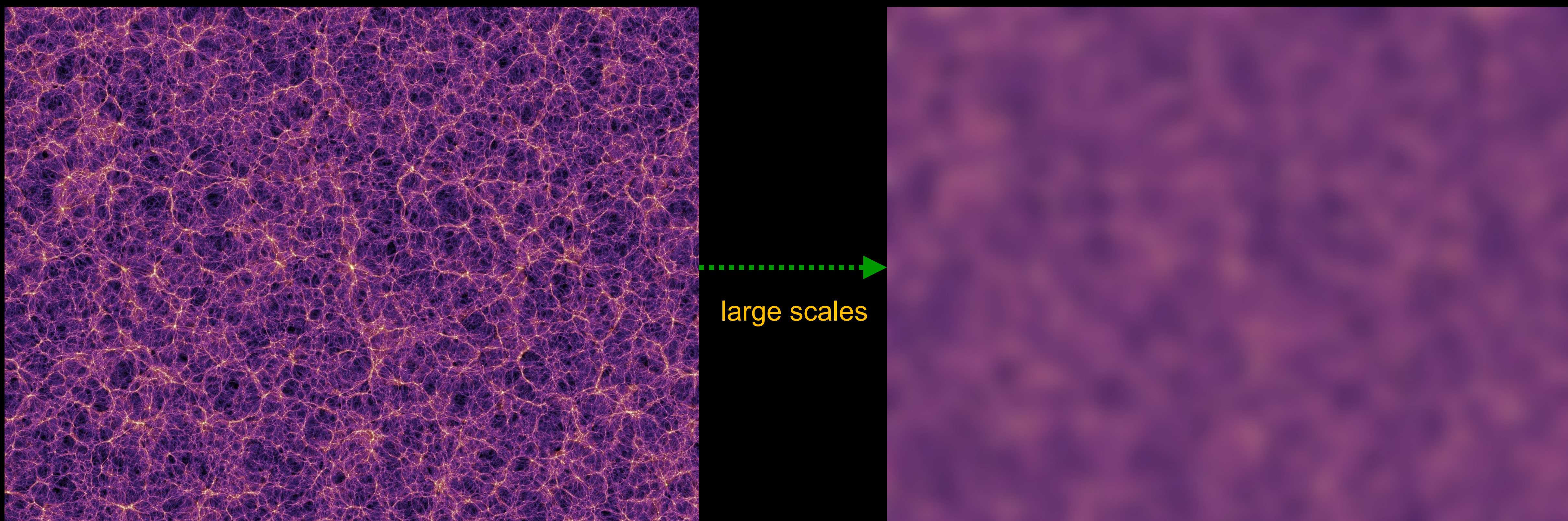




t ≈ 0 — 2 billion years — 14 billion years — time

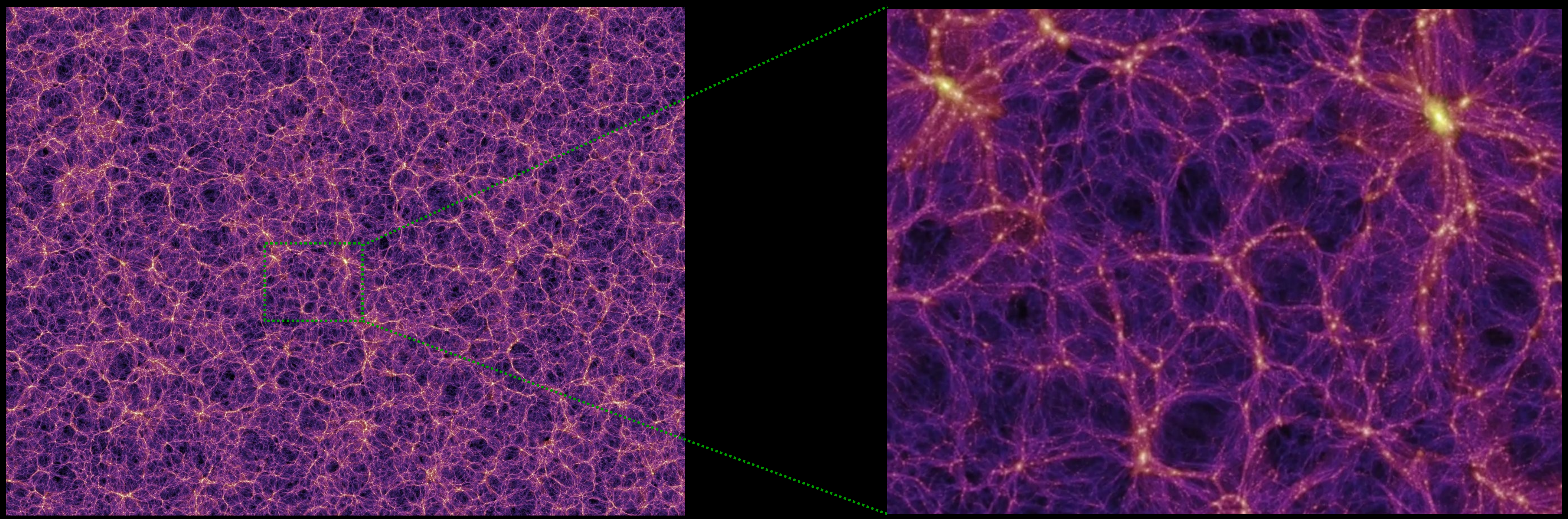
Dynamics of density fluctuations	linear		**non linear**
		→ **Gravity**	↓ induces
Statistical distribution	Gaussian		**non Gaussian**

