Large Scale Bayesian Analyses of Cosmological Datasets

Jens Jasche

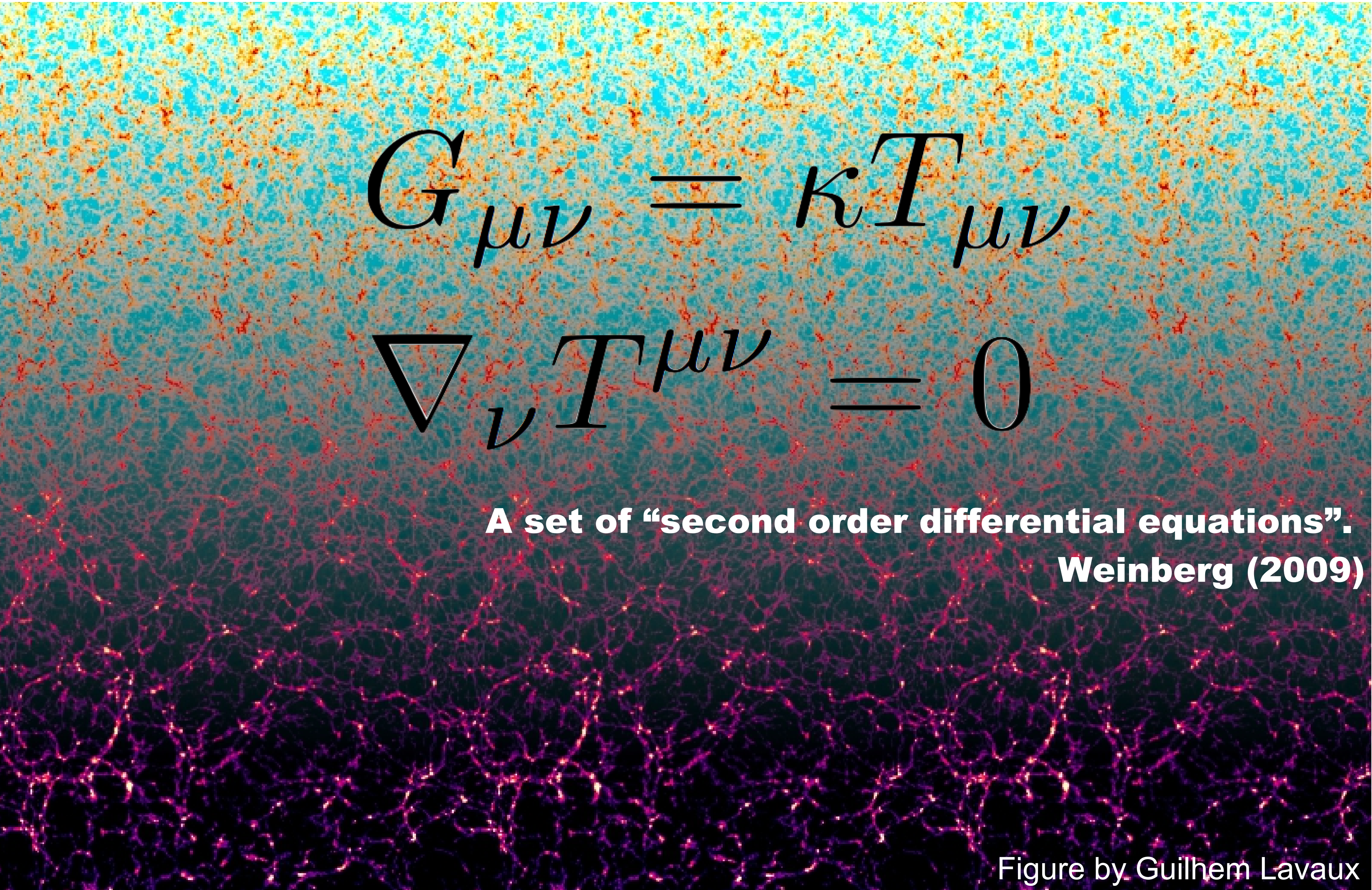
Guilhem Lavaux, Fabian Schmidt, Florent Leclercq, Natalia Porqueres, Doogesh Kodi Ramanah, Supranta Sarma Boruah, Dexter Bergsdal, Adam Andrews, Benjamin Wandelt

Artificial Intelligence in Astronomy

ESO, HQ, Garching, 26 July 2019

Visit us at: www.aquila-consortium.org

A pretty good model...



$$G_{\mu\nu} = \kappa T_{\mu\nu}$$

$$\nabla_{\nu} T^{\mu\nu} = 0$$

A set of "second order differential equations".

Weinberg (2009)

Figure by Guilhem Lavaux

A pretty good model...

$$G_{\mu\nu} = \kappa T_{\mu\nu}$$

$$\nabla_\nu T^{\mu\nu} = 0$$

A set of "second order differential equations".

Weinberg (2009)

What is gravity?

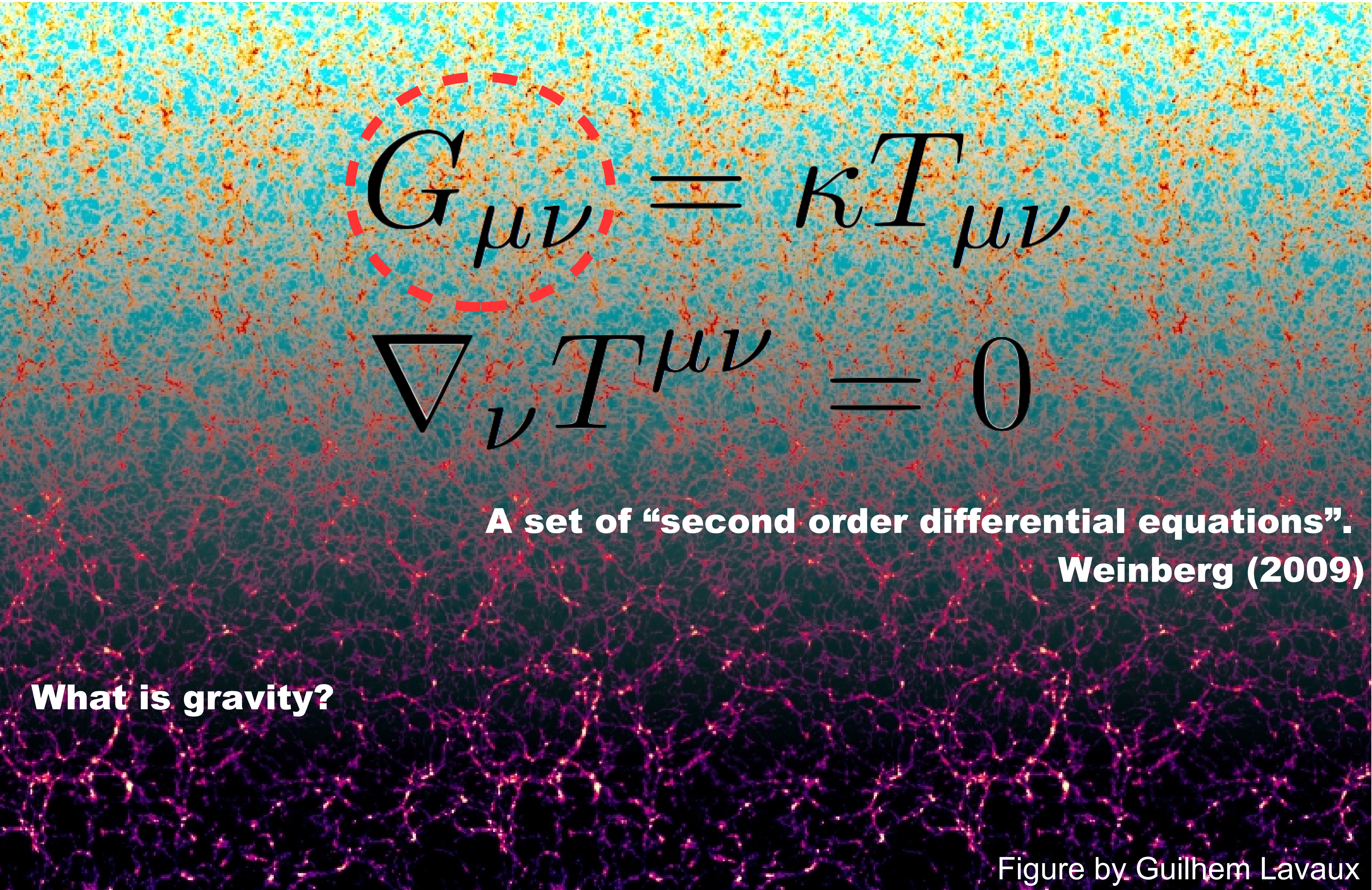
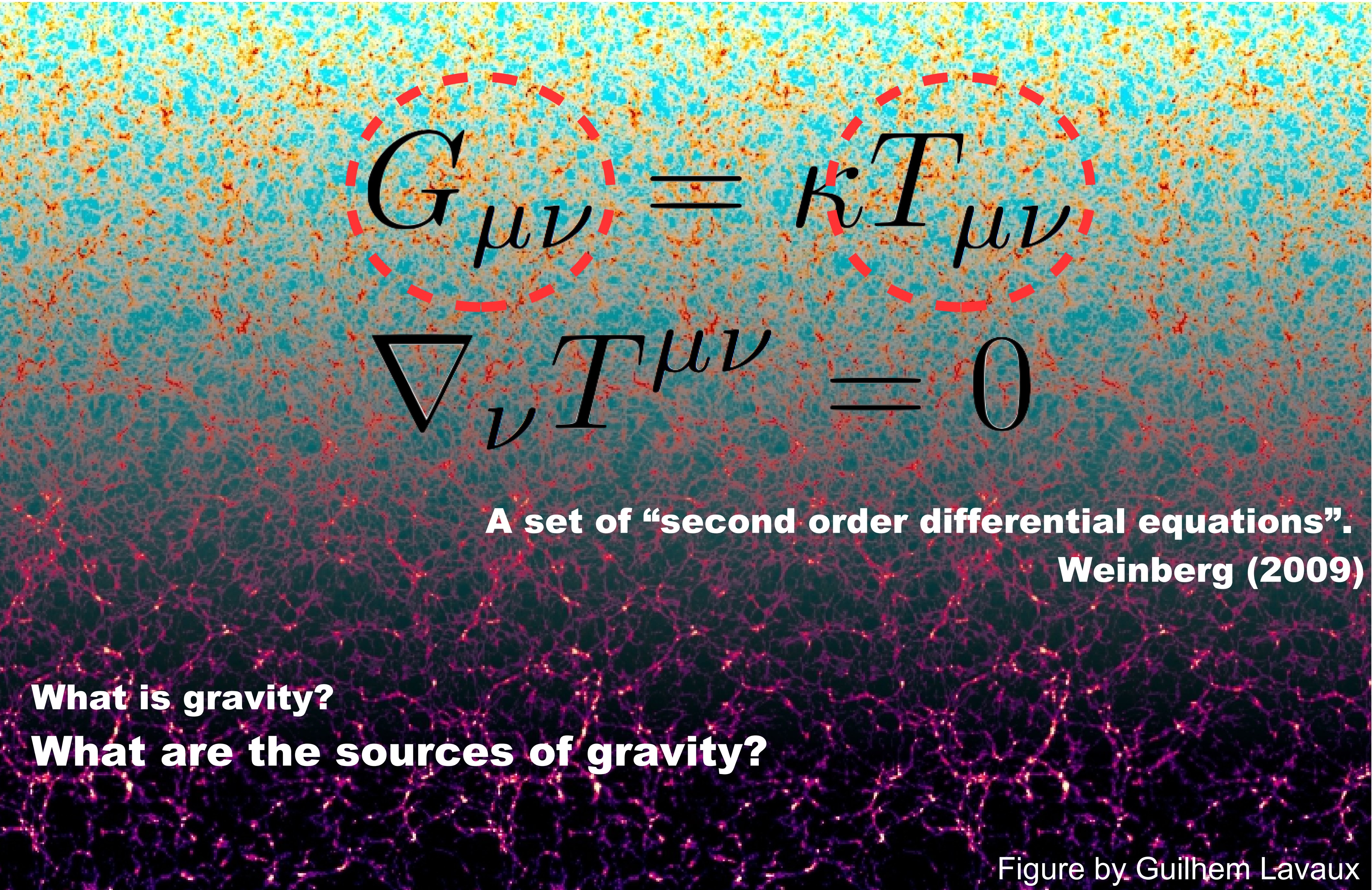


Figure by Guilhem Lavaux

A pretty good model...



