Data-driven Cosmology from 3D Light Cones

Yun-Ting Cheng
(Caltech)

In collaboration with
Benjamin Wandelt (IAP/CCA) **Tzu-Ching Chang** (JPL/Caltech) **Olivier Doré** (JPL/Caltech)

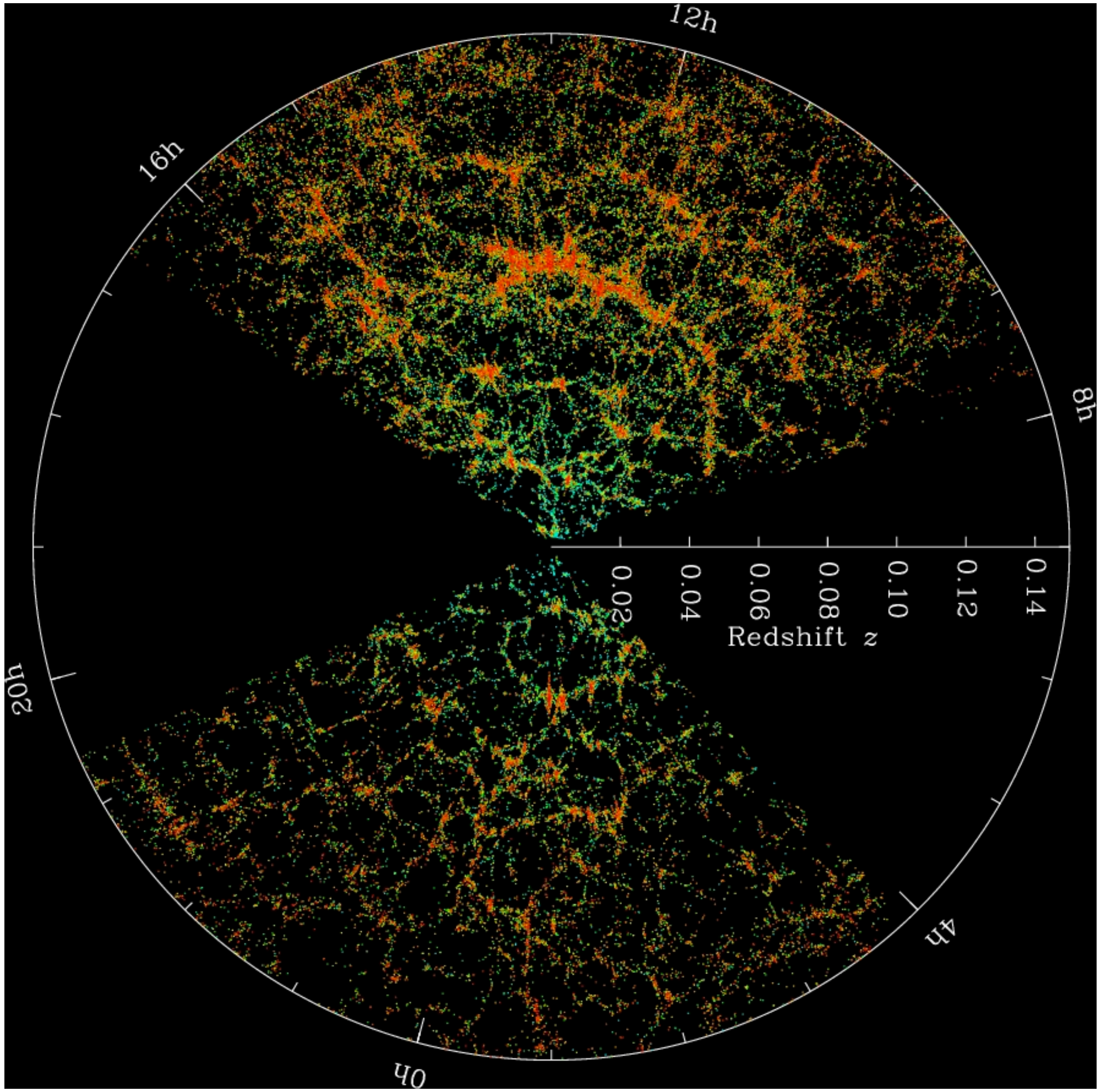
Cosmology from Home 2022

Cosmology from the Large-Scale Structure

- The large-scale structure (LSS) is a key probe of cosmological model.

- galaxy redshift surveys map the 3D LSS distribution (2dF, WiggleZ, SDSS, DES, etc)