Scales and (non-) linearity



Millenium simulation

Scales and (non-) linearity



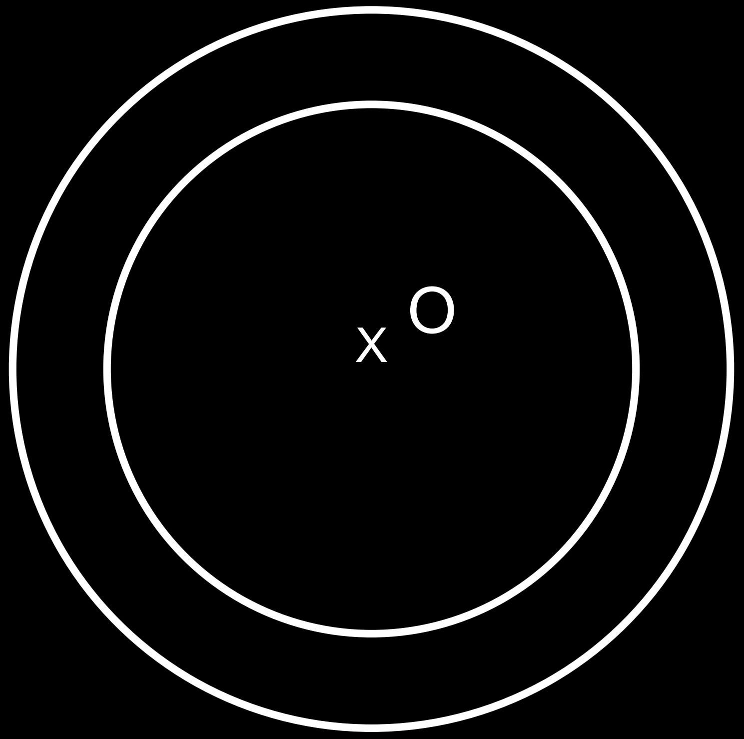
Millenium simulation

Photometric surveys
Euclid : 30 gals/arcmin^2
Rubin/LSST : 45 gals/arcmin^2

Radio surveys
SKA2 : +precise redshifts

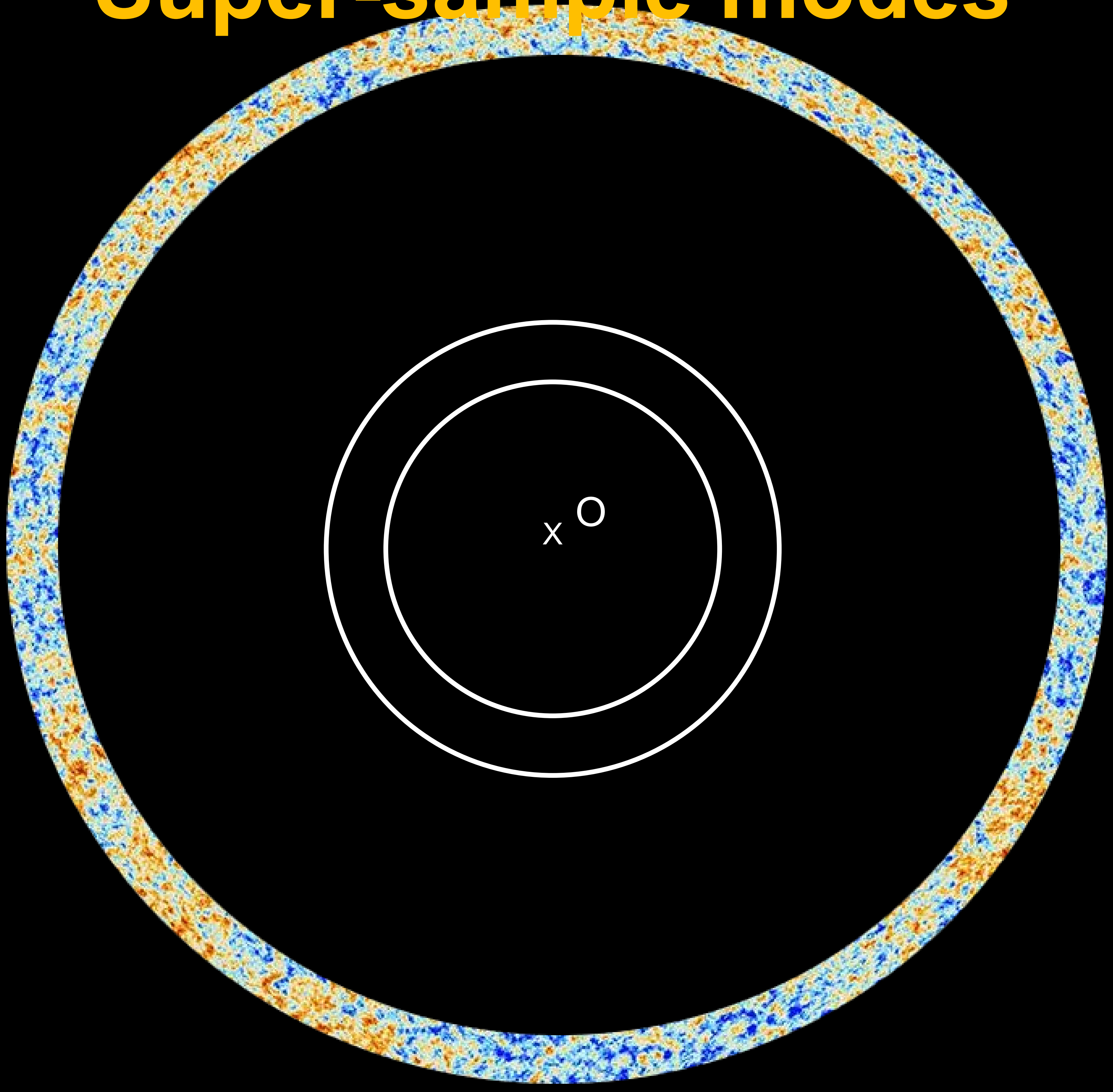
Super-sample modes and their impact

Super-sample modes



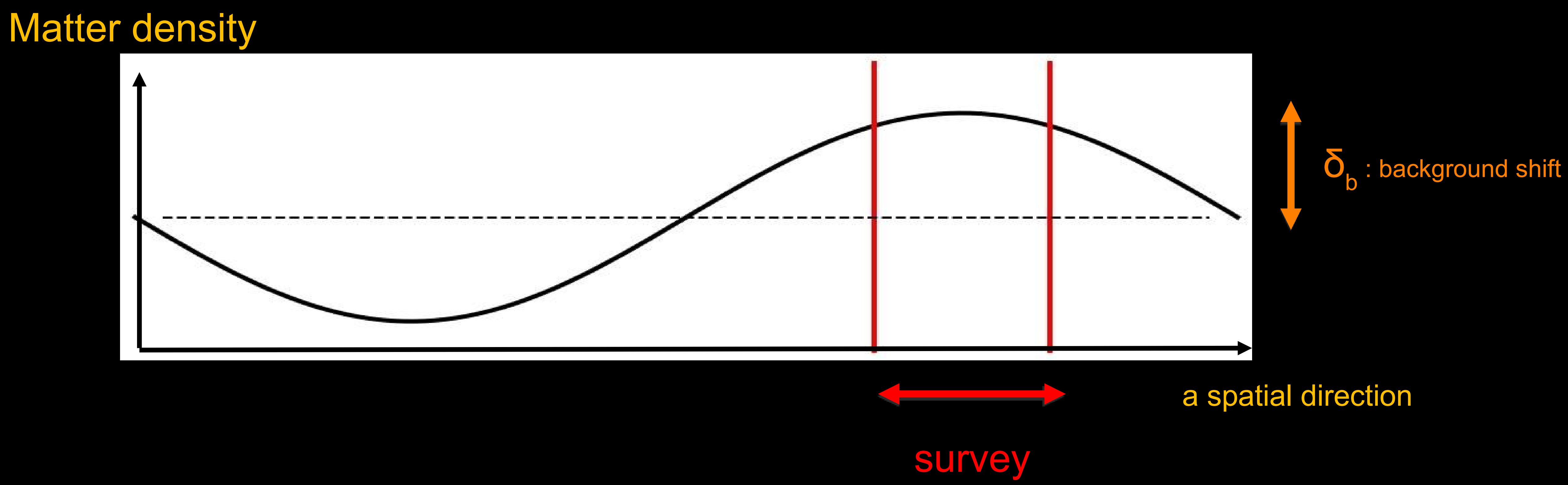
Finite z bin = finite volume
(even in full sky)

Super-sample modes

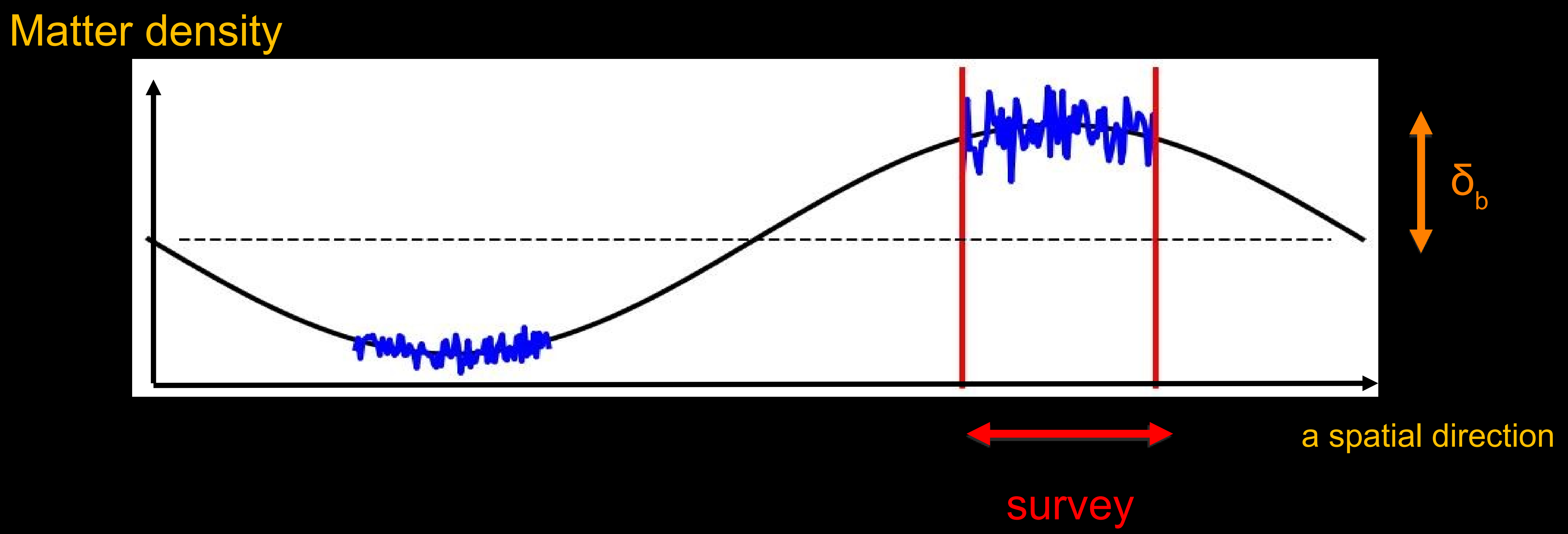


Super-sample fluctuations exist

Super-sample covariance (SSC)



Super-sample covariance (SSC)



Feedback/backreaction of long on short modes

Equivalent to evolution in separate Universe

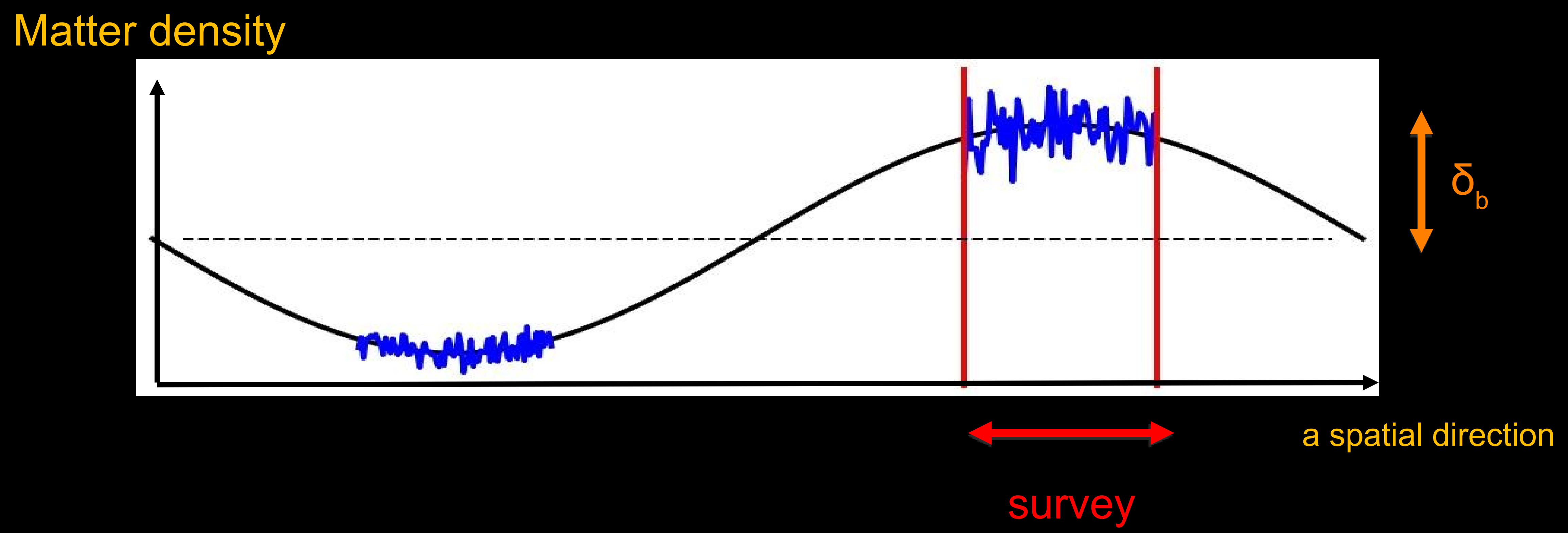
→ impacts all probes

Galaxies, clusters, lensing, HI…

→ impacts all statistics

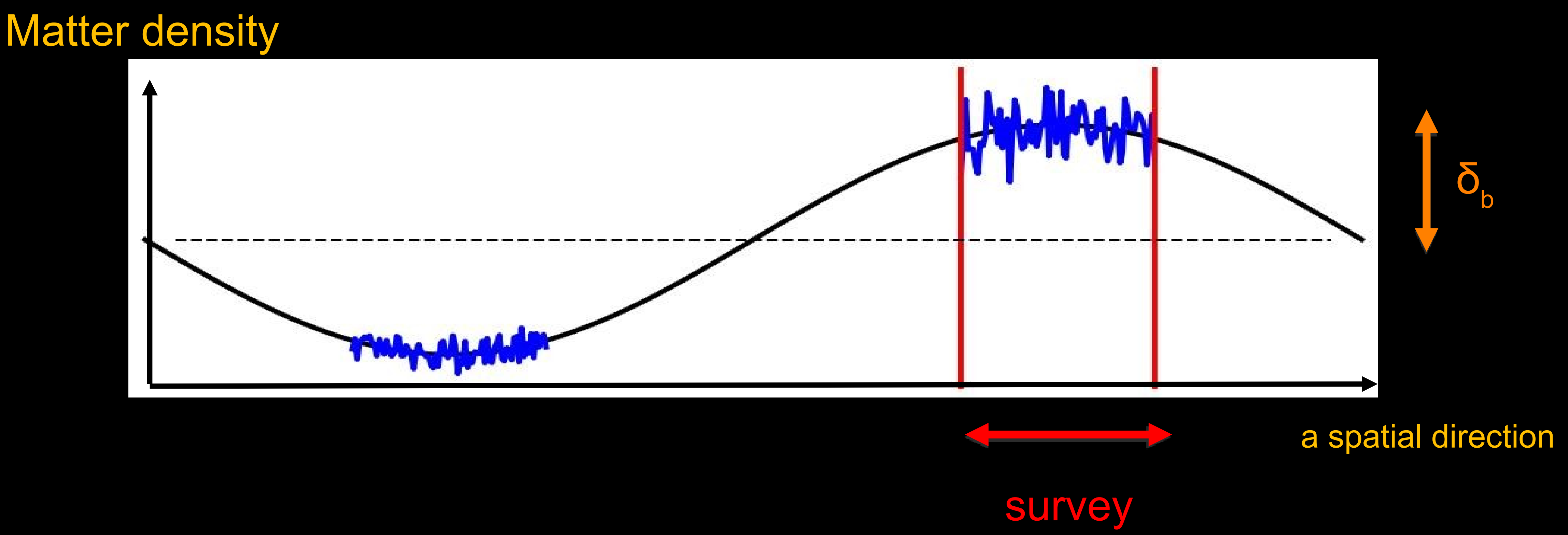
Power spectrum, bispectrum, machine learning, field level inference...

