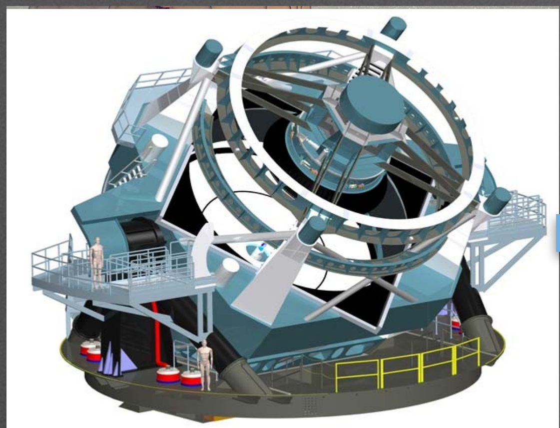
Covariances and Applications

Tim Eifler, University of Arizona

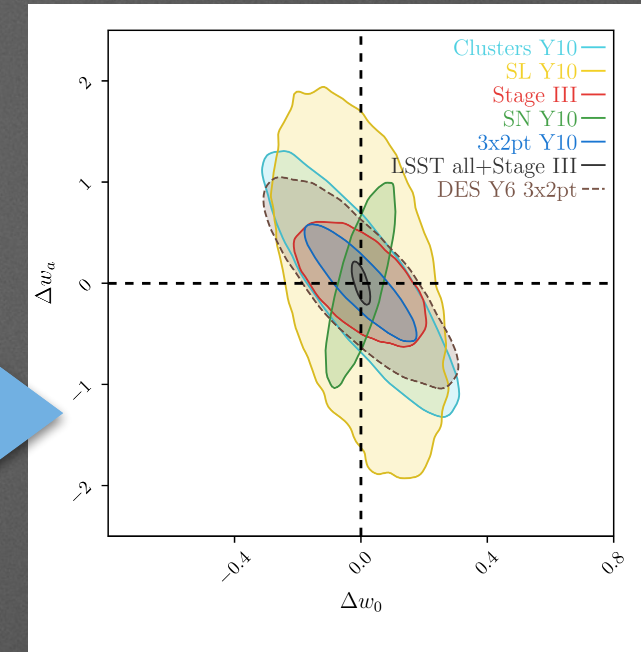
Accurate Lensing in the Era of Precision Cosmology
Berkeley 2019

The Challenge



reduced data and catalogs → summary statistics →

→



Independent probes
e.g., SN1a as priors

Large model vector (CMB+LSS)
Self-consistent modeling of all observables as a function of
1) cosmological parameters (~10)
2) nuisance parameters (XXX)

large data vector

posterior probability

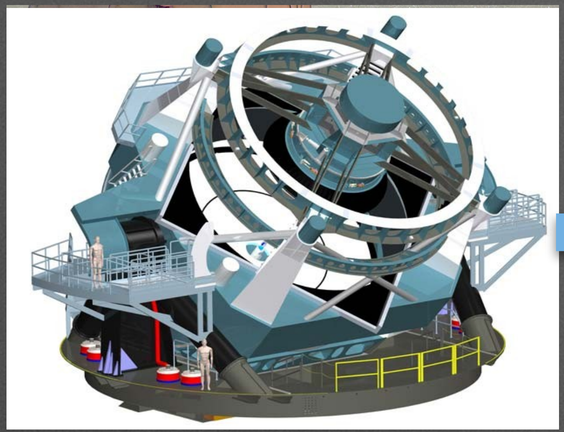
$$p\left(\boldsymbol{\pi}|\hat{\mathbf{d}}\right) \propto p\left(\boldsymbol{\pi}\right)\mathcal{L}\left(\hat{\mathbf{d}}|\mathbf{m}\left(\boldsymbol{\pi}\right), \boldsymbol{C}\right)$$

Enhanced modeling via
- Observations
- Simulations
- Theory

Statistics I - Likelihood function
- Multivariate Gaussian vs other parameterizations
- Non-parametric forms
- Approximate Bayesian Computation

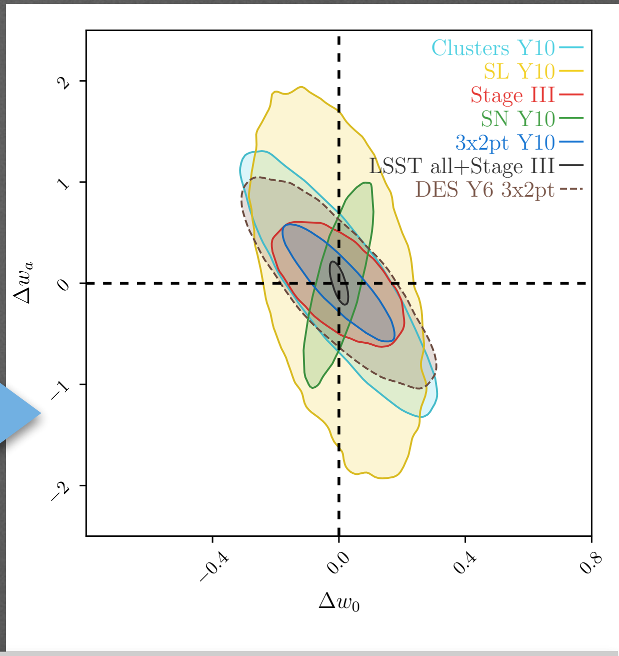
Statistics II - Covariances
- cosmology dependent Signal + constant Noise
- large and complicated, non-(block) diagonal
- different methods for derivation

The Challenge



reduced data and catalogs

summary statistics



Independent probes
e.g., SN1a as priors

Large model vector (CMB+LSS)
Self-consistent modeling of all observables as a function of
1) cosmological parameters (~10)
2) nuisance parameters (XXX)

large data vector

$$p\left(\pi|\hat{\mathrm{d}}\right) \propto p\left(\pi\right)\mathcal{L}\left(\hat{\mathrm{d}}|\mathbf{m}\left(\pi\right), C\right)$$

posterior probability

Enhanced modeling via
- Observations
- Simulations
- Theory

Statistics I - Likelihood function
- Multivariate Gaussian vs other parameterizations
- Non-parametric forms
- Approximate Bayesian Computation

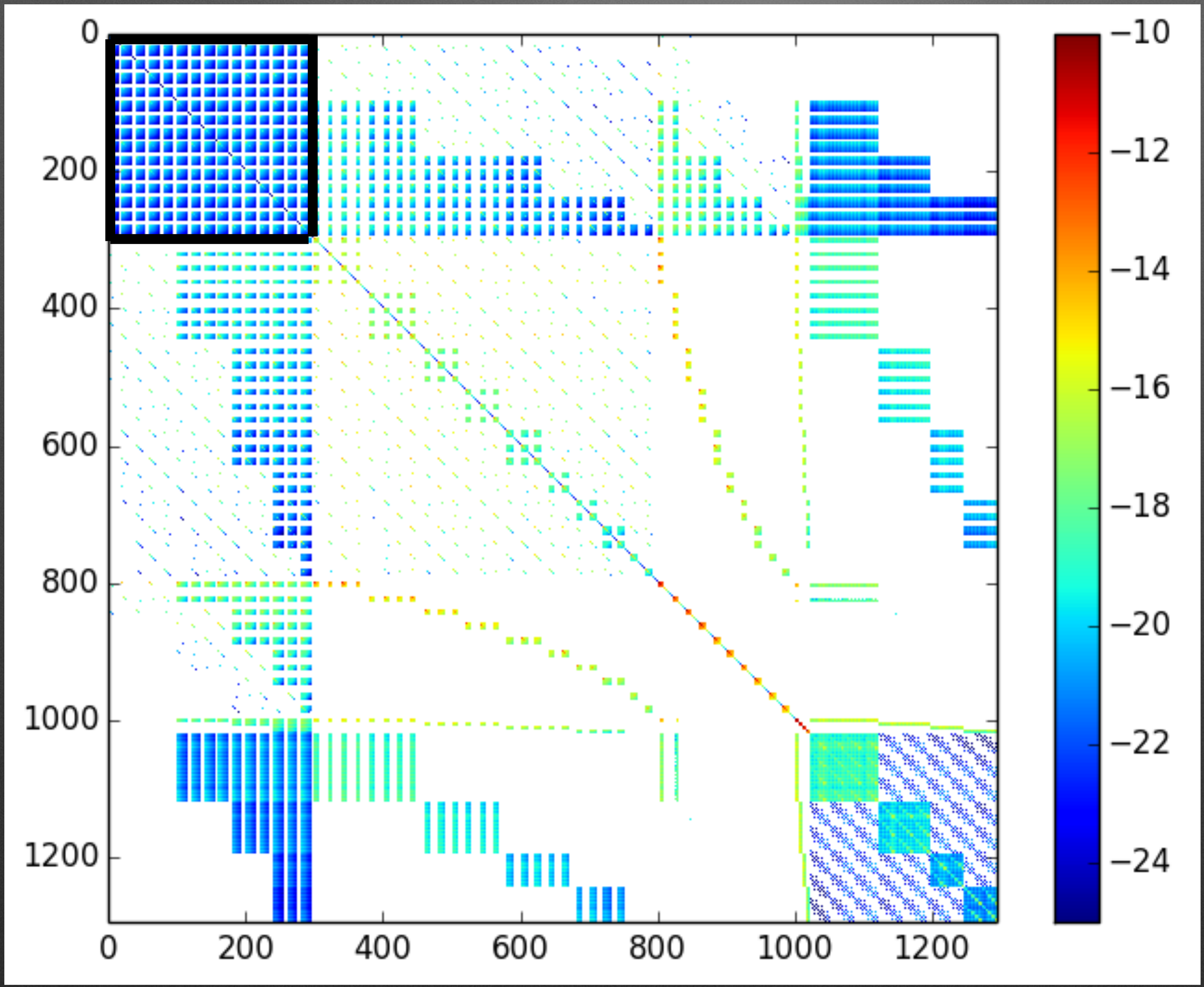
Statistics II - Covariances
- cosmology dependent Signal + constant Noise
- large and complicated, non-(block) diagonal
- different methods for derivation

What's the problem with Covariances…?

