Non-Gaussianity in data- and parameter space

Elena Sellentin

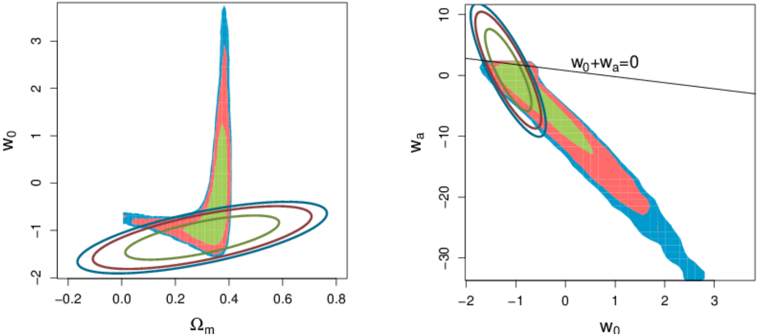
Institut für Theoretische Physik
Universität Heidelberg

Gravity on the Largest Scales

Code public on github:
http://Lnasellentin.github.io/DALI/

1000 SN 1a, vary Ω_m, w_0, w_a: $$F_{\alpha\beta} = \mu_{,\alpha} C^{-1} \mu_\beta$$

Wolz et al. 2012



Often, higher order derivatives will exist and encode the non-Gaussianity of the likelihood.

Approximate a Likelihood

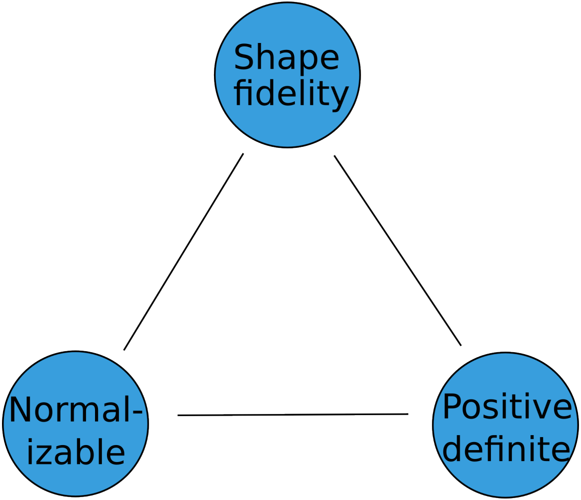
$$L = \exp(-\mathcal{L})$$

where

$$\mathcal{L} = \frac{1}{2}\mathrm{Tr}\left[\ln(\boldsymbol{C}) + (\boldsymbol{d}-\boldsymbol{\mu})\boldsymbol{C}^{-1}(\boldsymbol{d}-\boldsymbol{\mu})\right]$$

It's not just a usual Taylor approximation!
It's a Taylor approximation under the additional constraints of
normalizability and positive-definiteness.

Approximating a likelihood



Derivative expansion: DALI

Expand the *mean* or the *covariance matrix*:

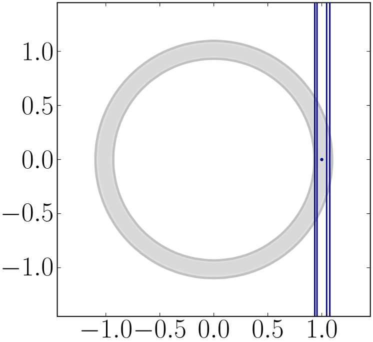
$$\boldsymbol{\mu}(\boldsymbol{p}) = \hat{\boldsymbol{\mu}} + \boldsymbol{\mu}_{,\alpha}\Delta p_\alpha + \frac{1}{2}\boldsymbol{\mu}_{,\alpha\beta}\,\Delta p_\alpha \Delta p_\beta + \frac{1}{3!}\boldsymbol{\mu}_{,\alpha\beta\delta}\Delta p_\alpha \Delta p_\beta \Delta p_\delta ...$$

$$\boldsymbol{C}(\boldsymbol{p}) = \hat{\boldsymbol{C}} + \boldsymbol{C}_{,\alpha}\,\Delta p_\alpha + \frac{1}{2}\boldsymbol{C}_{,\alpha\beta}\,\Delta p_\alpha \Delta p_\beta + \frac{1}{3!}\boldsymbol{C}_{,\alpha\beta\gamma}\,\Delta p_\alpha \Delta p_\beta \Delta p_\gamma ...$$

$$P \approx N \exp\left[-\frac{1}{2}(\boldsymbol{d} - \hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_{,\alpha}\Delta p_\alpha - \frac{1}{2}\boldsymbol{\mu}_{,\alpha\beta}\Delta p_\alpha \Delta p_\beta - ...)C^{-1}(\boldsymbol{d} - \hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_{,\alpha}\Delta p_\alpha - \frac{1}{2}\boldsymbol{\mu}_{,\alpha\beta}\Delta p_\alpha \Delta p_\beta - ...)\right]$$