Super-sample covariance (SSC)



$$\mathrm{Cov} = \mathrm{Cov}_{\mathrm{G}} + \mathrm{Cov}_{\mathrm{SSC}} + \mathrm{Cov}_{\mathrm{oNG}}$$

Super-sample covariance (SSC)



$$\mathrm{Cov} = \mathrm{Cov}_{\mathrm{G}} + \mathrm{Cov}_{\mathrm{SSC}} + \mathrm{Cov}_{\mathrm{oNG}}$$

$$\mathrm{Cov}_{\mathrm{SSC}}\left(\mathcal{O}_1, \mathcal{O}_2\right) = \int \mathrm{d}z_1\,\mathrm{d}z_2 \cdots \frac{\partial \mathfrak{o}_1}{\partial \delta_b}\,\frac{\partial \mathfrak{o}_2}{\partial \delta_b}\;\sigma^2(z_1, z_2)$$

Reaction of observable to background shift

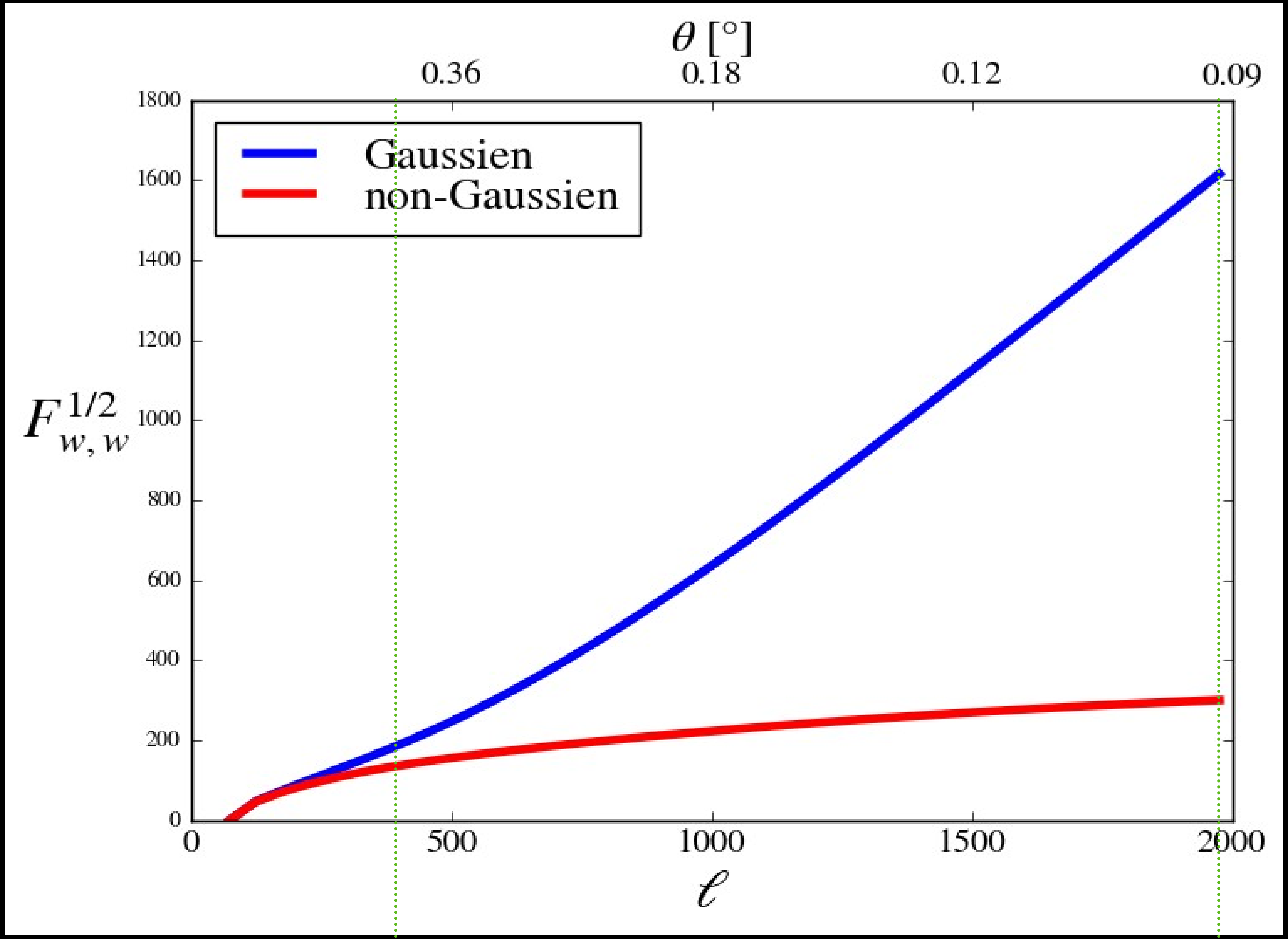
Variance of background shift

Lacasa & Rosenfeld et al. 2016

Euclid
and making SSC popular

Impact of SSC for Euclid : first look

Fisher information on Dark Energy (unmarginalized)
forecasts for photometric galaxy C_l



÷ 5

angular resolution of analysis

~ previous DES analysis

Euclid target

Fast and easy SSC

$$\mathrm{Cov}_{\mathrm{SSC}}\left(\mathcal{O}_1, \mathcal{O}_2\right) = \int \mathrm{d}z_1 \, \mathrm{d}z_2 \frac{\partial \mathfrak{o}_1}{\partial \delta_b} \frac{\partial \mathfrak{o}_2}{\partial \delta_b} \, \sigma^2(z_1, z_2)$$

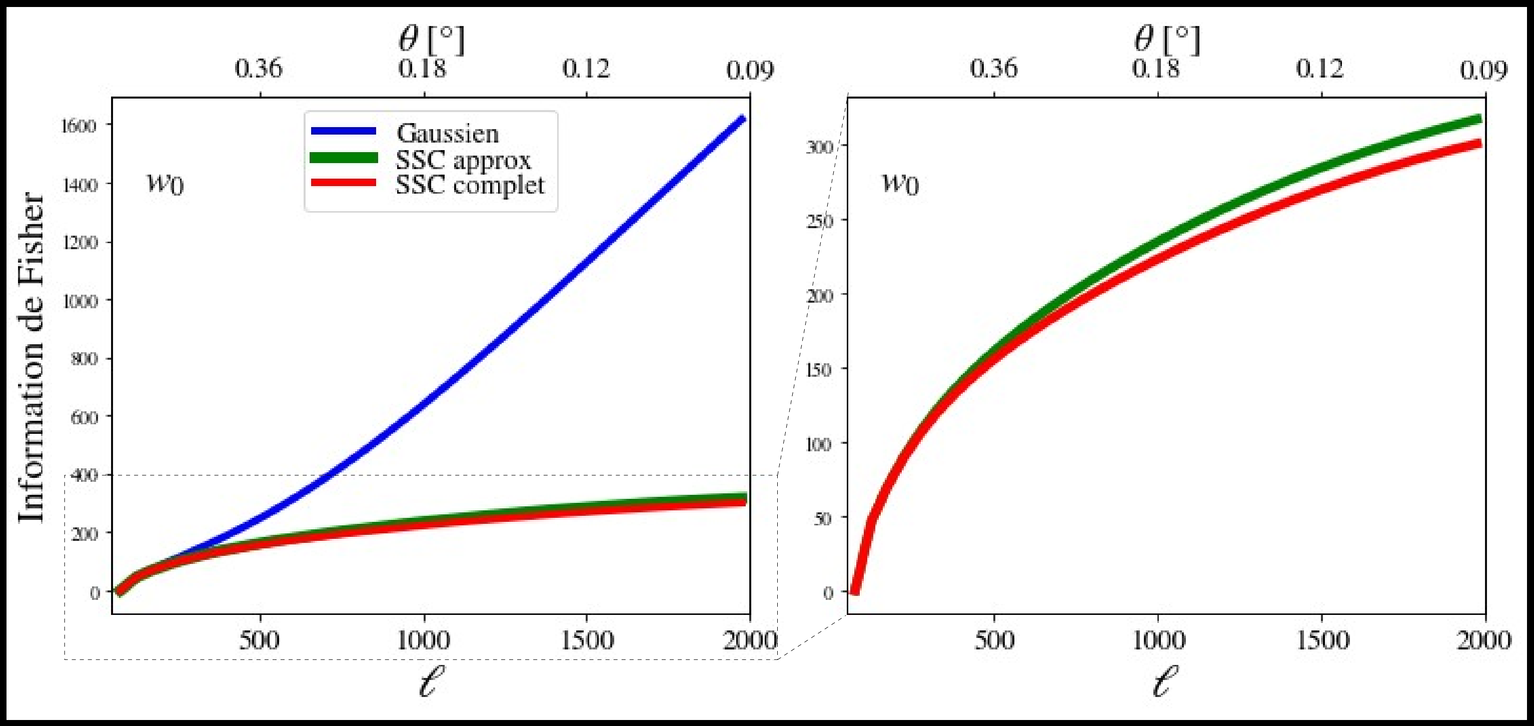
$$\mathrm{Cov}_{\mathrm{SSC}}\left(\mathcal{O}_1, \mathcal{O}_2\right) \approx R_1 \, \mathcal{O}_1 \;\; R_2 \, \mathcal{O}_2 \;\; S_{i,j}$$

Lacasa & Grain 2019

dimensionless response

can be computed with `PySSC`

https://github.com/fabienlacasa/PySSC
https://pyssc.readthedocs.io



5% precision

Validation of the SSC approximation: DE Fisher information

Impact on constraing Dark Energy

Euclid-like forecast for Dark Energy Figure of Merit (FoM)

with weak lensing (WL), photometric galaxy clustering (Gcph)
and their cross-correlation (XC)