- Analytic: flexible (scales redshifts), noise free, very stable even in multi-probe (Gaussian), straightfwd extension, "fast" to compute, unclear if sufficiently precise for LSST Y10
- Numerical sims: Potentially more precise than analytic, computationally expensive (prohibitively), somewhat inflexible (scales redshifts), hard to extend to other probes, noisy inversion -> expensive
- Data: Computationally inexpensive, "have all the effects included automatically", not enough realizations, variance of smaller patches don't resemble the full survey, noisy inversion

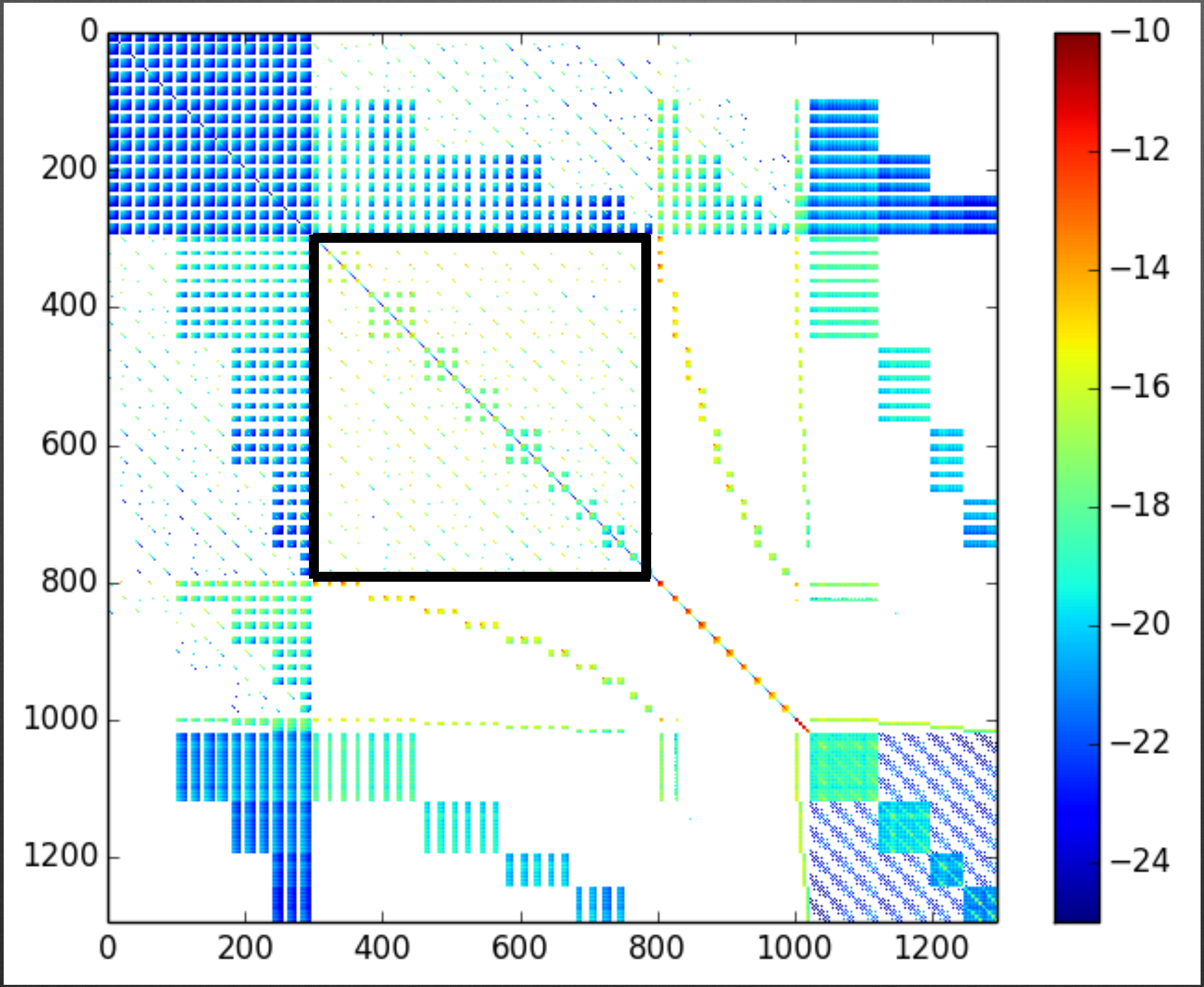
Covariances for LSST-DESC SRD

Cosmic Shear



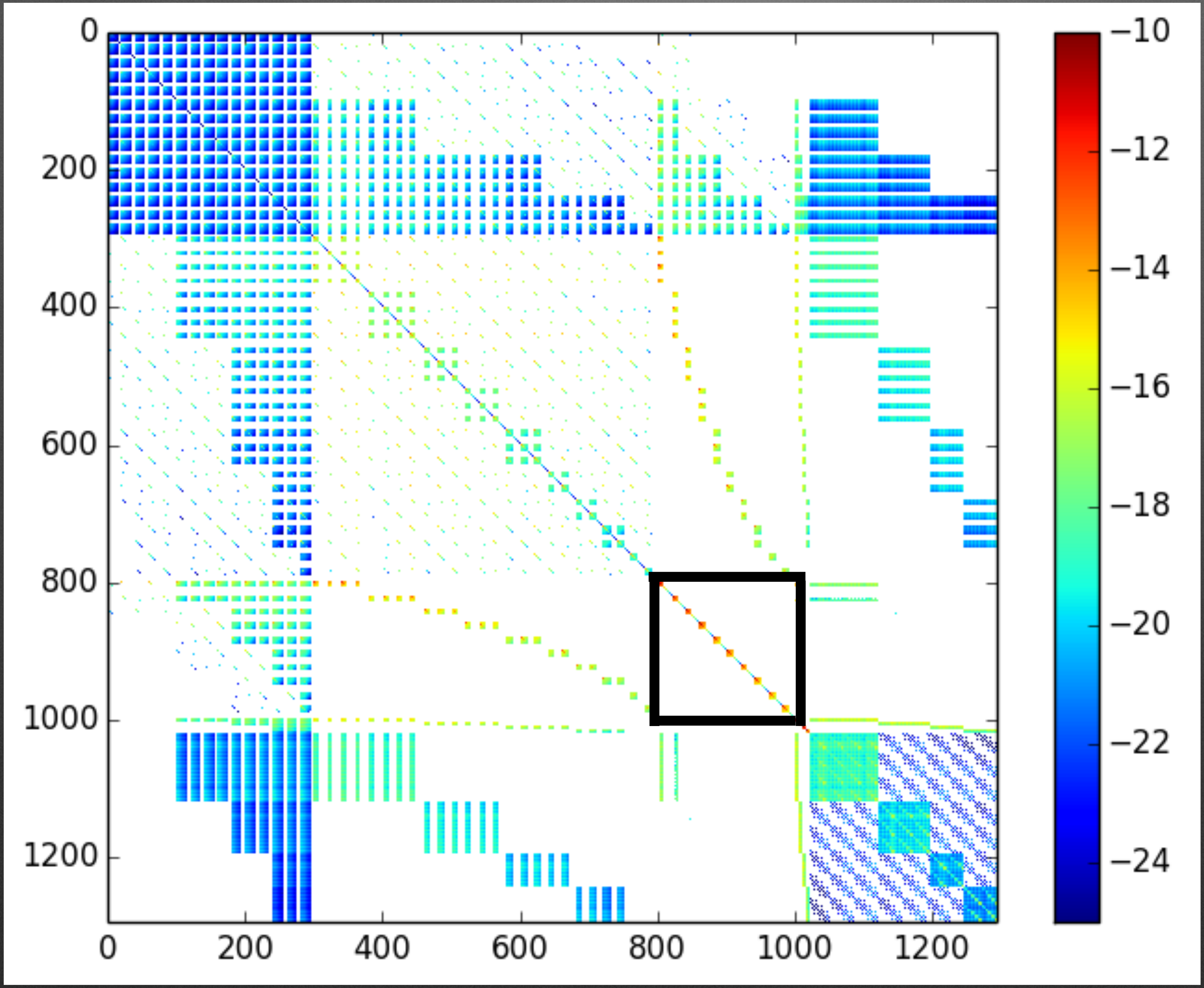
Covariances for LSST-DESC SRD

Galaxy-galaxy lensing



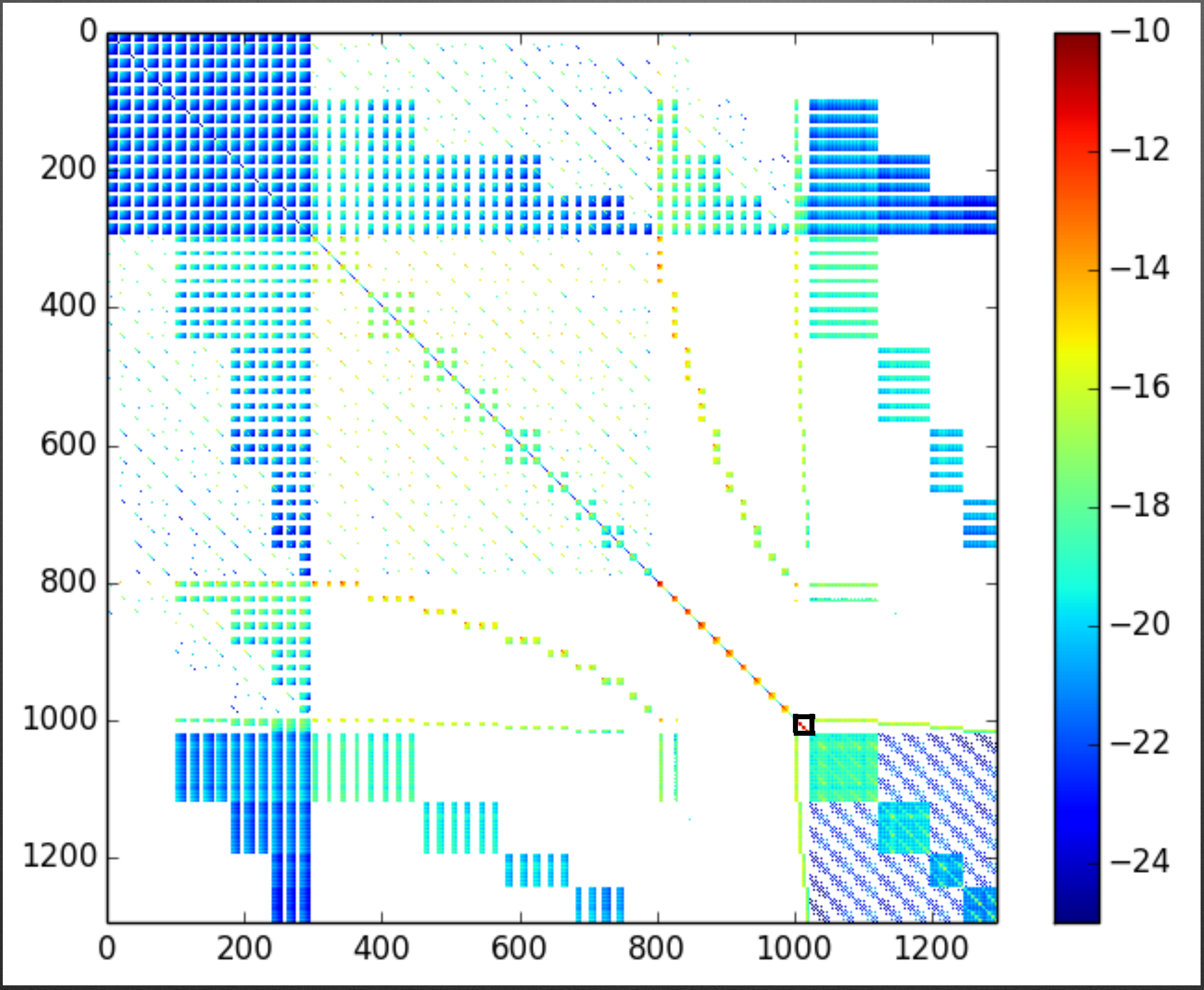
Covariances for LSST-DESC SRD

Galaxy Clustering



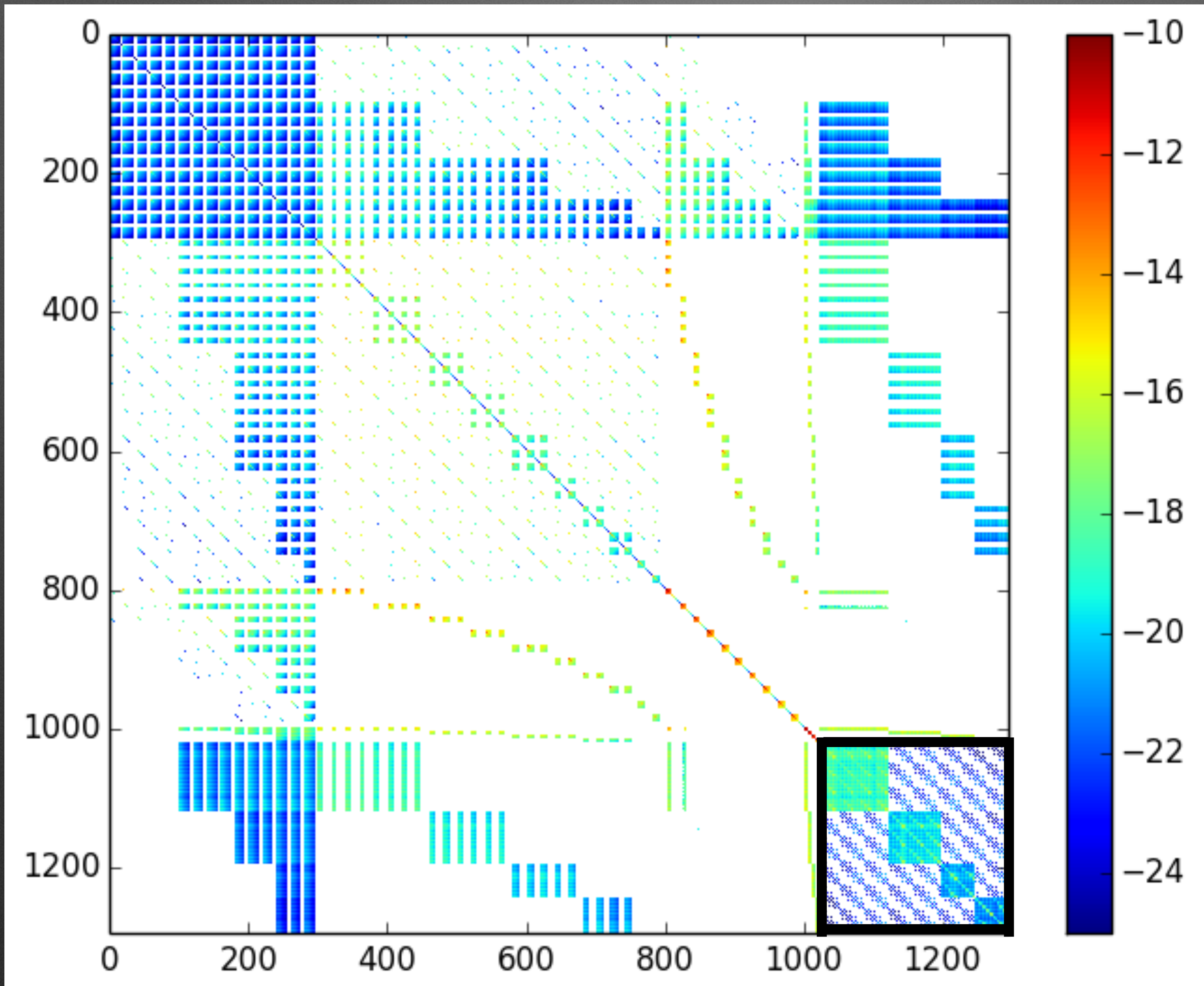
Covariances for LSST-DESC SRD

Cluster Number Counts



Covariances for LSST-DESC SRD

Cluster Weak Lensing



The LSST Dark Energy Science Collaboration (DESC) Science Requirements Document

The LSST Dark Energy Science Collaboration, Rachel Mandelbaum, Tim Eifler, Renée Hložek, Thomas Collett, Eric Gawiser, Daniel Scolnic, David Alonso, Humna Awan, Rahul Biswas, Jonathan Blazek, Patricia Burchat, Nora Elisa Chisari, Ian Dell'Antonio, Seth Digel, Josh Frieman, Daniel A. Goldstein, Isobel Hook, Željko