



Statistical Challenges in 21st Century Cosmology

Closing arguments
G. Bernstein (UPenn)

Thanks to...

Chairs of the SOC

Name	Institution	Country
Jean-Luc Starck	Commissariat à l'Energie Atomique	France
Alan Heavens	Imperial College London	UK

Members of the SOC

Name	Institution	Country
Bruce Bassett	University of Cape Town	South Africa
Vassilis Charmandaris	Nat. Obs. of Athens & Univ. of Crete	Greece
Eric Feigelson	Penn State University	USA
Alan Heavens	Imperial College London	United Kingdom
Joseph Hilbe	Arizona State University	United States
Alberto Krone-Martins	Lisbon University	Portugal
Rachel Mandelbaum	Carnegie Mellon University	USA
Vicent Martinez	Valencia Astronomical Observatory	Spain
Hiranya Peiris	University College of London	UK
Anais Bassat	Ecole Polytechnique Federale Lausanne	Switzerland
Jean-Luc Starck	Commissariat à l'Energie Atomique	France
Panagiotis Tsakalides	University of Crete & FORTH/ICS	Greece
Yanxia Zhang	National Astronomical Observatory	China

Members of the LOC

Name	Institution
Panagiotis Tsakalides	University of Crete & FORTH/ICS
Thanasis Rontogiannis	Nat. Obs. of Athens
Grigorios Tsagkatakis	FORTH/ICS
Michalis Zervakis	Technical University of Crete

Administrative Support

Name	Institution
Marili Chatzaki	University of Crete & FORTH/ICS

Statistical cosmologist employment bureaus

Megaprojects

- KiDS
- HSC
- DES
- JPAS

Gigaprojects

- DESI
- LSST
- Euclid
- WFIRST
- SKA




What's new this week?

- New cosmological theories to test?
- New experiments or hardware to collect data?
- New statistics to form from data that discriminate theories?
- New methods that improve precision of known tests?
- New methods that improve accuracy of known tests (as present/future precision demands)?
- New statistical/algorithmic methods that make known tests feasible at current/future scales?*

**1 MCPU-hrs costs ~50 microEuclids from Amazon EC2*

What will future cosmo-statistics
meetings feature?

Things we can and should stop “correcting” and just do it right in the first place starting now

- Inverses of covariance matrix estimates (propagate the uncertainty!)
- Eddington bias (in galaxy cluster mass functions) 
- Malmquist and other corrections in SN Hubble diagrams 
- “Noise bias” in shear measurements 

(Sellentin, Ma, Shariff, Mandel, Murray, Bernstein)

If we don't have huge progress on these by the next CosmoStat meeting, heads will roll...

- Photo-z's in the face of incomplete & biased training spectroscopy.
 - Mix empirical knowledge with physics to fill out the color space
 - Combine correlation info with color info
 - $n(z, type)$ as free parameter, and proper propagation of uncertainties here & in measurement into cosmology.
- Covariance matrices for long data vectors
 - Use analytic knowledge to reduce number of needed simulations?

(Malz, Leistedt, Amaro, Speagle,
Leauthaud, Joachimi, Zhang)

Maybe 2 Cosmostats until heads roll:

- Automated redshift estimation, quality assessment from ratty, confused spectra (Euclid, WFIRST grisms)
 - Machine learning, Gaussian processes, sparsity...
- Galaxy deblending (LSST, HSC, WFIRST)
- SKA imaging challenges...

(Joseph, Pratley, Beck, Mary, Wiaux, Jamal, Remy, Schmitt)