$$G_{\mu\nu} = \kappa T_{\mu\nu}$$

$$\nabla_\nu T^{\mu\nu} = 0$$

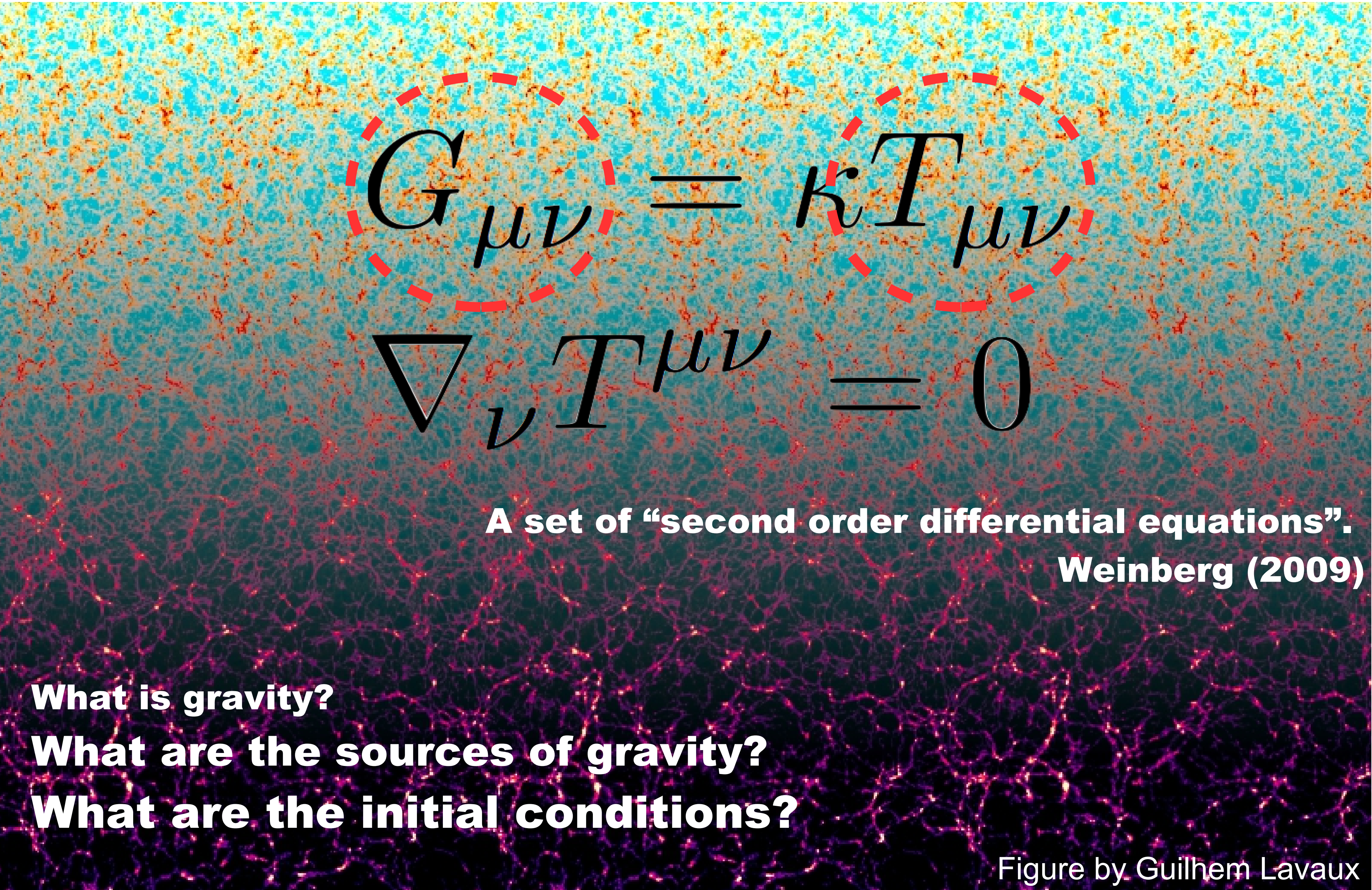
A set of "second order differential equations".
Weinberg (2009)

What is gravity?

What are the sources of gravity?

Figure by Guilhem Lavaux

A pretty good model...



$$G_{\mu\nu} = \kappa T_{\mu\nu}$$

$$\nabla_\nu T^{\mu\nu} = 0$$

A set of "second order differential equations".

Weinberg (2009)

What is gravity?

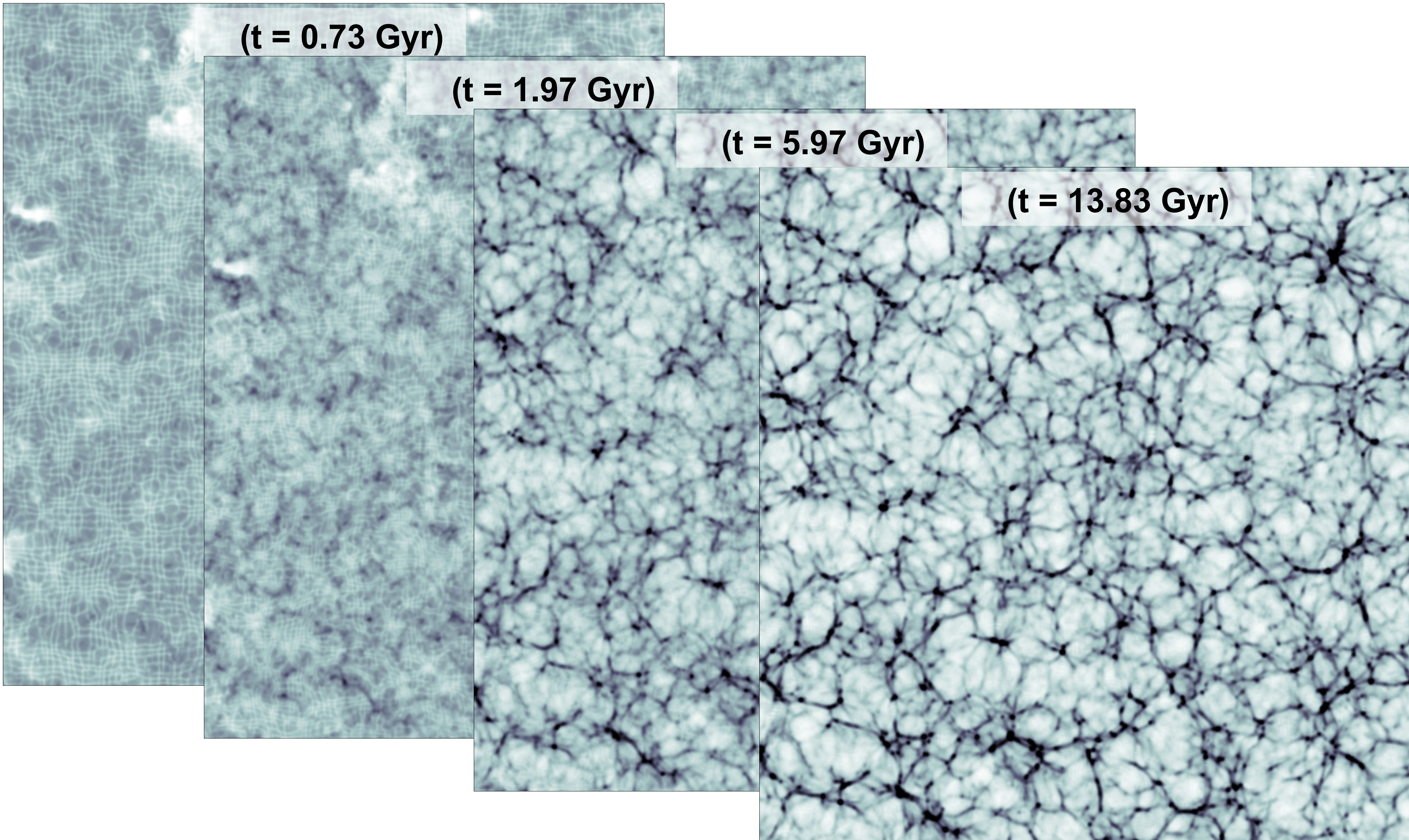
What are the sources of gravity?

What are the initial conditions?

Figure by Guilhem Lavaux

The cosmic large scale structure...

... A source of knowledge!



A large scale Bayesian inverse problem

Bayesian Forward modeling:

Jasche, Wandelt (2013)
Lavaux, Jasche (2016)
Jasche, Lavaux (2018)

Prior model

$$\mathscr{P}(\boldsymbol{s}|\boldsymbol{S}) = \frac{\mathrm{e}^{-\frac{1}{2}\boldsymbol{s}^{\mathrm{T}}\boldsymbol{S}^{-1}\boldsymbol{s}}}{\sqrt{\det(2\,\pi\,\boldsymbol{S})}}$$

Structure formation model

$$\mathscr{P}(\boldsymbol{\delta}|\boldsymbol{s}) = \prod_i \delta^D\left(\delta_i - G_i(\boldsymbol{s})\right)$$

$$\frac{\mathrm{d}\vec{x}}{\mathrm{d}a} = \frac{\vec{p}}{\dot{a}a^2}$$

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}a} = -\frac{3}{2}H_0^2\Omega_m\frac{\nabla^2\Phi}{Ha^2}$$

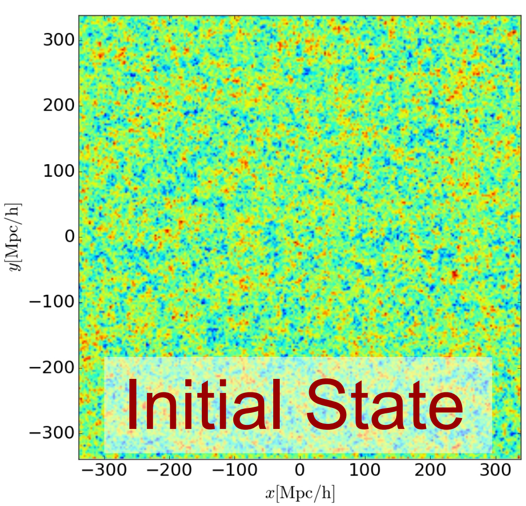
Data model

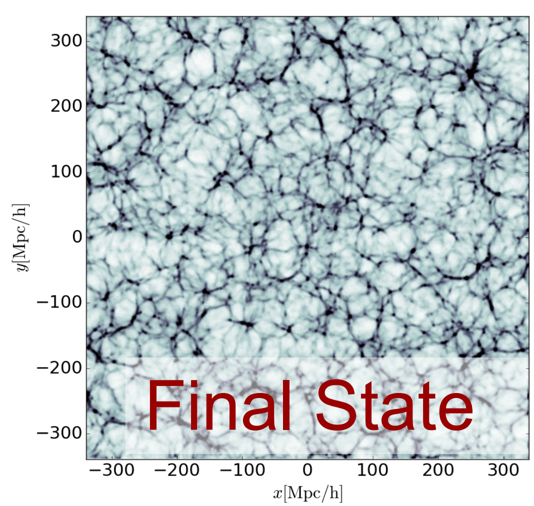
$$\mathscr{P}(\boldsymbol{N}|\boldsymbol{\lambda}(\boldsymbol{\delta})) = \prod_i \frac{\mathrm{e}^{-\lambda_i}\lambda_i^{N_i}}{N_i\,!}$$

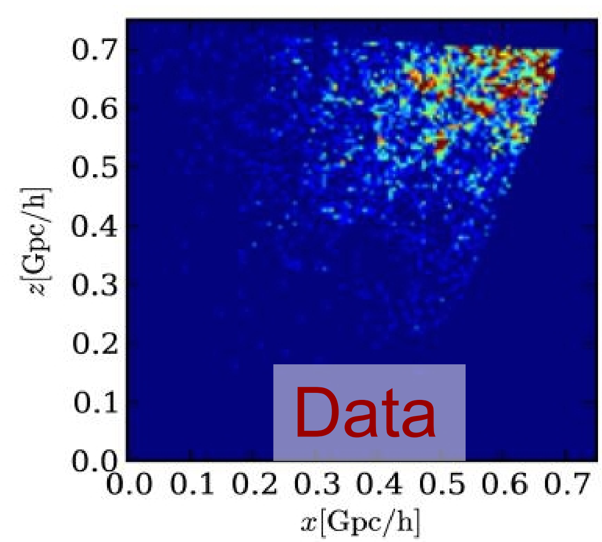
Galaxy bias model

$$\lambda_i = R_i\bar{N}\,(1+\delta)^{\beta}\,\mathrm{e}^{-\rho_g\,(1+\delta)^{-\epsilon_g}}$$

See e.g. Neyrinck et al. 2014
Ata et al. 2015
Lavaux & Jasche 2016







$$\dim(\boldsymbol{s}) \sim 10^7 \text{ parameters}$$

MCMC in high dimensions

HMC: Use Classical mechanics to solve statistical problems!

- The potential : $\psi(\mathbf{x}) = -\ln(\mathcal{P}(\mathbf{x}))$

see e.g. Duane et al. (1987)
Neal (2012)
Betancourt (2017)

MCMC in high dimensions

HMC: Use Classical mechanics to solve statistical problems!

- The potential : $\psi(\mathbf{x}) = -\ln(\mathcal{P}(\mathbf{x}))$
- The Hamiltonian : $H = \frac{1}{2}\mathbf{p}^T\mathbf{M}^{-1}\mathbf{p} + \psi(\mathbf{x})$

see e.g. Duane et al. (1987)
Neal (2012)
Betancourt (2017)

MCMC in high dimensions

HMC: Use Classical mechanics to solve statistical problems!

- The potential : $\psi(\mathbf{x}) = -\ln\left(\mathcal{P}(\mathbf{x})\right)$
- The Hamiltonian : $H = \frac{1}{2}\mathbf{p}^T\mathbf{M}^{-1}\mathbf{p} + \psi(\mathbf{x})$

Nuisance parameter!!!

see e.g. Duane et al. (1987)
Neal (2012)
Betancourt (2017)

MCMC in high dimensions

HMC: Use Classical mechanics to solve statistical problems!

- The potential : $\psi(\mathbf{x}) = -\ln\left(\mathcal{P}(\mathbf{x})\right)$
- The Hamiltonian : $H = \frac{1}{2}\mathbf{p}^T\,\mathbf{M}^{-1}\,\mathbf{p} + \psi(\mathbf{x})$

Nuisance parameter!!!

$$(\mathbf{x},\ \mathbf{p}) \Rightarrow \begin{aligned} \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} &= \frac{\partial H}{\partial \mathbf{p}} = \mathbf{M}^{-1}\mathbf{p} \\ \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} &= -\frac{\partial H}{\partial \mathbf{x}} = -\frac{\partial \psi(\mathbf{x})}{\partial \mathbf{x}} \end{aligned} \Rightarrow (\mathbf{x}',\ \mathbf{p}')$$

see e.g. Duane et al. (1987)
Neal (2012)
Betancourt (2017)

MCMC in high dimensions

HMC: Use Classical mechanics to solve statistical problems!

