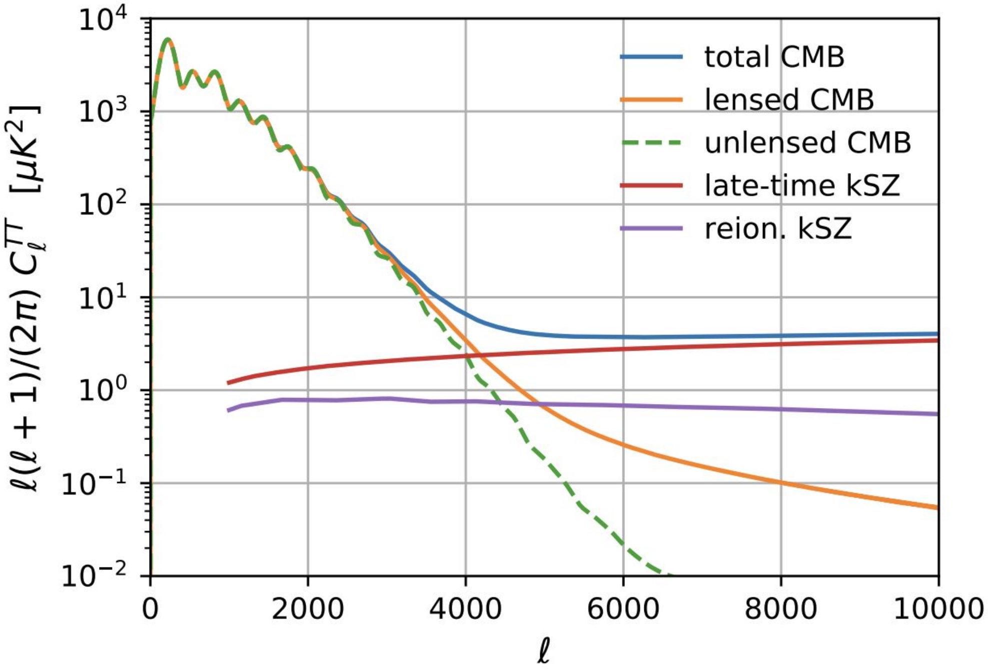
Reconstruction of the Radial Velocity Field of the Universe with a Joint CMB and Large Scale Structure Likelihood Analysis

Cosmology From Home - 2023
Yurii Kvasiuk (with Moritz Münchmeyer)
based on 2305.08903

1. Overview

Kinetic SZ

$$T(\hat{\mathbf{n}})\big|_{kSZ} = -\sigma_T \int d\chi \; a \; n_e(\hat{\mathbf{n}}, \chi) v_{\text{eff}}(\hat{\mathbf{n}}, \chi),$$



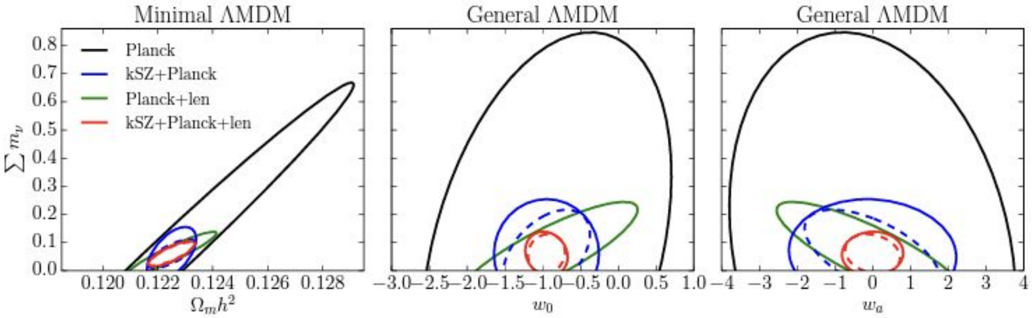
1810.13423

- Was proposed 1970 by Ya. Zeldovich and R. Sunyaev
- Was detected first with ACT and BOSS in 2012
- Depends on velocity, electron density
- **Is going to be well measured soon!**

What can be studied with kSZ?

$$\{\Omega_b h^2, \Omega_m h^2, \Omega_k, \Omega_\Lambda, w_0, w_a, n_s, \ln A_s, \gamma\}$$