Probe	Survey area [deg^2]	Gaussian		Full-sky SSC	
WL	5	0.014		0.009	
	15000	43.12	→	26.329	- 40%
GCph	5	0.035		0.029	
	15000	103.71	→	88.636	- 15%
GCph+WL+XC	5	0.346		0.153	
	15000	1038.13	→	454.59	- 55%

"full-sky" = rescaled by sky fraction
(as tradition for Gaussian covariance)

Gouyou Beauchamps et al. 2022

Fast and easy SSC, in partial sky

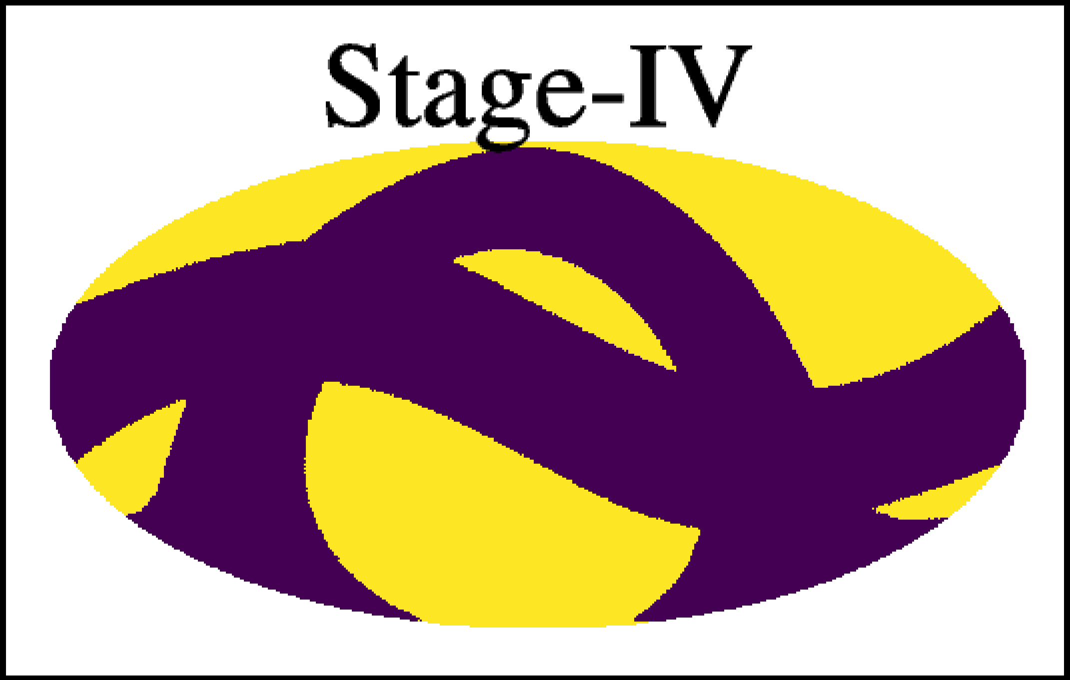
Team formed during Euclid France summer school 2019

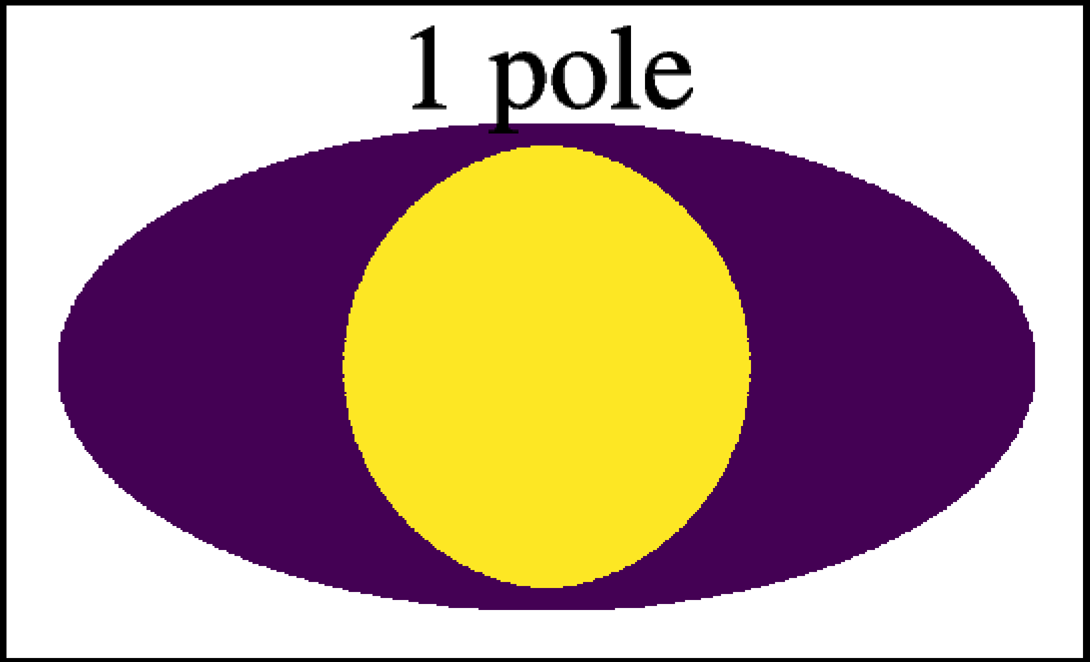
S. Gouyou Beauchamps, I. Tutusaus, M. Aubert, P. Baratta, A. Gorce

Formalism to account for survey mask in SSC

New `PySSC` release

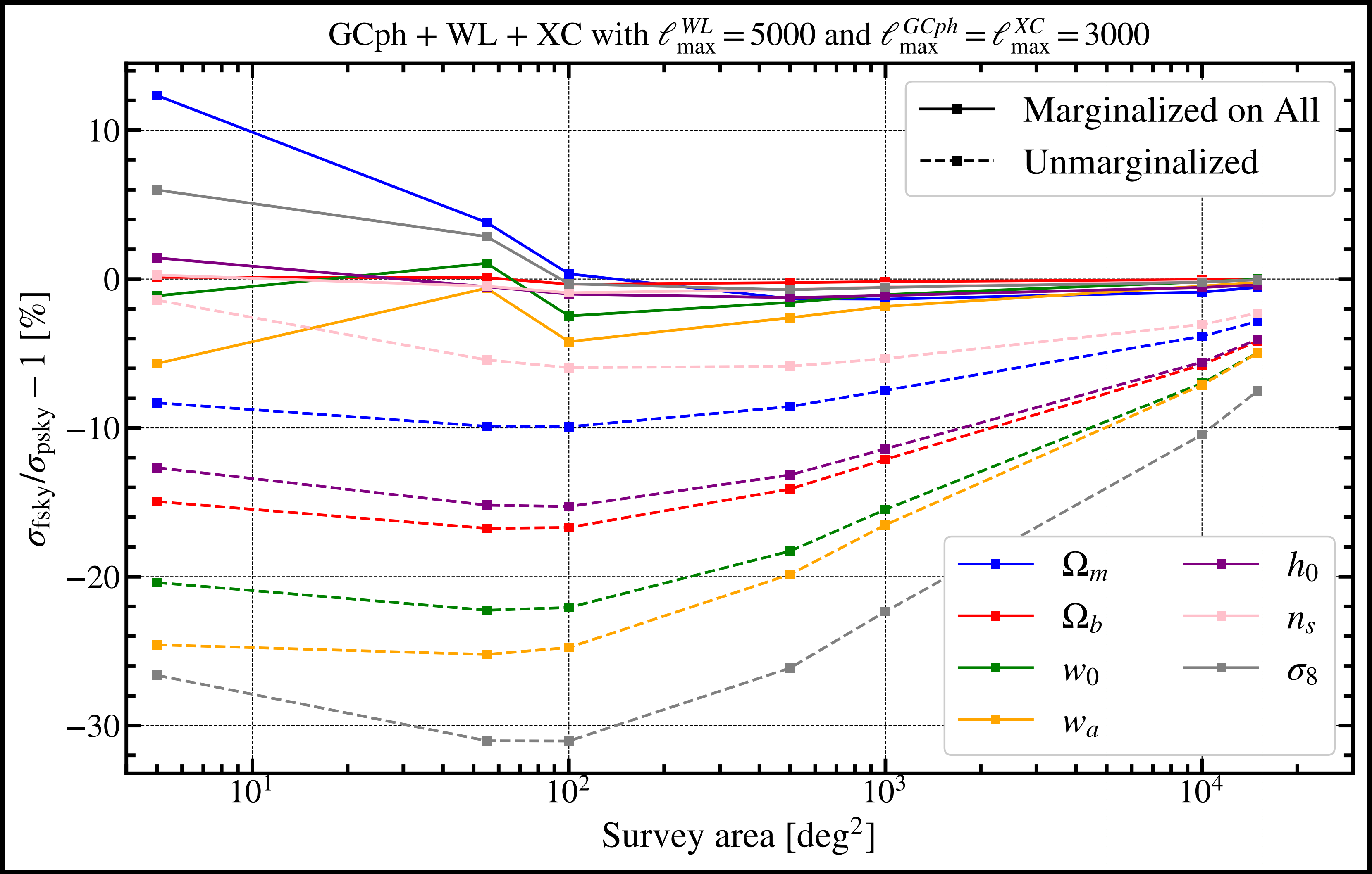
https://pyssc.readthedocs.io





Fast and easy SSC, in partial sky

Comparison of cosmo errors:
rigorous partial sky vs traditional rescaling



DES-like Euclid-like

Gouyou Beauchamps et al. 2022

A new approach to SSC

Problems of the trad approach (=numerical covariance)

Compared to Gaussian covariance, SSC matrix
- is highly off-diagonal
- has low rank

→ **Problems**
- memory and CPU cost is higher
- numerical inversion can be unstable (ill-conditioned problem)
- does not take advantage of a diagonal covariance

Solution: invert total covariance analytically (Woodbury identity)

→ analytical formulae for likelihood and Fisher

Lacasa et al. 2022

SSC directly for likelihood or Fisher

Likelihood is a correction to the noSSC case

$$\ln \mathcal{L} = \ln \mathcal{L}_{\mathrm{noSSC}} + \delta \ln \mathcal{L}$$

Can also get Fisher or inverse Fisher matrices

$$F = F_{\mathrm{noSSC}} - Y^T (1 + I(\delta_b)\, S)^{-1}\, Y$$

$$F^{-1} = F^{-1}_{\mathrm{noSSC}} + Z^T\, \Sigma^2(\delta_b)\, Z$$

$I(\delta_b)$: Fisher info on δ_b (unmarginalised). Y: contains probes responses.

