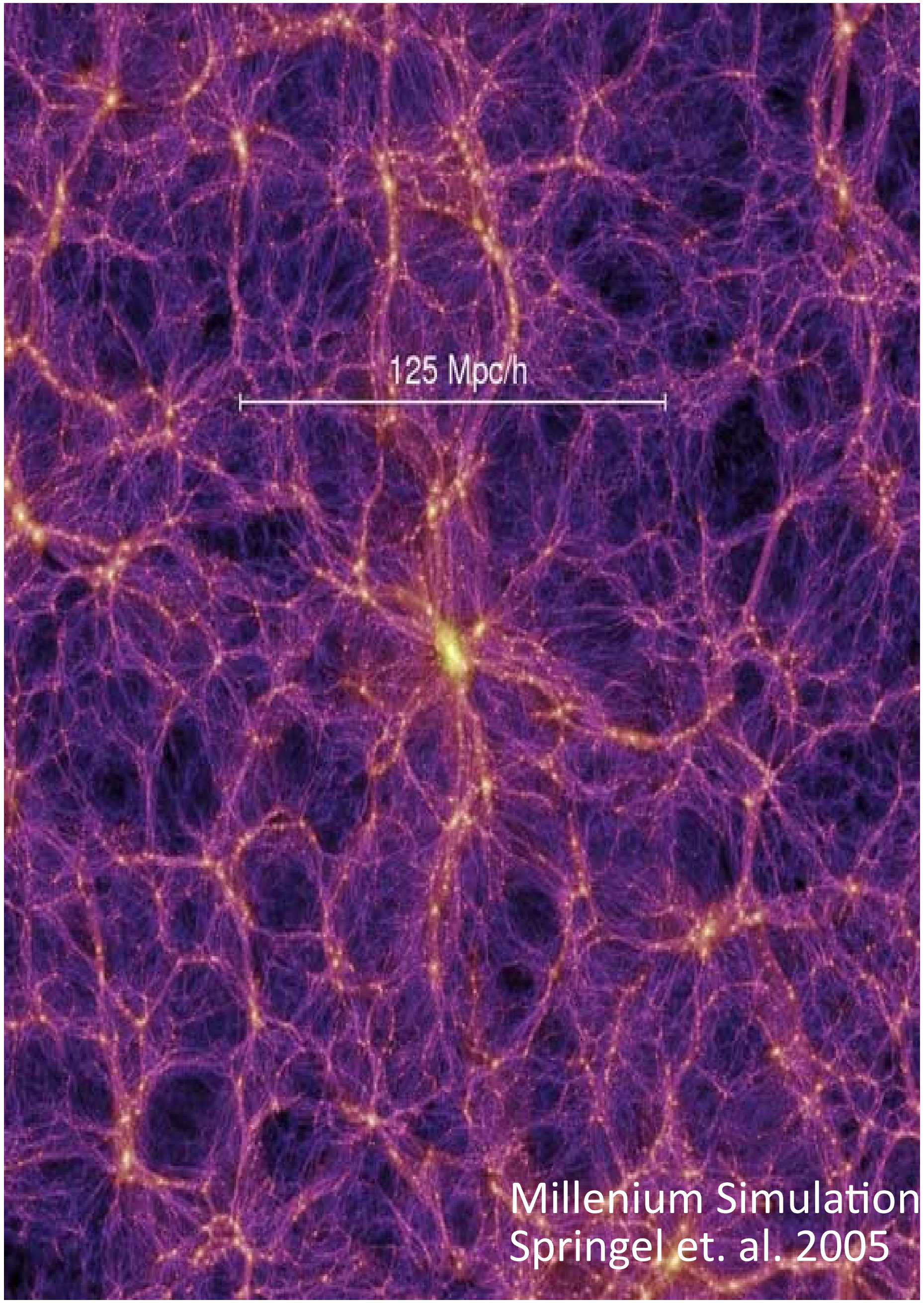
Bayesian reconstruction of the cosmic DM flow