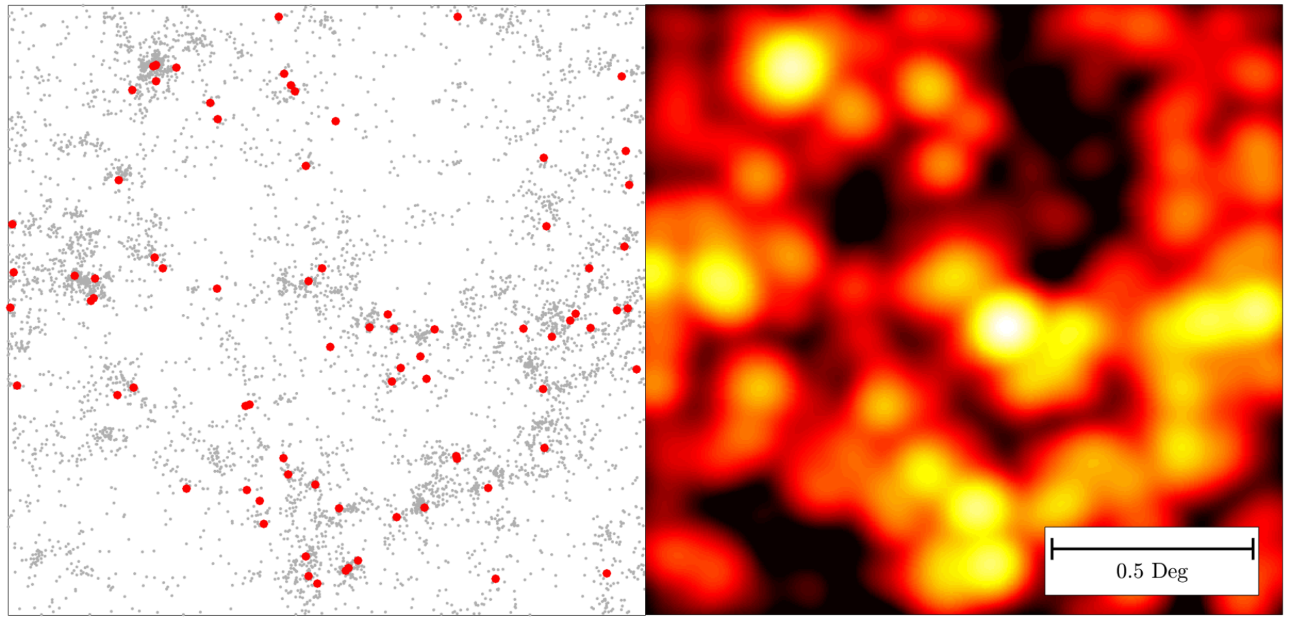
SDSS



Galaxy Redshift Surveys

1. Detect sources in the images
2. Extract their spectra
3. Fit the redshift (photo-z / spec-z)
4. Map the 3D LSS these sources

Credit: Patrick Breysse, Kovetz et al. 2017



- **Requires a SED library for redshift fitting**
- **Information from faint sources is discarded**