QuAAc: Quiz on Algorithm Acronyms

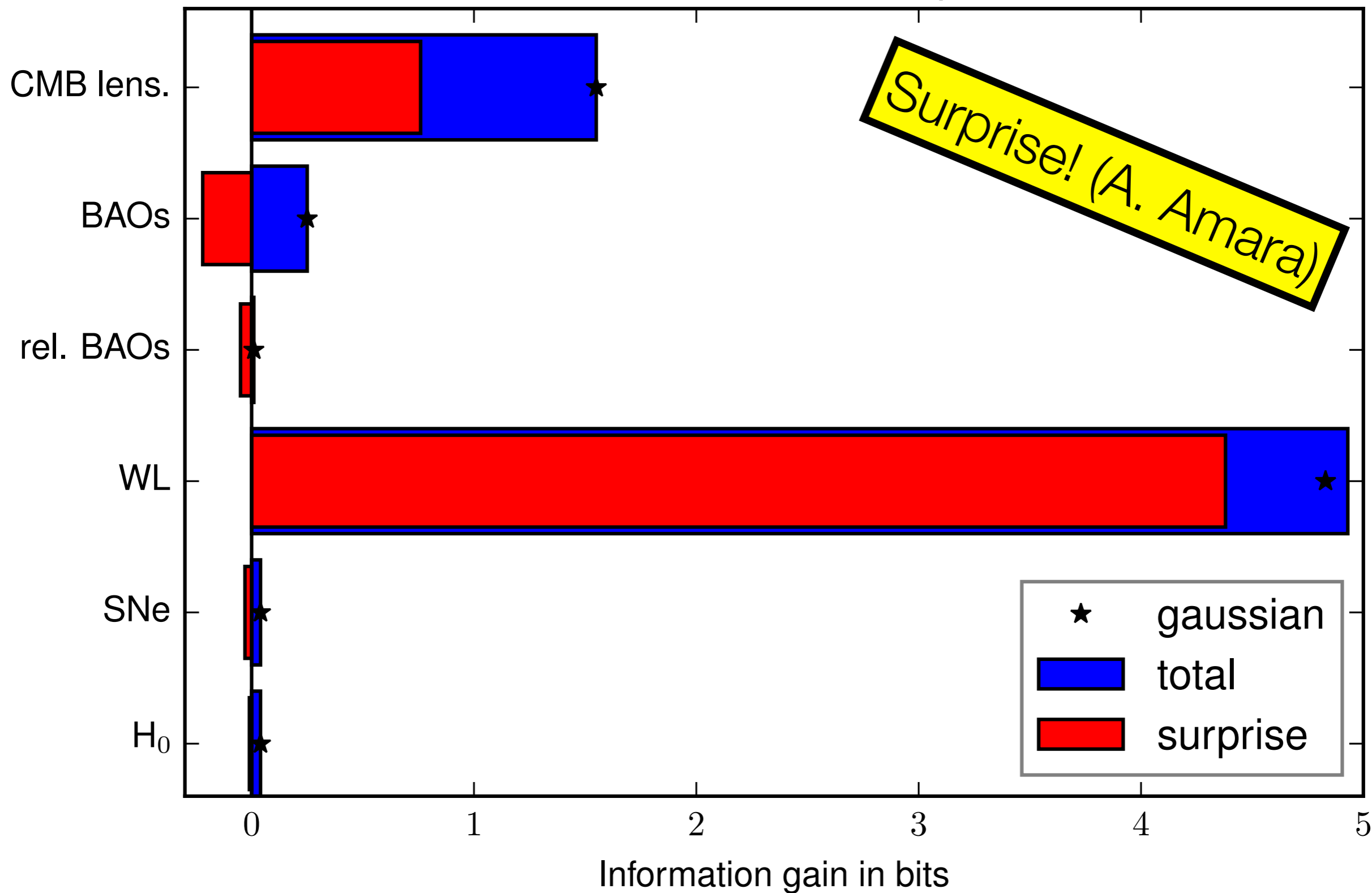
- NERCOME
- FISTA
- ABC
- BALROG
- ZOBOV
- HBM
- Nonparametric Eigenvalue-Regularized Precision or Covariance Matrix Estimator
- Fast Iterative Shrinkage-Thresholding Algorithm
- Approximate Bayesian Computation
- n/a!
- Zones Bordering on Voidness
- Bayesian Hierarchical Model (dialect)

The Non-Gaussian Frontier

- Avoid the heartbreak of information saturation and seek new fortunes in LSS with:
 - Voids, ridges/filaments, peaks, peak clipping, sliced correlations, phase statistics, counts in cells, Minkowski functionals, the BORG
 - (but the bispectrum was not invited to CosmoStat21!)**
- But if you thought modeling and assigning likelihoods to the power spectrum was hard, just wait...
 - Accurate analytic theory seems hopeless, never mind likelihoods. Total dependence on simulated Universes (incl. galaxy populations) + Approximate Bayesian Computation for model discrimination?
- A golden non-Gaussianity: CMB lensing

(Olenko, Ho, Neyrinck, Martinez, Lin, Pisani, Chen, Carron, Leclercq, Nadathur, Peel, Eggemeier, Demchenko, Giblin, Kovacs)

flat Λ CDM, Planck 15 prior



Getting (much) more out of data

- Multi-tracer large-scale structure (Abramo)
- 3d lensing mass reconstruction w/sparsity (Lanusse)
- Accurate mask functions (Aleksic)
- CMB lensing reconstruction at high S/N (Carron)
- Map component separation, scalar and otherwise (Umiltà, Wallis, McEwen, Rogers, Leistedt)
- Machine discovery, classification of galaxies, strong lenses, SNIa spectra (Fraix-Burnet, Ostrovski, Ishida, Petrillo, Beck, Caro, Lablanche)
- Interferometric image reconstruction, (Wiaux, Chan, Pratley)
- PSFs from stars (Farrens), stars from PSFs (Kuntzer)

See you all in Valencia

and until then, on the arXiv