- Properties of DE (eg. w, dw/da) [0511060, 0511061v1, 0712.0034]
- Modified gravity [1408.6248]
- Inflation (f_{NL} 1810.13424v1, CIP 2208.02829, etc)
- Neutrino mass
- Small-scale electron-galaxy spectrum



https://arxiv.org/abs/1412.0592

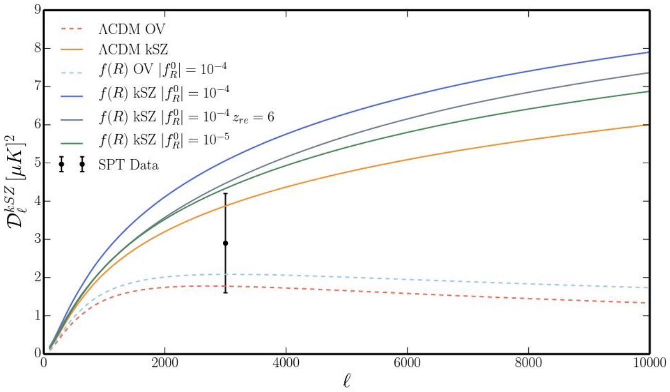


FIG. 3. The homogeneous kinetic Sunyaev-Zel'dovich power spectrum (solid lines) for standard ΛCDM and Hu and Sawicki models with $|f_R^0| = \{10^{-5}, 10^{-4}\}$ and $n = 1$ as a function of multipole ℓ. The dashed lines show the linear predictions, i.e. the OV effect. The black data band power $\mathcal{D}^{\rm kSZ}_{\ell=3000} = 2.9 \pm 1.3\mu{\rm K}^2$ (1σ confidence level) is taken from the South Pole Telescope (SPT) [32]). $z_{\rm re} = 9.9$ is assumed except where otherwise stated.

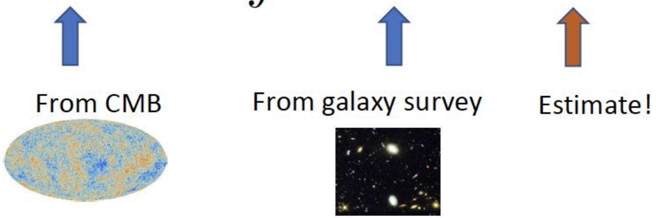
https://arxiv.org/abs/1510.08844

Summary of kSZ velocity reconstruction

- **Step 1**: estimate the radial velocity field from kSZ

Idea:

$$T_{kSZ}^{CMB} \sim \int dr\ \rho_e(r)\ v_r(r)$$


From CMB

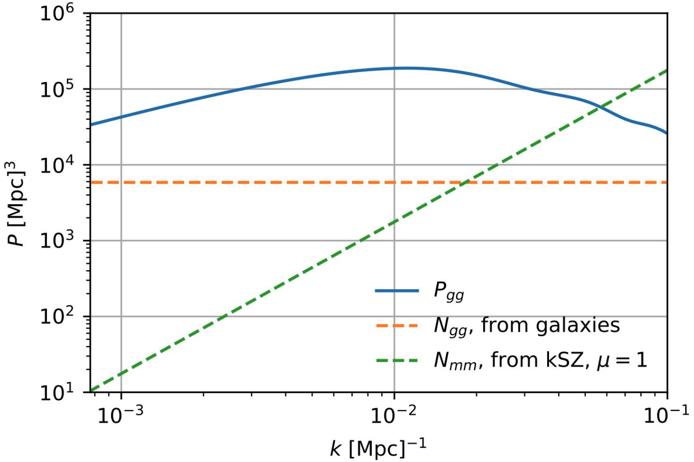
From galaxy survey

Estimate!

- **Step 2**: From reconstructed velocities, we can calculate the matter density perturbations (continuity equation).

$$\hat{v}_r(\mathbf{k}) \xrightarrow{\mathbf{v} \propto \frac{\delta_m}{k}} \hat{\delta}_r(\mathbf{k})$$

Radial matter modes are reconstructed with very high SNR.



This plot: forecast for DESI+SO

High signal-to-noise for upcoming experiments. Competitive probe for cosmology.