Florian Führer

IAP, Paris

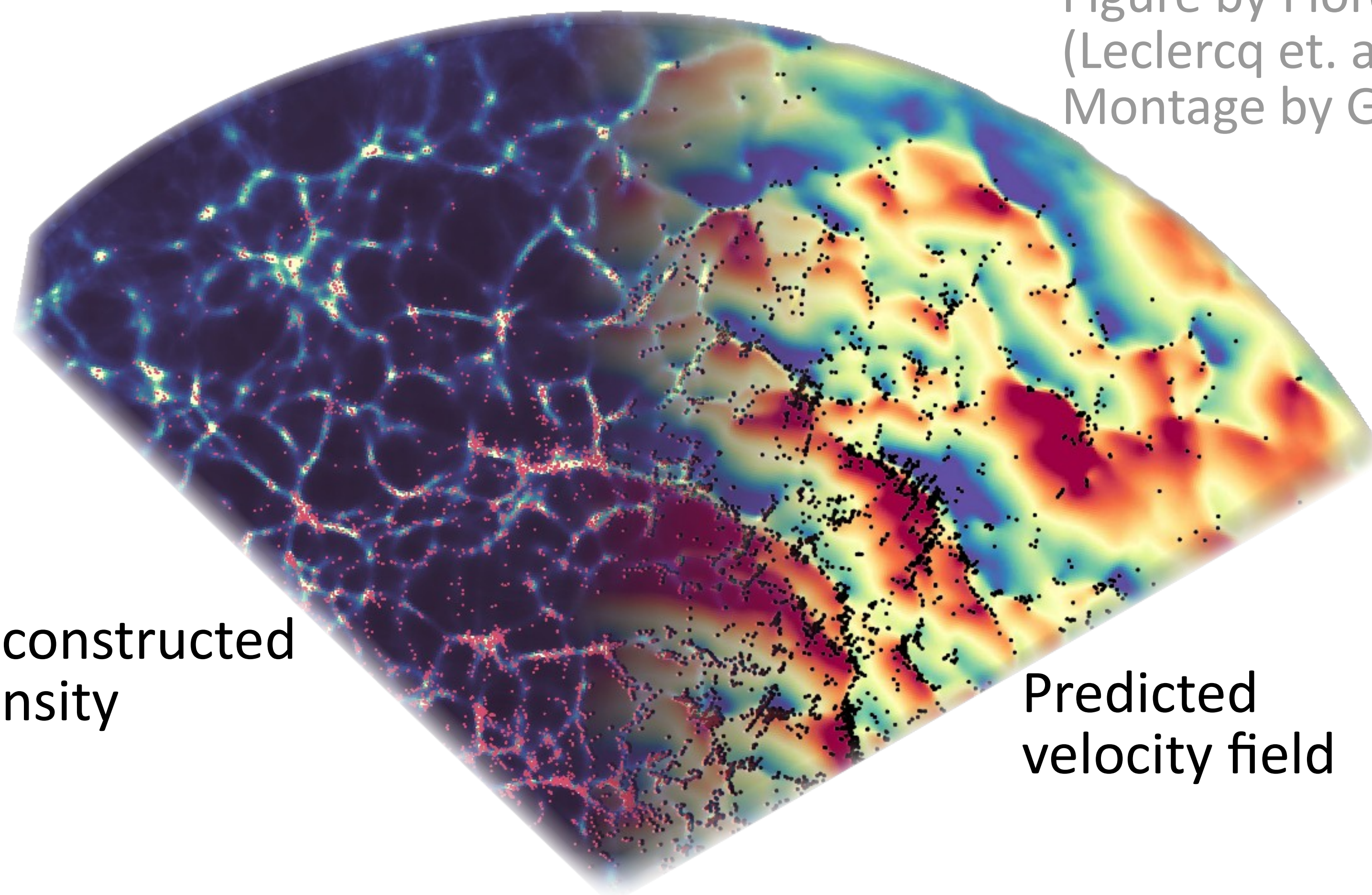
Cosmological large-scale structure

- Small initial conditions, approximate scale free and Gaussian
- Today still linear on large scales, non-linear on small scales
- Use to test of Dark Energy, Dark Matter, Gravity, Neutrino mass ...
- State-of-the-art analysis relies on the 2-point function
- Instead a reconstruction of the structures, allows to exploit (almost) all information