Interests:

- updating a pipeline at low cost
- fast if noSSC is diagonal

Lacasa et al. 2022

Conclusions

Conclusions

- Long wavelength modes are unobservables but backreact on observed modes
  $\longrightarrow$ super-sample covariance (SSC)

- SSC correlates all probes, even if independent otherwise

- SSC importance increases with statistical power
  Ex: push to high resolution, combine probes, combine surveys.
  => Impact for stage IV constraints on Dark Energy, σ_8, Ω_m … (modified gravity?)

- Fast & easy computation possible: public code `PySSC`
  including mask effect
  https://github.com/fabienlacasa/PySSC
  https://pyssc.readthedocs.io

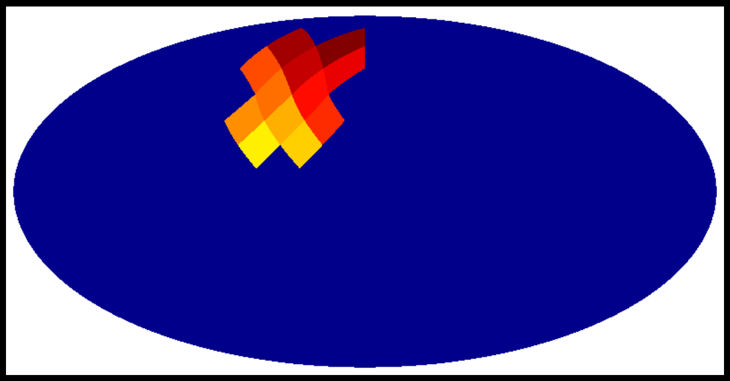
- New method directly at level of likelihood or Fisher
  $\longrightarrow$ easier, faster, robuster
  Currently implementing it for Euclid + CMB

Thank you

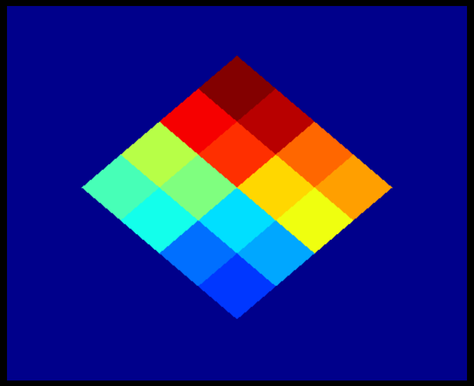
Additional slides

SSC and traditional covariance methods

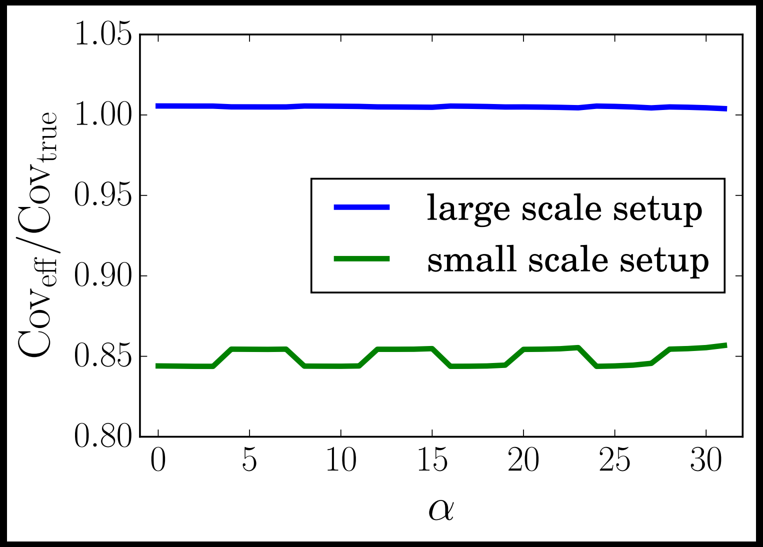
Method 1: survey footprint replicated at different points of a large simulation



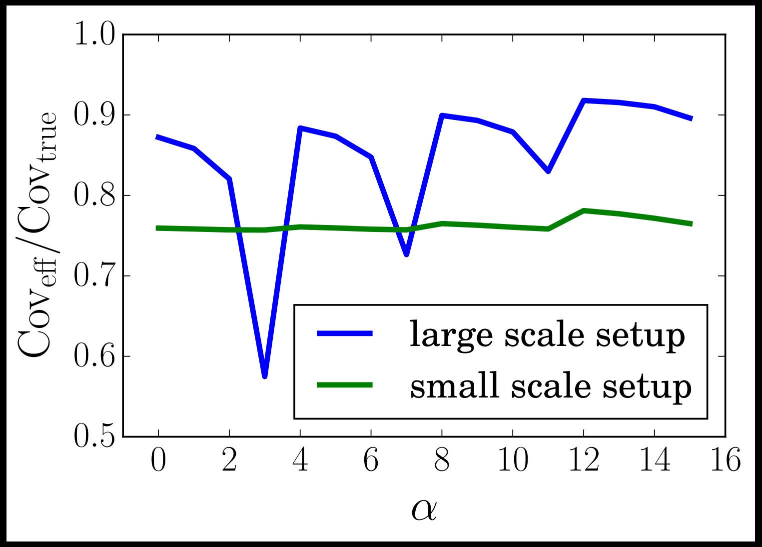
large scale setup



small scale setup



same redshifts



two different redshifts

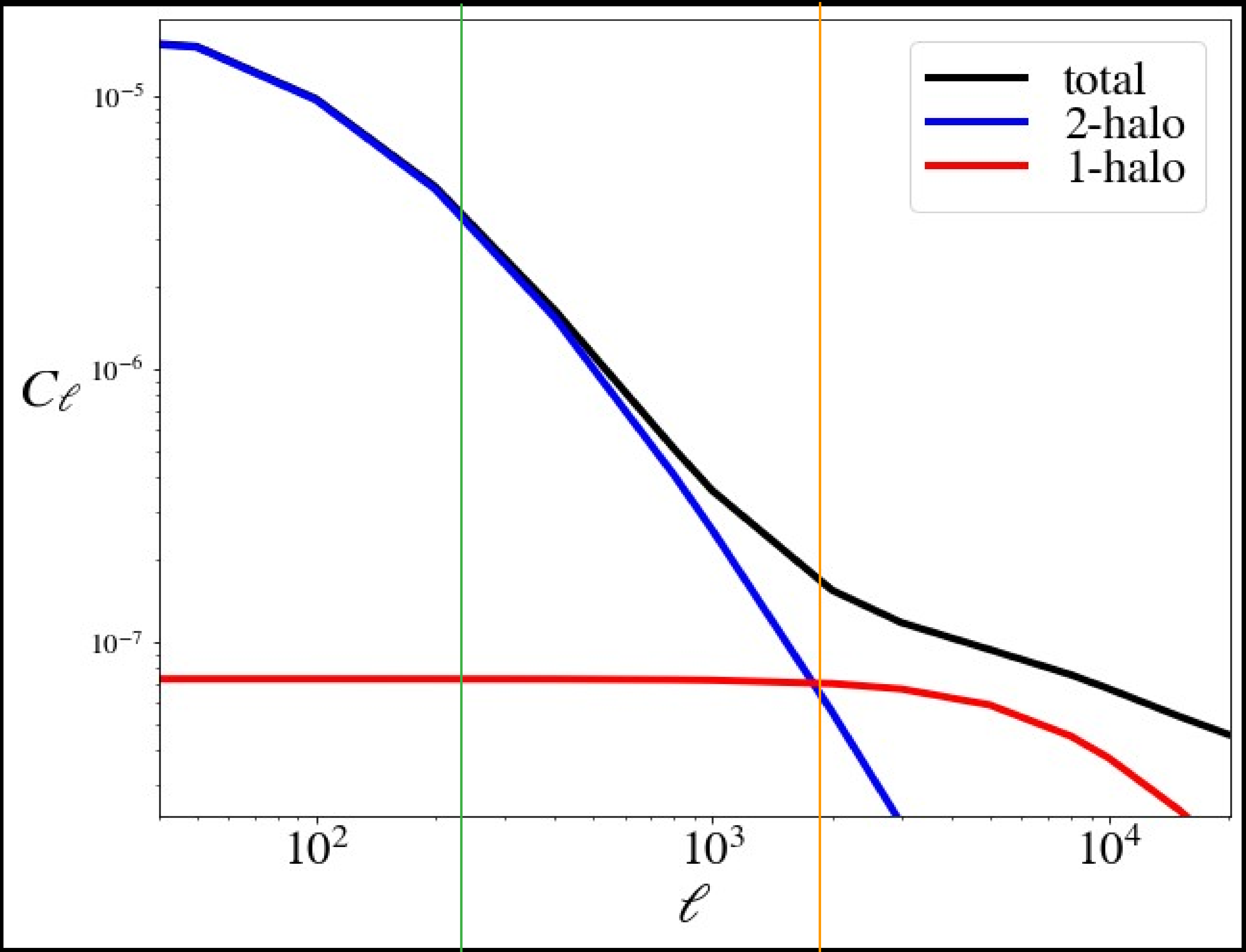
Method 2 (jackknife):

- Random subsamples removed from data → fake data realisations
- Apply canonic covariance estimator and rescale to survey area

Works for classic covariance terms. For SSC: biased by up to factor **5**.