2. Velocity Reconstruction with QE and Likelihood

Quadratic Estimator

- **An estimator constructed from a weighted product of observables - CMB and galaxy field.**
- Weights are obtained by minimizing the variance

$$\hat{v}_\alpha(\boldsymbol{l}) = \int d^2l_1 d^2l_2 W_\alpha(\boldsymbol{l_1},\boldsymbol{l_2})\tilde{\theta}(\boldsymbol{l_1})\tilde{\delta}^g_\alpha(\boldsymbol{l_2})\delta^{2D}(\boldsymbol{l}-\boldsymbol{l_1}-\boldsymbol{l_2})$$

Field-level MAP

- **A joint differentiable model of CMB and LSS to the combined field level data.**
- General idea: maximize Bayesian posterior wrt unknown physical field(s)
- Needs the forward model - model of observations dependent on parameters

$$\mathcal{P}(\phi|\phi^o) \propto \mathcal{L}(\phi|\phi^o) \times \mathcal{P}(\phi)$$

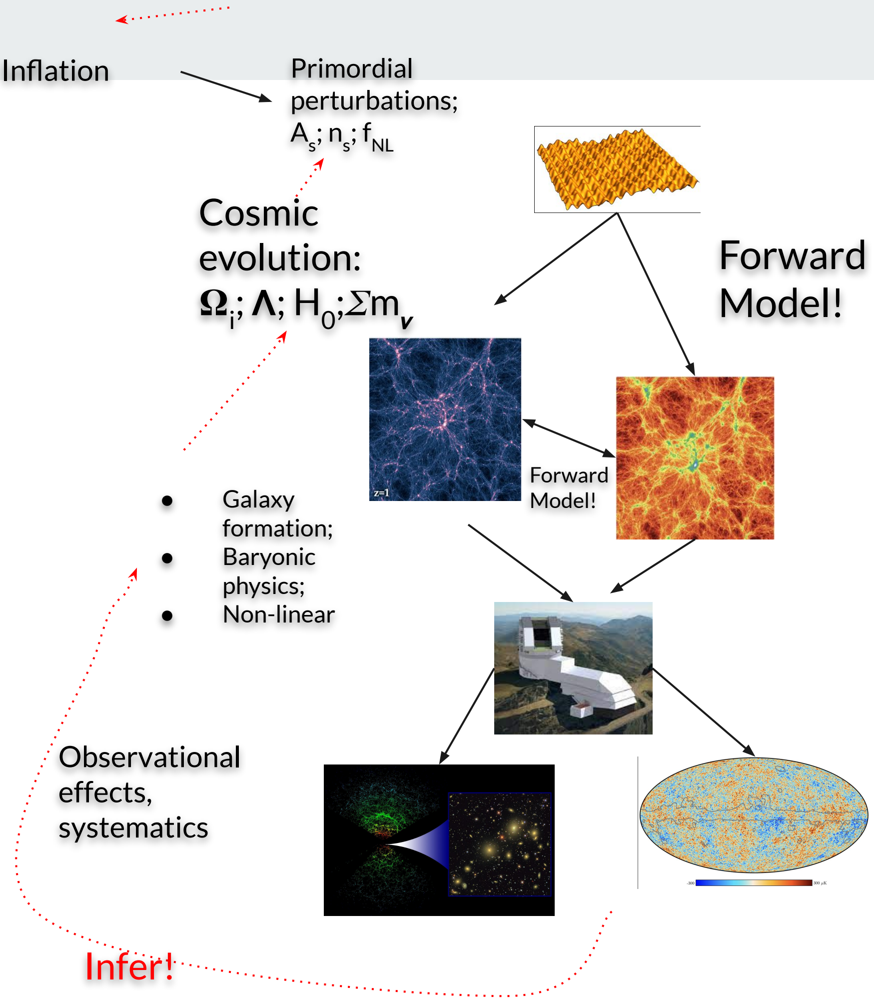
$$\hat{\phi}^{MAP} = \arg\max_{\phi} \mathcal{P}(\phi|\phi^o)$$

Forward Modelling Approach

- More capable than analytical approximation (easier to model than analytically calculate)
- Can be semi-analytical, numerical or ML-based
- Multiple effects can be included into the model
- Statistically more favorable framework (inherently Bayesian)

But

- Computationally intensive (can be mitigated)
- Needs to be differentiable



Inflation

Primordial perturbations; A_s; n_s; f_{NL}

Cosmic evolution: Ω_i; Λ; H_0;Σm_ν

Forward Model!

Forward Model!

- Galaxy formation;
- Baryonic physics;
- Non-linear

Observational effects, systematics

Infer!