Large-scale structure reconstruction

Figure by Florent Leclercq
(Leclercq et. al. 2017),
Montage by Guilhem Lavaux



Reconstructed
density

Predicted
velocity field

Reconstructed DM structures from observed galaxies

Bayesian velocity reconstruction

Density reconstruction relies on redshift measurements

$$b\, n_{gal}(z) = \delta(z) \rightarrow \delta(\mathbf{x})$$

Adding (noisy) distance data allows direct velocity reconstruction

This talk

$$z - Hd \sim v_{\mathrm{galaxy}} \rightarrow \mathbf{v}(\mathbf{x})$$

If can do both
→ test dynamics

$$\partial_t \delta + \nabla \cdot ((1+\delta)\mathbf{v}) = 0 \quad \nabla^2 \Phi \propto \delta$$

$$\partial_t \mathbf{v} + H\mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla \Phi$$

In both cases need full Posterior

$\mathcal{P}(\mathbf{v}|\mathcal{D}) \propto \mathcal{P}(\mathcal{D}|\mathbf{v})\mathcal{P}(\mathbf{v})$ where $dim(\mathbf{v}) \sim 10^6\text{-}10^7$

See also tutorial on IFT

Hamilton Monte Carlo

Reformulate as Hamiltonian particle system, with auxiliary variables (momentum) $\mathbf{p}$

Neal 2011
Jasche, Kitaura 2010

$$H = \frac{1}{2}\mathbf{p} \cdot \mathbf{M} \cdot \mathbf{p} - \log\left(\mathcal{P}\left(\theta | \mathcal{D}\right)\right)$$

Samples obtained by solving Hamiltonian e.o.m.

$$\dot{\theta} = \mathbf{M}^{-1} \cdot \mathbf{p} \qquad \dot{\mathbf{p}} = -\partial_\theta \log\left(\mathcal{P}\left(\theta | \mathcal{D}\right)\right)$$

Hamiltonian Monte Carlo is well suited to sample from high dimensional distributions

- Travel large distances in parameter space
- High acceptance rate
- Use gradient information

The forward model

We fit the following forward model

$$\mathbf{v}(d_{L,i}) \cdot \mathbf{n_i} = \frac{z_i - \bar{z}(d_{L,i})}{1 + \bar{z}(d_{L,i})}$$

Line of sight

The forward model

We fit the following forward model

$$\mathbf{v}(d_{L,i}) \cdot \mathbf{n_i} = \frac{z_i - \bar{z}(d_{L,i})}{1 + \bar{z}(d_{L,i})} + \epsilon_{\mathrm{NL},i}$$

Line of sight

Stochastic contribution

Galaxies are imperfect traces of the DM structures

Error-model:

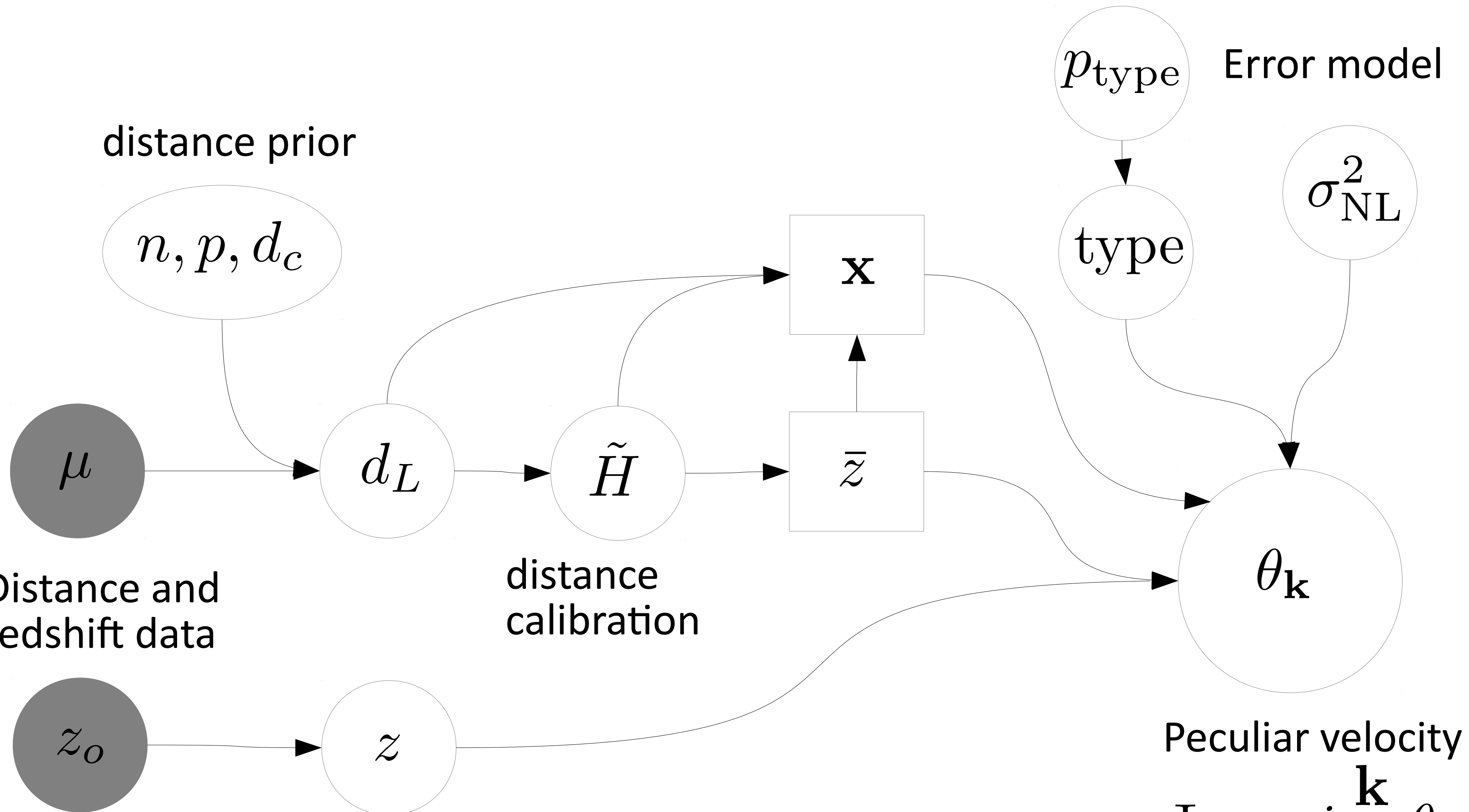
Assume ϵ_{NL} to be Gaussian

σ^2_{NL} can be different for different galaxies

→ classify galaxies into types with different

→ determine self-consistently

Statistical model and likelihood



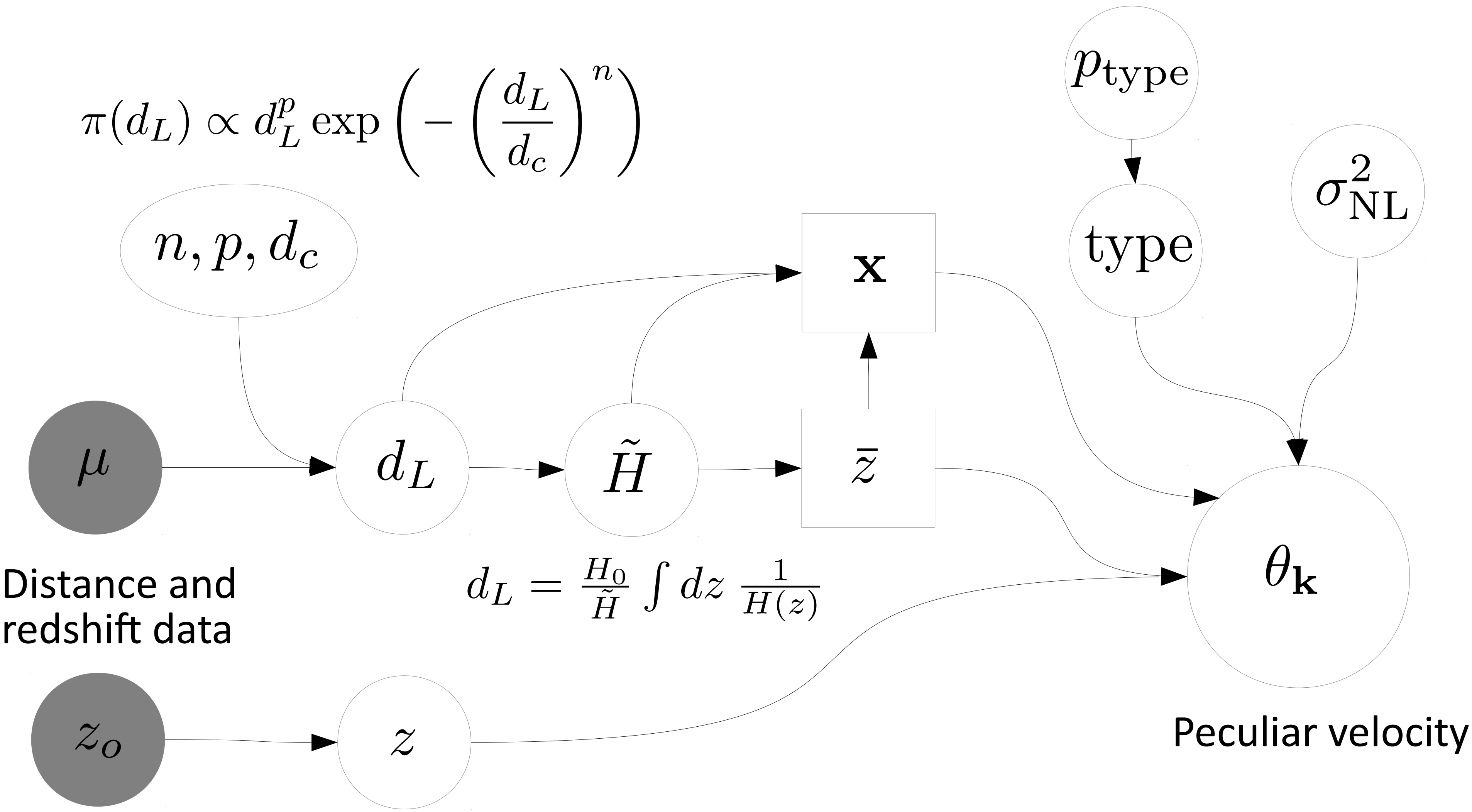
Peculiar velocity

$$\Psi_{\mathbf{k}} = i\frac{\mathbf{k}}{k^2}\theta_{\mathbf{k}}$$

$$\mathcal{L} \propto \prod_{\rm galaxies} \exp\left(-\frac{(v(z,\bar{z}) - Hf\Psi(\mathbf{x}))^2}{2\sigma^2_{\rm NL,type}}\right)\frac{1}{\sqrt{\sigma^2_{\rm NL,type}}}P(z|z_o)P(d_L|\mu)$$