- The potential : $\psi(\mathbf{x}) = -\ln\left(\mathcal{P}(\mathbf{x})\right)$
- The Hamiltonian : $H = \frac{1}{2}\mathbf{p}^T\,\mathbf{M}^{-1}\,\mathbf{p} + \psi(\mathbf{x})$

Nuisance parameter!!!

$$(\mathbf{x},\ \mathbf{p}) \Rightarrow \begin{aligned} \frac{\mathrm{d}\mathbf{x}}{\mathrm{dt}} &= \frac{\partial H}{\partial \mathbf{p}} = \mathbf{M}^{-1}\mathbf{p} \\ \frac{\mathrm{d}\mathbf{p}}{\mathrm{dt}} &= -\frac{\partial H}{\partial \mathbf{x}} = -\frac{\partial \psi(\mathbf{x})}{\partial \mathbf{x}} \end{aligned} \Rightarrow (\mathbf{x}',\ \mathbf{p}')$$

see e.g. Duane et al. (1987)
Neal (2012)
Betancourt (2017)

MCMC in high dimensions

HMC: Use Classical mechanics to solve statistical problems!

- The potential : $\psi(\mathbf{x}) = -\ln(\mathcal{P}(\mathbf{x}))$
- The Hamiltonian : $H = \frac{1}{2}\mathbf{p}^T\mathbf{M}^{-1}\mathbf{p} + \psi(\mathbf{x})$

Nuisance parameter!!!

$$(\mathbf{x},\ \mathbf{p}) \rightarrow \begin{aligned} \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} &= \frac{\partial H}{\partial \mathbf{p}} = \mathbf{M}^{-1}\mathbf{p} \\ \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} &= -\frac{\partial H}{\partial \mathbf{x}} = -\frac{\partial \psi(\mathbf{x})}{\partial \mathbf{x}} \end{aligned} \rightarrow (\mathbf{x}',\ \mathbf{p}')$$

Randomize $\mathbf{p}$ and accept $\mathbf{x}'$: $\alpha = \min\left[1, \mathrm{e}^{-(H'-H)}\right] = 1$

HMC beats the "curse of dimensionality" by:

- Exploiting gradients
- Using conserved quantities

see e.g. Duane et al. (1987)
Neal (2012)
Betancourt (2017)

The data model

BORG (**B**ayesian **O**rigin **R**econstruction from **G**alaxies)

- Incorporates physical model into Likelihood (2LPT / PM)
- Turn inference into initial conditions problem: Find $\mathbf{x}^0$!!!

$$\Psi(\mathbf{x}^0) \quad = \quad \Psi_{\mathrm{P}}(\mathbf{x}^0) + \Psi_{\mathrm{LH}}(\mathbf{x}^0)$$

Also see e.g. Jasche et al. 2015 (arXiv:1409.6308) / Wang et al. 2014 (arXiv:1407.3451)

The data model

BORG (**B**ayesian **O**rigin **R**econstruction from **G**alaxies)

- Incorporates physical model into Likelihood (2LPT / PM)
- Turn inference into initial conditions problem: Find $\mathbf{x}^0$!!!

$$\begin{aligned}\Psi(\mathbf{x}^0) &= \Psi_{\mathrm{P}}(\mathbf{x}^0) + \Psi_{\mathrm{LH}}(\mathbf{x}^0) \\ &= \frac{1}{2}\sum_{ij} x_i^0 S_{ij}^{-1} x_j^0 + \sum_i \lambda_i - N_i \ln(\lambda_i)\end{aligned}$$

Poisson Intensity: $\lambda_i = \bar{N} R_i \left(1 + x_i(\mathbf{x}^0)\right)^{\alpha}$

Also see e.g. Jasche et al. 2015 (arXiv:1409.6308) / Wang et al. 2014 (arXiv:1407.3451)

The data model

BORG (**B**ayesian **O**rigin **R**econstruction from **G**alaxies)

- Incorporates physical model into Likelihood (2LPT / PM)
- Turn inference into initial conditions problem: Find $\mathbf{x}^0$!!!

$$\begin{aligned}\Psi(\mathbf{x}^0) &= \Psi_{\mathrm{P}}(\mathbf{x}^0) + \Psi_{\mathrm{LH}}(\mathbf{x}^0) \\ &= \frac{1}{2}\sum_{ij} x_i^0 S_{ij}^{-1} x_j^0 + \sum_i \lambda_i - N_i \ln(\lambda_i)\end{aligned}$$

Poisson Intensity: $\lambda_i = \bar{N} R_i \left(1 + x_i(\mathbf{x}^0)\right)^{\alpha}$

$$\frac{\mathrm{d}\Psi(\mathbf{x}^0)}{\mathrm{d}x_k^0} = \sum_j S_{kj}^{-1} x_j^0 + \sum_i \left(1 - N_i \frac{1}{\lambda_i}\right) \frac{\partial \lambda_i}{\partial x_i} \frac{\mathrm{d}x_i}{\mathrm{d}x_k^0}$$

Also see e.g. Jasche et al. 2015 (arXiv:1409.6308) / Wang et al. 2014 (arXiv:1407.3451)

The data model

BORG (**B**ayesian **O**rigin **R**econstruction from **G**alaxies)

- Incorporates physical model into Likelihood (2LPT / PM)
- Turn inference into initial conditions problem: Find $\mathbf{x}^0$!!!

$$\begin{aligned}\Psi(\mathbf{x}^0) &= \Psi_{\mathrm{P}}(\mathbf{x}^0) + \Psi_{\mathrm{LH}}(\mathbf{x}^0) \\ &= \frac{1}{2}\sum_{ij} x_i^0 S_{ij}^{-1} x_j^0 + \sum_i \lambda_i - N_i \ln(\lambda_i)\end{aligned}$$

Poisson Intensity: $\lambda_i = \bar{N} R_i \left(1 + x_i(\mathbf{x}^0)\right)^\alpha$

$$\frac{\mathrm{d}\Psi(\mathbf{x}^0)}{\mathrm{d}x_k^0} = \sum_j S_{kj}^{-1} x_j^0 + \sum_i \left(1 - N_i \frac{1}{\lambda_i}\right) \frac{\partial \lambda_i}{\partial x_i} \frac{\mathrm{d}x_i}{\mathrm{d}x_k^0}$$

Model sensitivity

Also see e.g. Jasche et al. 2015 (arXiv:1409.6308) / Wang et al. 2014 (arXiv:1407.3451)

Adjoint coding

The problem: $x(x_0)$ is a computer program**, not analytic!!!**

Adjoint coding

The problem: $x(x_0)$ is a computer program, **not analytic!!!**

- **Finite differencing not feasible, too high-d!!!**

Adjoint coding

The problem: $x(x_0)$ is a computer program**, not analytic!!!**

- **Finite differencing not feasible, too high-d!!!**
- Any computer program is a sequence of elementary operations

$$x(x_0) = B_N\left(B_{N-1}\left(B_{N-2}\left(B_{N-3}\left(\ldots\left(B_0(x_0)\right)\ldots\right)\right)\right)\right)$$

Adjoint coding

The problem: $x(x_0)$ is a computer program, **not analytic!!!**

- **Finite differencing not feasible, too high-d!!!**
- Any computer program is a sequence of elementary operations

$$x(x_0) = B_N\left(B_{N-1}\left(B_{N-2}\left(B_{N-3}\left(...\left(B_0(x_0)\right)...\right)\right)\right)\right)$$

→ **Use chain rule**

$$\frac{\mathrm{d}x(x_0)}{\mathrm{d}x_0} = \frac{\partial B_N}{\partial B_{N-1}} \frac{\partial B_{N-1}}{\partial B_{N-2}} \frac{\partial B_{N-2}}{\partial B_{N-3}} ... \frac{\partial B_0}{\partial x_0}$$

Adjoint coding

The problem: $x(x_0)$ is a computer program**, not analytic!!!**

- **Finite differencing not feasible, too high-d!!!**
- Any computer program is a sequence of elementary operations

$$x(x_0) = B_N\left(B_{N-1}\left(B_{N-2}\left(B_{N-3}\left(\ldots\left(B_0(x_0)\right)\ldots\right)\right)\right)\right)$$

→ **Use chain rule**

$$\frac{\mathrm{d}x(x_0)}{\mathrm{d}x_0} = \frac{\partial B_N}{\partial B_{N-1}}\frac{\partial B_{N-1}}{\partial B_{N-2}}\frac{\partial B_{N-2}}{\partial B_{N-3}}\cdots\frac{\partial B_0}{\partial x_0}$$

line by line derivative of your computer code!!

BEWARE of if-switches!!!

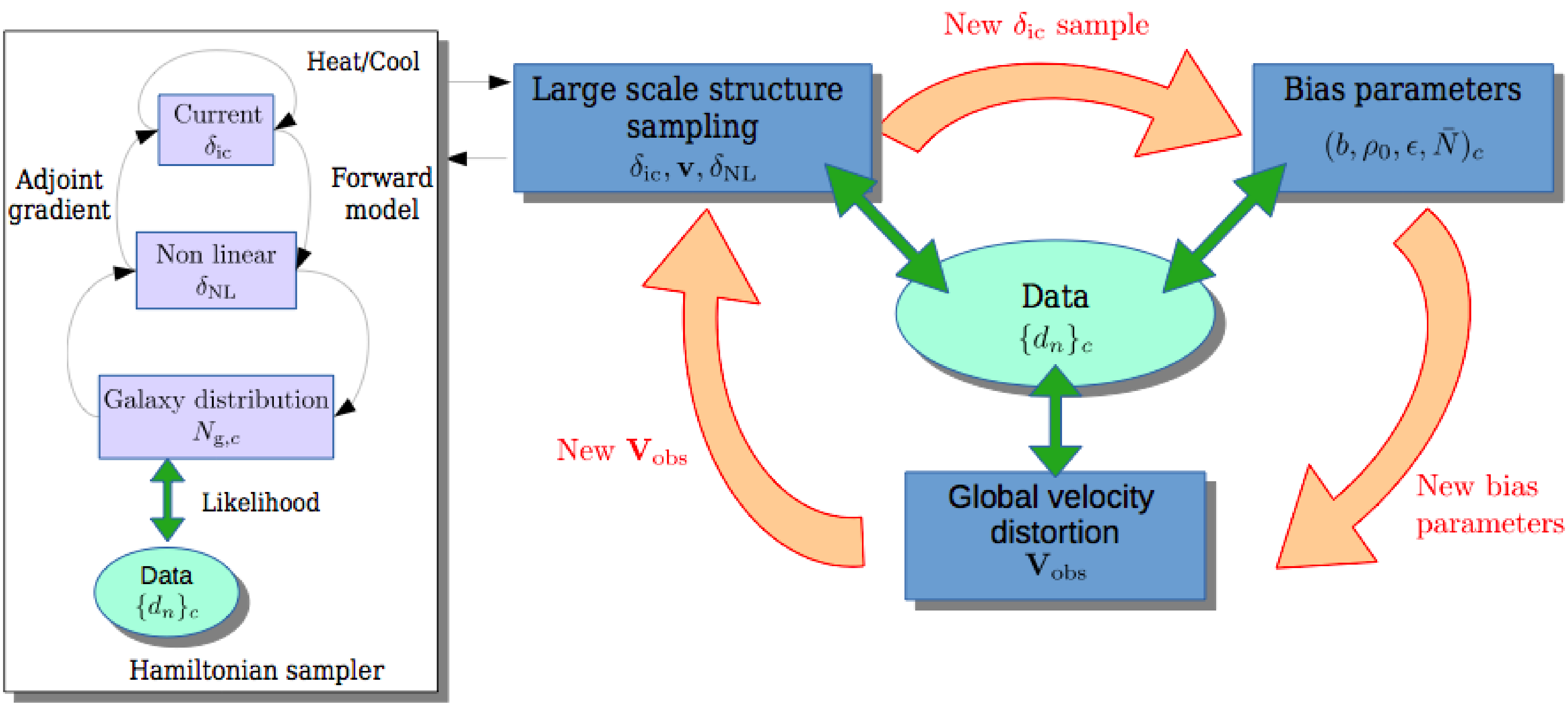
Result:

- Sensitivity Matrix of your computer model
- We need the adjoint, so do a transpose
- Matrix cannot be stored, so use operator formalism

BORG[3]: A Modular statistical programing engine

Build flexible data models

- Hierachical Bayes and block sampling



Jasche & Lavaux (2019, A&A)

A detailed and physically plausible model of The Nearby Universe

www.aquila-consortium.org

New insights into the nearby universe

Some samples from the Markov Chain...