Likelihood approach (CMB x LSS)

Related to 2205.15779
but optimization based.

- **Advantages of our new likelihood approach:**
  - As in case of lensing, there is a likelihood formulation that can in principle **improve over the QE in signal to noise**.
  - Conceptually straightforward to **combine multiple secondary anisotropies and matter probes.**
  - Include **cosmological and astrophysical parameters** (eg: f_nl, galaxy biases, velocity biases from optical depth degeneracy) in the fit.
  - Can **include machine learning elements to model non-Gaussian small-scale distributions** and improve SNR.
  - Can do **Bayesian analysis** in principle. Here: only **MAP**.
  - **Can combine several estimators in one analysis. E.g.** Likelihood naturally gives **partially "dekSZed" CMB**. Also includes "projected field estimator.

kSZ Likelihood

Abstract notation, bin and mode sum implicit.

Primary lensed CMB
+ unresolved kSZ

Some model of the kSZ given the
electron or galaxy distribution.

$$
\begin{aligned}
-2\ln\mathcal{L}\left(\Theta^{obs},\delta_g^{obs}|\Theta^{pCMB},\delta_g,v_r\right) = & \left(\Theta^{obs}-\Theta^{pCMB}-\Theta^{kSZ}(\hat{\delta}_e(\delta_g),v_r)\right)^T N_T^{-1} \\
& \times\left(\Theta^{obs}-\Theta^{pCMB}-\Theta^{kSZ}(\hat{\delta}_e(\delta_g),v_r)\right) \\
& +\left(\delta_g^{obs}-\delta_g\right)^T N_g^{-1}\left(\delta_g^{obs}-\delta_g\right)+\text{const.}
\end{aligned}
$$

Observe CMB and galaxy field

Reconstruct primary lensed CMB,
galaxy density, velocity field, and later
electron density

Assume Gaussian noise on Galaxies
and CMB temperature

Priors on the fields

To obtain a well-defined posterior we need priors on all the fields.

Simplest assumption: **all fields get Gaussian priors (common but suboptimal assumption)**

$$-2\ln\mathcal{P}(T^{pCMB},\delta_g,v_r) = (T^{pCMB})^T P_T^{-1} T^{pCMB} + v^T P_v^{-1} v + \delta_g^T P_{\delta_g}^{-1}\delta_g + \text{const.}$$

We use a **separation of scales**, i.e. velocities are on large scales only and galaxy field is on small scales.

More elaborate priors are possible, e.g. a **stochastic connection between electrons and galaxies**:

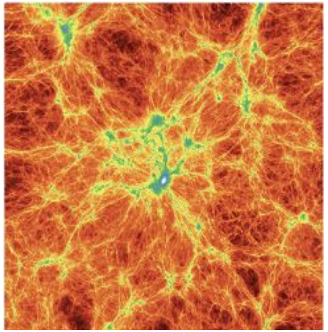
$$-2\ln\mathcal{P}(\delta_e,\delta_g) = \begin{pmatrix}\delta_e & \delta_g\end{pmatrix}\begin{pmatrix}P_{ee} & P_{ge}\\ P_{ge} & P_{gg}\end{pmatrix}^{-1}\begin{pmatrix}\delta_e\\ \delta_g\end{pmatrix} + \text{const.}$$

Key ingredient: Connection between kSZ and galaxies

kSZ equation: $T(\hat{n})\big|_{kSZ} = -\sigma_T \int d\chi \; a \; n_e(\hat{n},\chi) v_{eff}(\hat{n},\chi)$

N radial bins: $\theta(\boldsymbol{x})\big|_{kSZ}^{binned} = \sum_{\alpha=1}^{N_{bins}} \tau^{\alpha}(\boldsymbol{x}) v_r^{\alpha}(\boldsymbol{x})$

$$\tau^{\alpha}(\boldsymbol{x}) = f^{\alpha}(1 + \delta_e^{\alpha}(\boldsymbol{x}))$$



We need **a model that connects the electron density to the observed galaxy density**:

Simplest possibility (deterministic): $\hat{\delta}_e(k) = \frac{P_{\delta_e \delta_g}}{P_{\delta_g \delta_g}} \delta_g(k)$

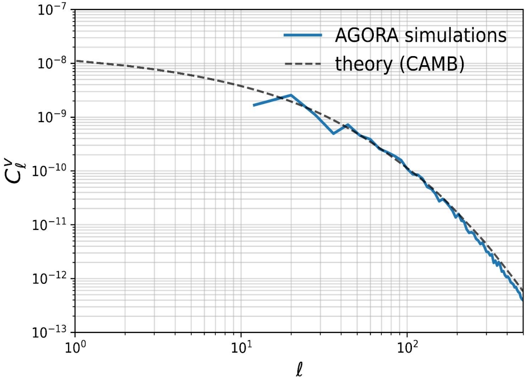
Also possible: **stochastic connection, machine learning based model, or stochastic machine learning based model.** See later in the talk.