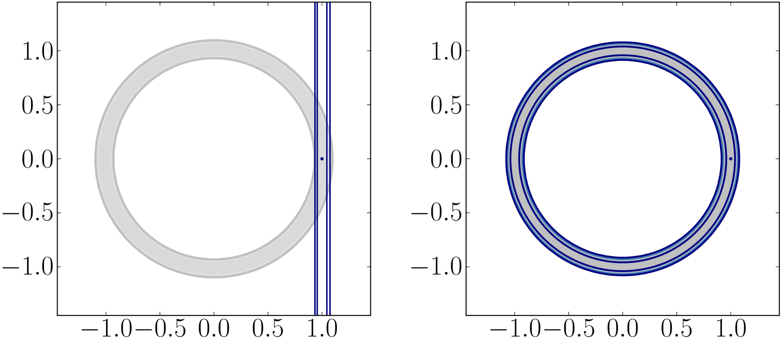
- E. Sellentin, M. Quartin, L. Amendola (2014),MNRAS, arXiv:1401.6892
- E. Sellentin (2015), MNRAS, arXiv: 1506.04866
- E. Sellentin, B.M. Schäfer, (2015), MNRAS, arXiv: 1506.05356
- C++ Code available on github

Extreme non-Gaussianities



$$C = x^2 + y^2$$

Extreme non-Gaussianities

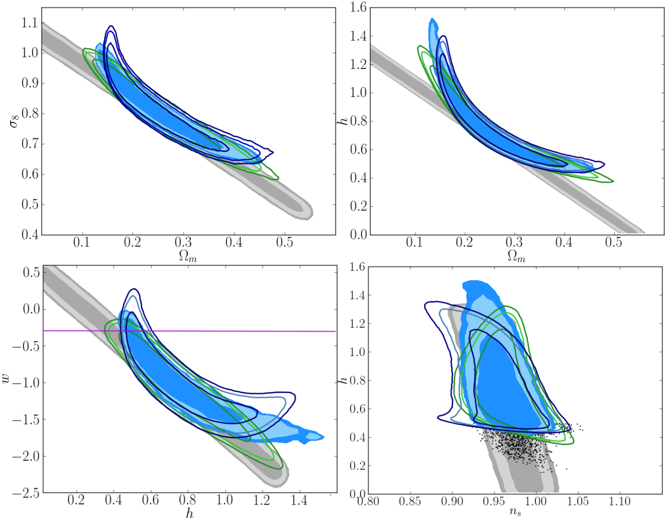


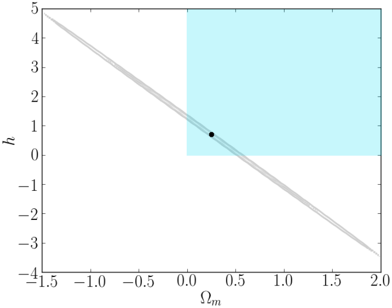
’Someone’ said:

”Is just a forecasting code! Just makes shapes funnier than the Fisher matrix.”

weak lensing @ Euclid precision

Without priors!





What about the Cramer-Rao inequality?

$$\mathrm{var}(\theta) \geq F_{\theta\theta}$$

$$\mathrm{var}(\theta) \geq F^{-1}_{\theta\theta}$$

Fisher matrix vs. Fisher information:

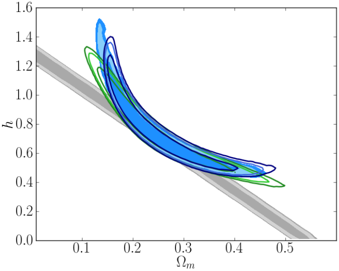
$$-\int f(x)\partial_\theta\partial_\theta \log\left[f(x)\right] \mathrm{d}x =$$

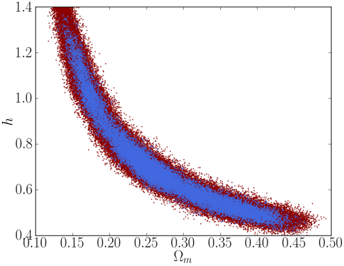
$$\int f(x)\left(\partial_\theta \log[f(x)]\right)^2 \mathrm{d}x - \int \partial_\theta\partial_\theta f(x)\mathrm{d}x$$

