Why do we need N-body simulations for WL I

- WL probes LSS on small, non-linear scales.
  Cosmic shear: down to sub-Mpc. Surveys sensitive to $k \sim 50h/\text{Mpc}$.
  Need theoretical prediction of non-linear power spectrum.
  (Semi-)analytical approaches go to $k \sim 0.5h/\text{Mpc}$.
- Shear field follows non-Gaussian distribution.
  Follows from the fact that δ in non-linear regime is non-Gaussian.
  Complex survey geometry modify distribution.
  At the least, need non-Gaussian covariance for likelihood.
  Difficult from (semi-)analytical models (see previous point).
- Baryonic physics modifies dark-matter halo properties (profile, concentration, ...).
  Model with hydro-dynamical simulations.
- Systematic effects that correlate to astrophysics or the LSS.
  Can use forward modelling for complex physical processes in N-body simulation.

Why do we need N-body simulations for WL II

Examples

- Blended galaxy images lead to deselection of galaxies in crowded fields, which are correlated to high-density regions, that are then under-represented. This leads to biases in inferred $n(z)$, cosmological parameters.
- ξ_{sys} = correlation between stars and PSF-corrected galaxies = measure of PSF residuals in galaxy shapes. But: Need to account for chance alignment between PSF and LSS.

- Test mathematical useful approximations: Born, neglecting lens-lens coupling, reduced shear $g = \gamma/(1-\kappa)$ versus shear γ. Most of these effects introduce higher-order correlations, again difficult to solve unless Gaussian limits.

Ray-tracing I

Principle of Ray-tracing

- Numerical evaluation of projection integral from particle distribution in N-body simulation.
- Most algorithms first project particles on *multiple lens planes* (Blandford & Narayan 1986), with $\Delta z/z$ of order 0.03 - 0.05.
  Corresponds to a finite-sum discretization of the projection integral.

$$\boldsymbol{\alpha} = \boldsymbol{\nabla}_\theta \psi = \frac{2}{c^2} \int_0^\chi \mathrm{d}\chi' \frac{\chi - \chi'}{\chi} \boldsymbol{\nabla}_\perp \Phi(\boldsymbol{x}(\chi'), \chi').$$

On each lens plane, compute Jacobi matrix $A_{ij} = \partial \beta_i / \partial_j = \delta_{ij} - \partial_i \partial_j \psi$.
Algorithm (Hilbert et al. 2009):

$$\mathsf{A}_{ij}^{(k)}(\boldsymbol{\theta}) = \delta_{ij} - \sum_{n=1}^{k-1} \frac{f_K^{(n,k)}}{f_K^{(k)}} \mathsf{U}_{ij}^{(n)}(\boldsymbol{\theta}). \qquad \mathsf{U}_{ij}^{(k)} = \frac{\partial^2 \psi^{(k)}(\boldsymbol{\beta}^{(k)})}{\partial \beta_i^{(k)} \partial \beta_j^{(k)}} = \frac{\partial \alpha_i^{(k)}(\boldsymbol{\beta}^{(k)})}{\partial \beta_j^{(k)}} \,.$$

Ray-shooting versus ray-tracing

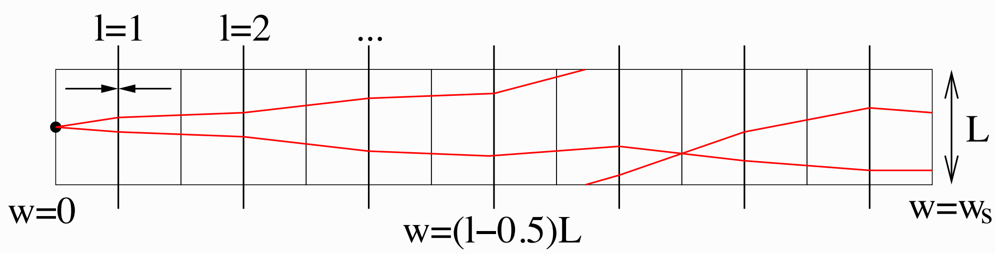
Two methods are common to propagate photons for the projection:

1. Ray shooting:
   Compute cumulative lensing potential ϕ on a grid. Light rays travel on (unperturbed) straight lines, corresponds to Born approximation.
2. Ray tracing:
   Additionally compute deflection angle $\boldsymbol{\alpha}$, change direction of light ray accordingly.
   Light rays travel on straight lines between lens planes, where they change direction.
   Start at observer and shoot backwards. Why?

Ray-shooting versus ray-tracing

Two methods are common to propagate photons for the projection:

1. Ray shooting:
   Compute cumulative lensing potential ϕ on a grid. Light rays travel on (unperturbed) straight lines, corresponds to Born approximation.
2. Ray tracing:
   Additionally compute deflection angle $\boldsymbol{\alpha}$, change direction of light ray accordingly.
   Light rays travel on straight lines between lens planes, where they change direction.
   Start at observer and shoot backwards. Why?

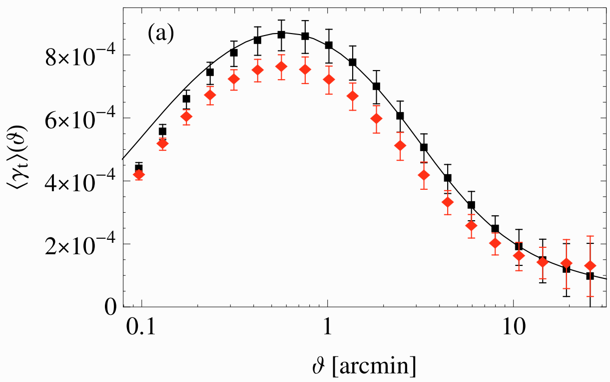


From Hartlap, PhD thesis 2005

Ray shooting

For cosmic shear ray-shooting is a very good (percent-level) approximation.

However, for galaxy-galaxy lensing this is not the case.



From (Hilbert et al. 2009).