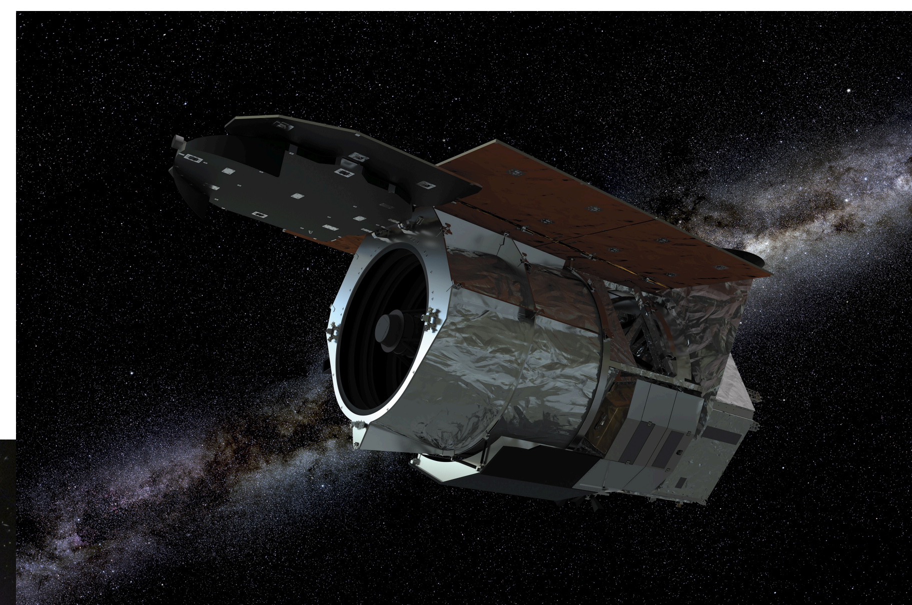
Future Cosmological Surveys

DESI

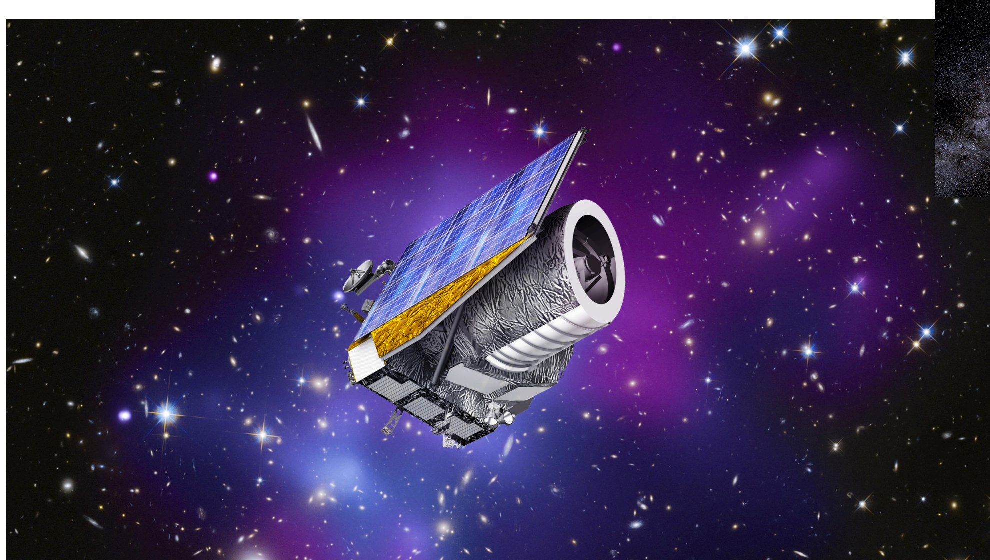
Rubin LSST

Legacy Survey of Space and Time

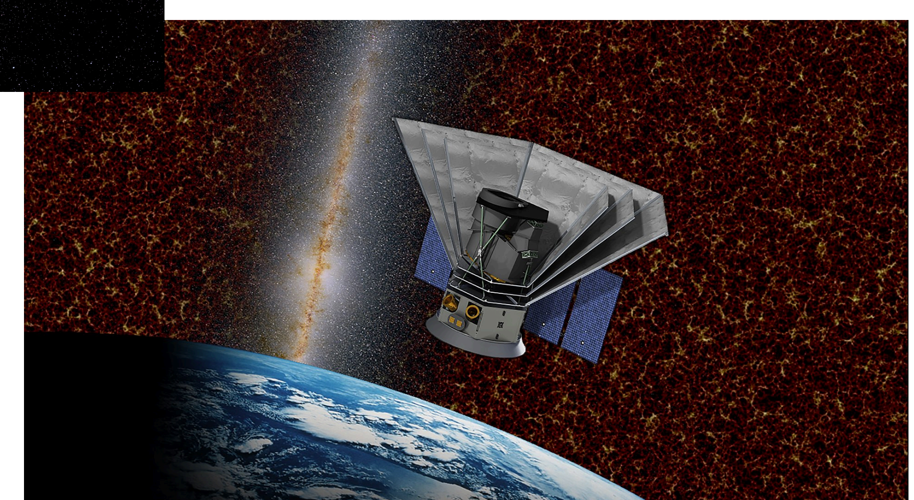
Roman



Euclid



SPHEREx

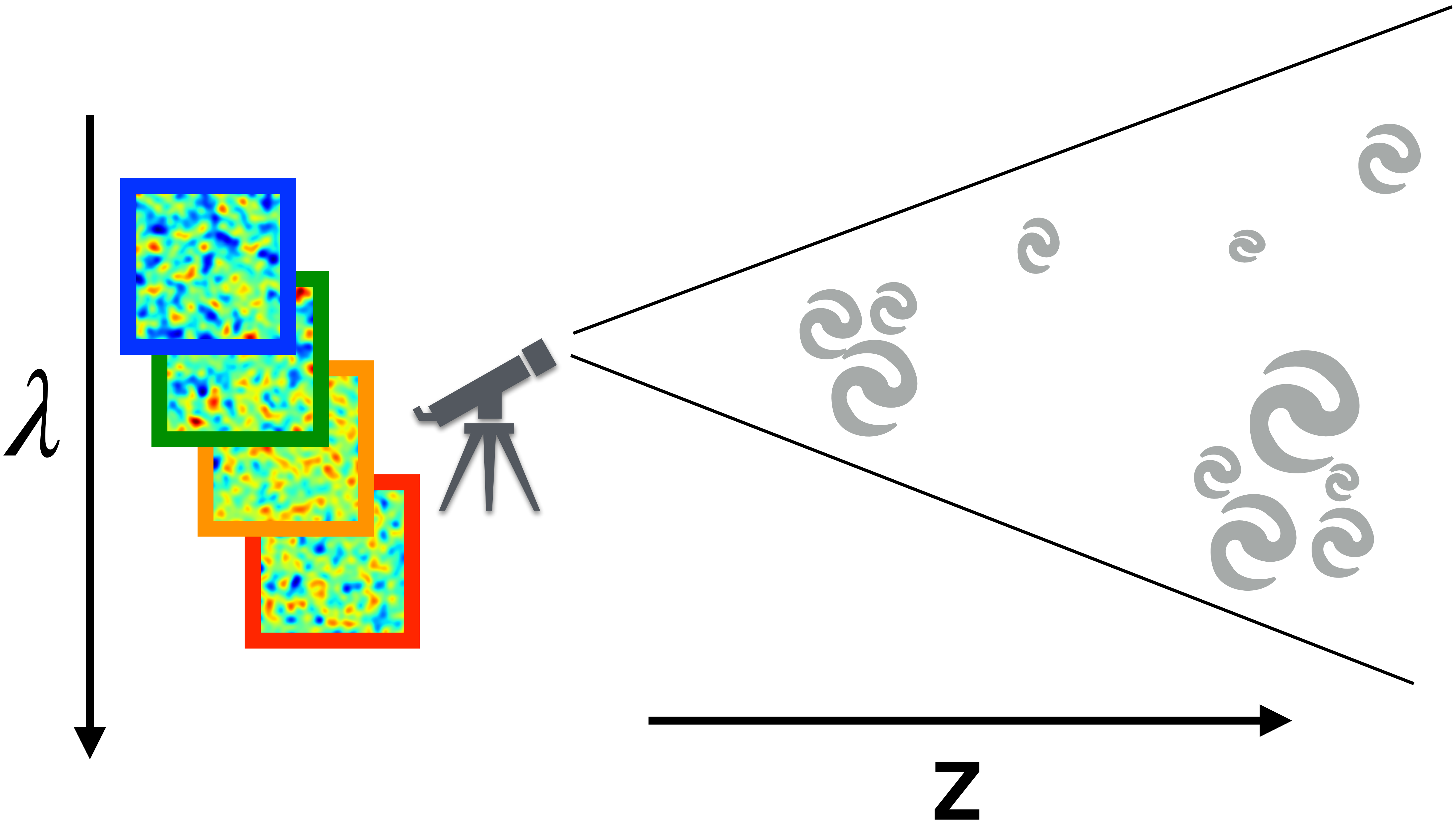


With these multi-frequency intensity maps, how to optimally extract the cosmological information?