Accelerating Hamiltonian Monte Carlo

- Use $\mathcal{L}$ as potential 'energy', add randomized kinetic energy

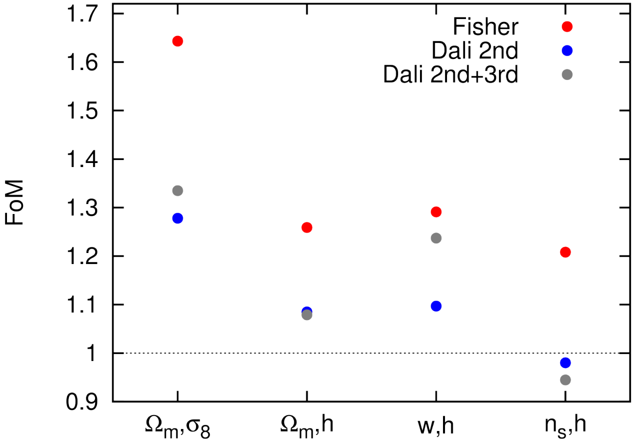
- Acceptance rate: 0.02 $\rightarrow 0.3 - 0.5$
- time/sample: $t_{leapfrog} \ll t_{MH}$ $\rightarrow$ speed up factor depends on leap frog steps
- Works also with real data!





weak lensing

DALI is better also quantitatively:



Parameter inference with estimated covariance matrices

Estimated covariance matrices

Data from a Gaussian distribution:

$$\boldsymbol{X_o} \sim \mathcal{G}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$L_{\mathcal{G}} = \frac{1}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\boldsymbol{X}_{\mathrm{o}} - \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}_{\mathrm{o}} - \boldsymbol{\mu})\right)$$

What if $\boldsymbol{\Sigma}$ cannot be calculated from first principles, but must be estimated?

$$\boldsymbol{S} = \frac{1}{N-1}\sum_{i=1}^{N}(\boldsymbol{X}_i - \bar{\boldsymbol{X}})(\boldsymbol{X}_i - \bar{\boldsymbol{X}})^T \quad (\text{for } N \to \infty \quad \boldsymbol{S} \to \boldsymbol{\Sigma})$$

→ Additional complications:

- $\boldsymbol{S}$ is a random variable, whereas $\boldsymbol{\Sigma}$ is not
- $\boldsymbol{S}^{-1}$ is biased wrt $\boldsymbol{\Sigma}^{-1}$

Estimated covariance matrices

The biased inverse:

$$\langle \boldsymbol{S}^{-1} \rangle = \frac{N-1}{N-p-2} \boldsymbol{\Sigma}^{-1} \tag{1}$$

The old recipe (Hartlap et al (2007)):

$$L_{\mathcal{G}} = \frac{1}{\sqrt{(2\pi)^p |\boldsymbol{S}|}} \exp\left(-\frac{1}{2} (\boldsymbol{X}_{\mathrm{o}} - \boldsymbol{\mu}) \frac{N-p-2}{N-1} \boldsymbol{S}^{-1} (\boldsymbol{X}_{\mathrm{o}} - \boldsymbol{\mu}) \right)$$

Still leads to biases of the inferred *parameters*!

Estimated covariance matrices

Correctly replacing $\boldsymbol{\Sigma}$ by $\boldsymbol{S}$:

$$P(\boldsymbol{X}_\mathrm{o}|\boldsymbol{\mu}, \boldsymbol{S}, N) = \int \mathrm{d}\boldsymbol{\Sigma}\, L_{\mathcal{G}}(\boldsymbol{X}_\mathrm{o}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) P(\boldsymbol{\Sigma}|\boldsymbol{S}, N) \tag{2}$$

How do you get $P(\boldsymbol{\Sigma}|\boldsymbol{S}, N)$?