Ivezić, Steven M. Kahn, Sowmya Kamath, David Kirkby, Thomas Kitching, Elisabeth Krause, Pierre-François Leget, Philip J. Marshall, Joshua Meyers, Hironao Miyatake, Jeffrey A. Newman, Robert Nichol, Eli Rykoff, F. Javier Sanchez, Anže Slosar, Mark Sullivan, M. A. Troxel

(Submitted on 5 Sep 2018)

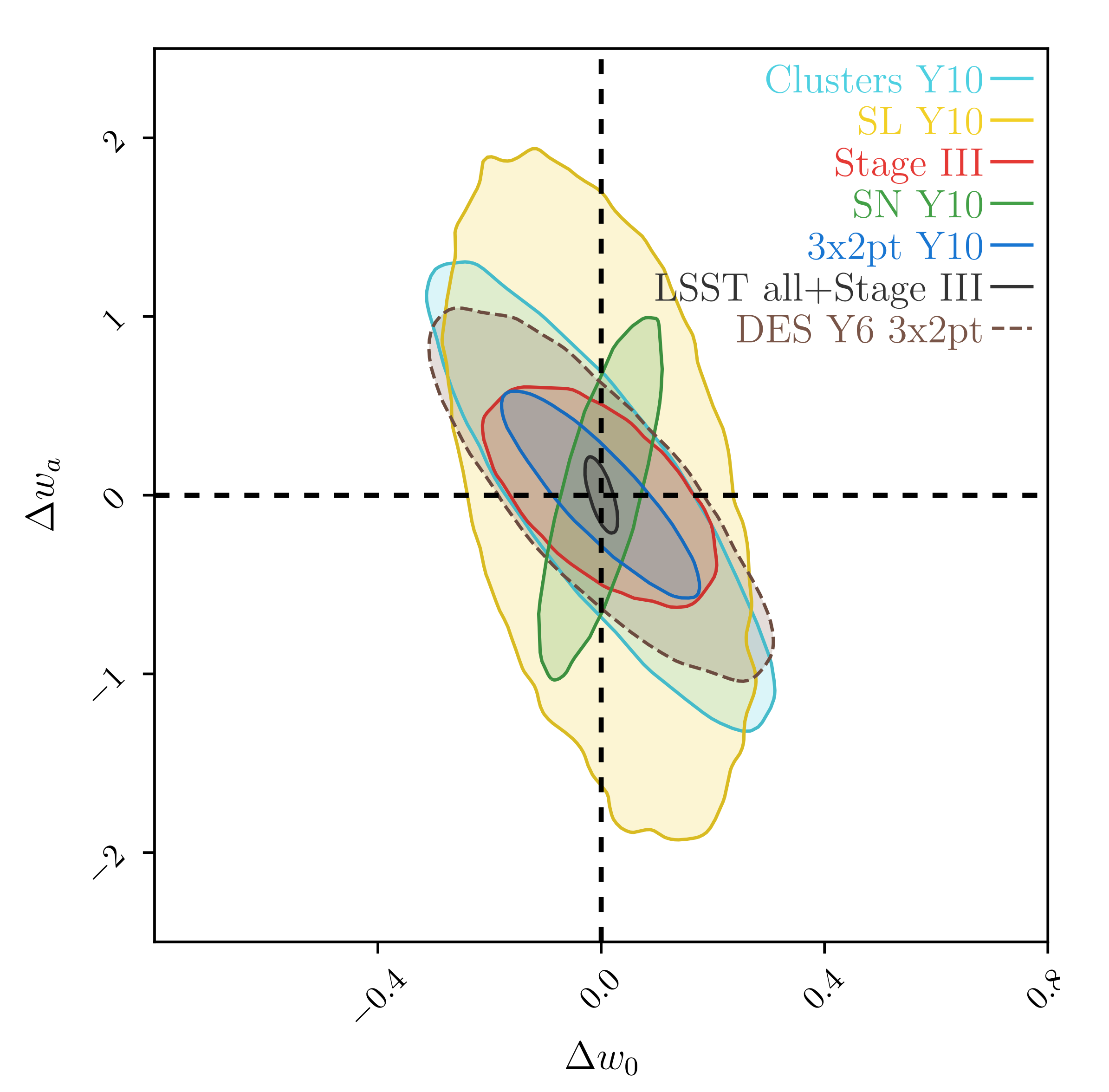
The Large Synoptic Survey Telescope (LSST) Dark Energy Science Collaboration (DESC) will use five cosmological probes: galaxy clusters, large scale structure, supernovae, strong lensing, and weak lensing. This Science Requirements Document (SRD) quantifies the expected dark energy constraining power of these probes individually and together, with conservative assumptions about analysis methodology and follow-up observational resources based on our current understanding and the expected evolution within the field in the coming years. We then define requirements on analysis pipelines that will enable us to achieve our goal of carrying out a dark energy analysis consistent with the Dark Energy Task Force definition of a Stage IV dark energy experiment. This is achieved through a forecasting process that incorporates the flowdown to detailed requirements on multiple sources of systematic uncertainty. Future versions of this document will include evolution in our software capabilities and analysis plans along with updates to the LSST survey strategy.