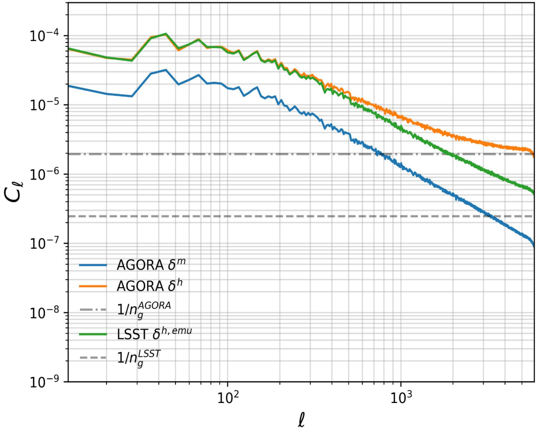
3. Implementation and results

Implementation

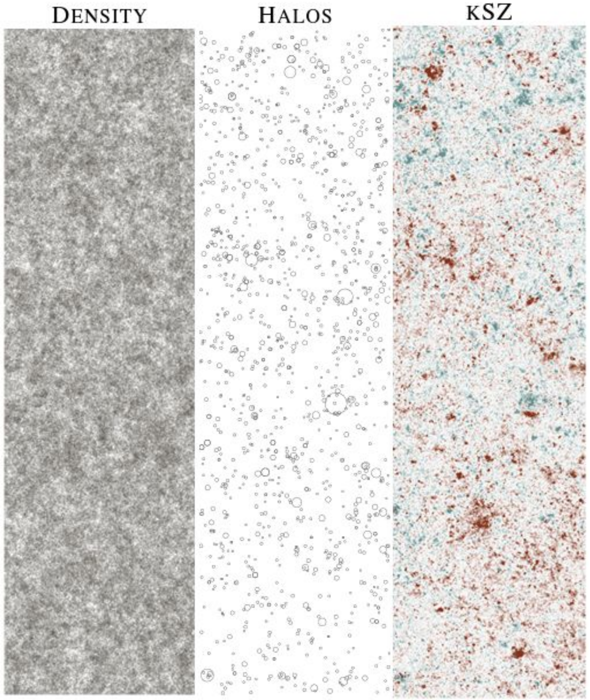
- (2n+1)xN^2; N pixels per side, n radial bins.
- Likelihood lives in pixel space, because noise covariance is diagonal there; priors live in Fourier space, where power spectra are diagonal -> we have to be able to switch effortlessly between them.
- With given resolution, for **z=0.5 to z=3, there are in total ~0.5 billion of parameters** to solve for - high dimensional!
- Tractable due to recent advances from the field AI - a highly-optimised GD (autodiff, multi-GPU)

Results on Agora simulations

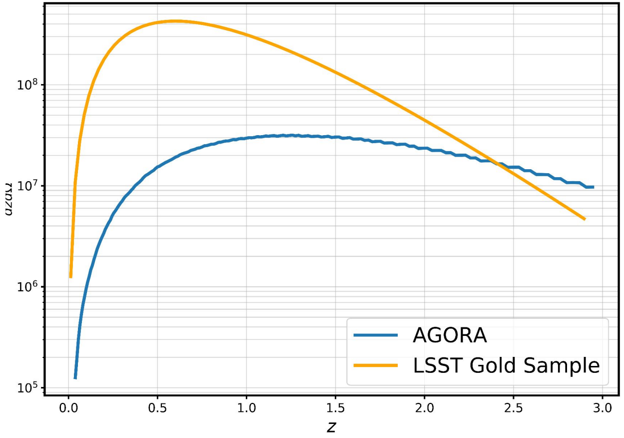




[Y. Omori, https://arxiv.org/abs/2212.07420]