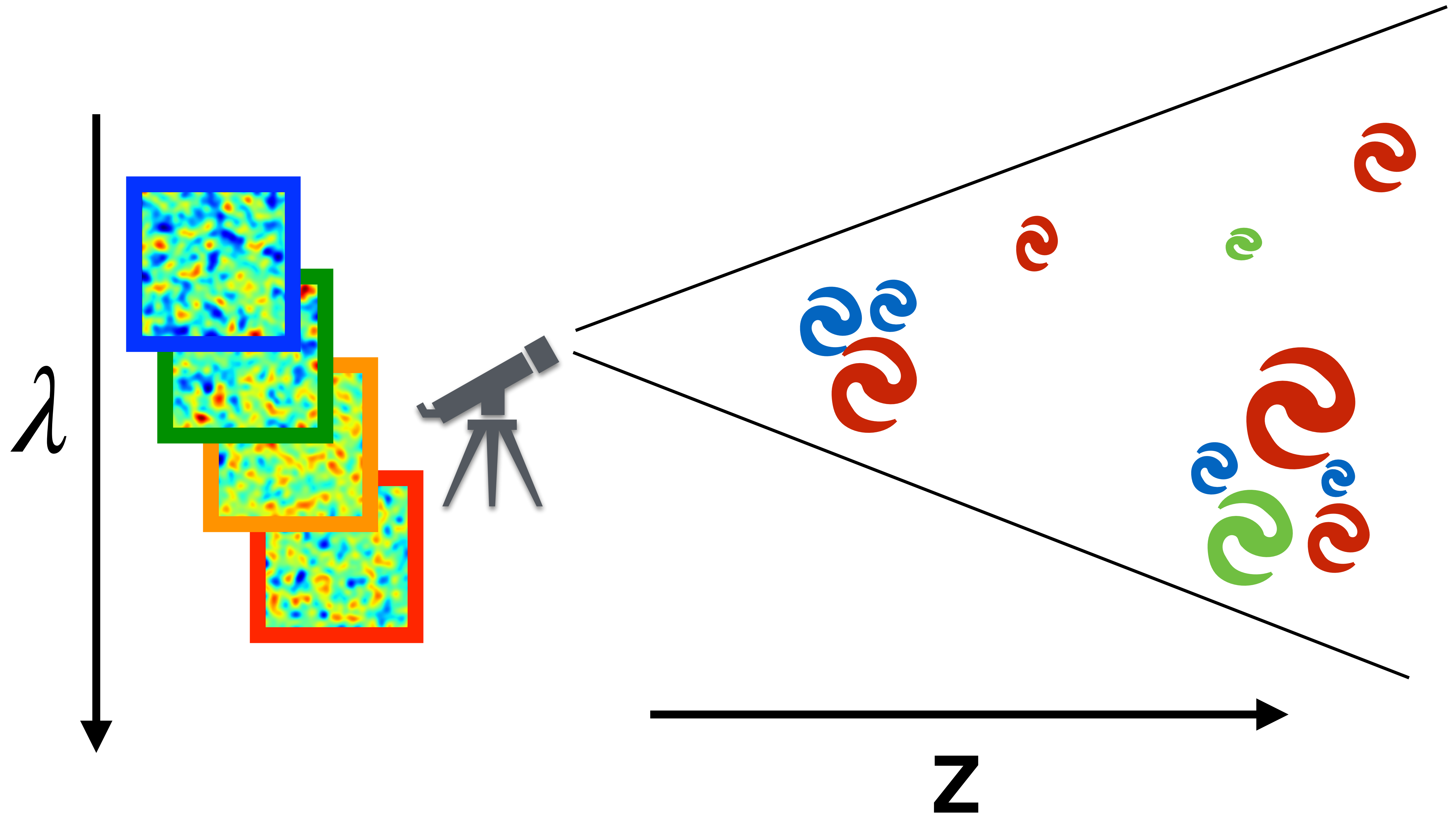
Data-driven Cosmology from 3D Light Cones

- Requires a SED library for redshift fitting
  Infer signals from data
- Information from faint sources is discarded
  Analysis on the intensity maps