Comments: 32 pages + 60 pages of appendices. This is v1 of the DESC SRD, an internal collaboration document that is being made public and is not planned for submission to a journal. Data products for reproducing key plots are available at the LSST DESC Zenodo community, this https URL; see "Executive Summary and User Guide" for instructions on how to use and cite those products

Subjects: **Cosmology and Nongalactic Astrophysics (astro-ph.CO)**

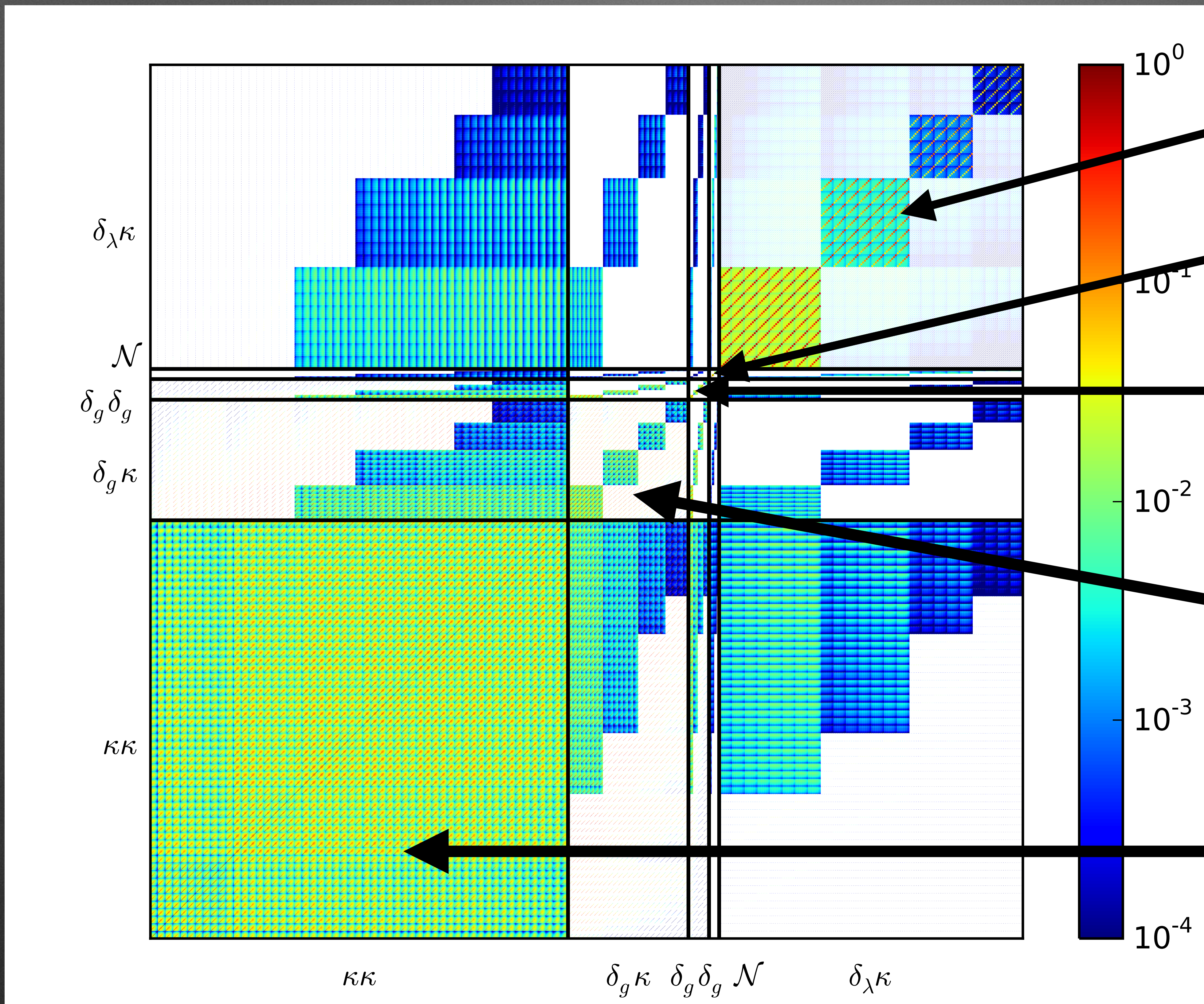
Cite as: arXiv:1809.01669 [astro-ph.CO]

(or arXiv:1809.01669v1 [astro-ph.CO] for this version)



- Some aspects to be improved:
- linear bias
- Baryon mitigation via scale cuts
- Gaussian photo-z