This is because relative distance between light rays from two bg galaxies for cosmic shear not much affected by coherent deflection.

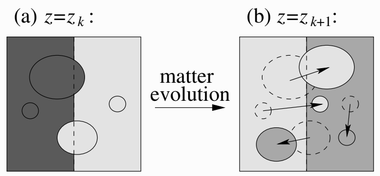
But distance between light ray from bg galaxy and fg galaxy position (impact parameter) is affected.

Ray-tracing approximations

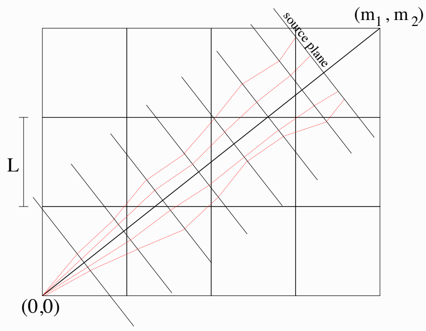
- Ray-tracing through N-body output snapshot boxes: Fixed cosmic time, neglecting LSS evolution during photon travel time through box. Limit box size to $L \lesssim 300$ Mpc.

  Larger boxes can be split and projected to more than one lens plane, but:
  - Avoid cutting through halos
  - Leads to loss of power on large scales

  Use snapshots at different output times to account for time evolution.
  If box size is small, boxes have to be concatenated. To avoid photons to encounter repeated structures at different epochs:
  - Rotate and translate randomly.
  - Shoot light rays under an skewed angle.



From (Hilbert et al. 2009).

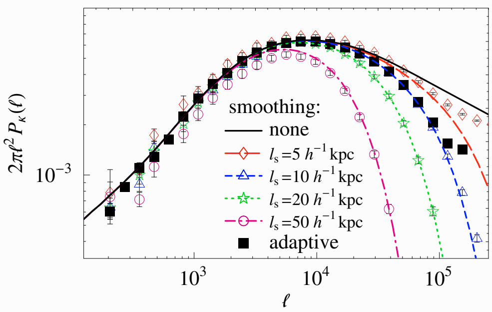


From Hartlap, PhD thesis 2005

Ray-tracing approximations

- Get shear and convergence by FFT, or finite differences in real space:
  Smoothing is necessary to reduce Poisson noise of N-body discrete particle distribution.

  However, other limitation is N-body resolution.



From (Hilbert et al. 2009).

- From Cartesian flat-sky simulations, lens planes are by construction parallel:
  - Neglects sky curvature.
  - Gradient of potential not orthogonal to light ray

  This limits simulated field of view to a few degreees.
  With convergence maps created on say grids of 1024^2 pixels $\rightarrow$ resolution of around 0.2 arcmin.
- Newtonian physics, neglects GR effects. Also, MoG simulations not possible under Newtonian approximation.

Further methods

- Compute lensing Jacobian on the fly while running N-body simulation (White & Hu 2000).
  Circumvents lens plane projections, allows for slightly higher time resolution.

  Easy for ray-shooting where photon tractories are known before hand, more difficult for ray tracing (Li et al. 2011).
- Store density field at different time steps on surface moving towards box center (= observer) with speed of light, use those after run ends for lensing projections (Teyssier et al. 2009).
- Full-sky simulations, for large upcoming surveys, CMB lensing.
  Create spherical concentric shells around observer on the fly, project onto *lens spheres*. (Fosalba et al. 2008, Das & Bode 2008, Teyssier et al. 2009, Becker 2013).
- General-relativity simulations.
- Modified gravity simulations.
  Take ~ 5 times compared to Newtonian ones.

Hydro-dynamical simulations I

Important processes to simulate:

- Gas pressure, $R \sim 1$ - 0.1 Mpc, suppression of structure formation, gas distribution is more diffuse than dark matter
- Baryonic cooling, $R < 0.1$ Mpc ($k > 10$/Mpc), gas condenses into stars and galaxies, more strongly clustered than dark matter
- AGN and SN feedback

Simulation methods

Dark matter usually simulated as (very massive) particles.

Hydrodynamic physics often simulated in cells on a grid (adaptive).

Non-resolved physical processes, effective treatment within cell ("sub-grid physics").

Hydrodynamical simulation can often not reproduce observational results, e.g. on AGN feedback. Need to calibrate simulations with observations.

Hydro-dynamical simulations II

Influence on WL

- Need to know total (dark + baryonic) power spectrum to 1-2% at k up to $10h/\mathrm{Mpc}$.
- Baryons (15% of total matter) behave differently than dark matter, but dark matter is influenced by this, e.g. slightly follows distribution of baryons
- P_κ strongly influenced for $\ell \geq 1000$ to 3000 (depending on statistical errors).

Mitigation of baryonic effects

- Removing small scales from survey analysis.
- Model baryonic effects e.g. with halo model. Fit to simulations, marginalise over nuisance parameters or different models.
- Self-calibration using combination of observations. E.g. additional observations of halo structure (Zentner et al. 2008), power spectrum and bi-spectrum (Semboloni et al. 2013).

WL covariance

General definition

Covariance of data vector $\boldsymbol{d} = \{d_i\}, i = 1 \ldots m$:

$$C_{ij} = \langle \Delta d_i \Delta d_j \rangle = \langle d_i d_j \rangle - \langle d_i \rangle \langle d_j \rangle,$$

Examples of $\boldsymbol{d}$:

$$d_i = P_\kappa(\ell_i); \qquad d_i = \xi_+(\vartheta_i); \qquad d_i = \langle M_{\mathrm{ap}}(\theta_i).$$

Case of data vector $= \hat{\xi}_\pm$

Recall the estimator for $\xi_\pm$:

$$\hat{\xi}_\pm(\theta) = \frac{\sum_{ij} w_i w_j \left(\varepsilon_{\mathrm{t},i}\varepsilon_{\mathrm{t},j} \pm \varepsilon_{\times,i}\varepsilon_{\times,j}\right)}{\sum_{ij} w_i w_j}.$$

Very roughly:

$$C \sim \langle \xi_\pm \xi_\pm \rangle \sim \langle \varepsilon\varepsilon\varepsilon\varepsilon \rangle.$$

With weak-lensing relation $\varepsilon = \varepsilon^{\mathrm{s}} + \gamma$:

$$C \sim \langle (\varepsilon^{\mathrm{s}})^4 \rangle + \langle (\varepsilon^{\mathrm{s}}\gamma)^2 \rangle + \langle \gamma^4 \rangle \equiv D + M + V$$

WL covariance components

- $D = \sigma_\varepsilon^4$: Poisson noise from intrinsic ellipticities, shape noise
- M: mixed term
- V: shear covariance, cosmic variance, if shear field approximated having Gaussian distribution (which it does not):
  $V \sim 3\langle\gamma^2\rangle$.