Leclercq et al. (2019, in prep)

Application: BORG – 2M++ (Lavaux & Hudson (2011, MNRAS))

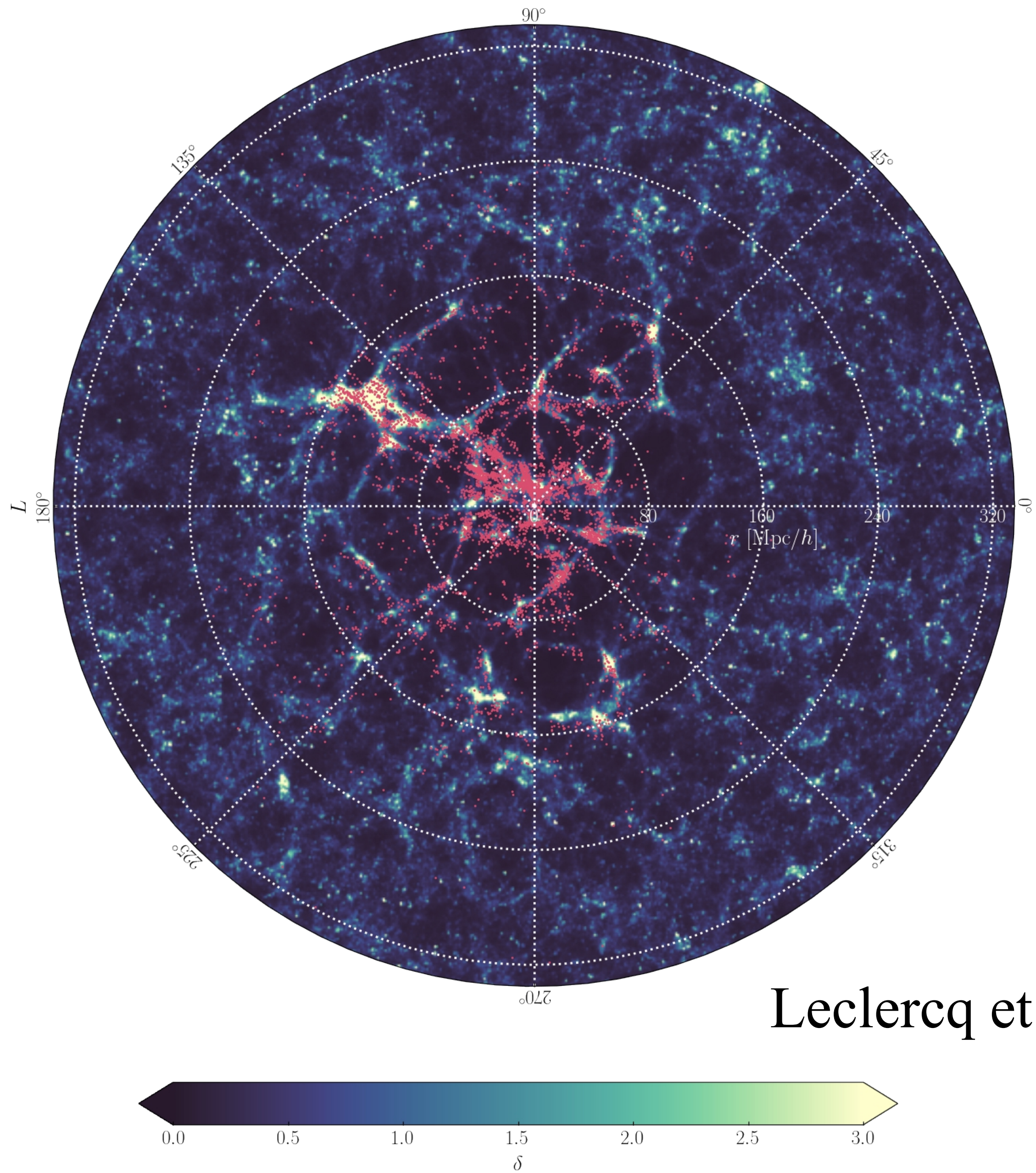
- Domain: $(677.7\ \mathrm{Mpc}/h)^3$
- IC fluctuation elements: 256^3
- Simulation particles: 512^3
- LSS model: Particle Mesh Solver

Jasche & Lavaux (2019, A&A)

Inferred mass density

Inferred mass density in super-galactic plane:

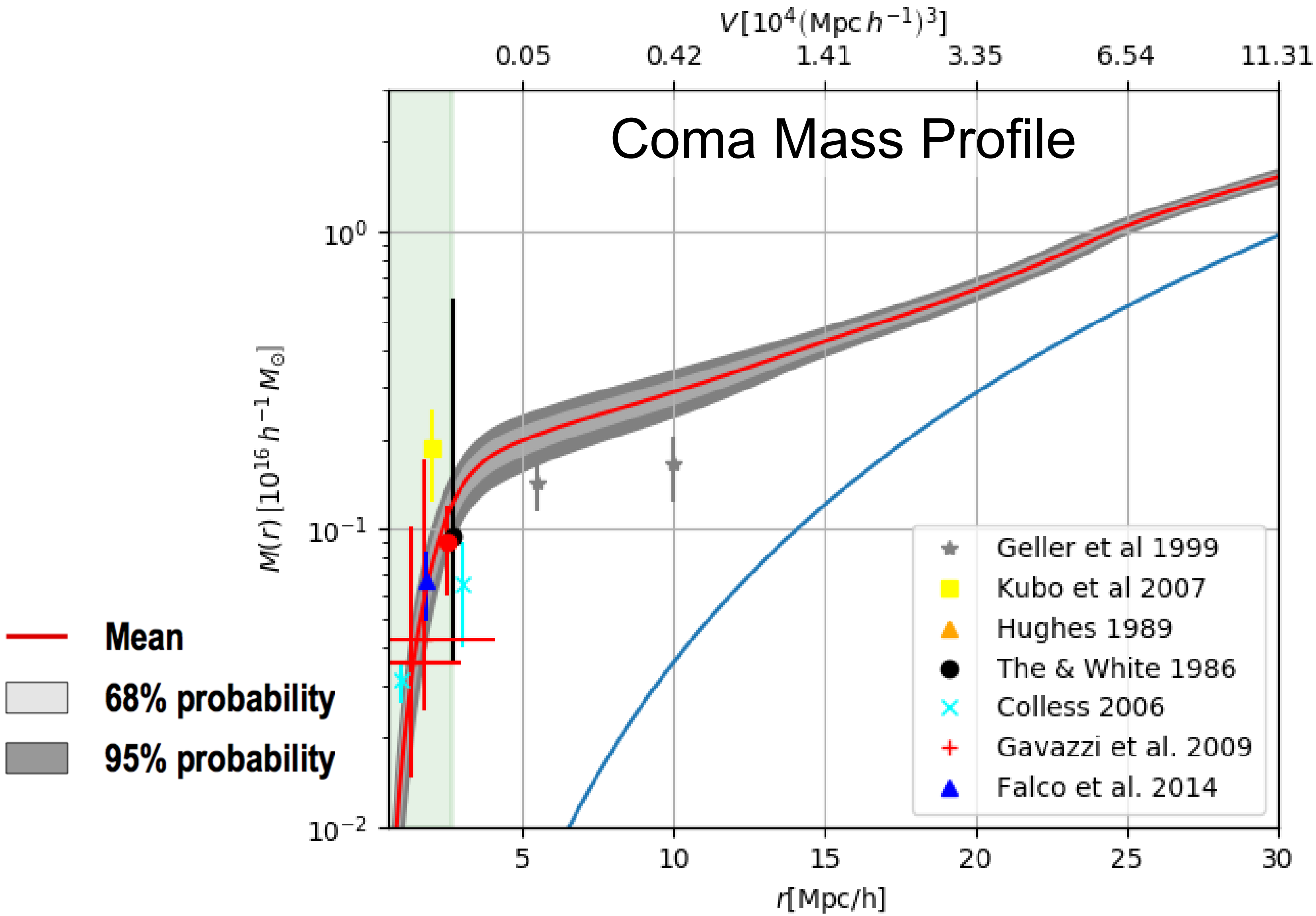
Preliminary results!



Leclercq et al. (2019, in prep)

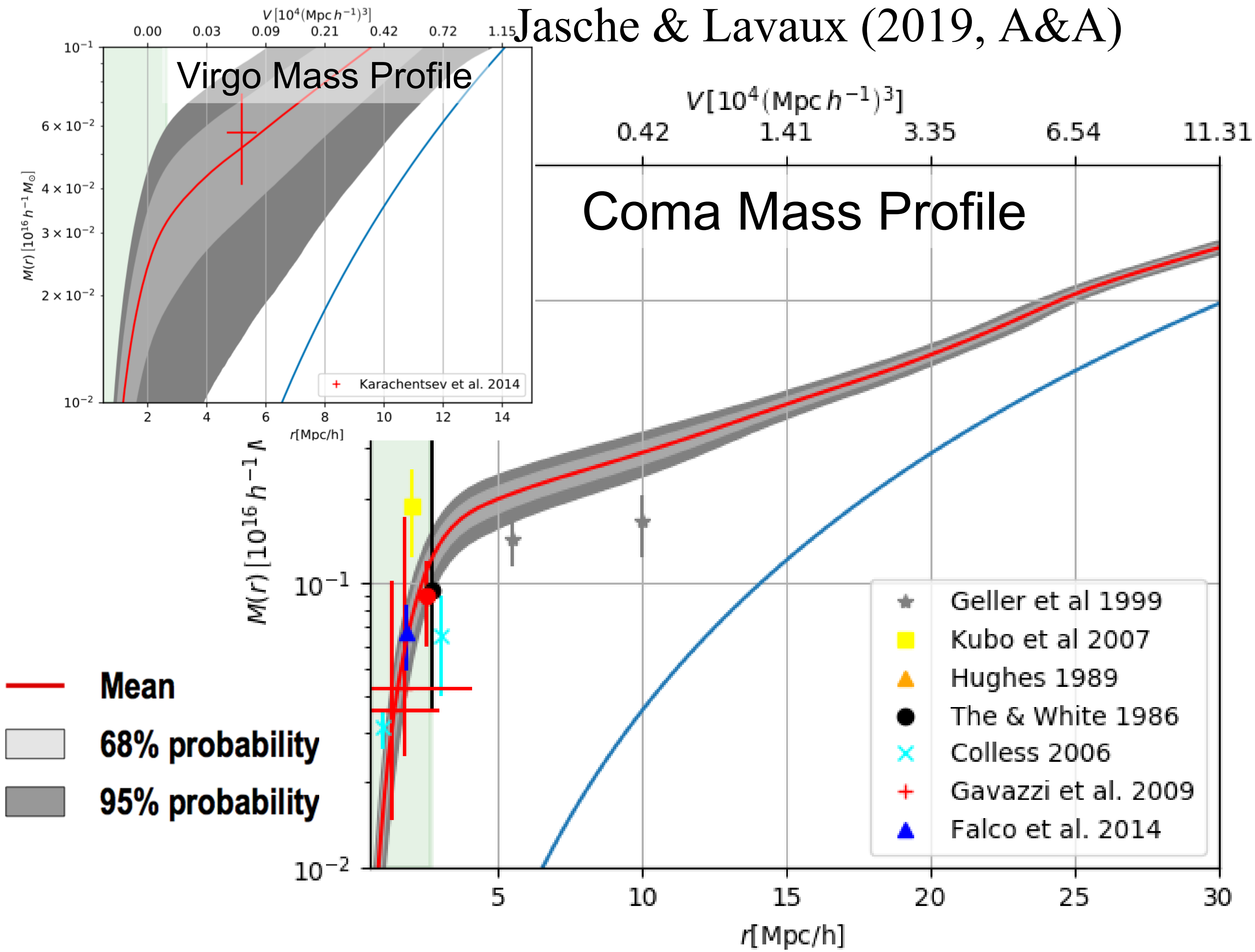
Estimating Cluster masses

Jasche & Lavaux (2019, A&A)



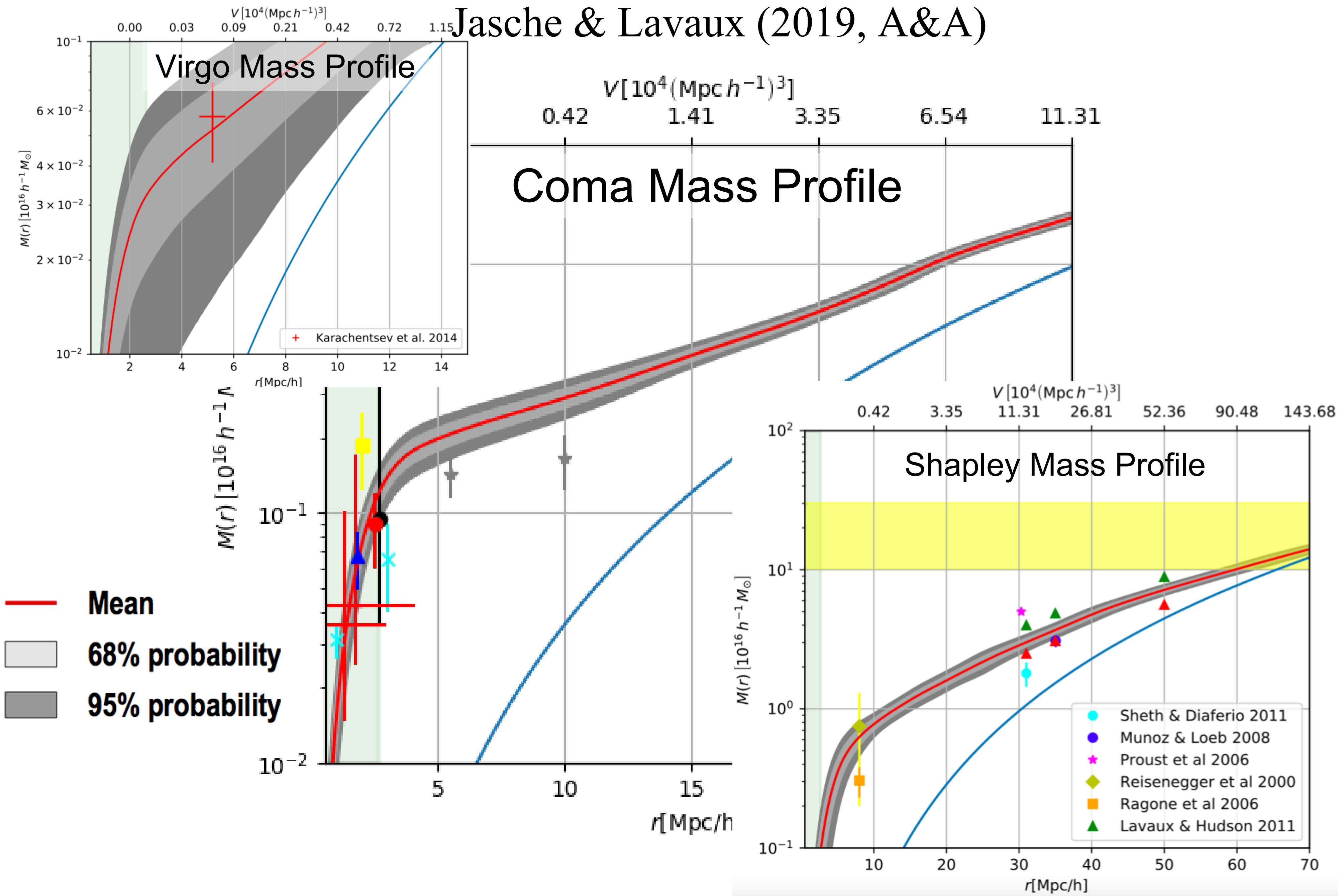
Estimating Cluster masses

Jasche & Lavaux (2019, A&A)