Lacasa & Kunz 2017

Scales



z=1

Regime

Linear

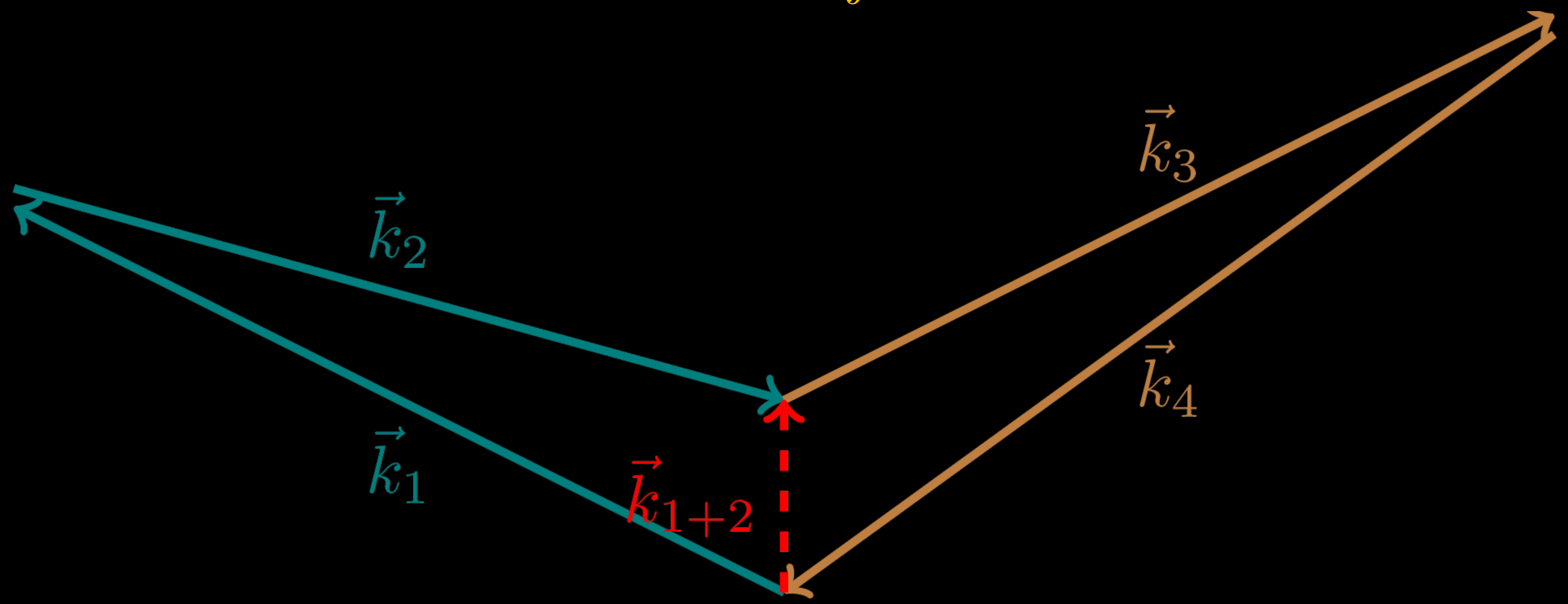
Mildly non-linear

Highly non-linear

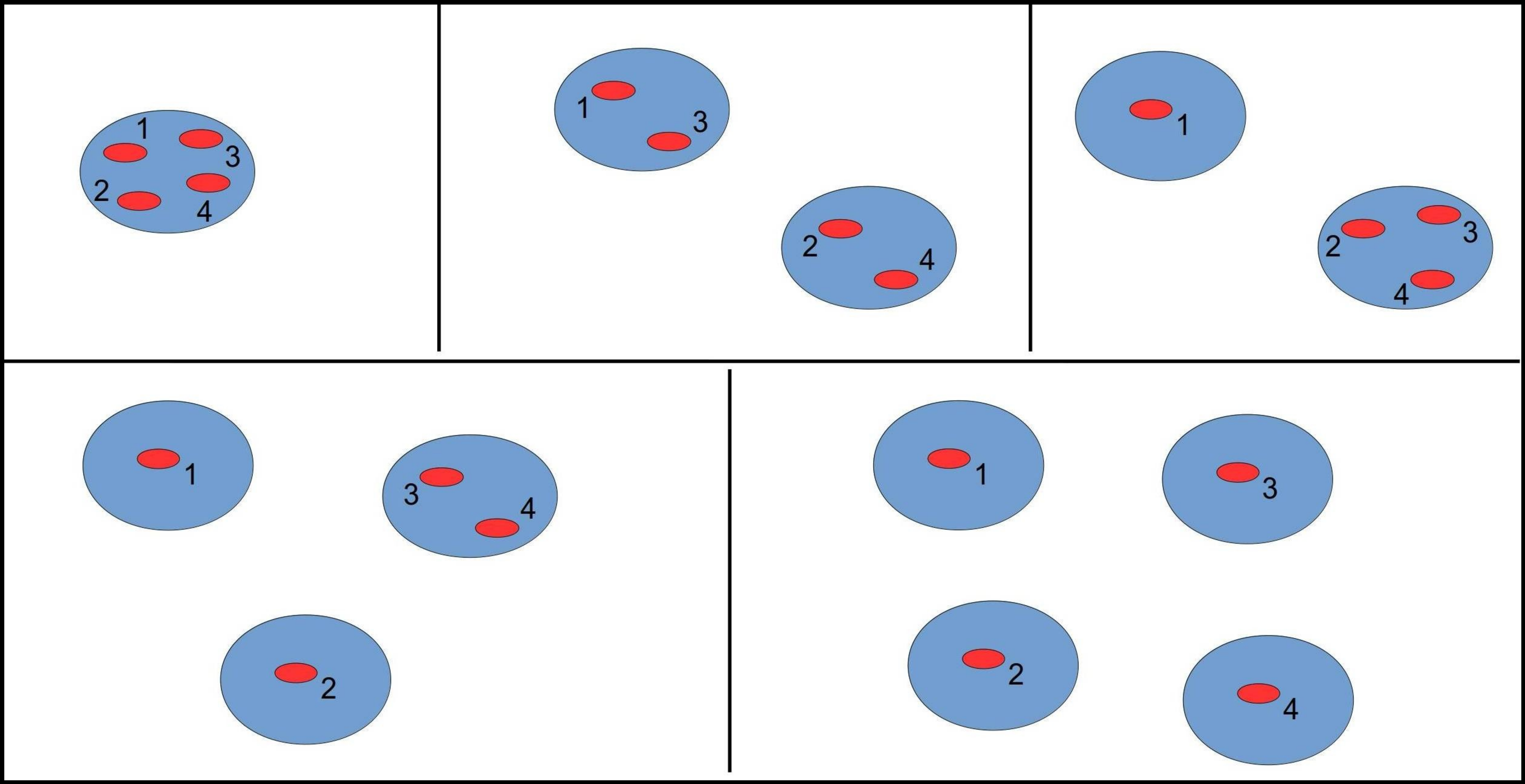
Covariance and 4-point function

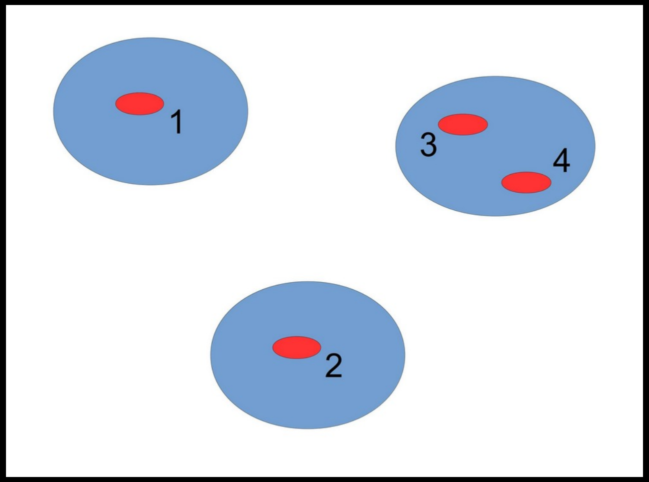
$$\mathrm{Cov}\left(C_\ell, C_{\ell'}\right) = \mathrm{Cov}_G + \mathrm{Cov}_{NG}$$

$$\mathrm{Cov}_{NG} \propto T_{\ell,\ell}^{\ell',\ell'}(L=0) \propto \int T_{\mathrm{gal}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4)$$

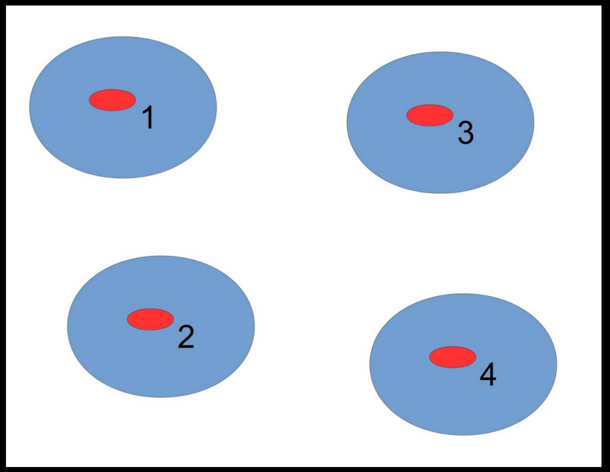


Diagrams for Cov(Cl)





Higher order terms



Context recall

$$\delta_h(\mathbf{x}) = b_1\,\delta_m(\mathbf{x}) + b_2\left(\delta_m(\mathbf{x})^2 - \langle\delta_m^2\rangle\right) + b_{s^2}\,s^2(\mathbf{x}) + \cdots = \sum_{\mathcal{O}} b_{\mathcal{O}}\,\mathcal{O}(\mathbf{x})$$

$$\delta_m = \delta_m^{(1)} + \delta_m^{(2)} + \delta_m^{(3)} \qquad \delta_m^{(n)}(\mathbf{k}) = \int \mathrm{d}^3\mathbf{q}_{1\cdots n}\delta_D(\mathbf{k}-\mathbf{q}_{1+\cdots+n})F_n(\mathbf{q}_{1\cdots n})\delta_m^{(1)}(\mathbf{q_1})\cdots\delta_{\mathbf{m}}^{(\mathbf{1})}(\mathbf{q_n})$$

3-halo term $$T_{3\mathrm{h}}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4) \propto B_{\mathrm{halo}}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3+\mathbf{k}_4) + 5\,\mathrm{perm.}$$

