Constraining Non-Gaussianity with Large Scale Structure

Anže Slosar, Brookhaven National Laboratory

Introduction

Plan for the talk:

- Pedagogical introduction to local non-Gaussianity
- Dalal effect and the question of φ bias
- Non-Gaussianity with alternative tracers
- Future constraints of f_{NL} from LSS.

What is non-Gaussianity in cosmological context

- Most physicist assume that non-Gaussianity means that the distribution of density fluctuations is distributed in a non-Gaussian way:
  - true, but wrong way to think about it
- For Gaussian fields:
  - Power spectrum gives everything
  - Odd correlators are zero, even correlators are given by Wick's theorem
  - 2nd cumulant is non-zero, all the others are zero
  - Primordial field that breaks this condition is non-Gaussian and non-Gaussianty is defined and probed in terms of correlators
- At the lowest order:
  - non-zero 3-point function of bi-spectrum

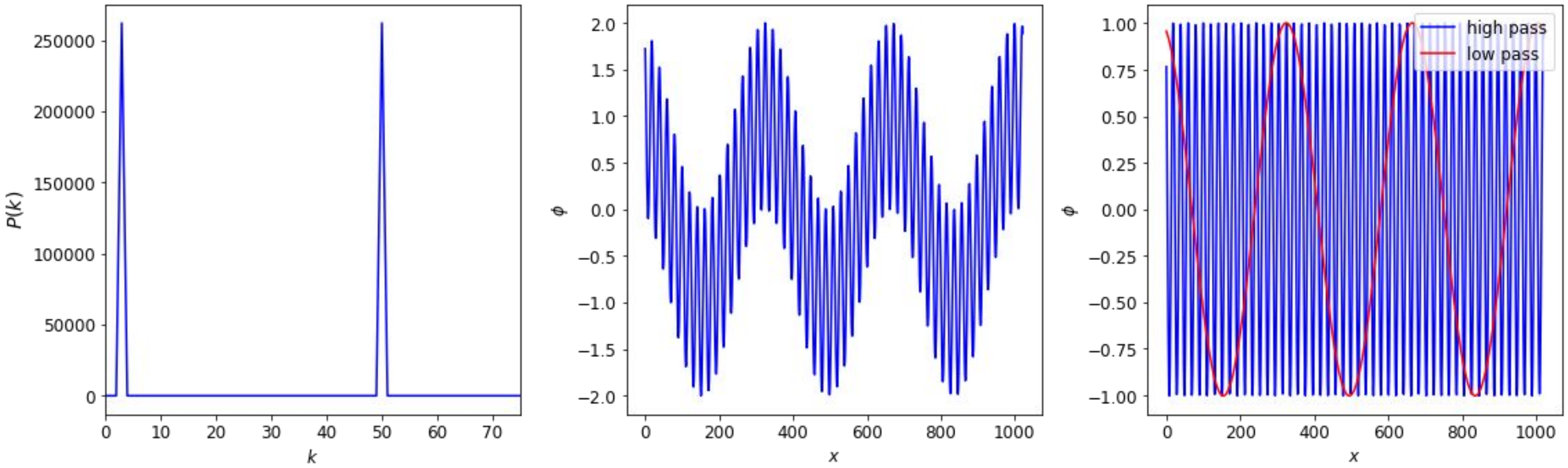
Local non-Gaussianity

Three statements are usually thrown around when talking about local non-Gaussianity:

1. : $$\phi = \phi_g + f_{\rm NL} \left(\phi_g^2 - \langle \phi_g^2 \rangle \right)$$
2. squeezed bispectrum shape
3. correlation between the large scale fluctuations and small scale power spectrum