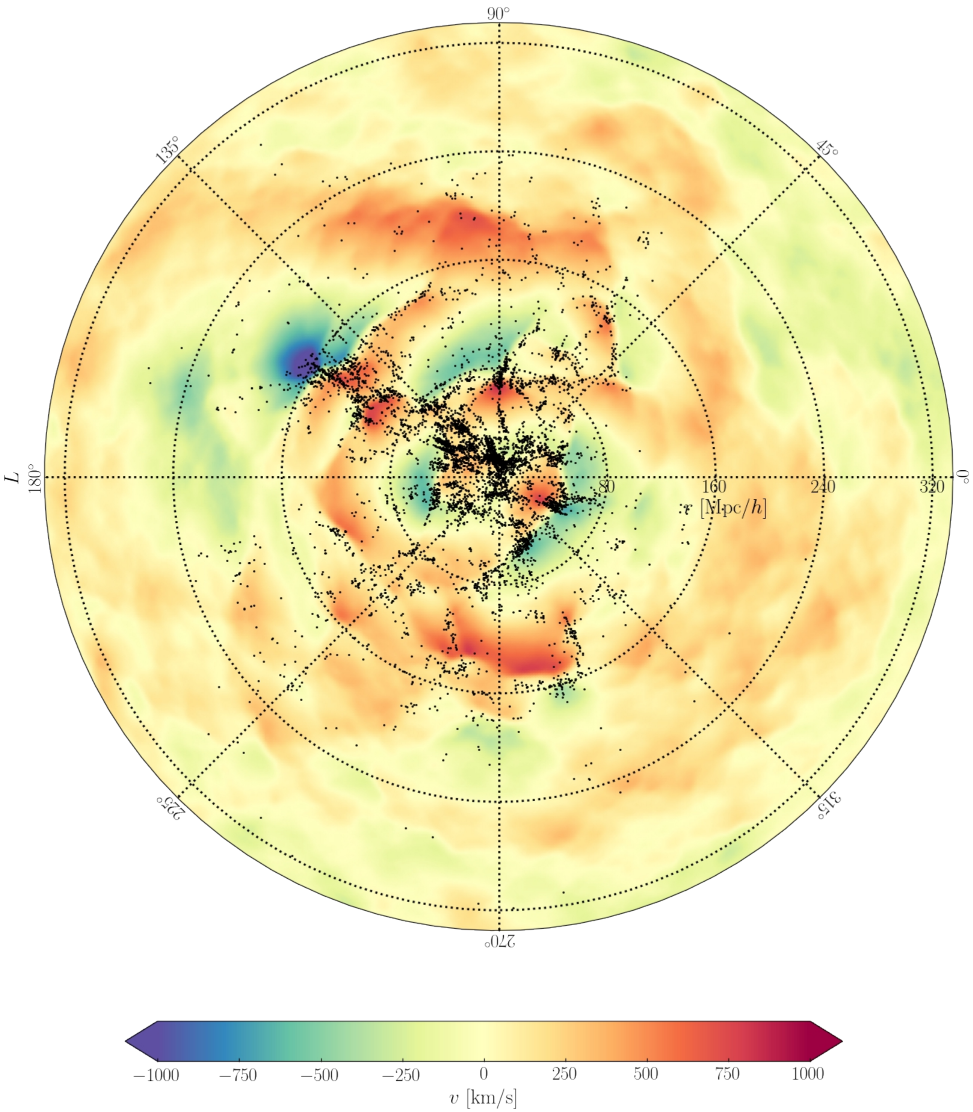
Estimating Cluster masses

Jasche & Lavaux (2019, A&A)

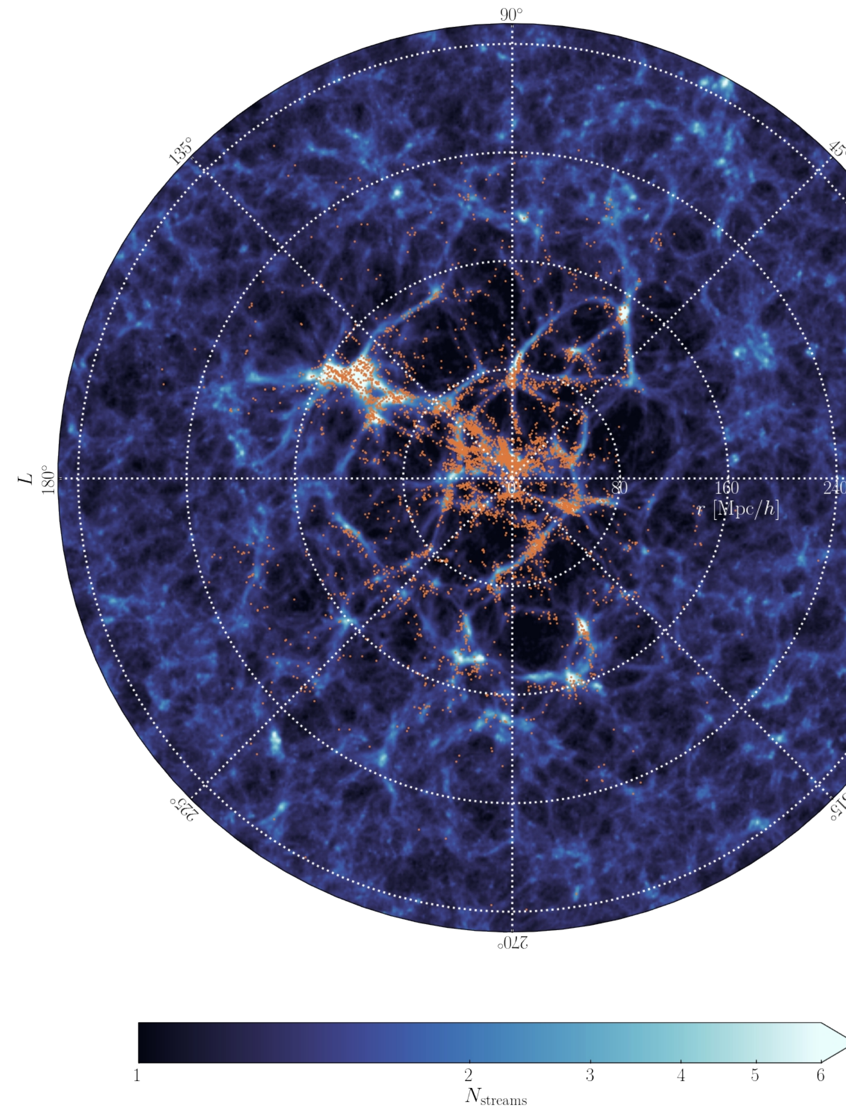


Dynamics of the Nearby Universe

Radial velocities



Counts of DM streams



Jasche & Lavaux (2019, A&A)

Leclercq et al. (2019, in prep)

Dynamics of the Nearby Universe

Going deeper:

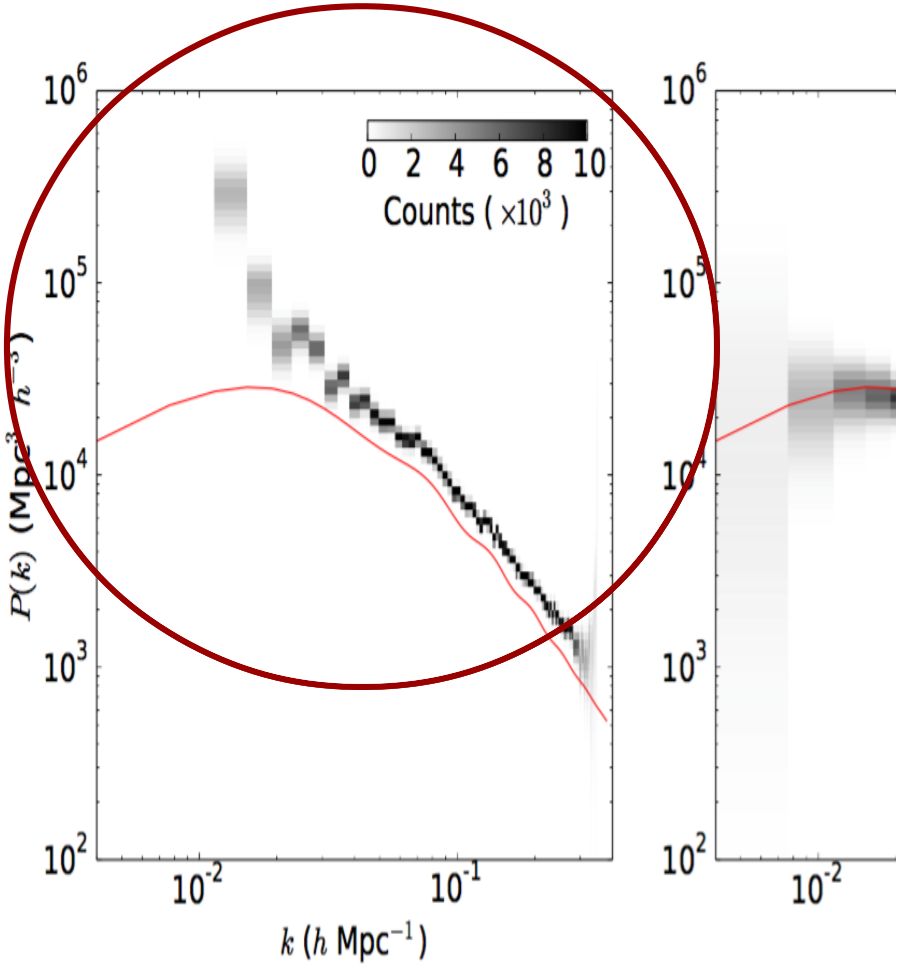
Analyzing

Sloan Digital Sky Survey III

DR12 ($z < 0.7$)