$$B_{\mathrm{halo}} = B_{2\mathrm{PT}} + B_{b_2} + B_{s^2}$$

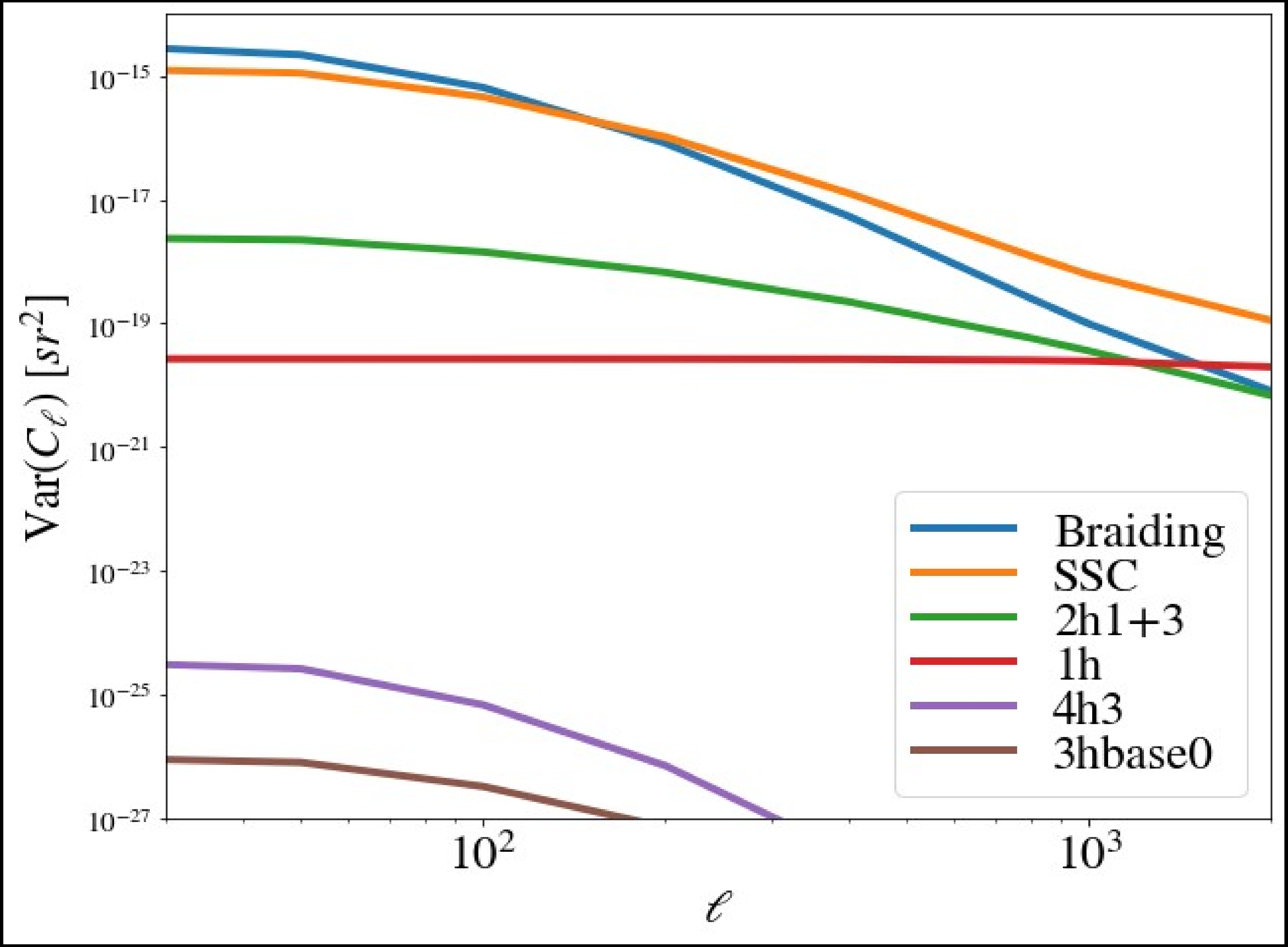
$$B_{2\mathrm{PT}}(k_{123}|M_{123}) = b_1(M_1)\,b_1(M_2)\,b_1(M_3)\,2!\,F_2(\mathbf{k}_1,\mathbf{k}_2)\,P(k_1)\,P(k_2) + 2\ \mathrm{perm.}$$

4-halo term $$T_{4\mathrm{h}}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4) \propto T_{\mathrm{halo}}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4) = T^{2\times2} + T^{1\times3}$$

$$T^{\mathrm{2PTx2PT}}(\mathbf{k}_{1234}|M_{1234}) = b_1(M_1)b_1(M_2)b_1(M_3)b_1(M_4)4\,F_2(\mathbf{k}_{1+3},-\mathbf{k}_1)F_2(\mathbf{k}_{1+3},\mathbf{k}_2) \times P(k_{1+3})\,P(k_1)\,P(k_2) + 11\,\mathrm{perm.}$$

$$T^{\mathrm{3PT}}(\mathbf{k}_{1234}|M_{1234}) = b_1(M_1)b_1(M_2)b_1(M_3)b_1(M_4)\,3!\,F_3(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)\,P(k_1)\,P(k_2)\,P(k_3) + 3\,\mathrm{perm.}$$

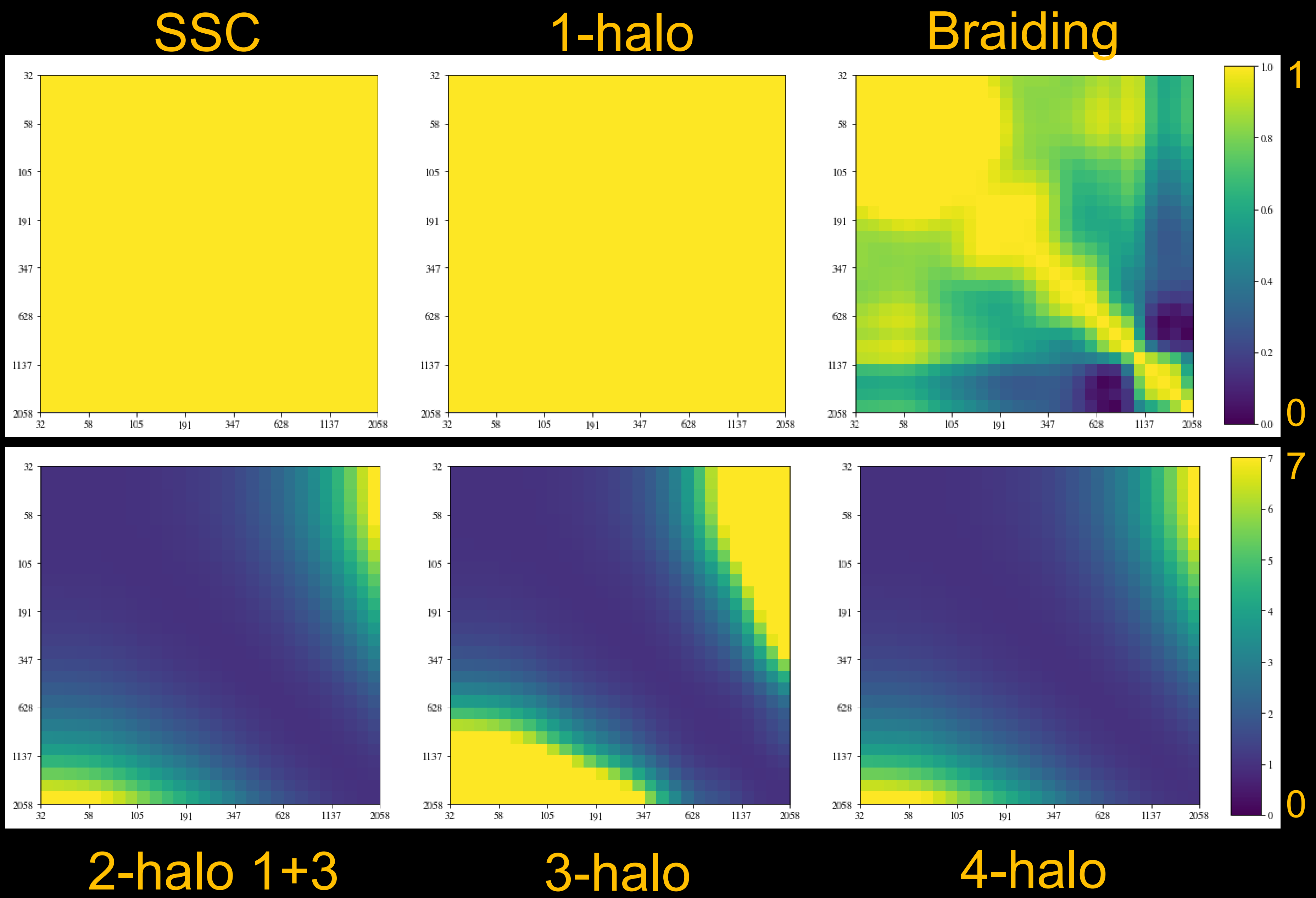
Measurement error bars



z=1

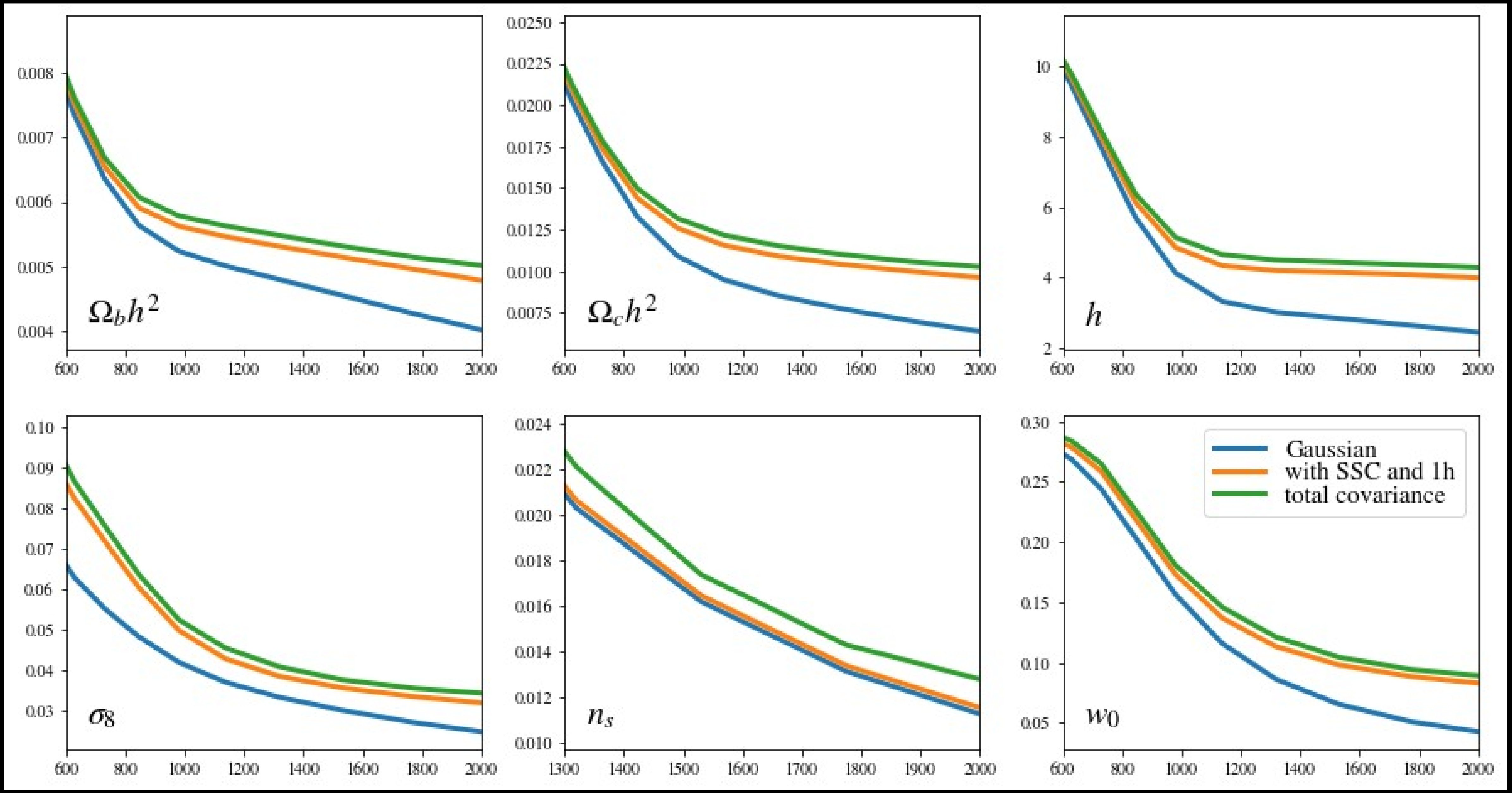
Covariance matrices : off-diagonal importance