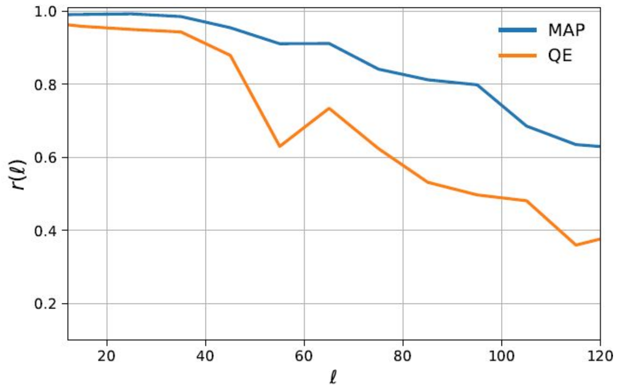
No hydro: $\delta_e \approx \delta_m$



AGORA galaxy densities are lower than those predicted for LSST

How do we quantify the results?

Cross-correlation coefficient:
the closer to 1 - the better



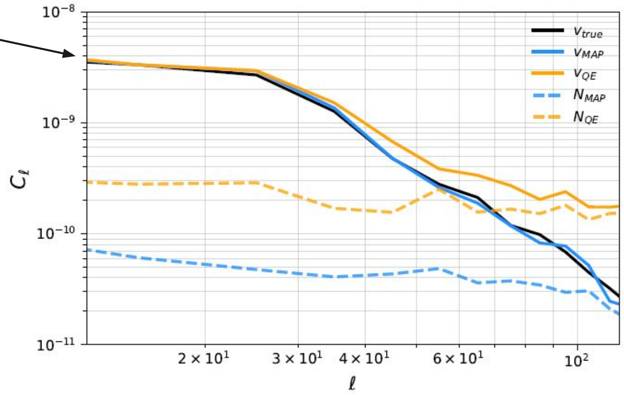
MAP

QE

$$r^{v,\hat{v}}(\ell) = \frac{C^{v,\hat{v}}(\ell)}{\sqrt{C^{v,v}C^{\hat{v},\hat{v}}(\ell)}}$$

$$C^{v}(\ell) = \langle v(\mathbf{l})v(\mathbf{l})\rangle$$

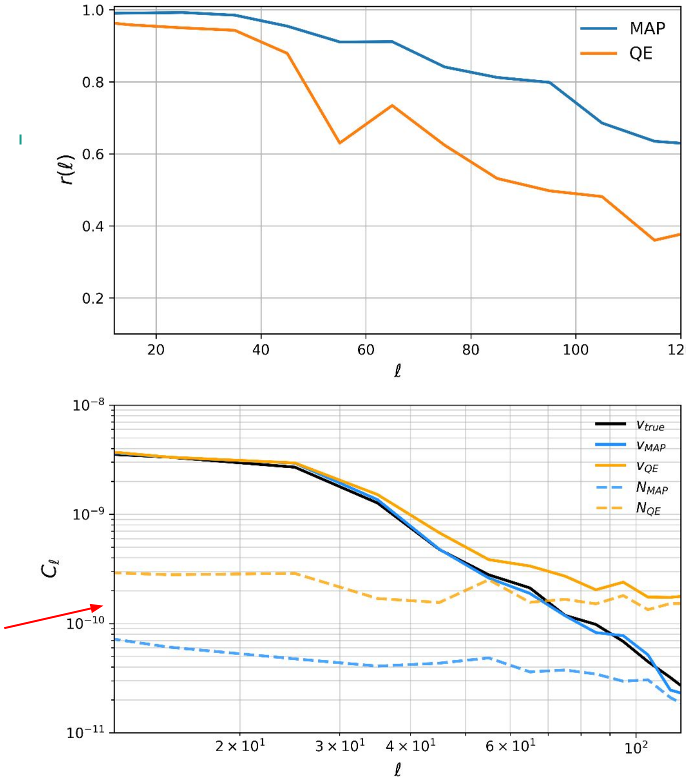
Power spectrum (2pt function)



Noise:
the lower - the better

$$C^{\epsilon}(\ell) = \langle |v^{true}(\mathbf{l}) - \hat{v}(\mathbf{l})|^2\rangle$$