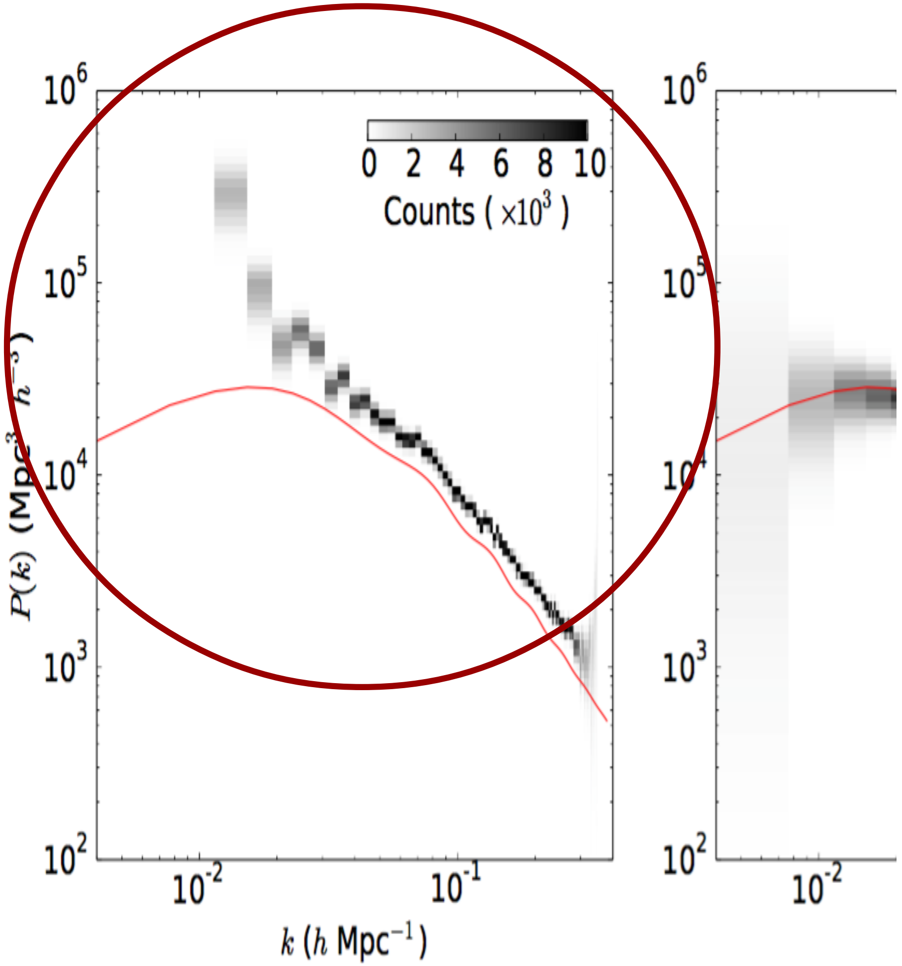
Application to SDSS III DR12

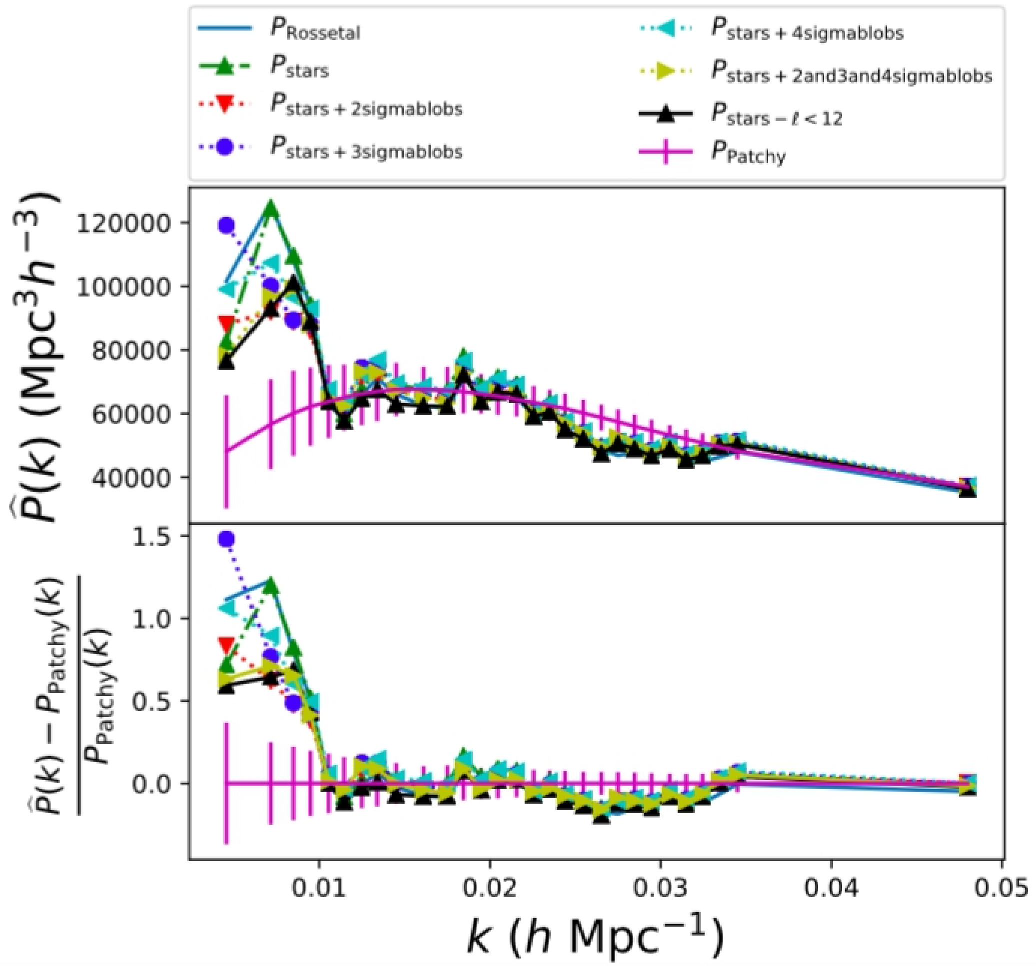
First Attempt!!!



Application to SDSS III DR12

First Attempt!!!

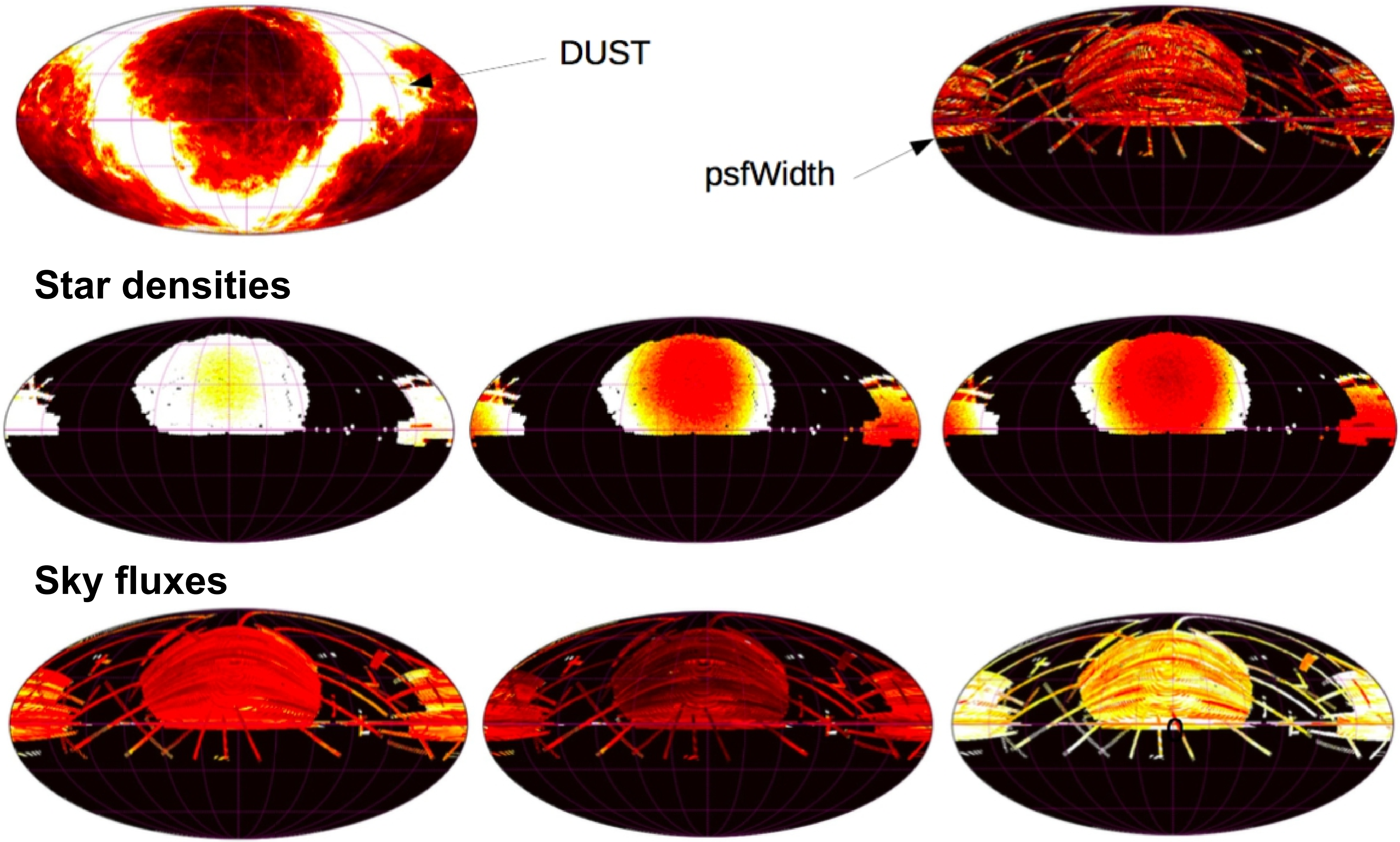




Kalus, Percival et al. (2018, MNRAS)

Foreground contaminations

Foregrounds hamper cosmological inference (see e.g. Leistedt & Peiris (2014))

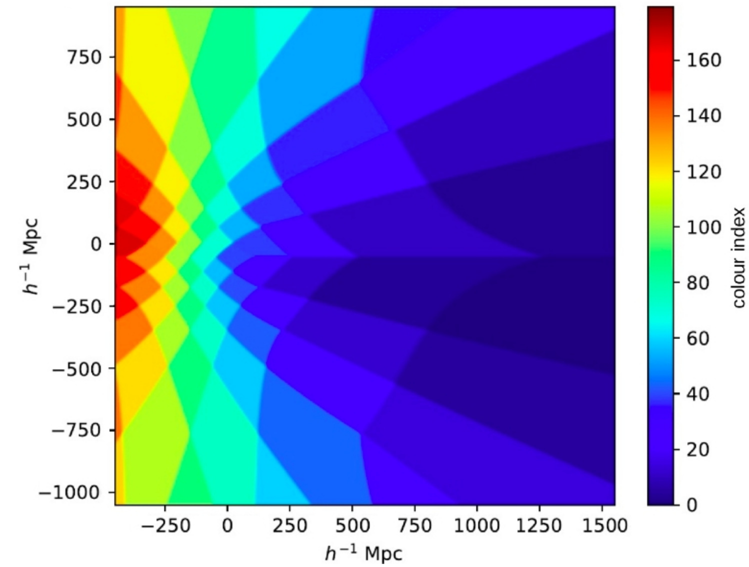


Jasche & Lavaux (2017, A&A)

Designing robust likelihoods

Map of patches on the sky

Extruded into 3d volume

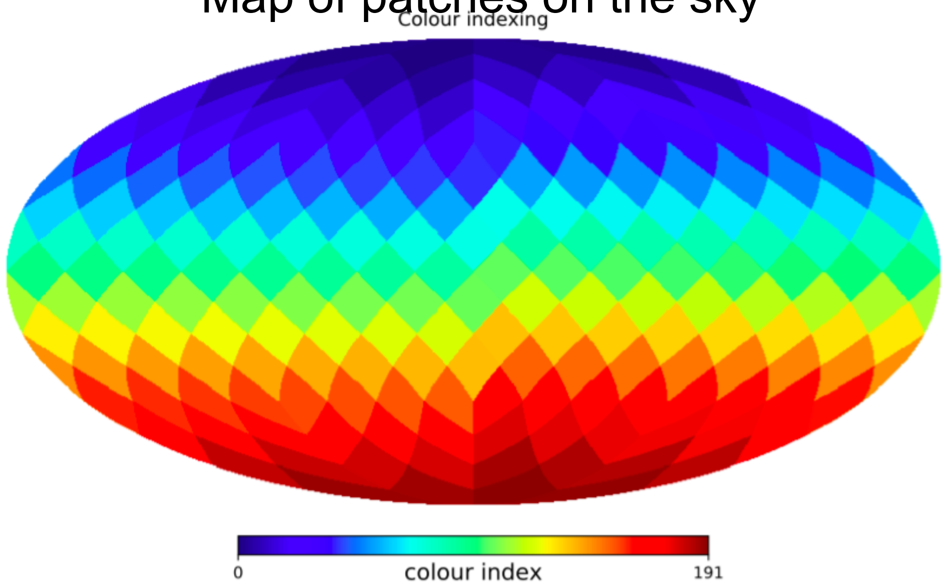


$$\mathcal{P}\left(N|\lambda(\delta)\,,A\right) = \prod_i \frac{e^{-A_i\,\lambda_i(\delta)}\left(A_i\,\lambda_i(\delta)\right)^{N_i}}{N_i}$$

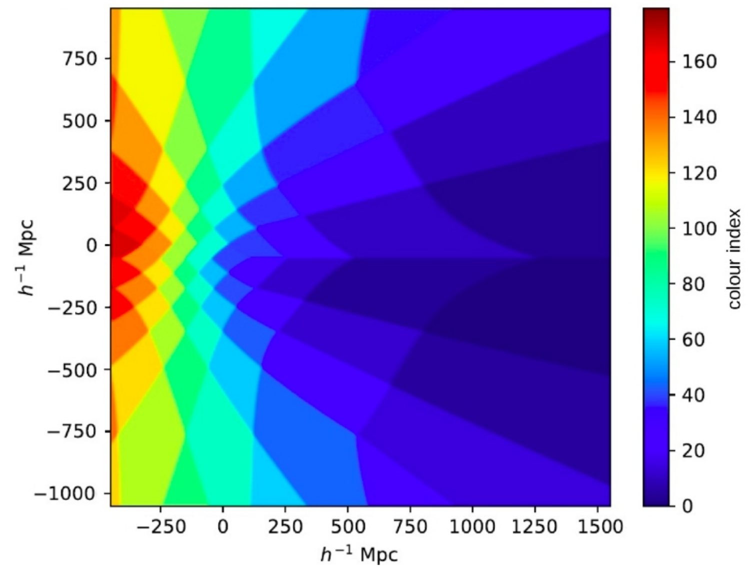
Porqueres et al (2019, A&A)

Designing robust likelihoods

Map of patches on the sky



Extruded into 3d volume



$$\mathcal{P}\left(N|\lambda(\delta), A\right) = \prod_i \frac{e^{-A_i \lambda_i(\delta)} \left(A_i \lambda_i(\delta)\right)^{N_i}}{N_i}$$

$$\mathcal{P}(N|\lambda(\delta)) = \int \mathrm{d}A \, \mathcal{P}(A) \, \mathcal{P}(N|A, \lambda(\delta))$$