  Otherwise, need to account for connected 4-pt (tri-spectrum) term.

Gaussian covariance of power spectrum P_κ

$$\langle(\Delta P_\kappa)^2\rangle(\ell) = \frac{1}{f_{\rm sky}(2\ell+1)} \left(\frac{\sigma_\varepsilon^2}{2\bar n} + P_\kappa(\ell) \right)^2 .$$

Non-Gaussian covariance

Mode coupling

- Couples different ℓ-modes, leads to saturation of information content on small scales
- Tri-spectrum coupling on small scales
- Coupling of small with large scales: *halo sample variance* (HSV), *beat couping*, *super-survey covariance* (SSC) (Takada & Hu 2013).
  SSC descreses faster with $f_{\rm sky}$ than other terms → sub-dominant for large surveys.

Modelling

- Tri-spectrum from halo model (+ PT)
- N-body simulations, but: difficult to include SSC
- From data, by spatial averaging over sub-fields, or Jackknife.

Spatial averaging: number of independent lines of sight n

For non-singular covariance of data vector with length m, need $n > m$.
For precision covariance (error bars on cosmo parameters of $< 5\%$, need $n > 10m$ (Taylor et al. 2013).

Likelihood function

Gaussian likelihood

$$L(\boldsymbol{d}|\boldsymbol{p}, M) = (2\pi)^{-m/2} |C(\boldsymbol{p}, M)|^{-1/2} \times \exp\left[-\frac{1}{2} \left(\boldsymbol{d} - \boldsymbol{y}(\boldsymbol{p}, M)\right)^{\mathrm{t}} C^{-1}(\boldsymbol{p}, M) \left(\boldsymbol{d} - \boldsymbol{y}(\boldsymbol{p}, M)\right) \right].$$

with $\boldsymbol{d}$ = data vector, C = covariance matrix, $\boldsymbol{y}$ = model, $\boldsymbol{p}$ = (cosmo) parameter vector, M = cosmological model.

But: True likelihood is non-Gaussian.

Model non-Gaussianity of observables:

- N-body simulations (very time-consuming)
- Transform data to be more Gaussian
- Approximate Bayesian Computation (ABC) sampling

Bayesian parameter inference

Bayes' theorem

Likelihood: probability of data given parameters and model

Prior

$$p(\boldsymbol{\pi}|\boldsymbol{x}, m) = \frac{L(\boldsymbol{x}|\boldsymbol{\pi}, m) P(\boldsymbol{\pi}|m)}{E(\boldsymbol{x}|m)}$$

Posterior: probability of parameters given data and model

Evidence

$\boldsymbol{\pi}$: parameters
$\boldsymbol{x}$: data
m : *model*

Parameter constraints = integrals over the posterior

$$\int \mathrm{d}^n\pi \, h(\boldsymbol{\pi}) p(\boldsymbol{\pi}|\boldsymbol{x}, m)$$

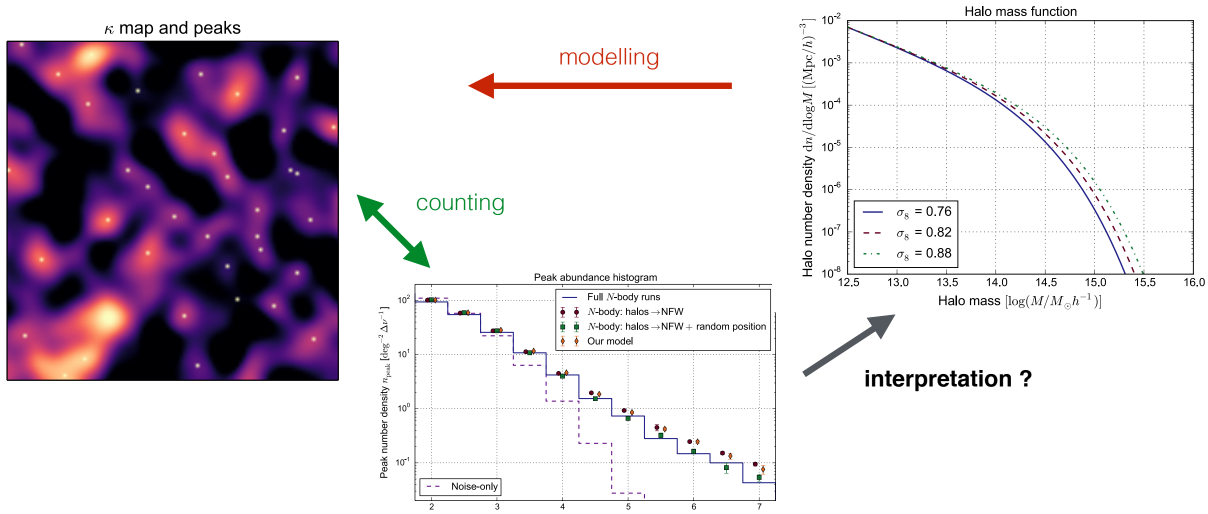
For example:

$h(\boldsymbol{\pi}) = \boldsymbol{\pi}$: mean
$h(\boldsymbol{\pi}) = 1_{68\%}$: 68% credible region

Approaches: Sampling (Monte-Carlo integration), Fisher-matrix approximation, frequentist evaluation, ABC, …

WL peak counts: Why do we want to study peaks?