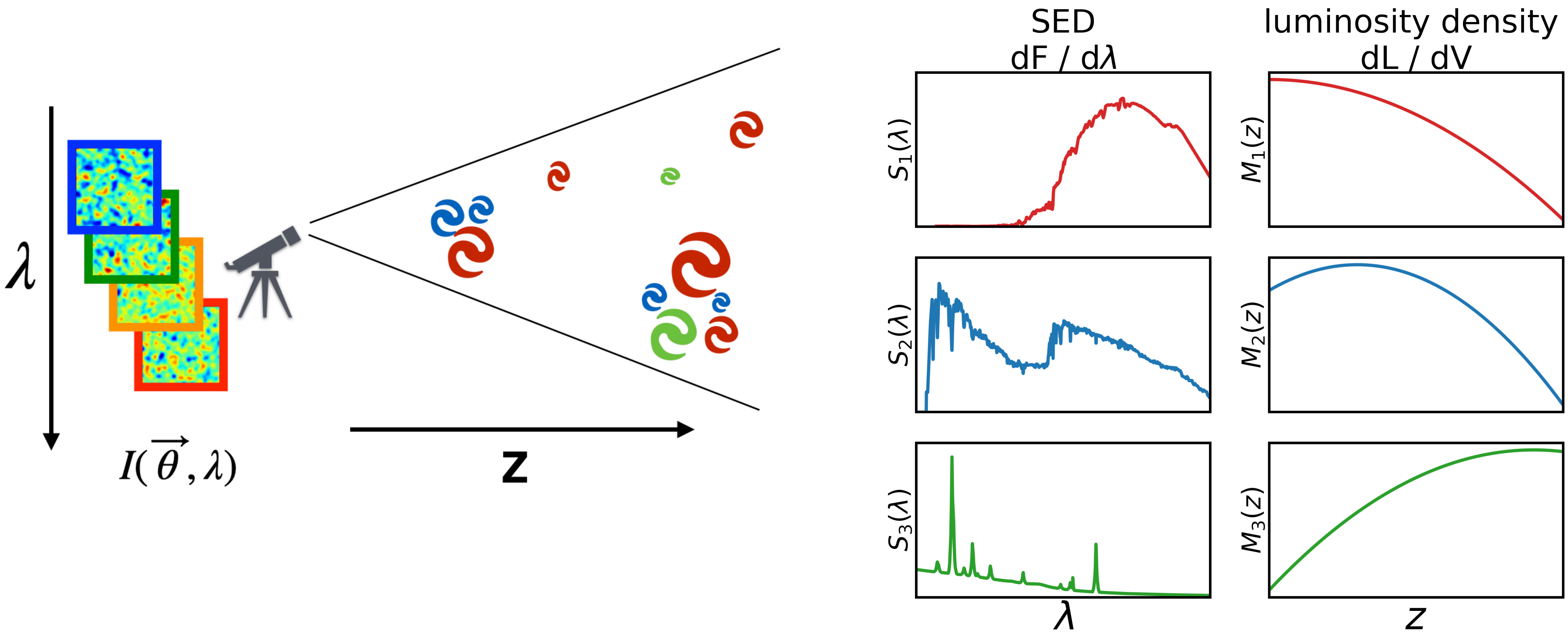
Signal in a Light Cone



Signal in a Light Cone

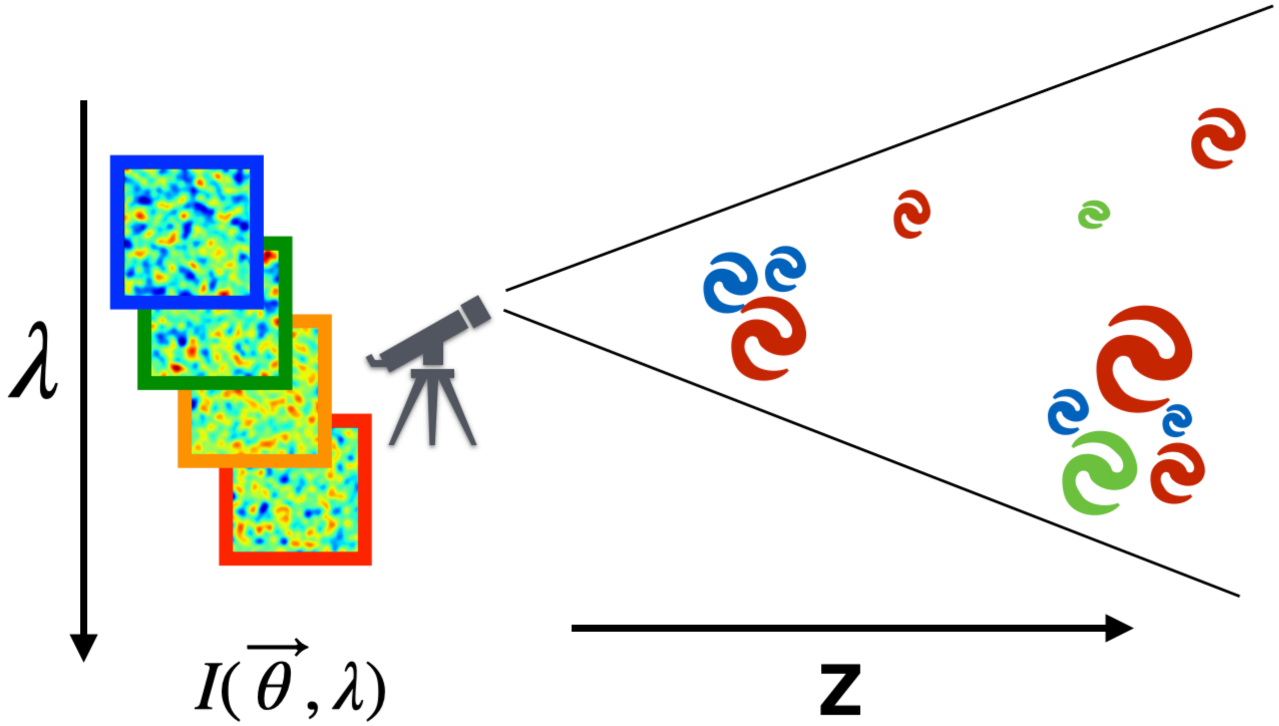


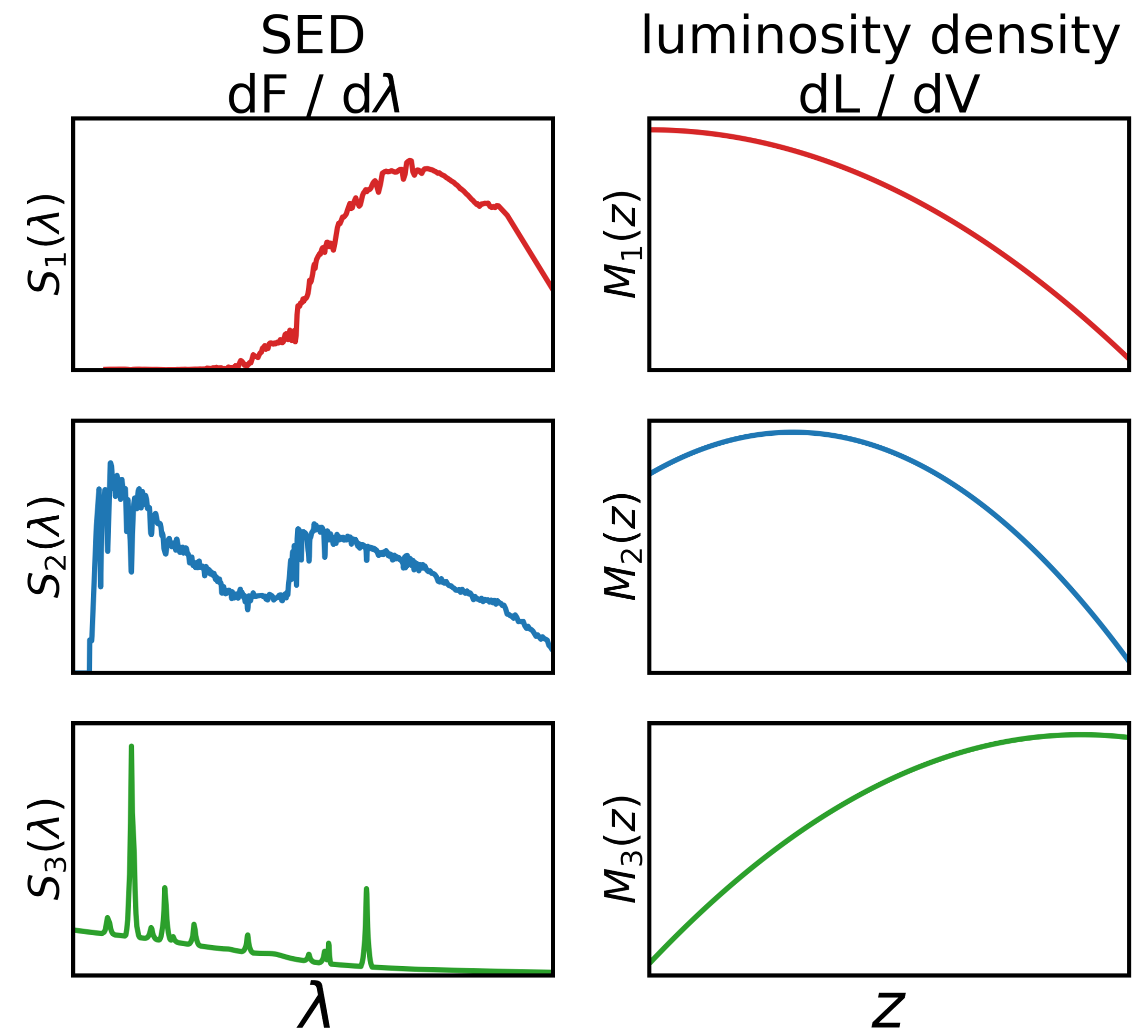
Signal in a Light Cone



- Source traces an underlying density field
- Emitting sources can be described by a few “classes”
- A “class” is described by a SED and a luminosity density
- This could be the PCA modes of SEDs in general

Signal in a Light Cone





$$I(\vec{\theta}, \lambda) = \int dz \, [\sum_{i=1,2,3} S_i(\frac{\lambda}{z+1}) \, M_i(z) \,] \, \delta_m(\vec{\theta}, z)$$

z integration

convolution of SED (S)
and luminosity density (M)

tracing matter
density field

Signal Covariance

$$I(\vec{\theta}, \lambda) = \int dz \, [\sum_{i=1,2,3} S_i(\frac{\lambda}{z+1}) \, M_i(z) \,] \, \delta_m(\vec{\theta}, z)$$

Signal Covariance

$$I(\vec{\theta}, \lambda) = \int dz \, [\sum_{i=1,2,3} S_i(\frac{\lambda}{z+1}) \, M_i(z) \,] \, \delta_m(\vec{\theta}, z)$$

$$< \widetilde{\delta}_m^*(\overrightarrow{k})\widetilde{\delta}_m(\overrightarrow{k}) > \sim P(k)$$

Signal Covariance

$$I(\vec{\theta}, \lambda) = \int dz \, [\sum_{i=1,2,3} S_i(\frac{\lambda}{z+1}) \, M_i(z) \,] \, \delta_m(\vec{\theta}, z)$$

$$\widetilde{I}(\ell, \lambda) = \left[\sum_{i=1,2,3} \mathbf{S}_i \mathbf{M}_i \mathbf{b}_i \right] \widetilde{\delta}_m(k)$$

$$< \widetilde{\delta}^*_m(\overrightarrow{k}) \widetilde{\delta}_m(\overrightarrow{k}) > \sim P(k)$$