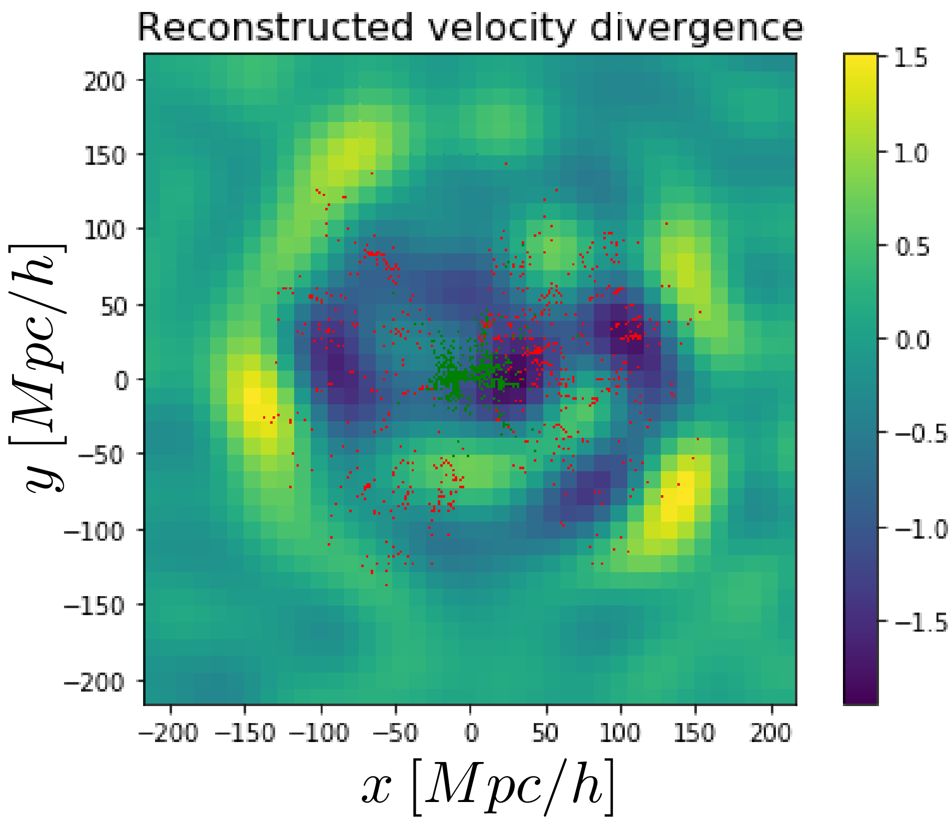
Statistical model and likelihood

$$\pi(d_L) \propto d_L^p \exp\left(-\left(\frac{d_L}{d_c}\right)^n\right)$$

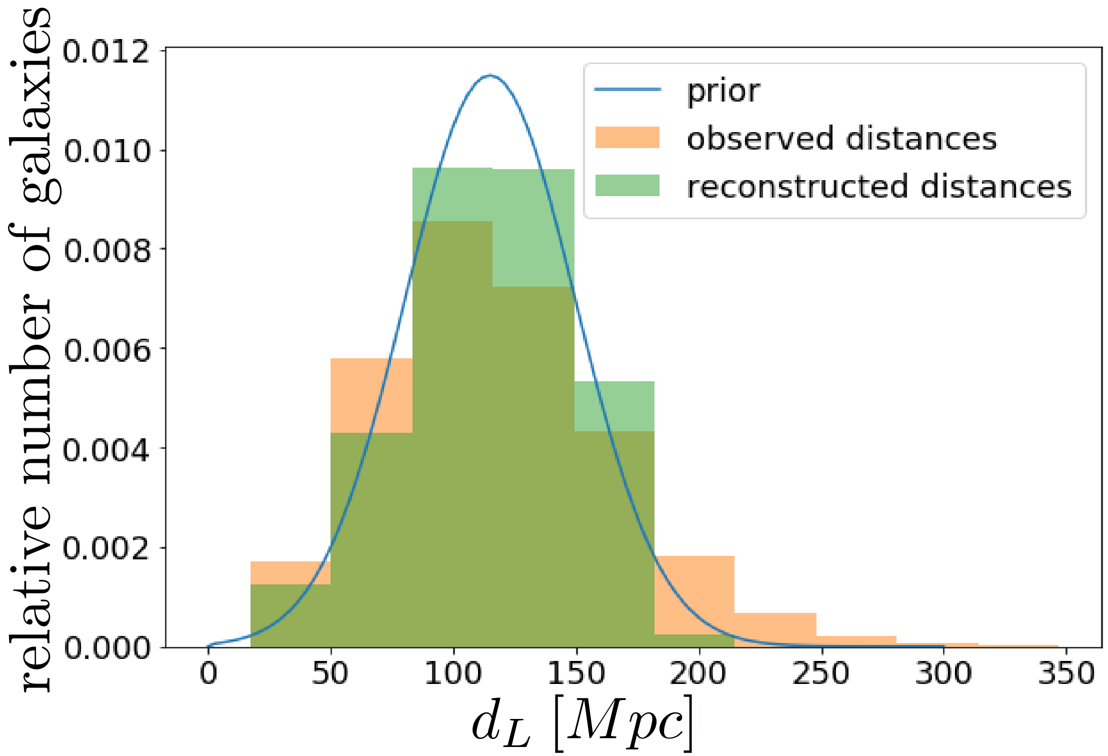


$$\mathcal{L} \propto \prod_{\text{galaxies}} \exp\left(-\frac{(v(z,\bar{z}) - Hf\Psi(\mathbf{x}))^2}{2\sigma^2_{\mathrm{NL,type}}}\right) \frac{1}{\sqrt{\sigma^2_{\mathrm{NL,type}}}} P(z|z_o) P(d_L|\mu)$$

Some preliminary results



First application of VIRBIuS to about 10000 galaxies from the 6dF and Spitzer catalogs



Reconstructed distances from 6dF catalog
Highlighting the effect of selection biases in distance measurements

Conclusion and Outlook

- Bayesian forward modeling combines physical modeling with Data Science See also tutorial on IFT
  - Signal reconstruction in very high-dimensions is feasible
  - Flexible data modeling via HMC and block sampling
  - Naturally accounts for systematics and uncertainties
- Will allow exhaustive use of available information
  - Use directly for parameter inference and test of models

    Kodi Ramanah et. al., 2018
- Open issues:

  Jasche, Lavaux 2018
  - Use of expensive forward models possible, but expensive
  - Model of selections related to instrument/data processing pipeline