- WL peaks probe high-density regions ↔ **non-Gaussian** tail of LSS
- **First-order** in observed shear: less sensitive to systematics, circular average!
- High-density regions ↔ **halo mass function**, but **indirect** probe:
  - Intrinsic ellipticity **shape noise**, creating false positives, up-scatter in S/N
  - **Projections** along line of sight



WL peak counts. What are peaks good for?

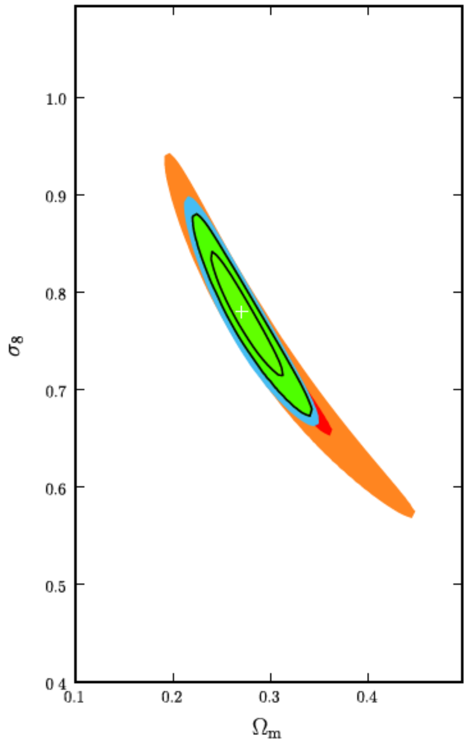
What do we gain from peak counting?

- Additional and complementary information and constraints compared to 2nd order shear
- Non-Gaussian information

Figure from Dietrich & Hartlap 2010

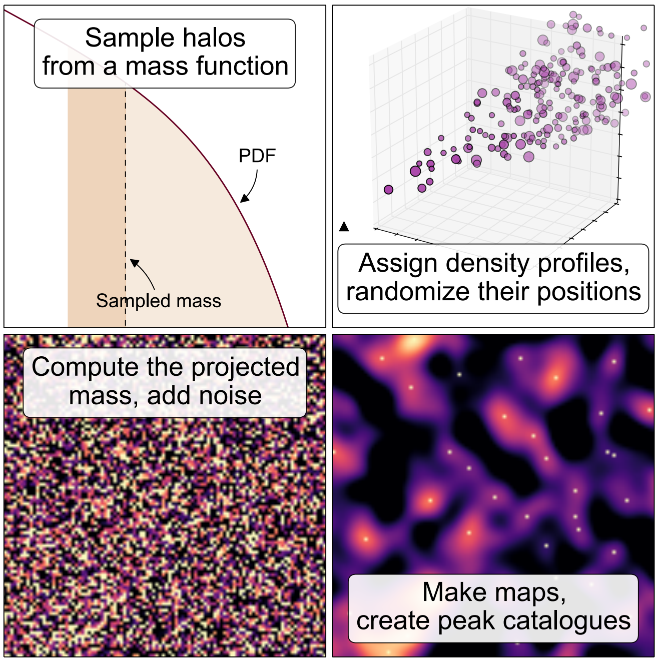
red/orange: cosmic shear

green: shear & peak



WL peaks: A fast stochastic model

Replace N-body simulations by Poisson distribution of halos



Simulating
1-halo term

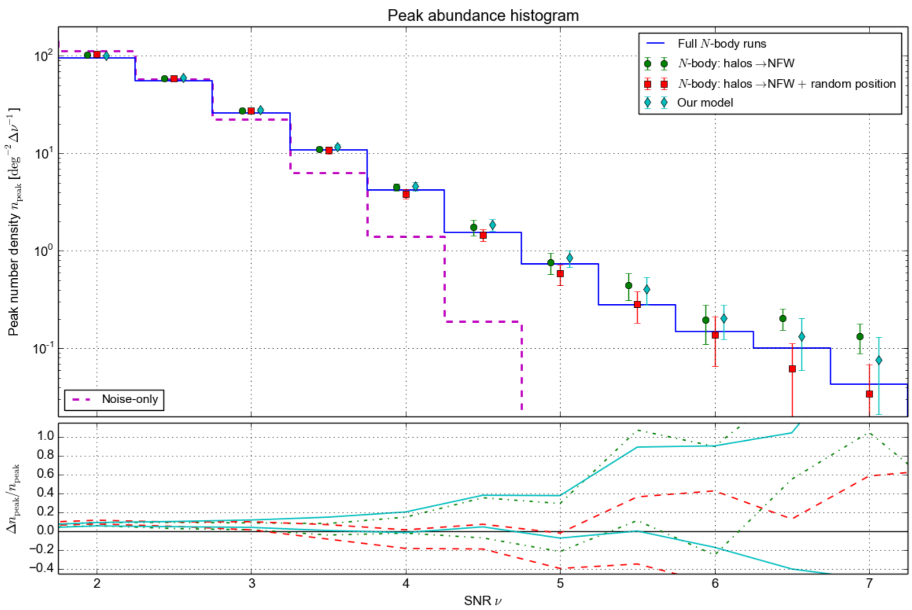
Lin, MK & Pires 2016

WL peaks: histograms

Hypotheses:

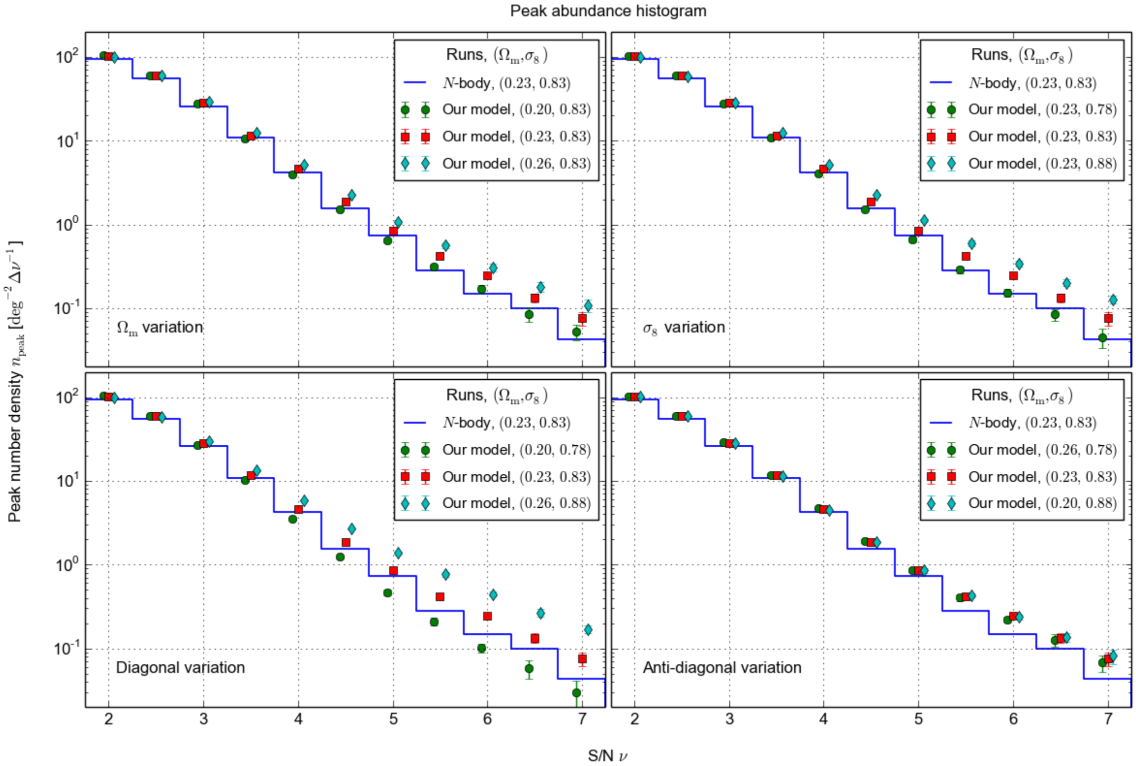
1. Clustering of halos not important for counting peaks (along los: Marian et al. 2013)
2. Unbound LSS does not contribute to WL peaks

Test:



Field of view = 54 deg^2; 10 halo redshift bins from z = 0 to 1; galaxies on regular grid, z_s = 1.0

WL peaks: cosmological parameters



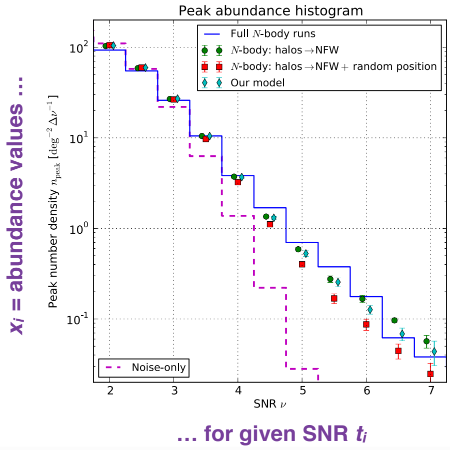
Lin & Kilbinger (2015a)

WL peaks: data vector choices

Data vector $x = x(t_i)$. Different cases:

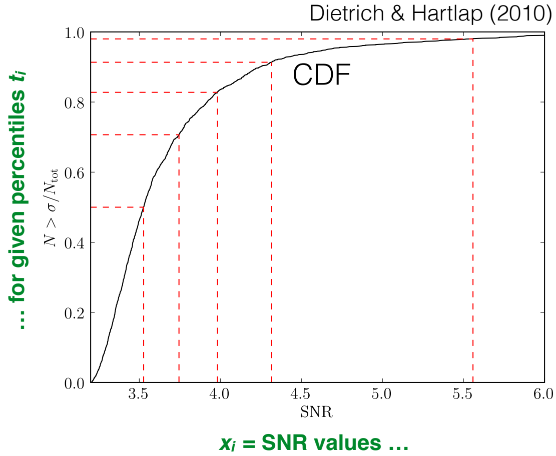
- **Abundance** of peaks n_i as fct. of SNR ν (PDF; binned histogram) **or**
- **SNR values** ν_i at some percentile values of peak CDF)
  - with or without lower cut ν_{min}.

Lin & MK 2015a

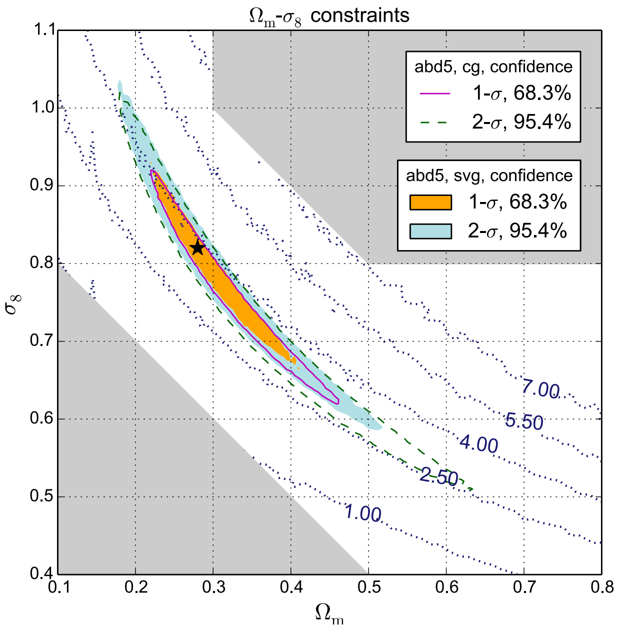


or

Dietrich & Hartlap (2010)