large scale (low-k), scale-independent bias b(k)
degenerate with M

Signal Covariance

$$I(\vec{\theta}, \lambda) = \int dz \, [\sum_{i=1,2,3} S_i(\frac{\lambda}{z+1}) \, M_i(z) \,] \, \delta_m(\vec{\theta}, z)$$

$$\widetilde{I}(\ell, \lambda) = \left[\sum_{i=1,2,3} \mathbf{S}_i \mathbf{M}_i \mathbf{b}_i \right] \widetilde{\delta}_m(k)$$

$$< \widetilde{\delta}^*_m(\overrightarrow{k}) \widetilde{\delta}_m(\overrightarrow{k}) > \sim P(k)$$

$$C_\ell(\lambda, \lambda') \sim < \widetilde{I}^*(\ell, \lambda) \widetilde{I}(\ell, \lambda') > = \left[\sum_{i=1,2,3} \mathbf{S}_i \mathbf{M}_i \right] \mathbf{P} \left[\sum_{i'=1,2,3} \mathbf{S}_{i'} \mathbf{M}_{i'} \right] + C_{\ell,\mathrm{noise}}$$

data covariance

signal **power spectrum** **signal** **noise (Instrument, foregrounds)**

Signal Covariance

$$I(\vec{\theta},\lambda) = \int dz \, [\sum_{i=1,2,3} S_i(\frac{\lambda}{z+1}) \, M_i(z) \,] \, \delta_m(\vec{\theta}, z)$$

$$\widetilde{I}(\ell,\lambda) = \left[\sum_{i=1,2,3} \mathbf{S}_i \mathbf{M}_i \mathbf{b}_i\right] \widetilde{\delta}_m(k)$$

$$< \widetilde{\delta}^*_m(\vec{k})\widetilde{\delta}_m(\vec{k}) > \sim P(k)$$

$$C_\ell(\lambda,\lambda') \sim < \widetilde{I}^*(\ell,\lambda)\widetilde{I}(\ell,\lambda') > = \left[\sum_{i=1,2,3} \mathbf{S}_i \mathbf{M}_i\right] \mathbf{P} \left[\sum_{i'=1,2,3} \mathbf{S}_{i'} \mathbf{M}_{i'}\right] + C_{\ell,\mathrm{noise}}$$