Also had some other covs for LSST (many)…
This was Take 1



Cluster Lensing

Clusters

Galaxy Clustering

Galaxy-Galaxy Lensing

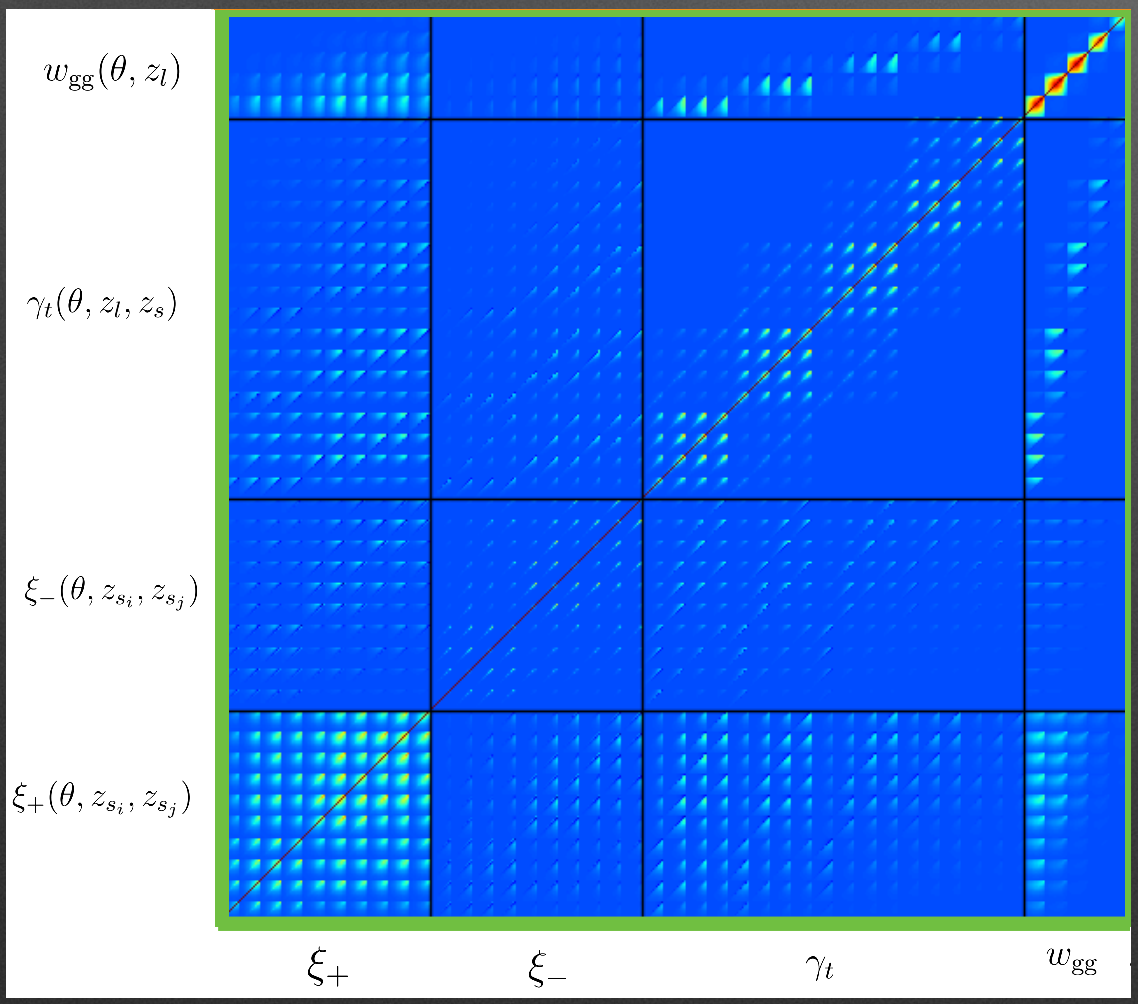
Cosmic Shear

7+ million elements

details: Krause&TE '17

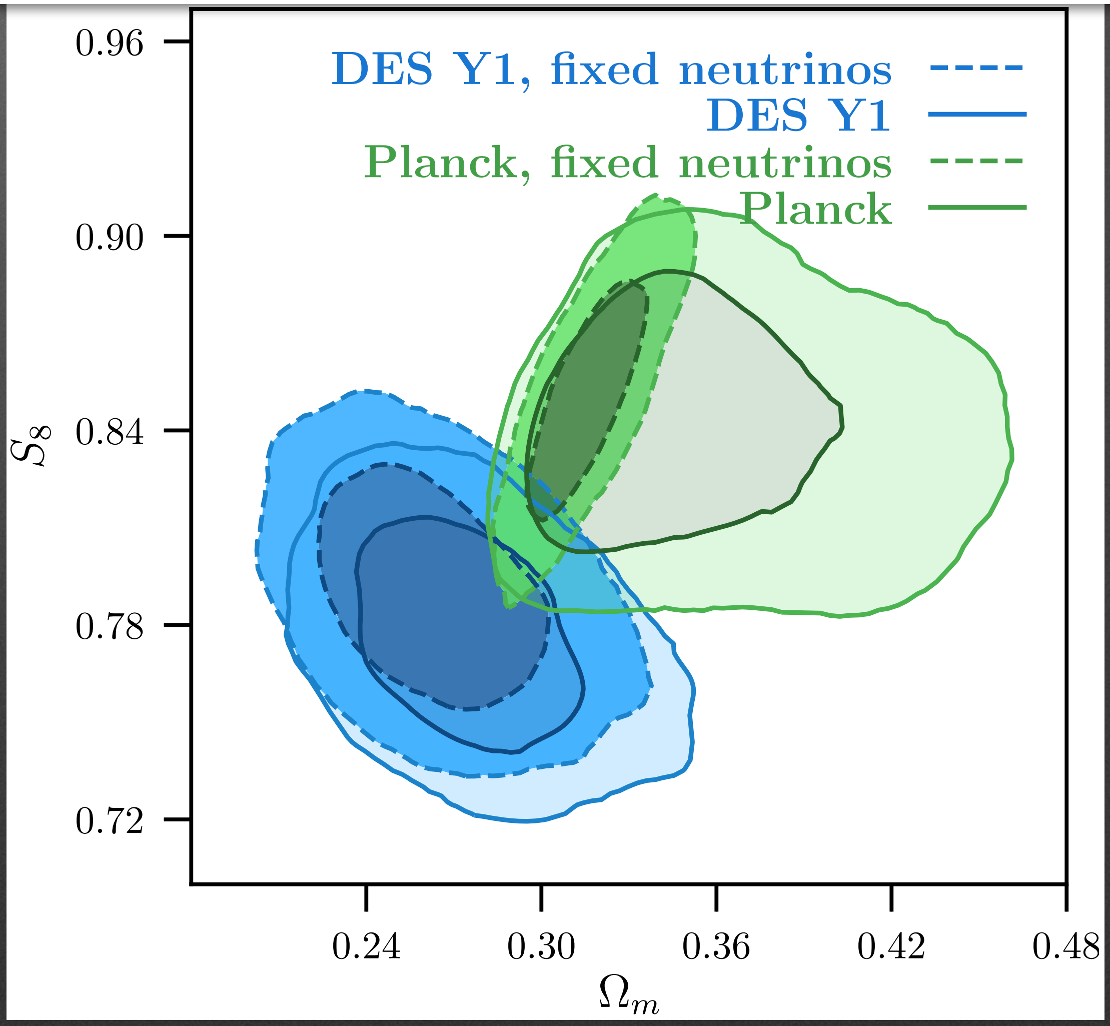
DES Covariances... Real Space

3x2pt



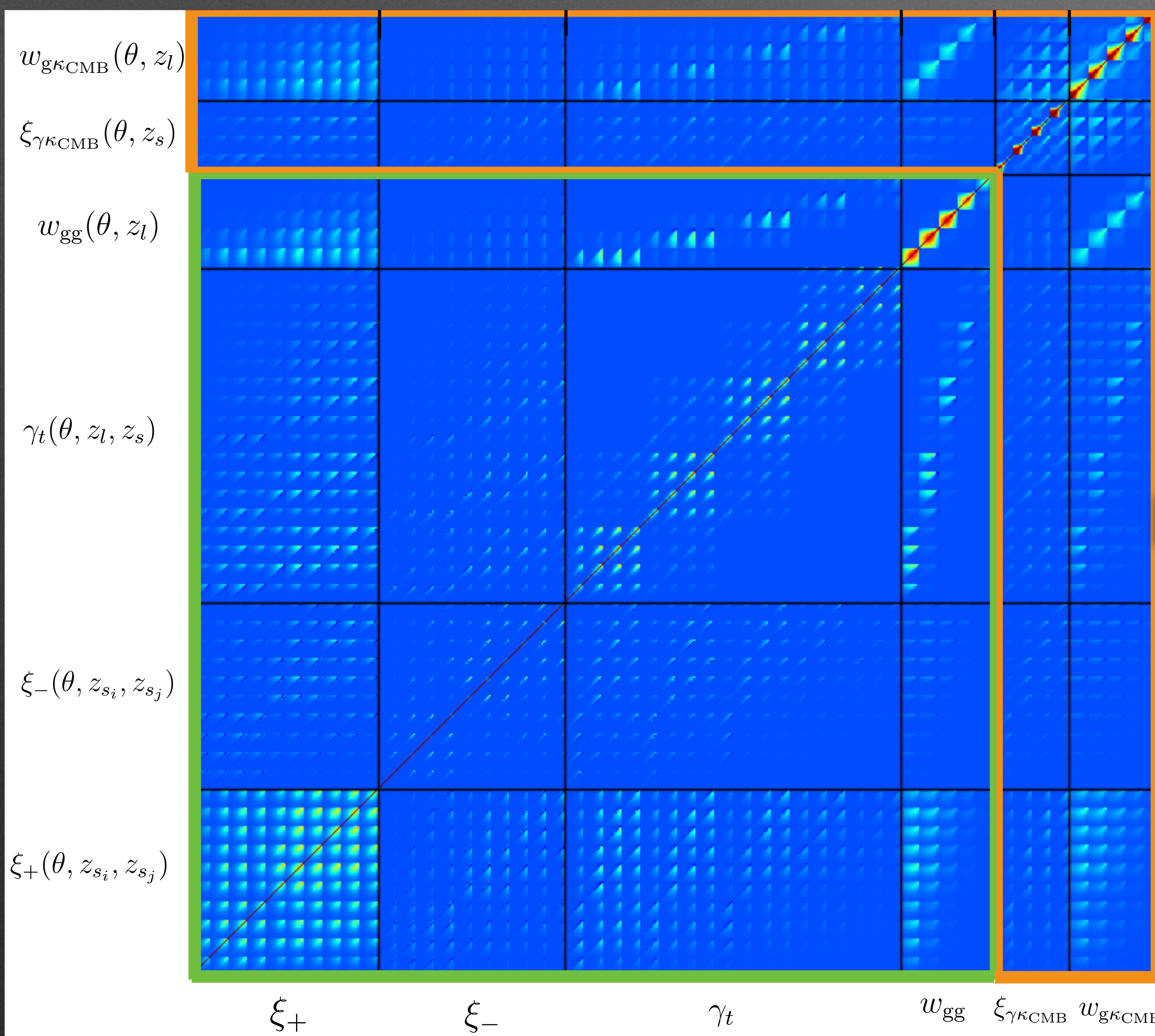
Abbott et al 2017

Dark Energy Survey Year 1 Results:
Cosmological Constraints from Galaxy Clustering and Weak Lensing



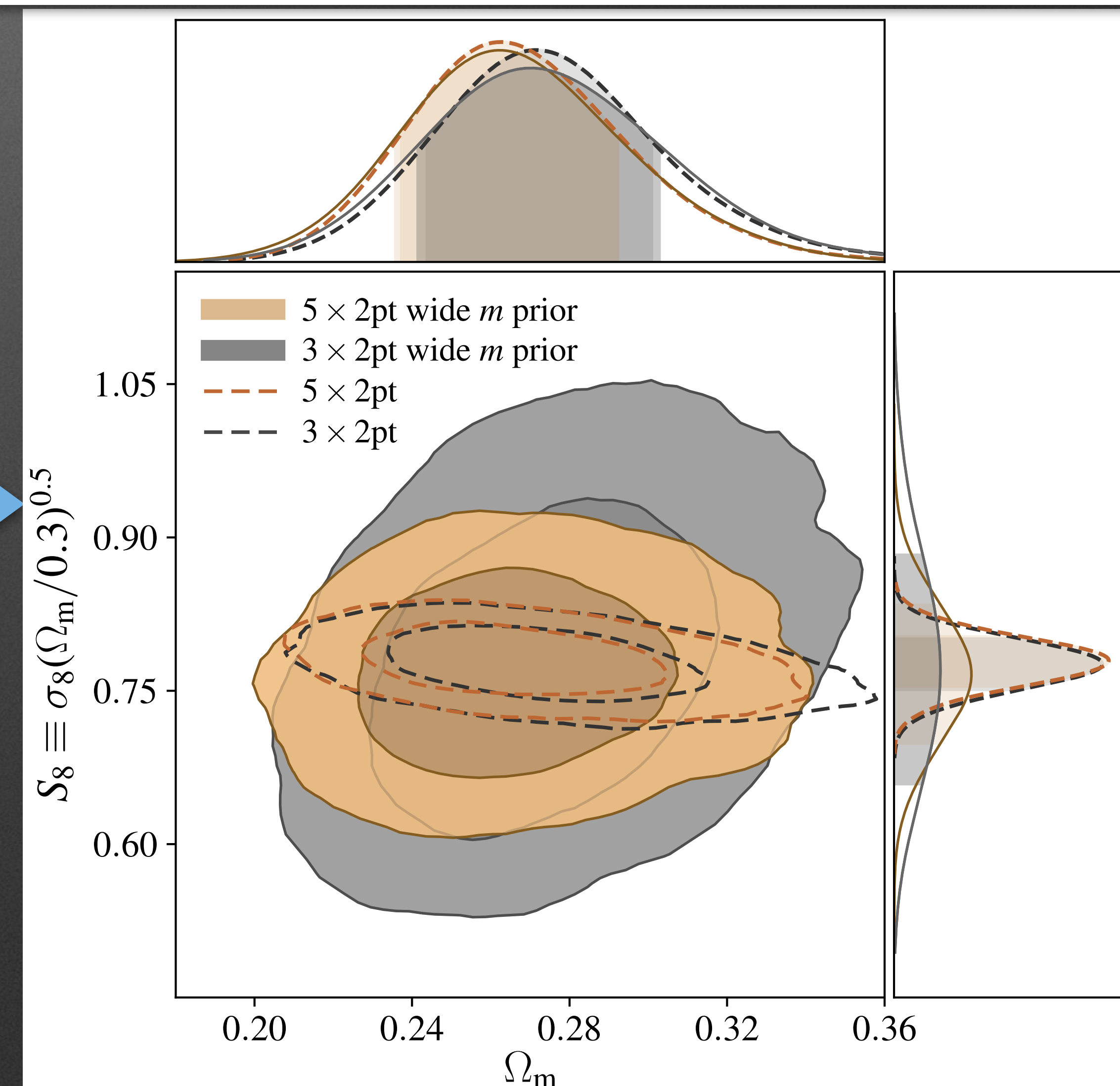
DES Covariances… Real Space

5x2pt



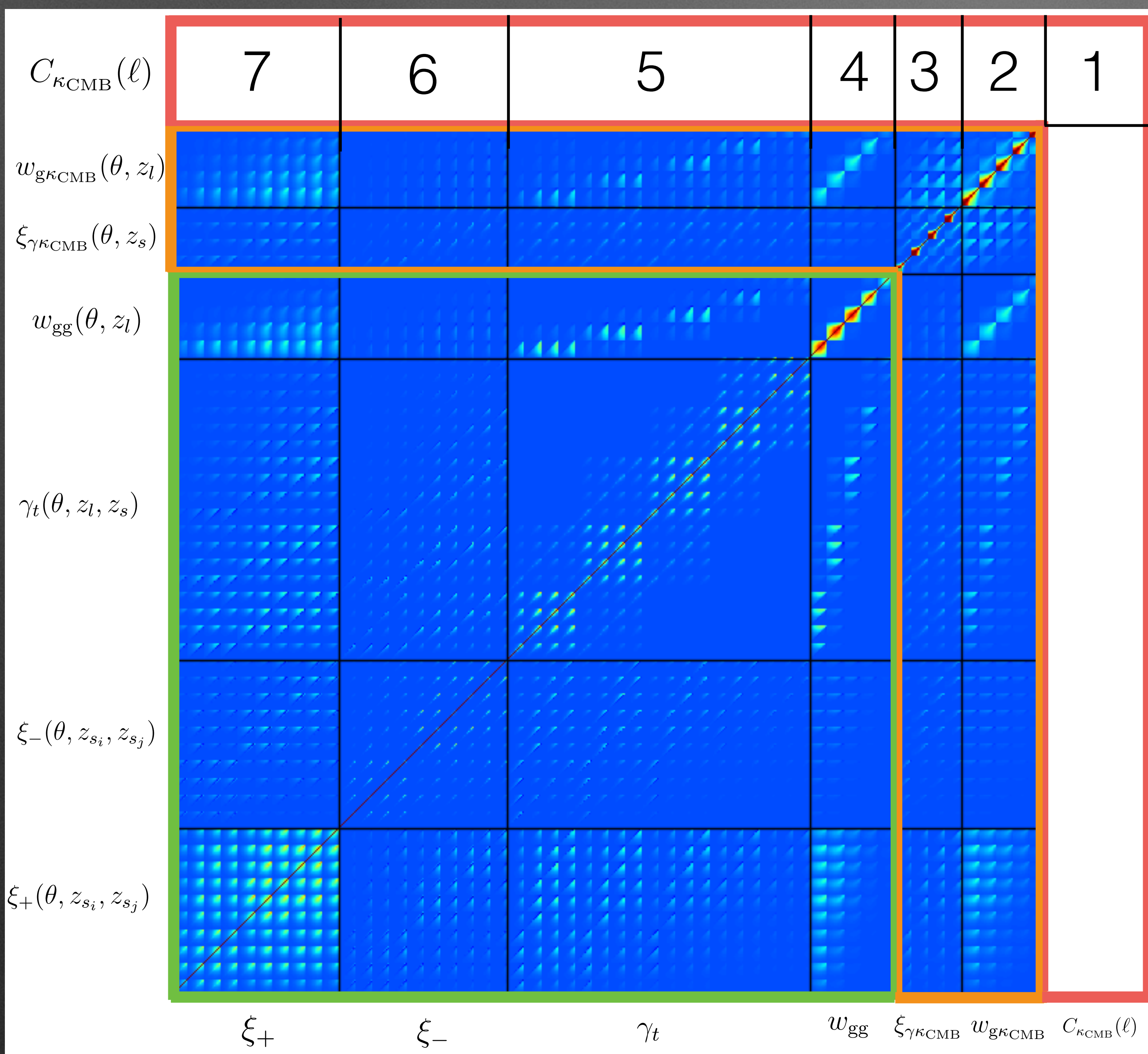
Abbott et al 2018

Dark Energy Survey Year 1 Results: Joint Analysis of Galaxy Clustering, Galaxy Lensing, and CMB Lensing Two-point Functions