WL peaks: Gaussian likelihood



$$L_{\rm cg} \equiv \Delta \boldsymbol{x}^T(\boldsymbol{\pi})\, \widehat{\boldsymbol{C}^{-1}}(\boldsymbol{\pi}^{\rm obs})\, \Delta \boldsymbol{x}(\boldsymbol{\pi}),$$

$$L_{\rm svg} \equiv \Delta \boldsymbol{x}^T(\boldsymbol{\pi})\, \widehat{\boldsymbol{C}^{-1}}(\boldsymbol{\pi})\, \Delta \boldsymbol{x}(\boldsymbol{\pi}), \quad \text{and}$$

$$L_{\rm vg} \equiv \ln\left[\det \widehat{\boldsymbol{C}}(\boldsymbol{\pi})\right] + \Delta \boldsymbol{x}^T(\boldsymbol{\pi})\, \widehat{\boldsymbol{C}^{-1}}(\boldsymbol{\pi})\, \Delta \boldsymbol{x}(\boldsymbol{\pi}).$$

Cosmology-dependent covariance [(s)vg] reduces
error area by 20%.

ABC: Approximate Bayesian Computation I

Likelihood:
probability of data given
parameters and model

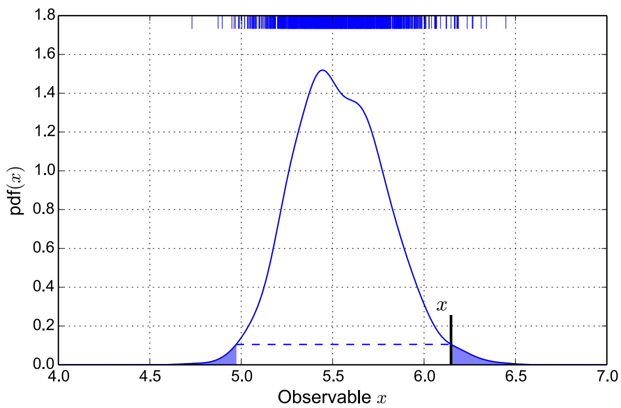
$$p(\boldsymbol{\pi}|\boldsymbol{x}, m) = \frac{L(\boldsymbol{x}|\boldsymbol{\pi}, m) P(\boldsymbol{\pi}|m)}{E(\boldsymbol{x}|m)}$$

$\boldsymbol{\pi}$: parameters
$\boldsymbol{x}$: data
m : *model*

Likelihood: how likely is it that model prediction $\boldsymbol{x}^{\rm mod}(\boldsymbol{\pi})$ reproduces data $\boldsymbol{x}$?

Classical answer: evaluate function L at $\boldsymbol{x}$.

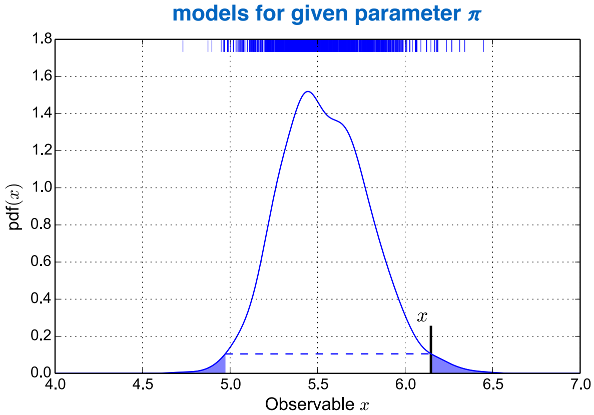
Alternative: compute fraction of models that are equal to the data $\boldsymbol{x}$.



ABC: Approximate Bayesian Computation II

Probability = p/N in frequentist sense.

Magic: Don't need to sample N models. **One** per parameter π is sufficient with accept-reject algorithm.



ABC can be performed if:

- it is possible and easy to sample from L

ABC is useful when:

- functional form of L is unknown
- evaluation of L is expensive
- model is intrinsically stochastic

Part II day 3: Cosmological parameter estimation | Higher order statistics: peak counts

ABC: Approximate Bayesian Computation III

Example: let's make soup.

Goal: Determine ingredients from final result.
Model physical processes? Complicated.

ABC: Approximate Bayesian Computation IV

Example: let's make soup.

Goal: Determine ingredients from final result.
Model physical processes? Complicated.
Easier: Make lots of soups with different ingredients, compare.

ABC: Approximate Bayesian Computation V

Example: let's make soup.

Questions:

- What aspect of data and simulations do we compare? (**summary statistic**)
- How do we compare? (**metric, distance**)
- When do we accept? (**tolerance**)

ABC: Approximate Bayesian Computation VI

Parameter constraints: ABC

- Summary statistic

  $\boldsymbol{s} = \boldsymbol{x}$ (data vector for 2 cases)

- Metric D: two cases

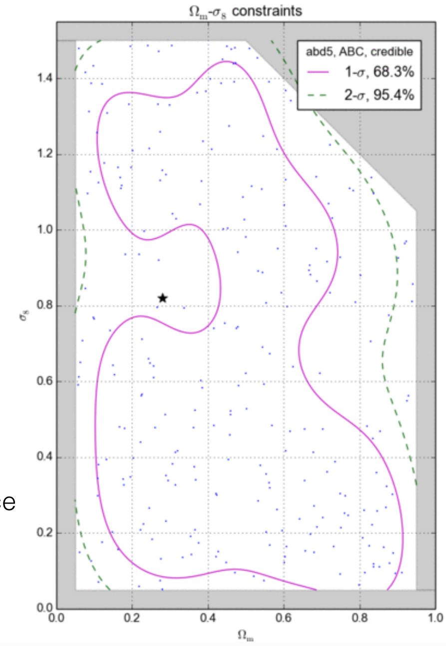
$$D_1\left(\boldsymbol{x}, \boldsymbol{x}^{\rm obs}\right) \equiv \sqrt{\sum_i \frac{\left(x_i - x_i^{\rm obs}\right)^2}{C_{ii}}},$$

$$D_2\left(\boldsymbol{x}, \boldsymbol{x}^{\rm obs}\right) \equiv \sqrt{(\boldsymbol{x} - \boldsymbol{x}^{\rm obs})^T \boldsymbol{C}^{-1} (\boldsymbol{x} - \boldsymbol{x}^{\rm obs})},$$