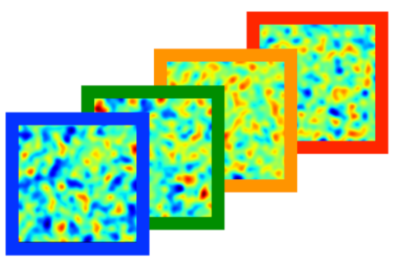
data covariance — **signal** — **power spectrum** — **signal** — **noise (Instrument, foregrounds)**

Lossless representation of the data on large scales

power spectrum encodes the full information in a homogenous and isotropic Gaussian field

Analysis Procedure

$$C_\ell(\lambda,\lambda') \sim < \widetilde{I}^*(\ell,\lambda)\widetilde{I}(\ell,\lambda') > = \left[\sum_{i=1,2,3} \mathbf{S}_i\mathbf{M}_i\right] \mathbf{P} \left[\sum_{i'=1,2,3}\mathbf{S}_{i'}\mathbf{M}_{i'}\right] + C_{\ell,\mathrm{noise}}$$



$I^{\mathrm{data}}(\vec{\theta},\lambda)$ ➔ $C_\ell^{\mathrm{data}}(\lambda,\lambda')$ ➔ **Inference** ➔

- source SED (**S**)
- source density (**M**)
- LSS clustering (**P**)
- noise

*** overall scaling of S, M, P is degenerate

Analysis Procedure

$C_\ell^{\text{data}}(\lambda, \lambda')$ → **Inference** →

- source SED (**S**)
- source density (**M**)
- LSS clustering (**P**)
- noise

astrophysics

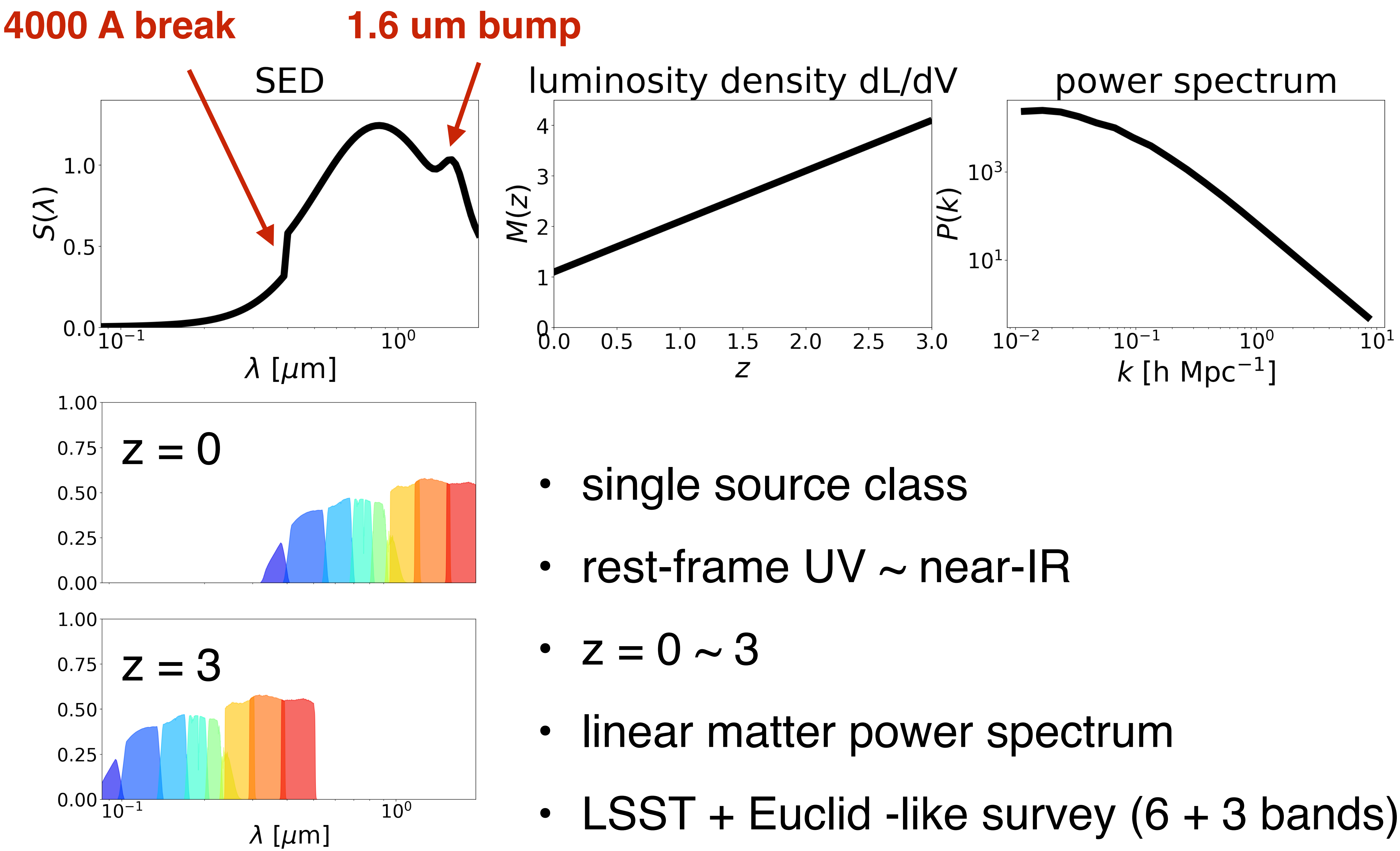
cosmology

nuisance

- simultaneously solve for signal and noise from the data covariance
- no assumptions required for signal and noise
- blindly infer information from the covariance multi-band images