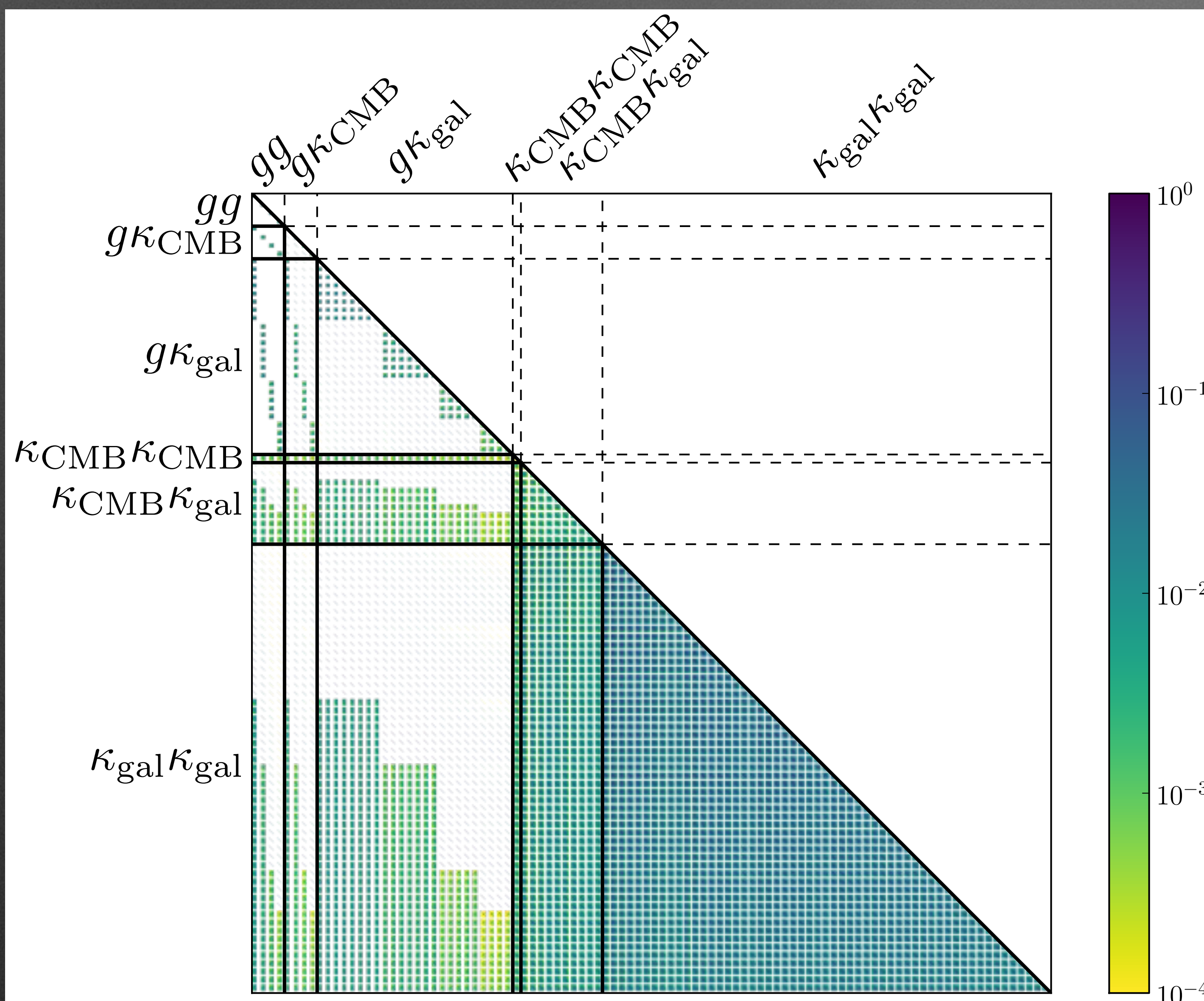
DES Covariances… Real Space

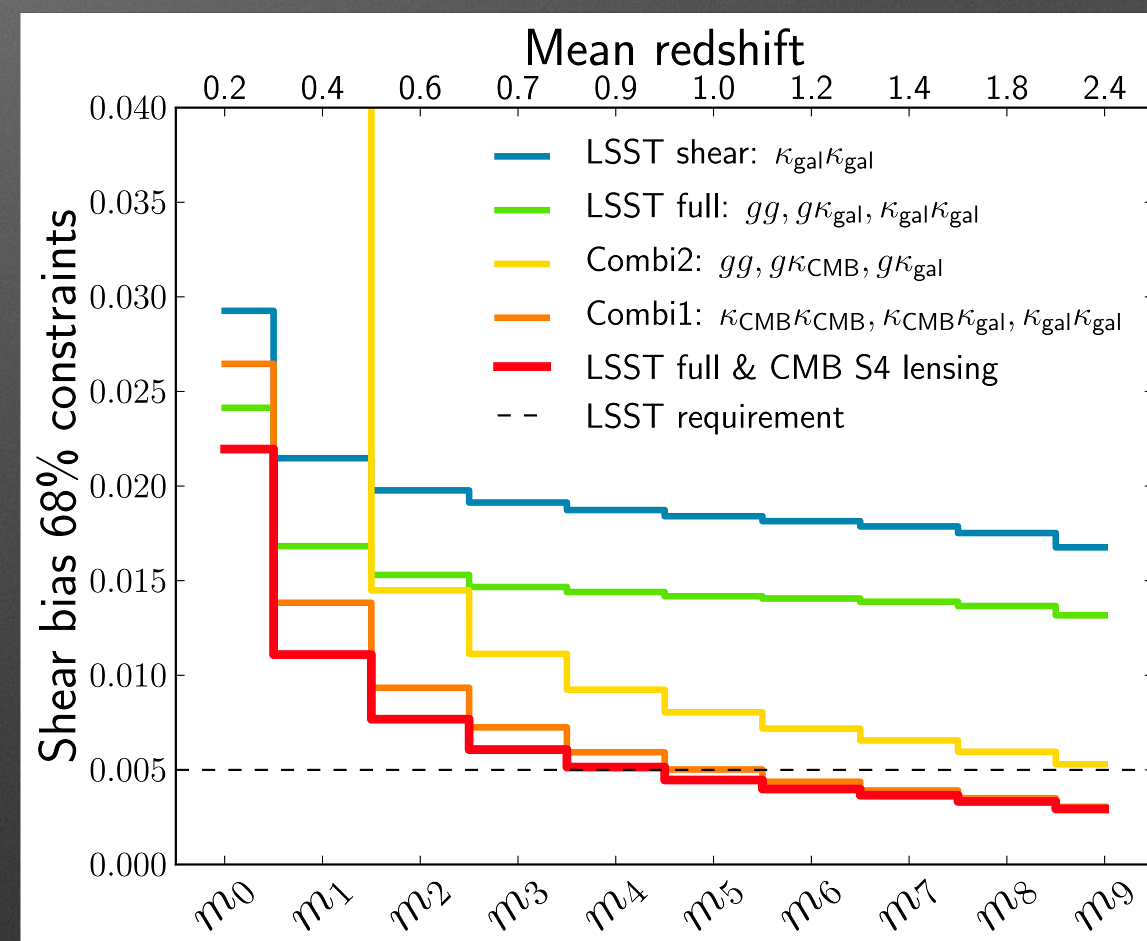
5x2pt



??

6x2pt in Fourier Space has been done (LSST x CMBS4)





details: Schaan, Krause, TE et al 2017

Precision Matrix Expansion:

Reducing Super-Computing needs…

“Precision matrix expansion - efficient use of numerical simulations in estimating errors on cosmological parameters”

Friedrich & TE 2018

The Problem: Inverse Covariance Estimation

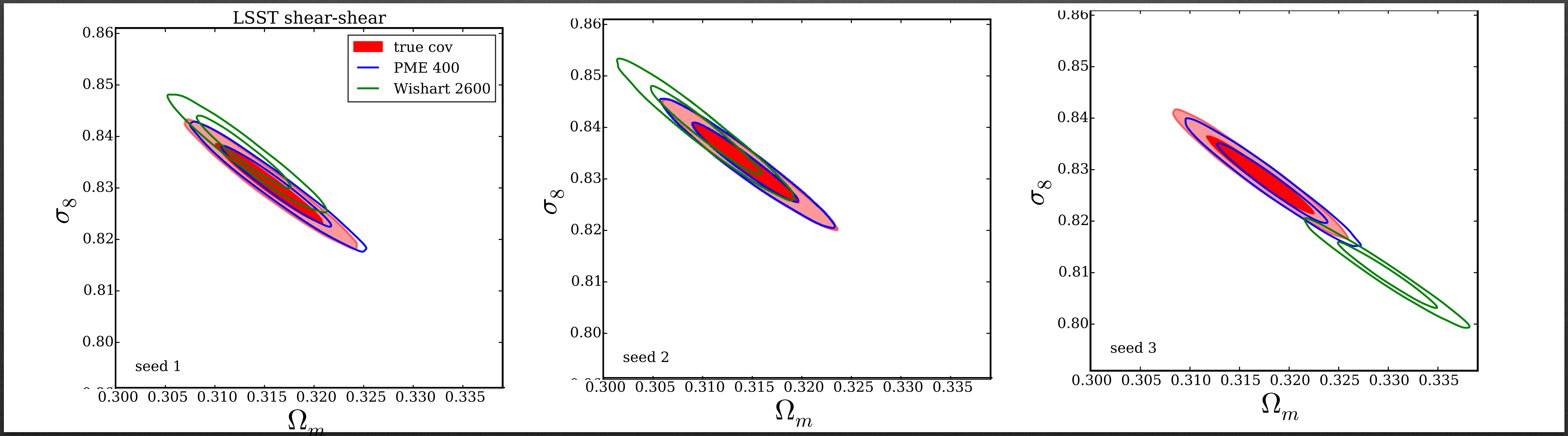
- Analytical covariance model relies on approximations that might be too imprecise for an LSST Y10 data set