Toy example

- Imagine a universe with 2 modes: a large scale mode and a small scale mode

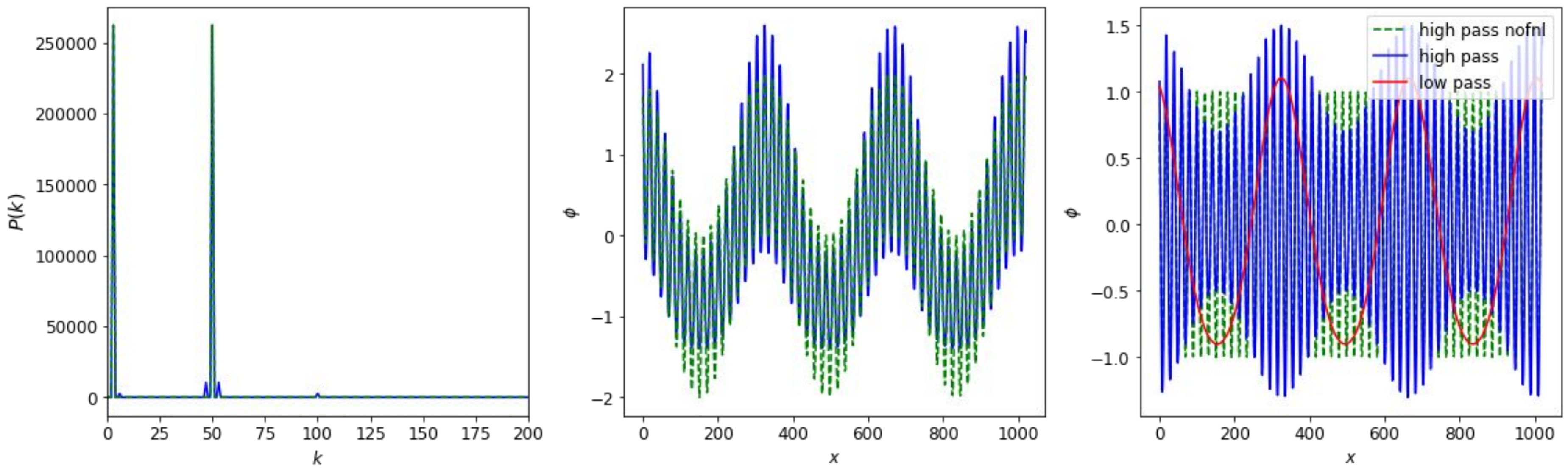


$$\phi = \cos(k_L x) + \cos(k_S x)$$

$$\phi = \phi_g + f_{\mathrm{NL}} \left(\phi_g^2 - \langle \phi_g^2 \rangle\right)$$

$$\phi = \phi_g + \boxed{2 f_{\mathrm{NL}} \cos(k_L x) \cos(k_S x)} + \frac{f_{\mathrm{NL}}}{2} \left(\cos(2k_L x) + \cos(2k_S x)\right)$$

- Modification of power spectrum
- Coupling of large and small-scale modes



More realistic case

Split fields into large scale, approx const over a patch and small-scale:

$$\phi = \phi_L + \phi_S$$

Since $\delta \propto \nabla^2 \phi$, in a small patch one gets

$$\delta_{\rm NG} = \delta_L + \delta_S(1 + 2f_{\rm NL}\phi_L)$$

Net result in a patch of cosmic volume, the statistics of the small scale field are modulated by

- Value of large scale overdensity which raises/lowers the mean
- Value of large scale potential, which modulates small scale power
- (modifications to the shape of power spectrum at f_{NL} level)

Putting it all together

- Local non-Gaussianity, is a local transformation in φ
- It causes correlation of large scale potential with a small scale amplitude
- It generates bispectrum of the kind $\langle \delta_L \delta_S \delta_S \rangle \propto f_{\rm NL} P_L P_S$
- This particular shape is a squeezed triangle shape

Common question:

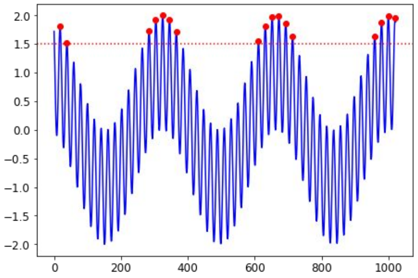
- If δ and φ are connected by Poisson equation, how can we treat them as independent variable?