$$\frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}$$



z=1

Cosmological error bars

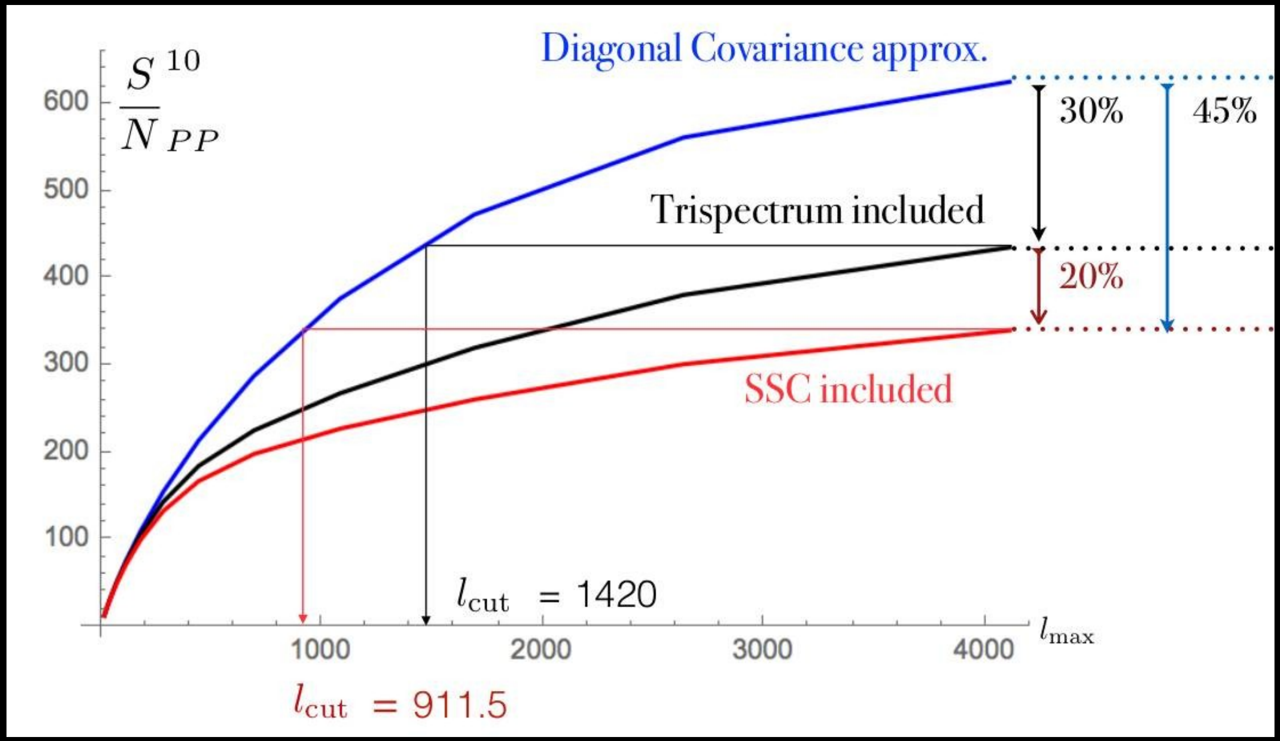


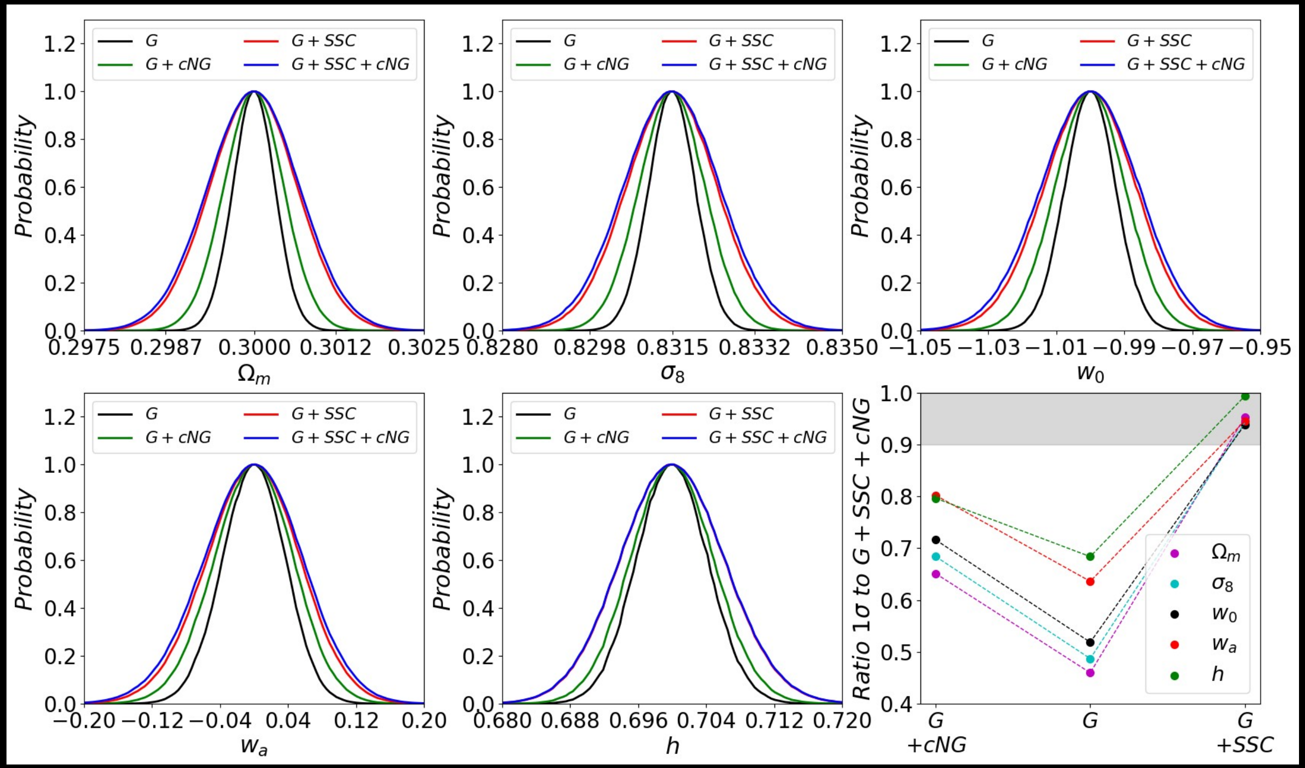
$\ell_{\rm max}$ $\ell_{\rm max}$ $\ell_{\rm max}$

all z

Is SSC important ?

Weak lensing : **yes**





Rizzato et al. 2018

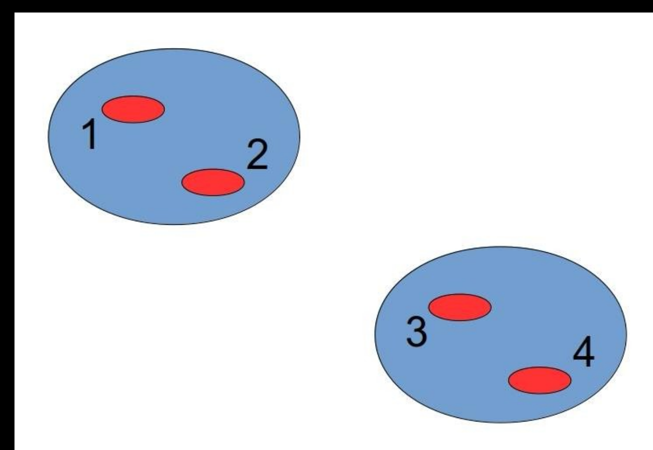
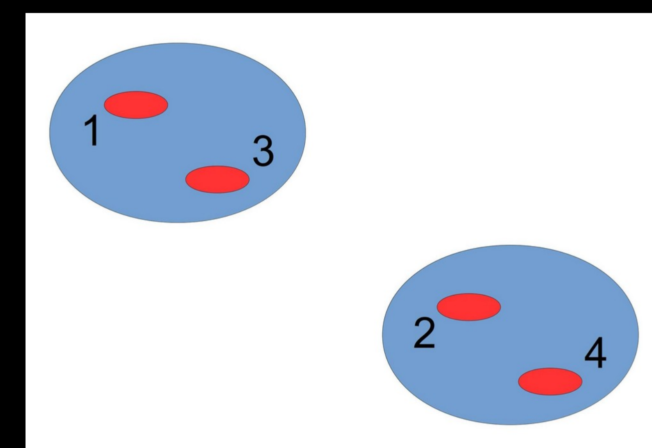
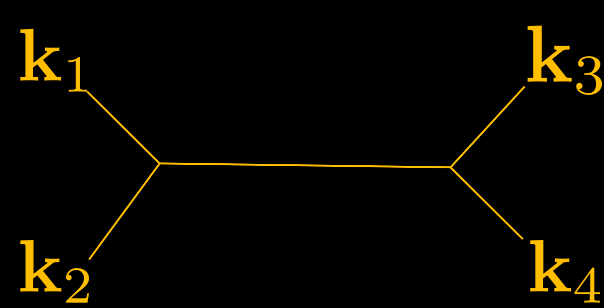
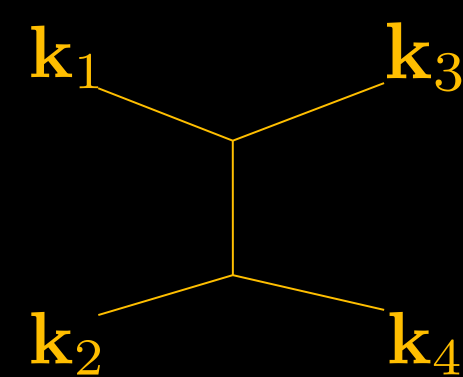
Euclid : decrease of S/N by factor ~2

Barreira et al. 2018

Euclid : error bars increase +30% to +110%

DE, σ_8 and Ω_m particularly affected

SSC vs Braiding

	SSC	Braiding
Terms in the trispectrum	$\propto P(\mathbf{k}_1 + \mathbf{k}_2)$	$\propto P(\mathbf{k}_1 + \mathbf{k}_3)$
Diagram example (2-halo)		
Feynman diagram (4-halo, 2x2)		
Physical interpretation	Modulation of the two spectra by a long wavelength (super-survey) matter mode	Modulation of each mode by an intra-survey matter mode ?