  D_1 in Lin & MK 2015b

  D_1 + D_2 in Lin, MK & Pires 2016

- ABC algorithm: iterative importance sampling (PMC) with decreasing tolerance



ABC: Approximate Bayesian Computation VII

ABC's accept-reject process is actually a sampling under P_ϵ (green curve):

$$P_\epsilon(\pi|x^{\rm obs}) = A_\epsilon(\pi) P(\pi),$$

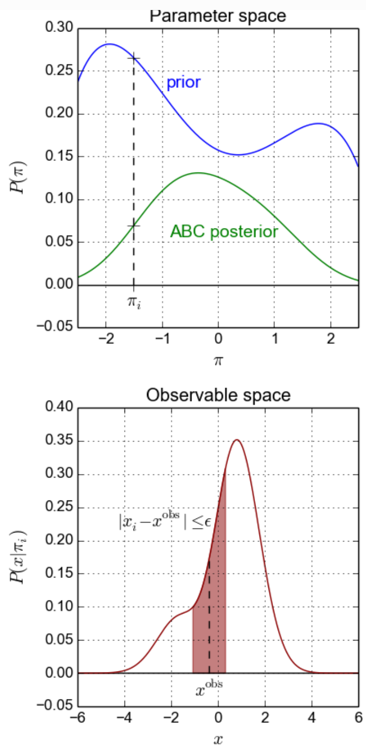
where $P(\pi)$ stands for the prior (blue curve) and

$$A_\epsilon(\pi) \equiv \int {\rm d}x \; P(x|\pi) \mathbb{1}_{|x - x^{\rm obs}| \leq \epsilon}(x),$$

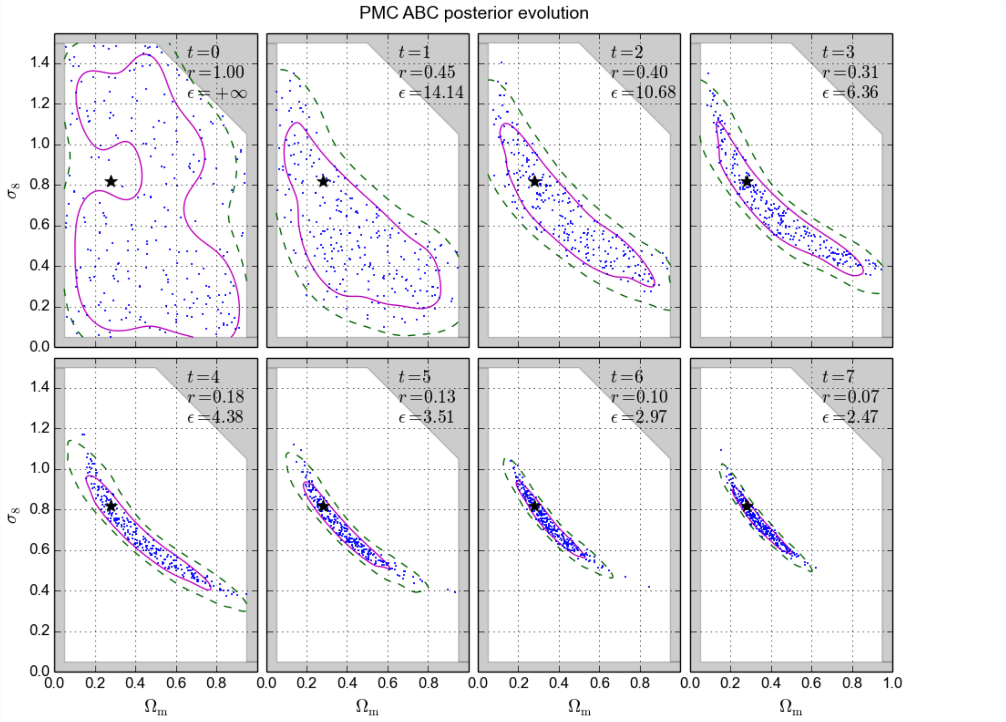
is the accept probability under π (red area). One can see that

$$\lim_{\epsilon \to 0} A_\epsilon(\pi_0)/\epsilon = P(x^{\rm obs}|\pi_0) = \mathcal{L}(\pi_0),$$

so P_ϵ is proportional to the true posterior when $\epsilon \to 0$.

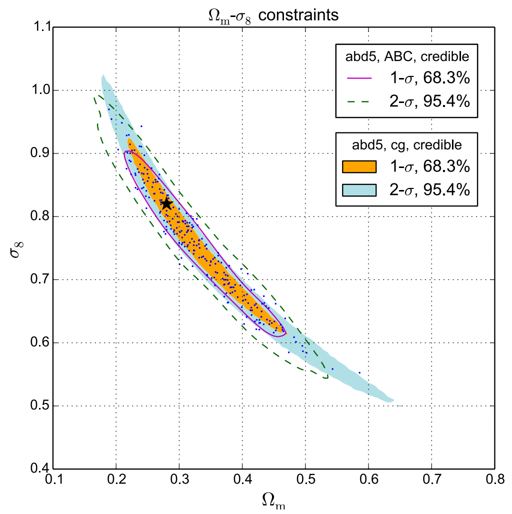


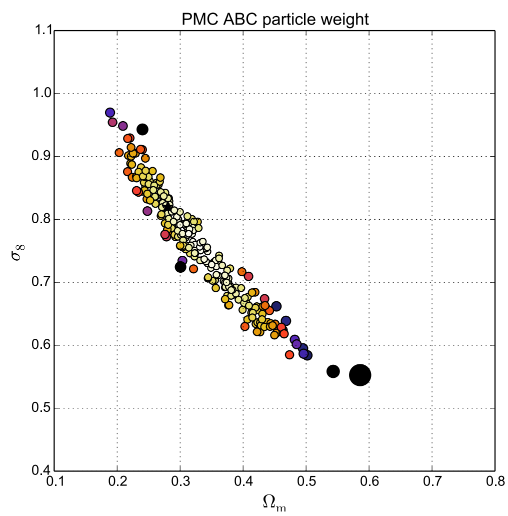
ABC: Approximate Bayesian Computation VIII



Lin & Kilbinger (2015b)

ABC: Approximate Bayesian Computation IX





ABC wider but less elongated and less bent contours than Gaussian with const cov.
KDE smoothing effect?