$$\phi(0) \propto \int \phi_k W(k) d^3k$$

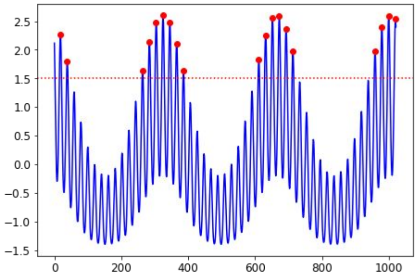
$$\delta(0) \propto \int k^2 \phi_k W(k) d^3k$$

Going to tracers

- In a Press-Schecter picture, objects happen when fluctuations go over a barrier



Gaussian



non-Gaussian

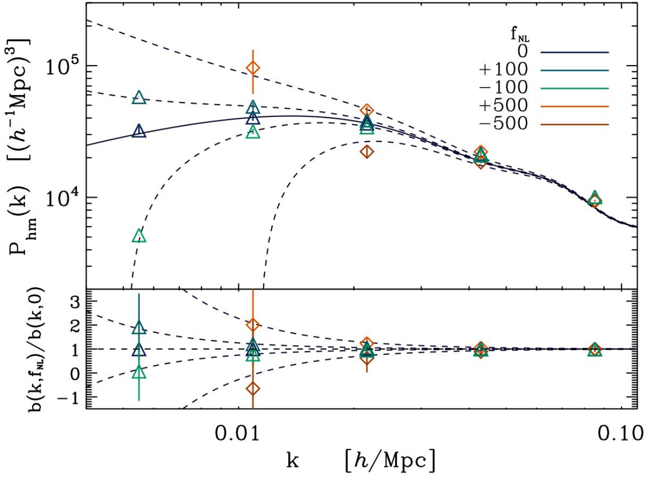
- Collapsed objects form whenever local density is sufficiently high:
  - it can be helped over the barrier by a large scale over-density mode
  - or a large scale φ mode

$$\delta_t = b_1 \delta + b_\phi \phi$$

Dalal (2007) effect

- For a universal mass function and vanilla HOD with no assembly bias

$$\Delta b(M,k) = 3f_{\mathrm{NL}}(b-1)\delta_c \frac{\Omega_m}{k^2 T(k) D(z)} \left(\frac{H_0}{c}\right)^2$$

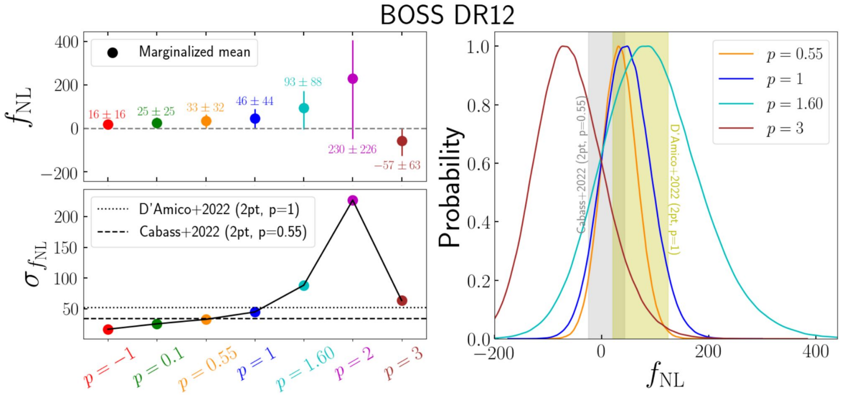


This was unexpected at the time

- Since $\varphi = O(10^{-5})$, then $f_{NL} = 1$ corresponds to 10^{-5} correction
- The effect is so large in power spectrum because the number density is exponentially dependent on the power spectrum amplitude
- Broadband corrections to the power spectrum shape are very very small
- There are other shapes of bispectrum:
    - See massive review by Desjacques, Jeong & Schmidt (2016)
    - b_φ can is proportional to
        - k^{-2} for local NG
        - k^{-1} for orthogonal NG
        - k^{0} for most NGs
- If you loose Dalal effect, it is going to be hard:
    - need to differentiate bispectrum from non-linearities vs primordial contribution