- Estimation the covariance from numerical simulations (brute force), requires 10^5-10^6 realizations of an LSST Year 10 (WL only) like survey to shield against noise in the estimator

- Why?
  - The estimated inverse covariance is not the inverse of the estimated covariance
  - High-dimensionality of the data vector -> many elements in the covariance

How bad can it be …

- LSST Y10 data vector, reasonable parameter space
- Estimate Covariance assume Krause & Eifler 17 LSST cov (halo model) as truth and drawing from Wishart distribution (2600 realizations)
- Draw 3 different data vectors from noisy covariance
- Run likelihood analysis



Blue is truth, green is biased, red is the PME estimator

Idea: Estimate the inverse directly

$$p(\boldsymbol{\pi}|\hat{\boldsymbol{\xi}}) \sim \exp\left(-\frac{1}{2}\chi^2\left[\boldsymbol{\pi} \mid \hat{\boldsymbol{\xi}}, \mathbf{C}\right]\right) p(\boldsymbol{\pi})$$

$$\chi^2\left[\boldsymbol{\pi} \mid \hat{\boldsymbol{\xi}}, \mathbf{C}\right] = \left(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}]\right)^T \mathbf{C}^{-1} \left(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}]\right)$$

Idea: Estimate the inverse directly

$$p(\boldsymbol{\pi}|\hat{\boldsymbol{\xi}}) \sim \exp\left(-\frac{1}{2}\chi^2\left[\boldsymbol{\pi} \mid \hat{\boldsymbol{\xi}}, \mathbf{C}\right]\right) p(\boldsymbol{\pi})$$

$$\chi^2\left[\boldsymbol{\pi} \mid \hat{\boldsymbol{\xi}}, \mathbf{C}\right] = \left(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}]\right)^T \mathbf{C}^{-1} \left(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}]\right)$$

Standard Estimator

$$\hat{\boldsymbol{\Psi}} = \frac{\nu - N_d - 1}{\nu} \hat{\mathbf{C}}^{-1}$$

Standard estimator

$$\hat{\boldsymbol{\Psi}} = \frac{\nu - N_d - 1}{\nu} \hat{\mathbf{C}}^{-1}$$

$$\hat{\mathbf{C}} := \frac{1}{\nu} \sum_{i=1}^{N_\mathrm{s}} \left(\hat{\boldsymbol{\xi}}_i - \bar{\boldsymbol{\xi}} \right) \left(\hat{\boldsymbol{\xi}}_i - \bar{\boldsymbol{\xi}} \right)^T$$

Inverting quantities with "hats" is dangerous

Idea: Estimate the inverse directly

$$p(\boldsymbol{\pi}|\hat{\boldsymbol{\xi}}) \sim \exp\left(-\frac{1}{2}\chi^2\left[\boldsymbol{\pi} \mid \hat{\boldsymbol{\xi}}, \mathbf{C}\right]\right) p(\boldsymbol{\pi})$$

$$\chi^2\left[\boldsymbol{\pi} \mid \hat{\boldsymbol{\xi}}, \mathbf{C}\right] = \left(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}]\right)^T \mathbf{C}^{-1} \left(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}]\right)$$

Standard Estimator

$$\hat{\boldsymbol{\Psi}} = \frac{\nu - N_d - 1}{\nu} \hat{\mathbf{C}}^{-1}$$

New idea: Include theory information into estimator

$$\mathbf{C} = \mathbf{A} + \mathbf{B}$$

A: Stuff you know perfectly

B: Stuff you have some model for

Idea: Estimate the inverse directly

$$p(\boldsymbol{\pi}|\hat{\boldsymbol{\xi}}) \sim \exp\left(-\frac{1}{2}\chi^2\left[\boldsymbol{\pi} \mid \hat{\boldsymbol{\xi}}, \mathbf{C}\right]\right) p(\boldsymbol{\pi})$$

$$\chi^2\left[\boldsymbol{\pi} \mid \hat{\boldsymbol{\xi}}, \mathbf{C}\right] = \left(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}]\right)^T \mathbf{C}^{-1} \left(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}]\right)$$

Standard Estimator

$$\hat{\boldsymbol{\Psi}} = \frac{\nu - N_d - 1}{\nu} \hat{\mathbf{C}}^{-1}$$

New idea: Include theory information into estimator

$$\mathbf{C} = \mathbf{A} + \mathbf{B}$$

$$\mathbf{M} = \mathbf{A} + \mathbf{B}_m$$

$$\mathbf{C} = \mathbf{M} + (\mathbf{B} - \mathbf{B}_m)$$

$$\mathbf{X} := (\mathbf{B} - \mathbf{B}_m)\ \mathbf{M}^{-1}$$

$$\mathbf{C} = (\mathbb{1} + \mathbf{X})\ \mathbf{M}$$

Idea: Estimate the inverse directly

$$p(\boldsymbol{\pi}|\hat{\boldsymbol{\xi}}) \sim \exp\left(-\frac{1}{2}\chi^2\left[\boldsymbol{\pi} \mid \hat{\boldsymbol{\xi}}, \mathbf{C}\right]\right)\, p(\boldsymbol{\pi})$$

$$\chi^2\left[\boldsymbol{\pi} \mid \hat{\boldsymbol{\xi}}, \mathbf{C}\right] = \left(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}]\right)^T \mathbf{C}^{-1}\left(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}]\right)$$

Standard Estimator

$$\hat{\boldsymbol{\Psi}} = \frac{\nu - N_d - 1}{\nu}\hat{\mathbf{C}}^{-1}$$

New idea: Include theory information into estimator

$$\mathbf{C} = \mathbf{A} + \mathbf{B}$$

$\mathbf{M} = \mathbf{A} + \mathbf{B}_m$ →

$$\mathbf{C} = \mathbf{M} + (\mathbf{B} - \mathbf{B}_m)$$

$\mathbf{X} := (\mathbf{B} - \mathbf{B}_m)\,\mathbf{M}^{-1}$ →

$$\mathbf{C} = (\mathbb{1} + \mathbf{X})\,\mathbf{M}$$

Invert and expand as power series ↓

$$\begin{aligned}\boldsymbol{\Psi} &= \mathbf{M}^{-1}\left(\sum_{k=0}^{\infty}(-1)^k\mathbf{X}^k\right) \\ &= \mathbf{M}^{-1}\left(\mathbb{1} - \mathbf{X} + \mathbf{X}^2 + \mathcal{O}\left[\mathbf{X}^3\right]\right)\end{aligned}$$

← Build Estimator

$$\begin{aligned}\hat{\boldsymbol{\Psi}}_{\mathrm{2nd}} = \; & \mathbf{M}^{-1} + \mathbf{M}^{-1}\mathbf{B}_m\mathbf{M}^{-1}\mathbf{B}_m\mathbf{M}^{-1} \\ & -\mathbf{M}^{-1}\left(\hat{\mathbf{B}} - \mathbf{B}_m\right)\mathbf{M}^{-1} \\ & -\mathbf{M}^{-1}\hat{\mathbf{B}}\mathbf{M}^{-1}\mathbf{B}_m\mathbf{M}^{-1} \\ & -\mathbf{M}^{-1}\mathbf{B}_m\mathbf{M}^{-1}\hat{\mathbf{B}}\mathbf{M}^{-1} \\ & +\mathbf{M}^{-1}\frac{\nu^2\hat{\mathbf{B}}\mathbf{M}^{-1}\hat{\mathbf{B}} - \nu\hat{\mathbf{B}}\,\mathrm{tr}\left(\mathbf{M}^{-1}\hat{\mathbf{B}}\right)}{\nu^2 + \nu - 2}\mathbf{M}^{-1}\end{aligned}$$

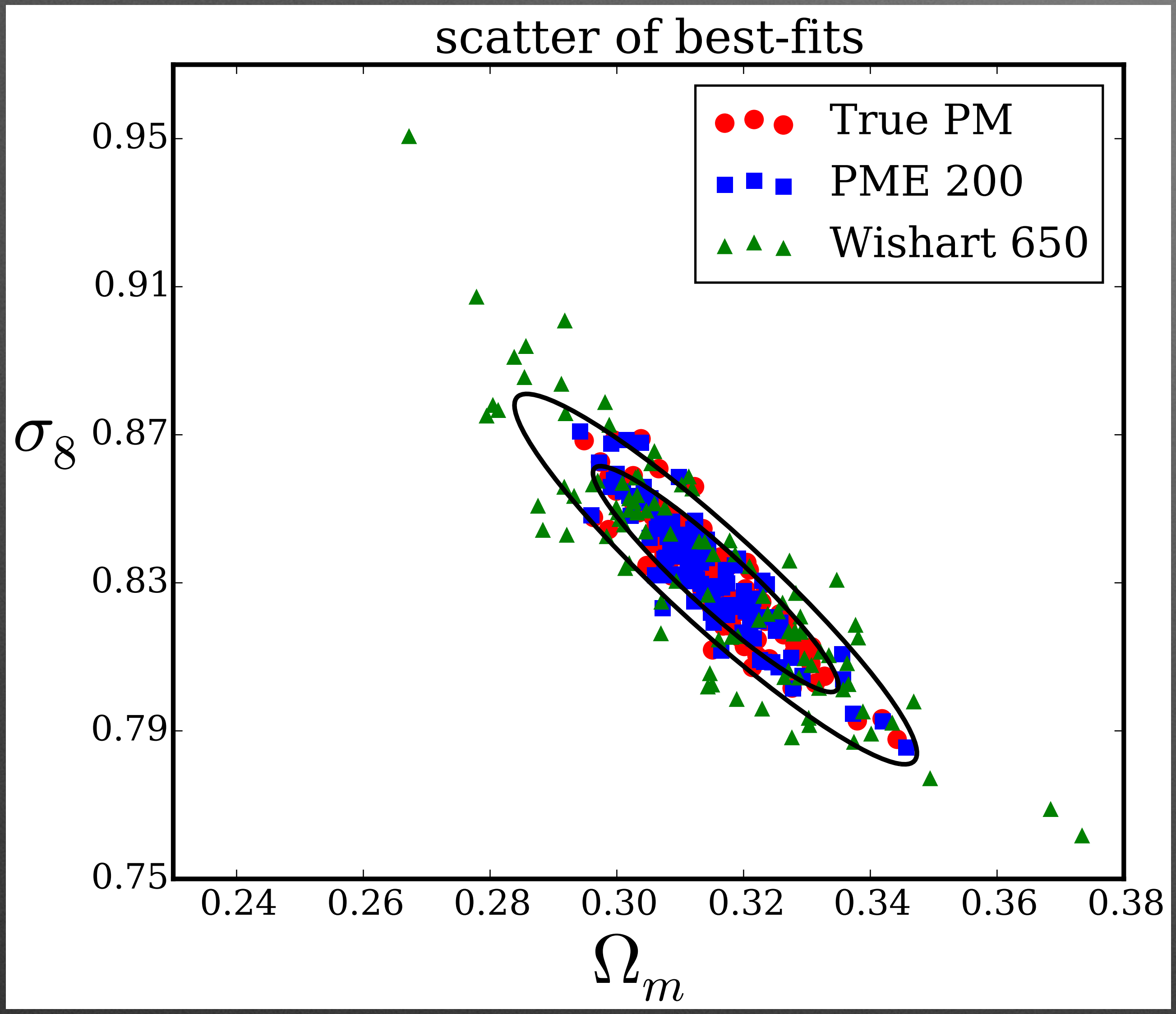
Only matrix multiplication, no inversion of estimated quantities