Bibliography I

Asgari M, Schneider P & Simon P 2012 *A&A* **542**, A122.

Baldauf T, Smith R E, Seljak U & Mandelbaum R 2010 *Phys. Rev. D* **81**(6), 063531.

Beaulieu J P, Bennett D P, Fouqué P, Williams A, Dominik M & al. 2006 *Nature* **439**, 437–440.

Becker M R 2013 *MNRAS* **435**, 115–132.

Bernardeau F, Bonvin C, Van de Rijt N & Vernizzi F 2012 *Phys. Rev. D* **86**(2), 023001.

Blake C, Joudaki S, Heymans C, Choi A, Erben T & al. 2016 *MNRAS* **456**, 2806–2828.

Blandford R & Narayan R 1986 *ApJ* **310**, 568–582.

Castro P G, Heavens A F & Kitching T D 2005 *Phys. Rev. D* **72**(2), 023516.

Chon G, Challinor A, Prunet S, Hivon E & Szapudi I 2004 *MNRAS* **350**, 914–926.

Clowe D, Bradač M, Gonzalez A H, Markevitch M, Randall S W & al. 2006 *ApJ* **648**, L109–L113.

Bibliography

Bibliography II

Coupon J, Kilbinger M, McCracken H J, Ilbert O, Arnouts S & al. 2012 *A&A* **542**, A5.

Das S & Bode P 2008 *ApJ* **682**, 1–13.

Eifler T, Schneider P & Krause E 2010 *A&A* **510**, A7.

Fosalba P, Gaztañaga E, Castander F J & Manera M 2008 *MNRAS* **391**, 435–446.

Fu L & Kilbinger M 2010 *MNRAS* **401**, 1264–1274.

Fu L, Semboloni E, Hoekstra H, Kilbinger M, van Waerbeke L & al. 2008 *A&A* **479**, 9–25.

Heymans C, Van Waerbeke L, Bacon D, Berge J, Bernstein G & al. 2006 *MNRAS* **368**, 1323–1339.

Hilbert S, Hartlap J, White S D M & Schneider P 2009 *A&A* **499**, 31–43.

Hildebrandt H, Viola M, Heymans C, Joudaki S, Kuijken K & al. 2017 *MNRAS* **465**, 1454–1498.

Hoag A, Bradac M, Trenti M, Treu T, Schmidt K B & al. 2017 *Nature Astronomy* **1**, 0091.

Bibliography III

Hoekstra H, Viola M & Herbonnet R 2016 *ArXiv e-prints* .

Hoekstra H, Yee H K C & Gladders M D 2001 *ApJ* **558**, L11–L14.

Hu W 2000 *Phys. Rev. D* **62**(4), 043007.

Kaiser N, Squires G, Fahlman G & Woods D 1994 *in* 'Clusters of galaxies, Proceedings of the XIVth Moriond Astrophysics Meeting, Méribel, France' p. 269.

Kilbinger M, Fu L, Heymans C, Simpson F, Benjamin J & al. 2013 *MNRAS* **430**, 2200–2220.

Kilbinger M, Heymans C, Asgari M, Joudaki S, Schneider P & al. 2017 *ArXiv e-prints* .

Leonard A, Pires S & Starck J L 2012 *MNRAS* **423**, 3405–3412.

Li B, King L J, Zhao G B & Zhao H 2011 *MNRAS* **415**, 881–892.

Mandelbaum R, Rowe B, Armstrong R, Bard D, Bertin E & al. 2015 *MNRAS* **450**, 2963–3007.

Bibliography IV

Reyes R, Mandelbaum R, Seljak U, Baldauf T, Gunn J E & al. 2010 *Nature* **464**, 256–258.

Schneider P 1996 *MNRAS* **283**, 837.

Schneider P, Eifler T & Krause E 2010 *A&A* **520**, A116.

Schneider P & Kilbinger M 2007 *A&A* **462**, 841–849.

Schneider P, Van Waerbeke L, Jain B & Kruse G 1998 *MNRAS* **296**, 873–892.

Schneider P, Van Waerbeke L, Kilbinger M & Mellier Y 2002 *A&A* **396**, 1–19.

Semboloni E, Hoekstra H & Schaye J 2013 *MNRAS* **434**, 148–162.

Simon P, Hetterscheidt M, Schirmer M, Erben T, Schneider P & al. 2007 *A&A* **461**, 861–879.

Starck J L, Pires S & Réfrégier A 2006 *A&A* **451**, 1139–1150.

Takada M & Hu W 2013 *Phys. Rev. D* **87**(12), 123504.

Taylor A, Joachimi B & Kitching T 2013 *MNRAS* .

Teyssier R, Pires S, Prunet S, Aubert D, Pichon C & al. 2009 *A&A* **497**, 335–341.

Bibliography V

Velander M, van Uitert E, Hoekstra H, Coupon J, Erben T & al. 2014 *MNRAS* **437**, 2111–2136.

White M & Hu W 2000 *ApJ* **537**, 1–11.

Zentner A R, Rudd D H & Hu W 2008 *Phys. Rev. D* **77**(4), 043507–+.