Note that this is in general not true

- General tracers do not need to respond to change in small scale power like this
- It is a fundamental problem, since we cannot not know b_φ any other way
- You cannot escape this effect, Alex will find you

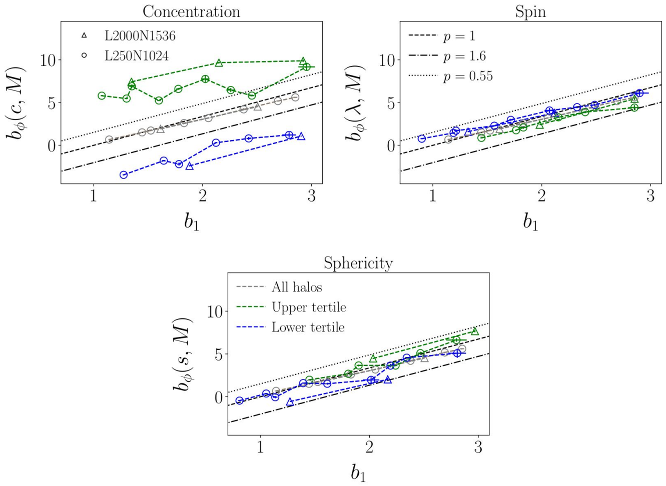


Barreira 2022

Justin Bieber after writing f_{NL} forecasting paper

Is this a problem

- If we detect something, mostly an annoyance:
  - detection significance not impacted
  - when detected in multiple surveys can start to disentangle
- If we don't, it is a major problem:
  - cannot compare upper limits apples to apples
  - cannot really trust forecasts
- Is there a way out?
  - It will have to involve simulations
  - With sufficient work we can probably reduce this systematic error to ~20% level from O(1) level, but not eliminate it



From Lazeyras et al. 2022

What about other tracers?

- For a general tracer
  - If you have a fluctuating field F so that $$\delta(x) = \frac{F(x)}{\bar{F}}$$
  - Then

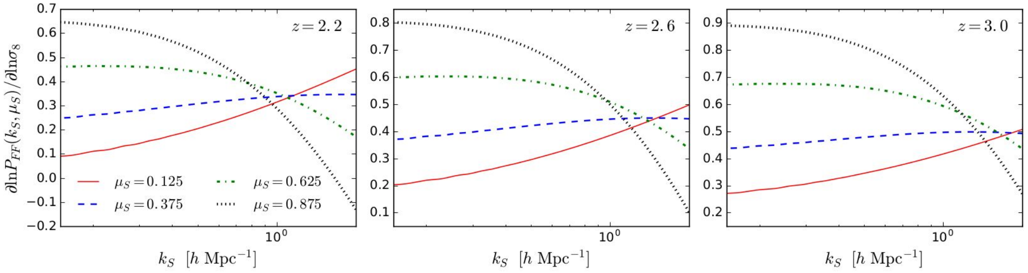
$$b_1 = \frac{d\ln \bar{F}}{d\delta_L}$$

$$b_\phi = 4\frac{d\ln \bar{F}}{d\sigma_8^2}$$

  - For galaxy density, we have $\bar{F} = \bar{n}$, but in general, one can use **anything,** as long as you can calculate its response to change in small scale power spectrum.
  - Can rely on simulations