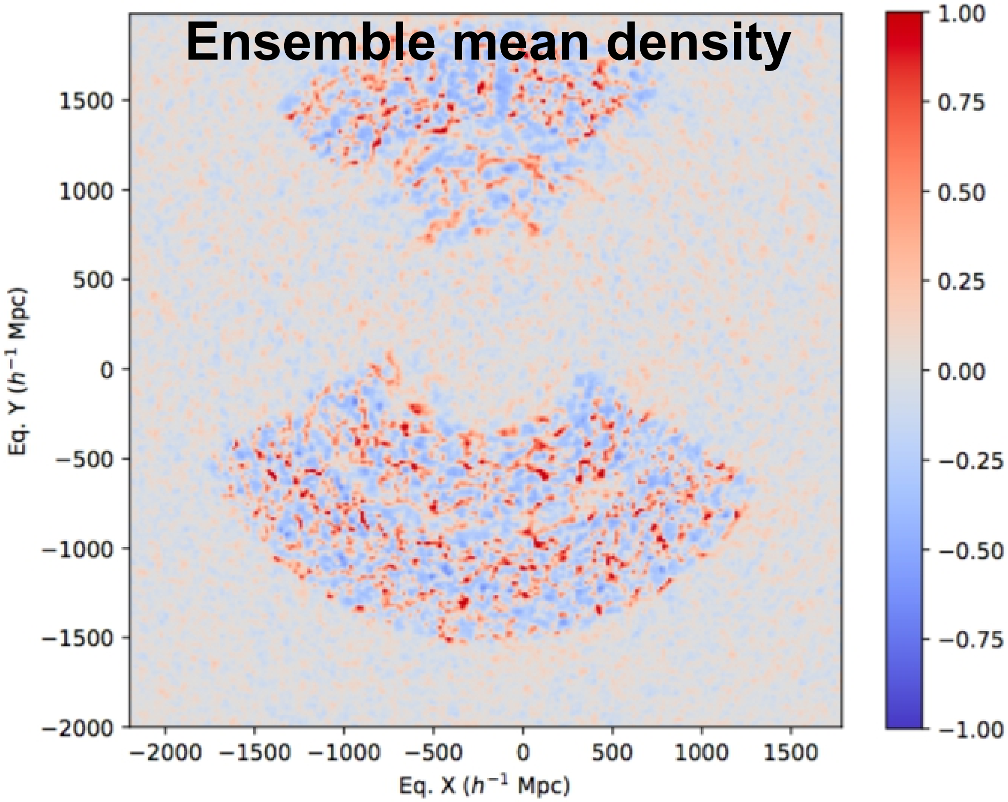
$$= \prod_{\text{patch}} \prod_{i \in \text{patch}} \left(\frac{\lambda_i(\delta)}{\sum_{j \in \text{patch}} \lambda_j(\delta)} \right)^{N_i}$$

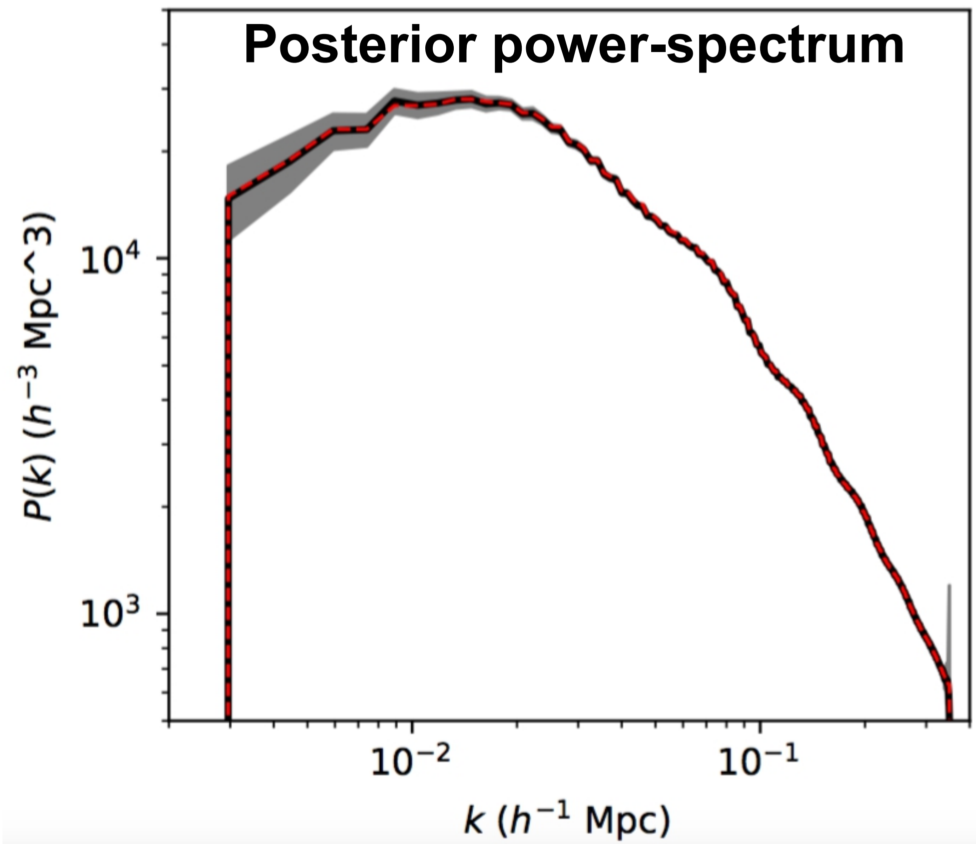
Porqueres et al (2019, A&A)

Application to SDSS III DR12

Preliminary results!

Real data application LOWZ + CMASS





Application: BORG – SDSS III

- Domain: (4000 Mpc/h)^3
- IC f luctuation elements: 256^3
- Simulation particles: 512^3
- LSS model: Lagrangian Perturbation Theory

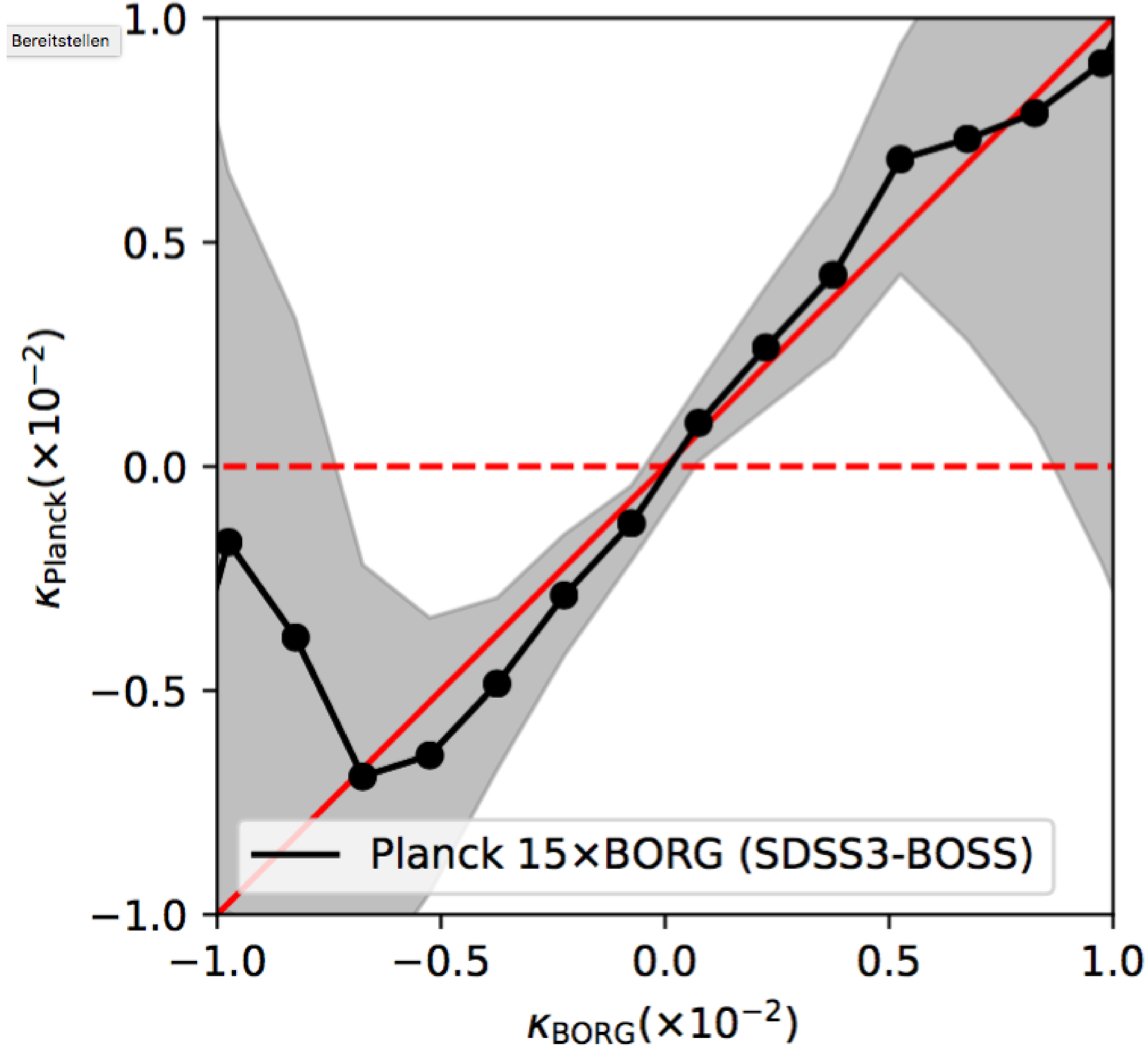
Lavaux et al (2019, in prep)

Independent test of inferred mass

CMB Lensing:

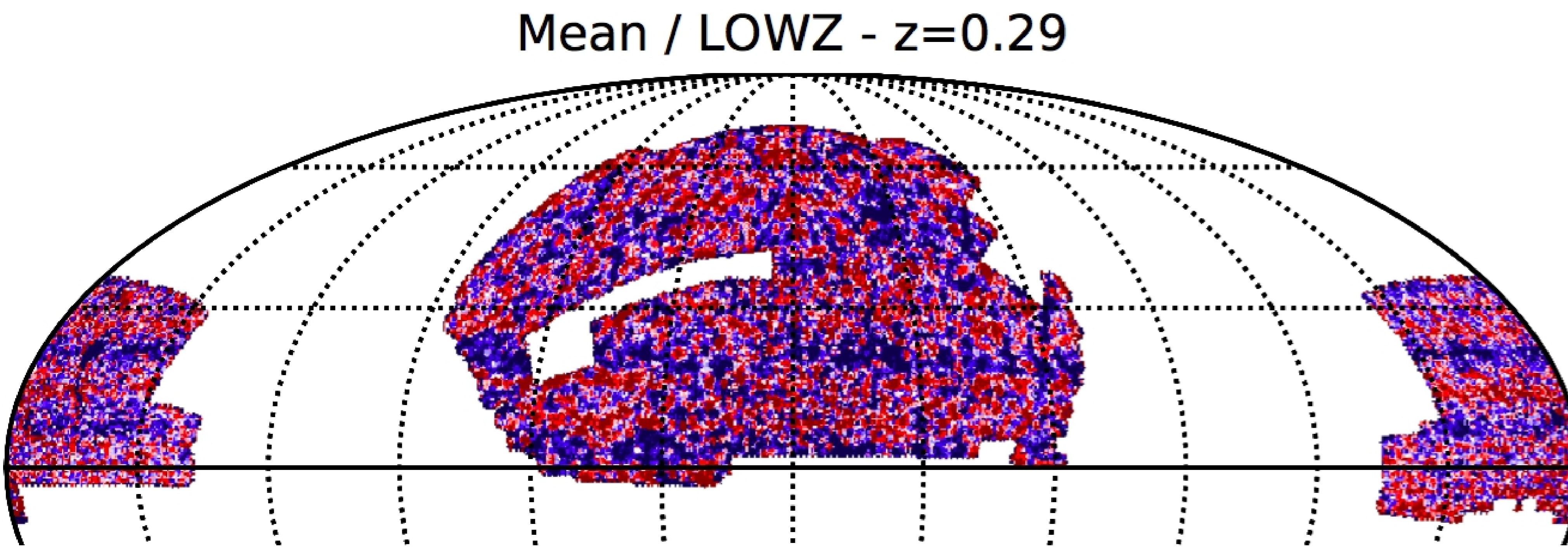
Preliminary results!

- Correlation Planck 15 vs. BORG SDSS3 convergence map



Jasche et al (2019 in prep)

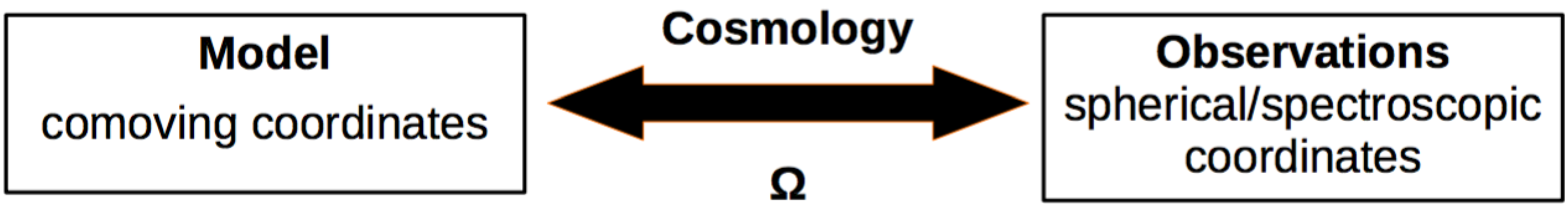
What are the foregrounds?