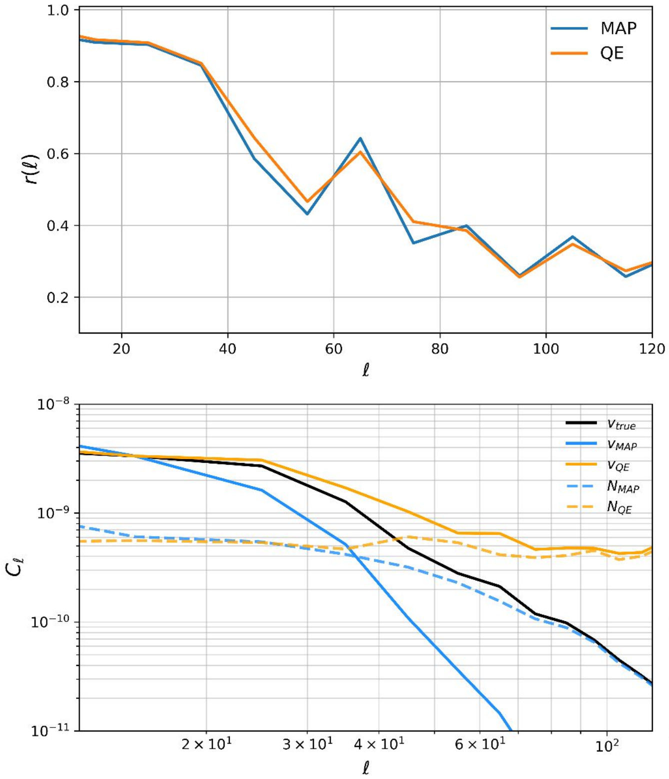
Futuristic noises - Realistic noises



3-4 times noise reduction

$N_{cmb} = N_{hd}$;
$n_g = 10*n_{g,LSST}$

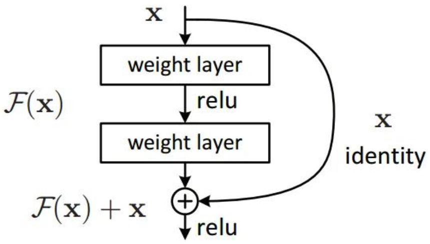
Good improvement over QE!



$N_{cmb} = N_{S4}$
AGORA halos

Equivalent to QE

Local ResNet for electron density estimation

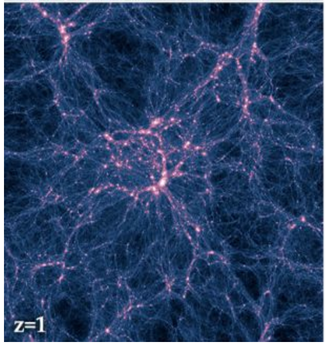


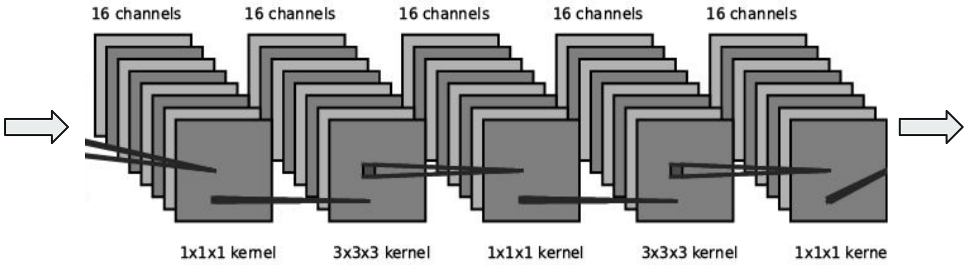
Credits: https://arxiv.org/abs/1512.03385

Motivation:

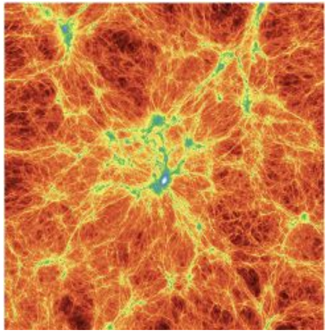
$$(N_\ell^v)^{-1} \propto \int dll \frac{(P_\alpha^{g\tau}(l))^2}{\tilde{P}^{\theta\theta}(l)\tilde{P}^{gg}(l)} = \int dll \frac{f_\tau^2 r^2(l) P^{ee}}{\tilde{P}^{\theta\theta}(l)}, \quad r^2(l) = \frac{(P^{eg})^2}{\tilde{P}^{gg} P^{ee}}$$