Example

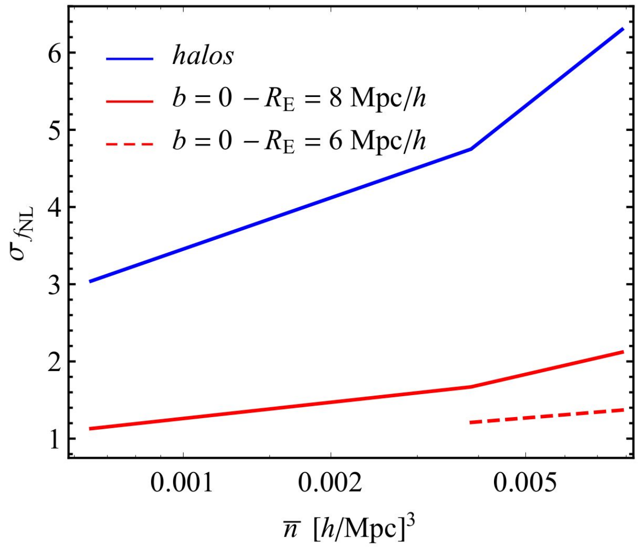
- In Chiang et al 2017, we did this for the Lyman-α forest



- Here F is small scale power spectrum
- This is in the deeply non-linear regime, but doesn't matter
- Results were somewhat disappointing with DESI sensitivity to local non-Gaussianity of only ~60.

Example 2:

- Castorina et al 2018
- They engineer a sample that has b_1=0. For sufficiently low-shot noise the sample variance “cancels”

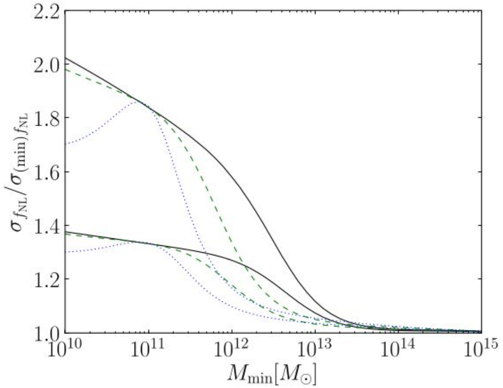


Example 3:

- If you have a sample of galaxies for which you have a secondary information, for example an estimate of an individual host halo mass from stellar mass, you can create a new sample by more optimal weighting:
- From 0808.0044
- These results probably wrong, because they assume independent Poisson shot noise
- Something that is probably worth revisiting for the next-generation of experiments

$$\alpha(M) = \frac{b(M) - b_{\min}}{b_{\max} - b_{\min}}$$

$$\beta(M) = \frac{b_{\max} - b(M)}{b_{\max} - b_{\min}},$$

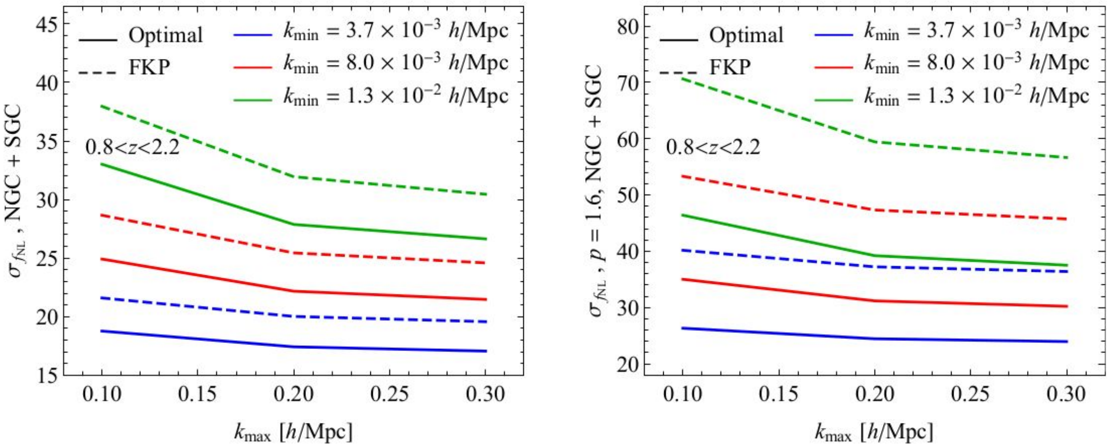


Engineering tracers

- I think there is a lot to explore in this direction
- A specially engineered tracers can help with:
    - one can use any information:
        - secondary typing beyond redshift in spectroscopic
        - shape and color information in photometric
        - local environment, etc
    - systematic effects
    - knowledge of phi bias
    - maximization of SNR