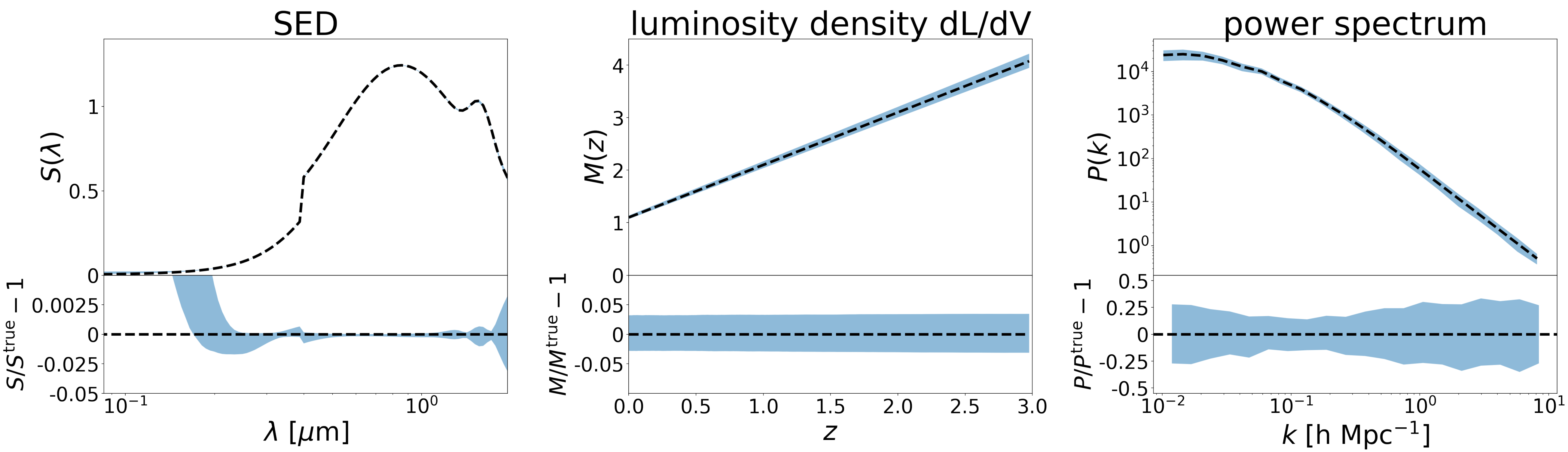
—> **no need for source detection and photo-z / spec-z**

Example on a simulated observation

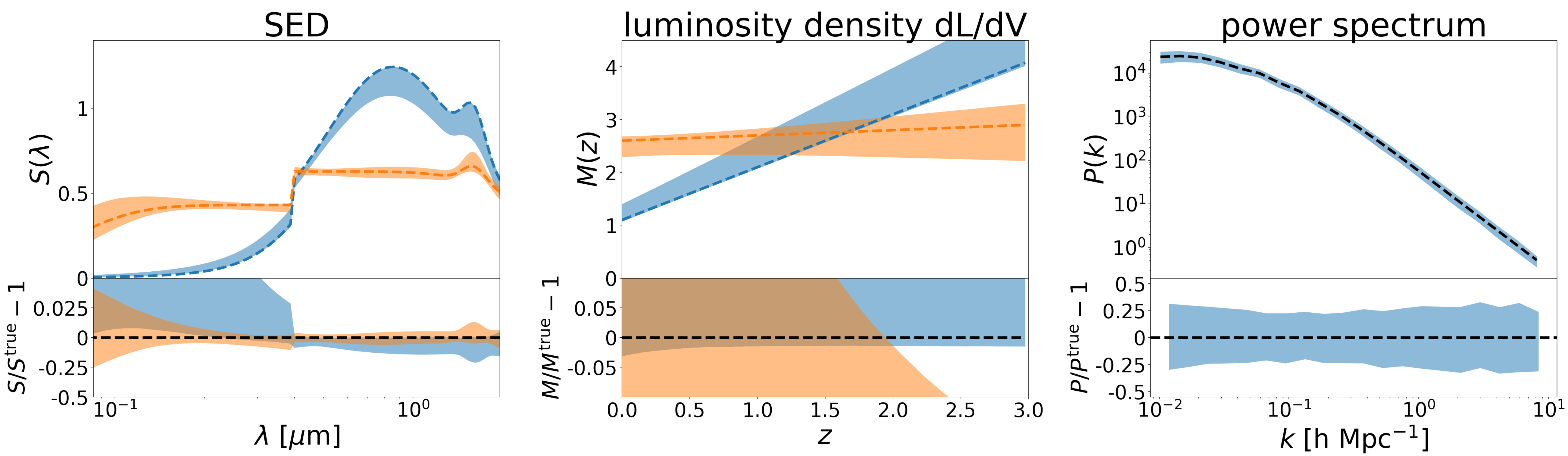


- single source class
- rest-frame UV ~ near-IR
- z = 0 ~ 3
- linear matter power spectrum
- LSST + Euclid -like survey (6 + 3 bands)

Preliminary Results



Preliminary Results - two classes



Summary

- We develop a data-driven method to analyze large-scale multi-frequency intensity map
- Our method simultaneously extracts source signal, LSS, and noise from the data covariance
- Our method does not require any prior assumption or source detection
- This can be applied future cosmological surveys

Y.-T. Cheng, B. D. Wandelt, T.-C. Chang, O. Doré 2022 in prep.