Observed Galaxy Density





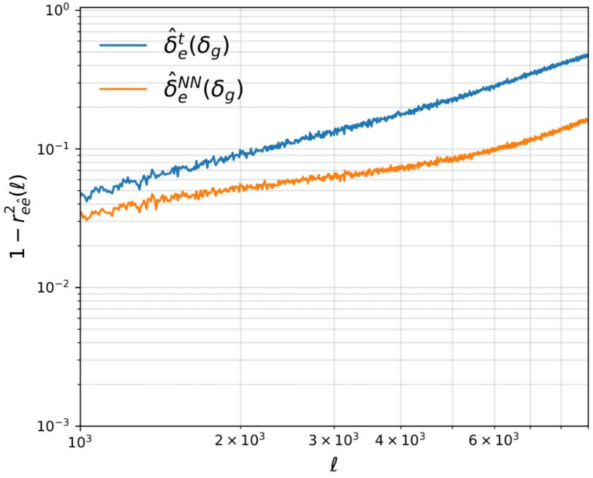
Electron density



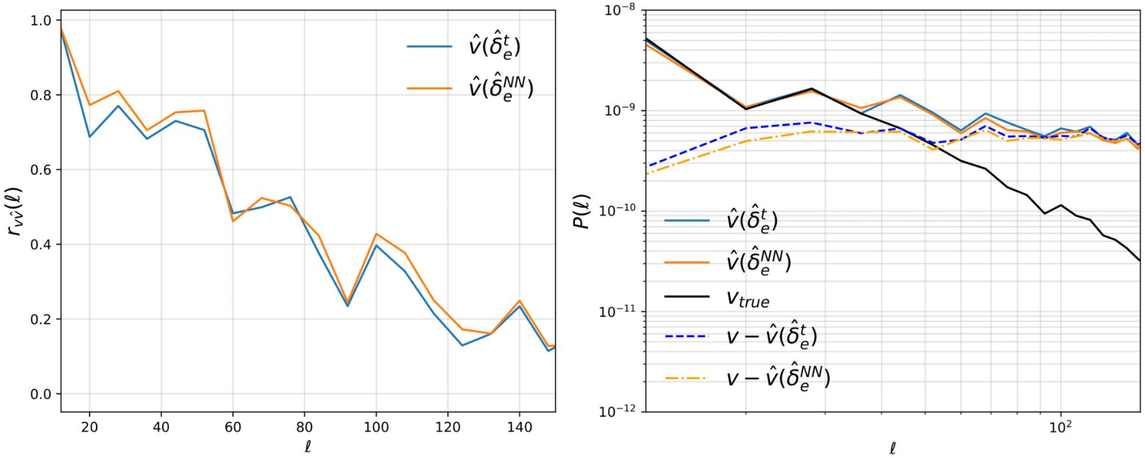
Credits:
https://arxiv.org/abs/2205.12964
Note: for illustrative purpose only

Credit: https://www.illustris-project.org/media/

Res-Net for better estimator of electron density (QE case)



Improvement of cross-correlation of the estimator for electron density field



Effect on kSZ velocity reconstruction with QE.
Motivates to try on hydro sims.

Fitting model parameters (work in progress)

The posterior depends on **large-scale cosmological parameters** (Λ_L) and **small scale astrophysical parameters** (Λ_S) such as the kSZ optical depth.

In full generality the posterior is:

$$\mathcal{P}\left(T^{pCMB}, \delta_g, \delta_e, v_r, \Lambda_S, \Lambda_L | T^{obs}, \delta_g^{obs}\right) \propto \mathcal{L}\left(T^{obs}, \delta_g^{obs} | T^{pCMB}, \delta_g, \delta_e, v_r\right) \times \mathcal{P}(T^{pCMB}, \delta_e, \delta_g, v_r, \Lambda_S, \Lambda_L)$$

This is the **joint posterior**. We explore possibilities to integrate out fields to obtain **marginalized posteriors** and estimate error bars. One way to proceed is to use MUSE [2112.09354]

Conclusions and outlook

- kSZ is a cool new cosmological probe with a lot of applications.
- We found a velocity MAP with forward modeling of kSZ and numerical optimization
- This formalism is flexible and versatile tool with a possibility of inclusion of other effects.
- Novel optimization and computation methods open the window for building more broad forward-modelling based Bayesian analysis approaches