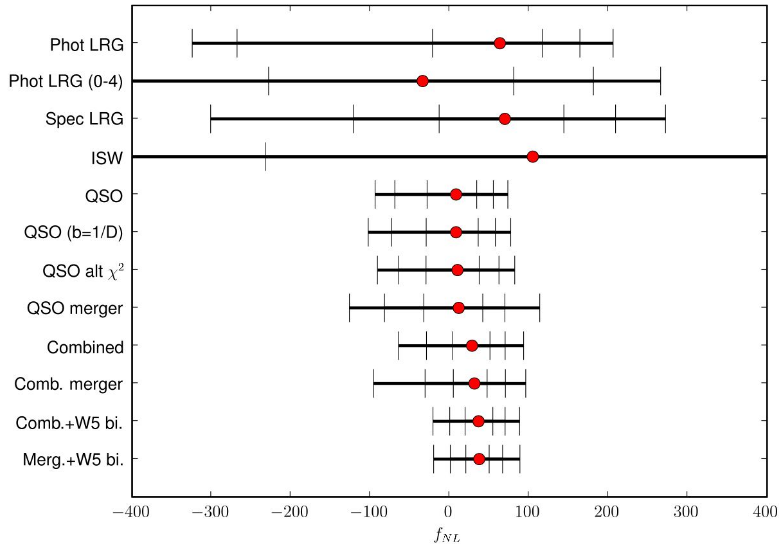
Redshift weighting

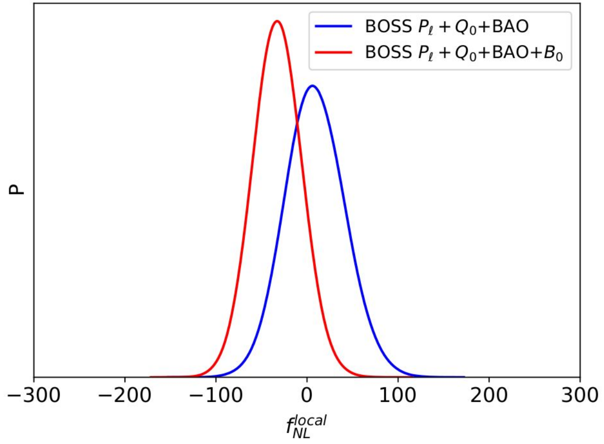
- Within sample, you can optimize by weighting with ~halo mass
- Across redshift, you can optimally weight wrt to required SNR
- E.g. Castorina et al, 2019, signal for BOSS quasars improves by ~40%:



Current constraints

CMB constraints still the strongest by some margin!





Cabass et al 22: $f_{\rm NL}^{\rm local} = -33 \pm 28$

2008: $-29\ (-65) < f_{\rm NL} < +70\ (+93)$

Compared to e.g. BAO, constraints have improved surprisingly little: SDSS photometric QSOs remain by far the largest volume tracer

See also d'Amico et al 2022 , Castorina et al, 2019, Leistedt et al 2015, etc. Note that these constraints contain bispectrum, which improves results by ~20-50%.