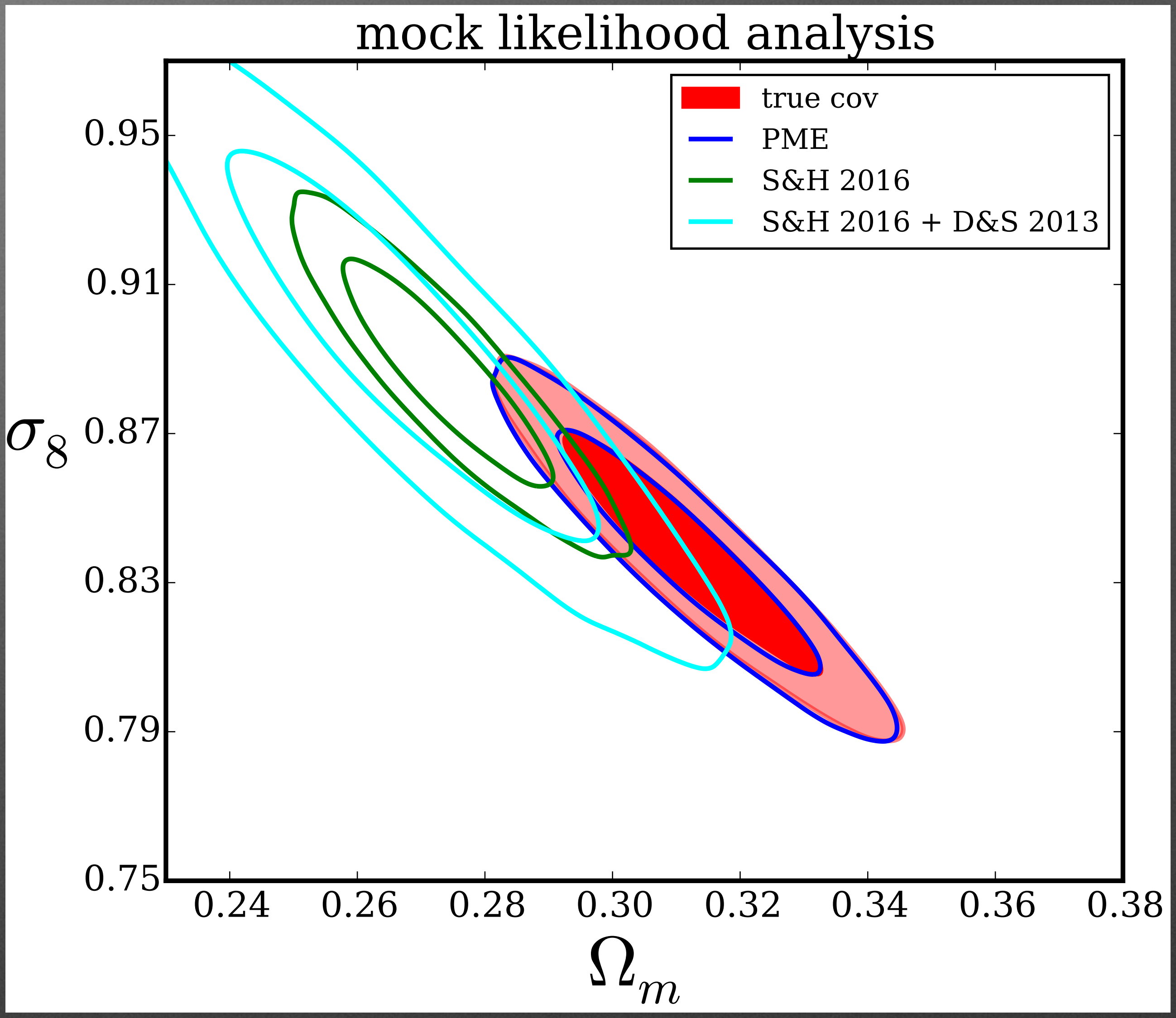
New estimator

$$
\begin{aligned}
\hat{\boldsymbol{\Psi}}_{\rm 2nd} \quad = \quad & \mathbf{M}^{-1} + \mathbf{M}^{-1}\mathbf{B}_m\mathbf{M}^{-1}\mathbf{B}_m\mathbf{M}^{-1} \\
& -\mathbf{M}^{-1}\left(\hat{\mathbf{B}} - \mathbf{B}_m\right)\mathbf{M}^{-1} \\
& -\mathbf{M}^{-1}\hat{\mathbf{B}}\mathbf{M}^{-1}\mathbf{B}_m\mathbf{M}^{-1} \\
& -\mathbf{M}^{-1}\mathbf{B}_m\mathbf{M}^{-1}\hat{\mathbf{B}}\mathbf{M}^{-1} \\
& +\mathbf{M}^{-1}\,\frac{\nu^2\hat{\mathbf{B}}\mathbf{M}^{-1}\hat{\mathbf{B}} - \nu\hat{\mathbf{B}}\,\mathrm{tr}\left(\mathbf{M}^{-1}\hat{\mathbf{B}}\right)}{\nu^2 + \nu - 2}\mathbf{M}^{-1}
\end{aligned}
$$

No more inversion of “hat quantities”…

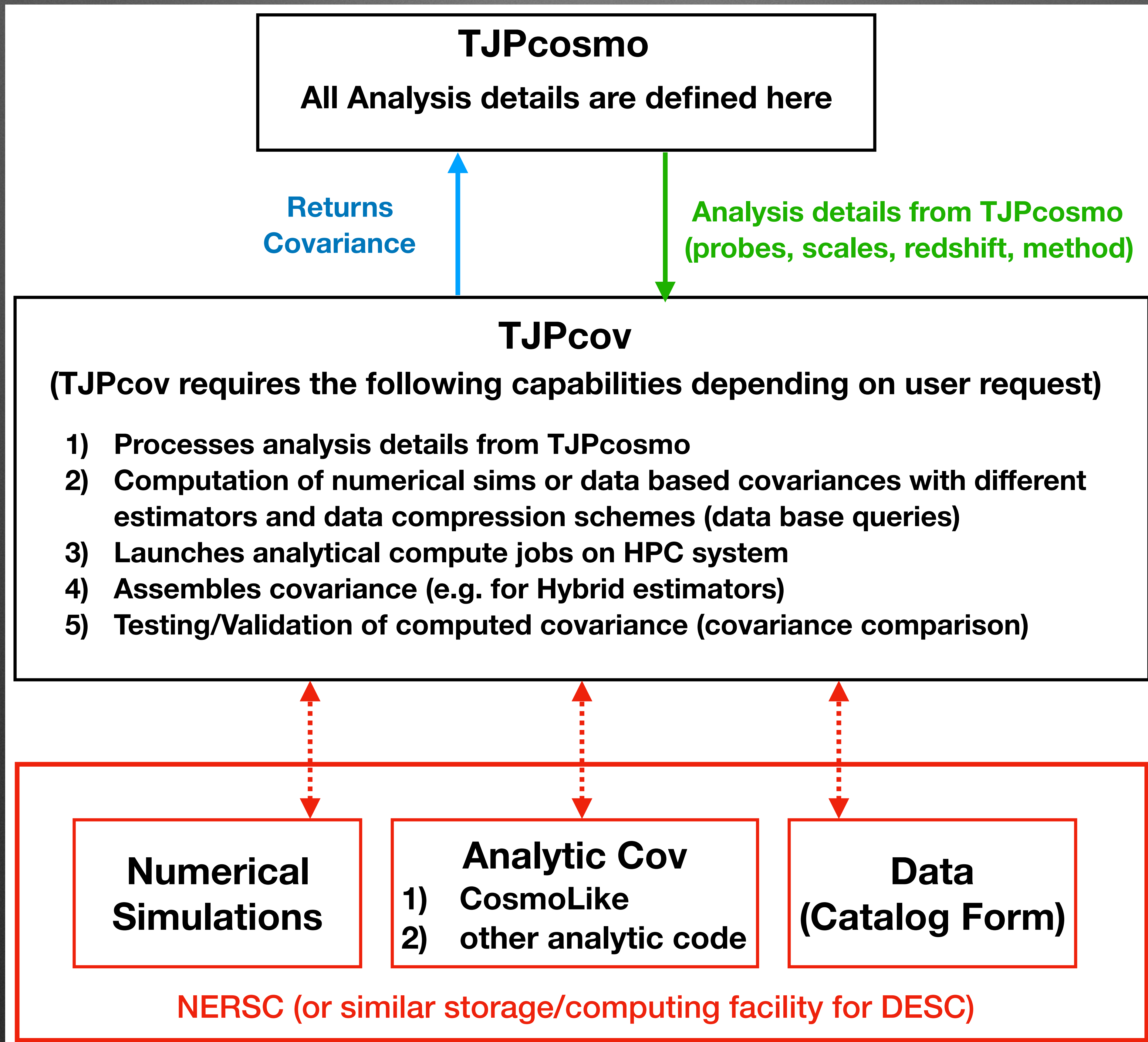
New estimator performance





- Instead of >10^5 our new estimator only requires ~2000 numerical simulations (LSST case)
- Given that 1 sim is 1M CPUh, at 1c/CPUh
- New method reduces cost $1B to $20M (-> fund theorists!)
- Next step: data compression

Implementation for LSST



Conclusions

- Multi-probe covariances are hard (think which probes to add)
- Covariance topic won't go away as Gaussian is probably ok as likelihood + ABC so far not convincing (please disagree…)
- Hybrid estimators are very interesting to reduce computational requirements iff analytic proves insufficient

- New project for the community: "covfefe: covariances for evaluating future experiments"
  - We are computing+collecting several covariances for DES, LSST, WFIRST and will make them available on our group's server "amypond"
  - Useful for all that could use covariances