Lavaux et al (2019, in prep)

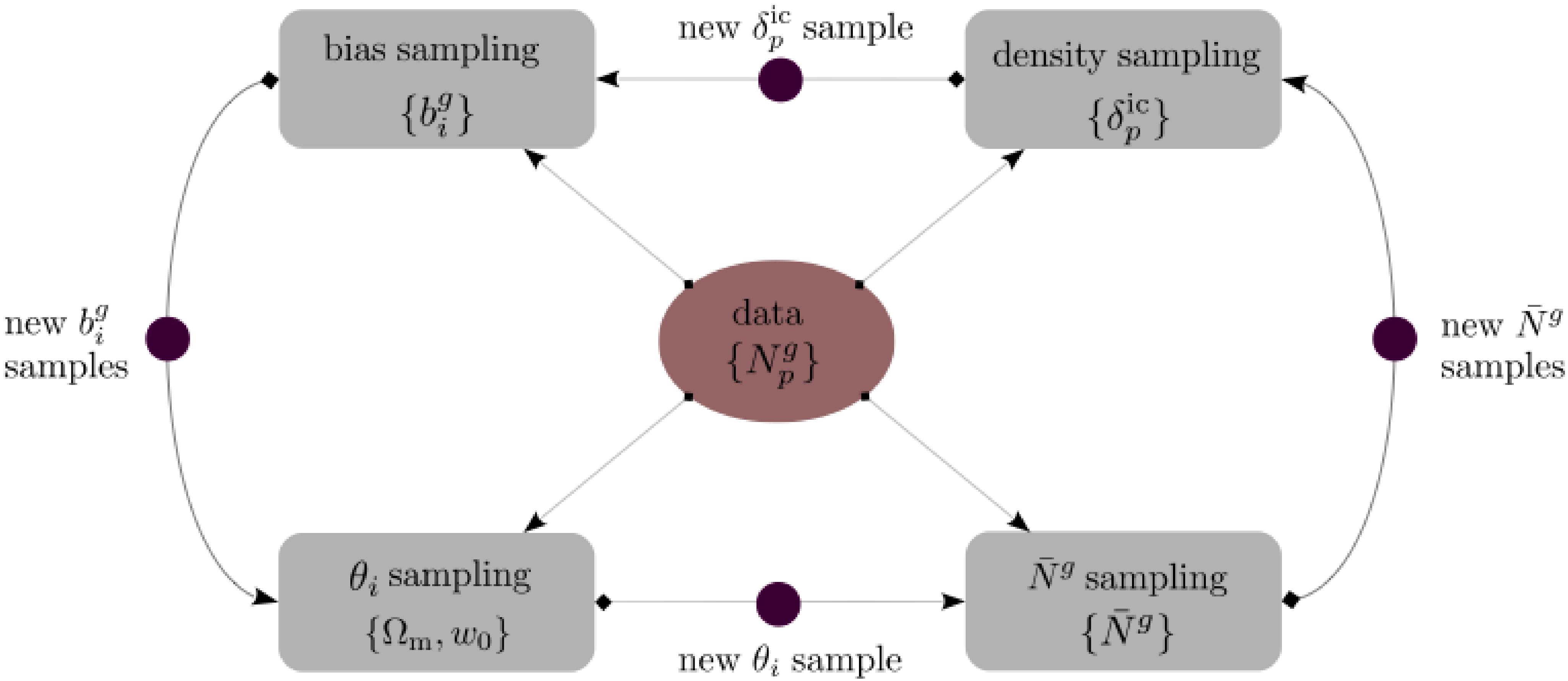
Cosmological parameter via the A/P test

Using the Alcock-Paczynski cosmological test



Kodi Ramanah et al (2019, A&A)

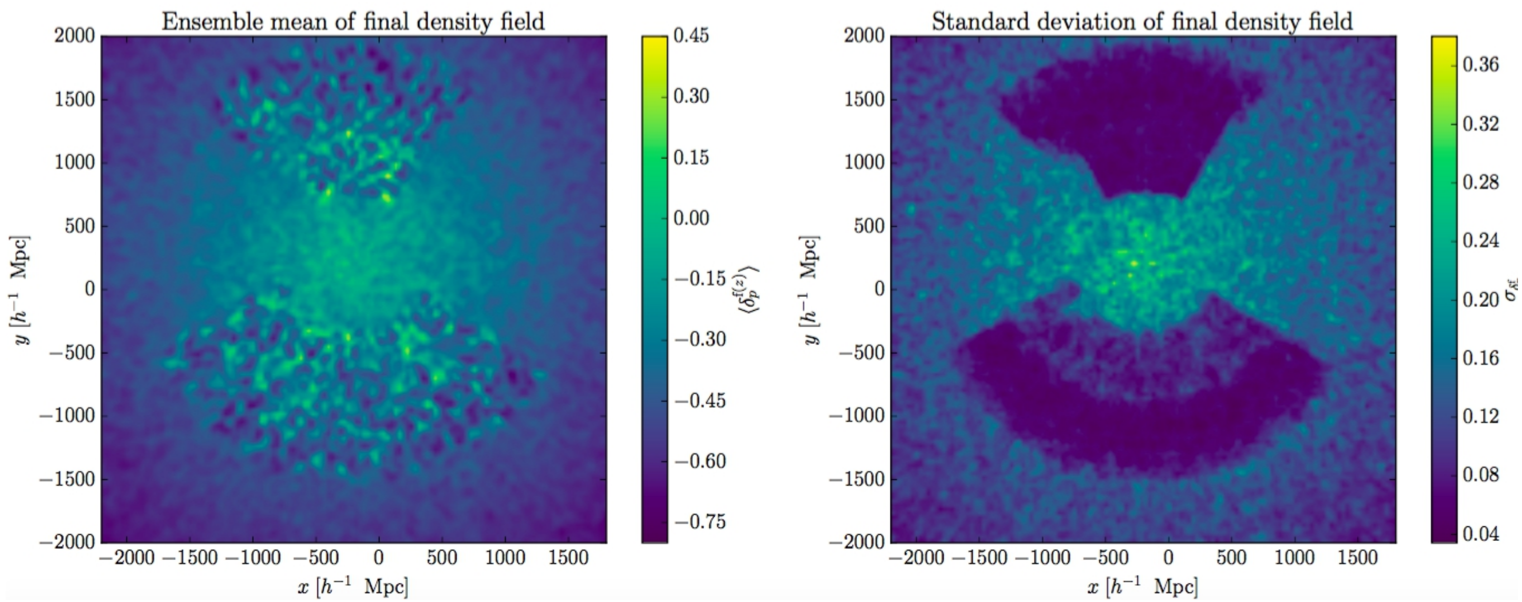
Cosmological parameter via the A/P test

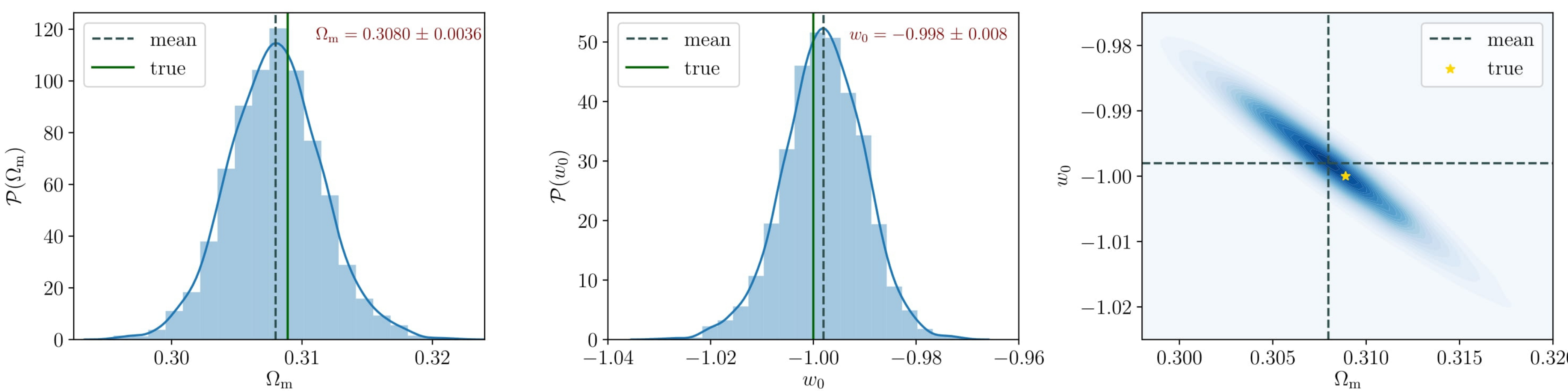


Kodi Ramanah et al (2019, A&A)

Cosmological parameter via the A/P test

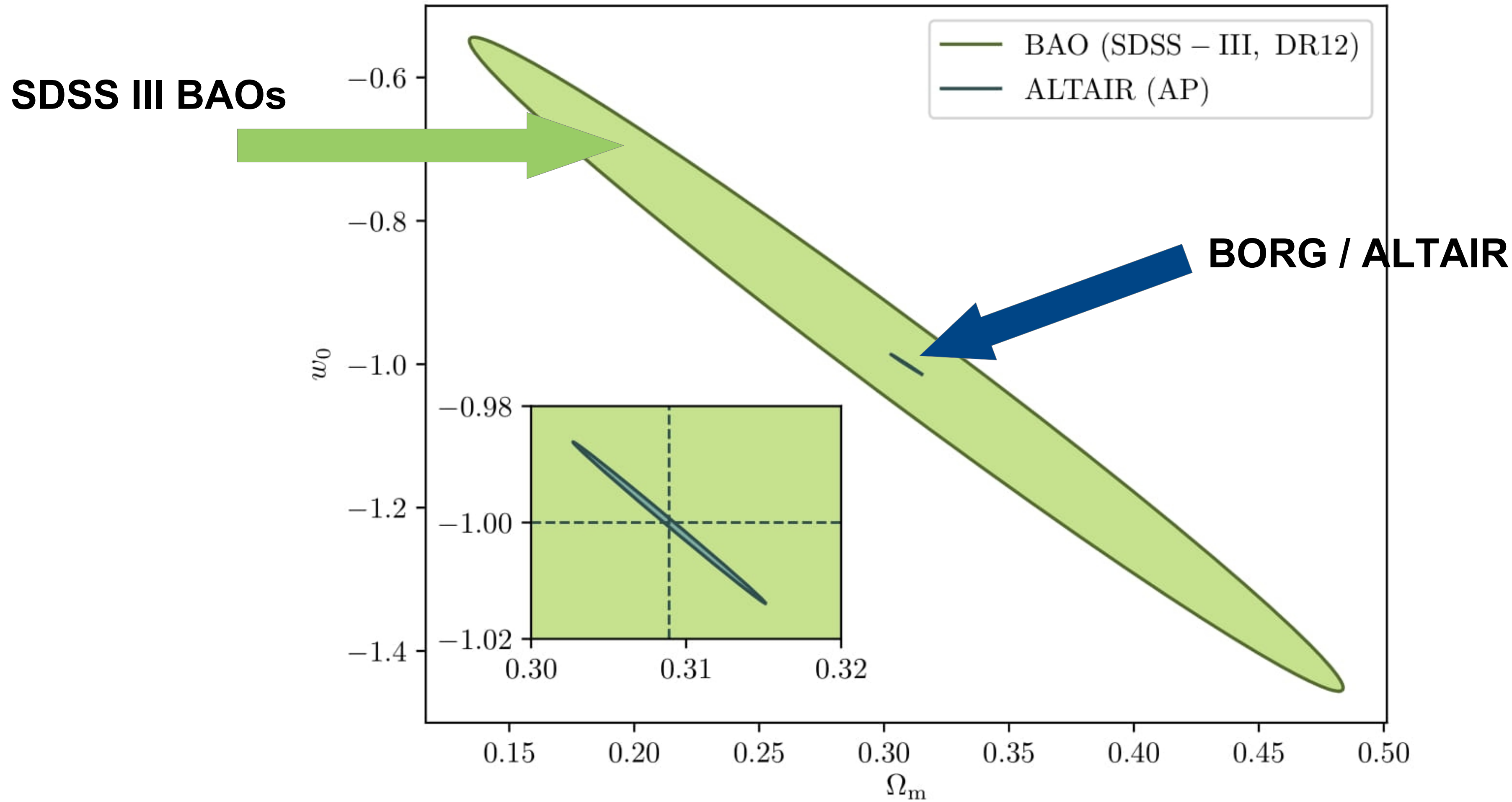
Application to SDSS III mock data:





Kodi Ramanah et al (2019, A&A)

Cosmological parameter via the A/P test



Source of Information (Kodi Ramanah et al (2019, in prep))**:**

- Complete use of modes
- Exploitation of higher order statistics

Kodi Ramanah et al (2019, A&A)

Summary & Conclusion

BORG combines physical modeling with data science:

- Dynamical modeling accounts for non-Gaussian statistics
- Flexible data modeling via HMC and block sampling
- Solves complex high dimensional statistics problems

Scientific results:

- Characterization of initial conditions
- Accurate & Detailed reconstructions of the DM field
- Complementary mass estimates
- Dynamical reconstructions
- Inference of cosmological parameter

The end...

Thank You!

Selected Bibliography:

Porqueres et al A&A 624, 115 (2019).

Ramanah et al A&A 621, 69 (2019).

Jasche, J. & Kitaura, F. S. *MNRAS* **407,** 29–42 (2010).

Jasche, J., Kitaura, F. S., Li, C. & Enßlin, T. A. *MNRAS* **409,** 355–370 (2010).

Jasche, J. & Wandelt, B. D. *MNRAS* **425,** 1042–1056 (2012).

Jasche, J. & Wandelt, B. D. *MNRAS* **432,** 894–913 (2013).

Jasche, J., Leclercq, F. & Wandelt, B. D. *JCAP* **01,** 036 (2015).

Lavaux, G. & **Jasche**, J. *MNRAS* **455,** 3169–3179 (2016).

Visit us at: www.aquila-consortium.org