Constraints from auto power spectrum

- In auto power spectrum, systematics can only add power (if correlated than not a systematic)
- Therefore

$$f_{\mathrm{NL}}^{\mathrm{meas}} = f_{\mathrm{NL}} + f_{\mathrm{NL}}^{\mathrm{sys}}$$

- Any contamination is always positive
- Upper limits on non-Gaussianity are robust

Going forward

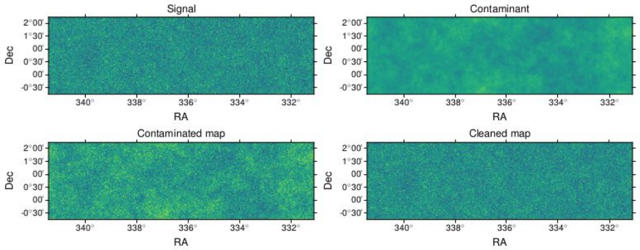
- To measure non-Gaussianity one needs to go to the largest scales
- The largest scales are systematically most uncertain
- Many years ago we had “detections” from BOSS quasars
- The most robust way is in **cross-correlation of two tracers that have least in common**:
  - either tracer-tracer power spectrum on large scales or bispectrum with long modes from survey #1 vs two short modes from survey #2
  - CMB kSZ cross quasars (Münchmeyer et al 2018), also sample variance cancellation
  - lensing cross quasars (Giannantonio & Percival 2014, etc.)
  - 21cm small scale power spectrum cross quasars (write a paper!)
- Note that two galaxy samples is better than one, but could easily share large-scale systematics (reddening etc)

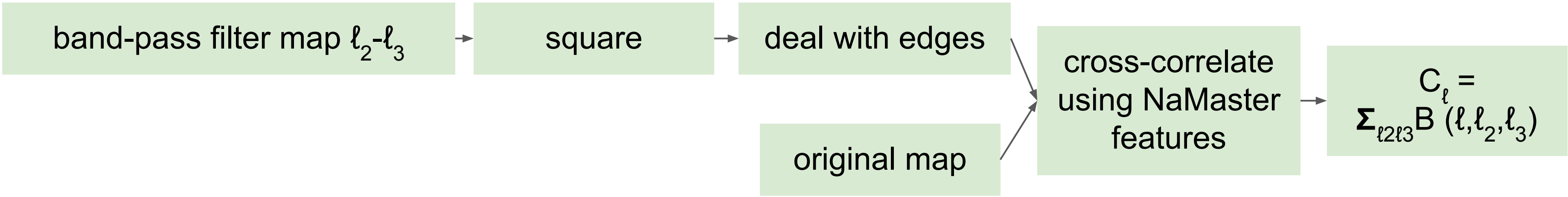
Future Surveys

- Current surveys:
    - Lots of work ongoing in BOSS/DESI, but VRO LSST could see more effort:
        - Need to write general pipelines
        - Work on Alex's problem
        - Think of non-canonical tracers (e.g. galaxy morphology)
- Future Surveys:
    - In the US Snowmass process, Stage 5 spectroscopic facility seems to have won:
        - A dedicated large-volume survey at z>2 focusing on large N_{lin} mode science
        - Focus on inflationary science: non-Gaussianity, but also features
        - Implementation not decided yet:
            - MegaMapper or some version thereof might be the most likely candidate:
                - 6m telescope, 24000 fibers
            - Alternatives typically assume a bigger telescope
    - A general goal is to reach f_{NL}~1

Extending MASTER algorithm for bispectrum

- NaMaster: David Alonso's child: a very robust power spectrum calculating machine
- Implements MASTER algorithm, but lots of features:
    - spin-0,1 and 2
    - Careful window treatment
    - template subtraction
    - Full Gaussian covariance matrix
    - ,,
- We want to use these features to calculate bispectrum
- Long history of "position dependent power spectrum" (Komatsu, Chiang cca 2015)
- A continuous version of this idea





Conclusions

- Local non-Gaussianity causes correlations between large scale potential modes and small scales power spectra
- For tracers it leads to a large-scale power spectrum corrections that can lead to detection of non-Gaussianity, but with a poorly know constant of proportionality
- Real-life constraints have improved relatively modestly over the past decade, but constraints should improve a lot in the coming decade
- We are reaching limits of easy constraints