Estimated covariance matrices

Construct $P(\boldsymbol{\Sigma}|\boldsymbol{S}, N)$:

- $\boldsymbol{S} = \frac{1}{N-1}\sum_{i=1}^{N}(\boldsymbol{X}_i - \bar{\boldsymbol{X}})(\boldsymbol{X}_i - \bar{\boldsymbol{X}})^T$ follows a Wishart distribution $\mathcal{W}(\boldsymbol{S}|\boldsymbol{\Sigma}/(N-1), N-1)$
- Invert with Bayes' Theorem: $P(\boldsymbol{\Sigma}|\boldsymbol{S}, N)\pi(\boldsymbol{S}) = \mathcal{W}(\boldsymbol{S}|\boldsymbol{\Sigma}/n, n)\pi(\boldsymbol{\Sigma})$ by using some priors $\pi(\boldsymbol{\Sigma})$

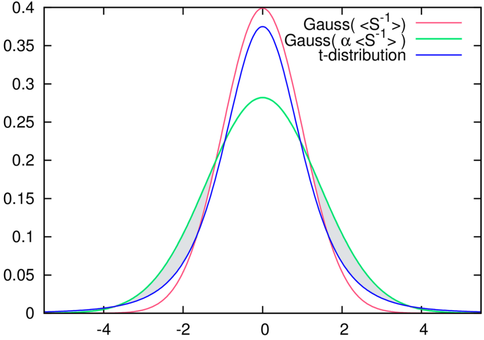
The prior discussion...

- We use an independence-Jeffreys prior $\pi(\boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{p+1}{2}}$.

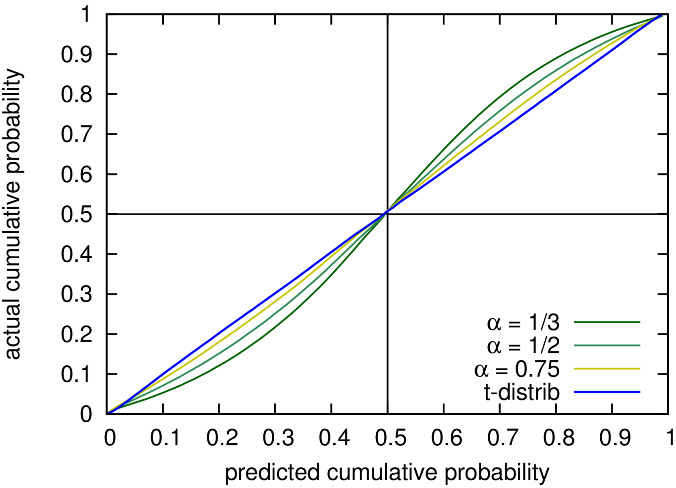
Estimated covariance matrices

Solve $P(\boldsymbol{X}_\text{o}|\boldsymbol{\mu},\boldsymbol{S},N) = \int \mathrm{d}\boldsymbol{\Sigma}\, L_\mathcal{G}(\boldsymbol{X}_\text{o}|\boldsymbol{\mu},\boldsymbol{\Sigma}) P(\boldsymbol{\Sigma}|\boldsymbol{S},N)$:

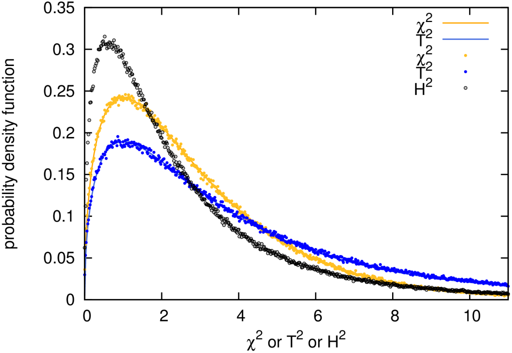
$$P(\boldsymbol{X}_\text{o}|\boldsymbol{\mu},\boldsymbol{S},N) = \frac{\bar{c}_p |\boldsymbol{S}|^{-1/2}}{\left[1 + \frac{(\boldsymbol{X}_\text{o}-\boldsymbol{\mu})^T \boldsymbol{S}^{-1} (\boldsymbol{X}_\text{o}-\boldsymbol{\mu})}{N-1}\right]^{\frac{N}{2}}}$$



Estimated covariance matrices



Estimated covariance matrices



$$\chi^2 = (\boldsymbol{X}_\mathrm{o} - \boldsymbol{\mu})^T \, \boldsymbol{\Sigma}^{-1} \, (\boldsymbol{X}_\mathrm{o} - \boldsymbol{\mu})$$

$$H^2 = (\boldsymbol{X}_\mathrm{o} - \boldsymbol{\mu})^T \, \alpha \boldsymbol{S}^{-1} \, (\boldsymbol{X}_\mathrm{o} - \boldsymbol{\mu})$$

$$T^2 = (\boldsymbol{X}_\mathrm{o} - \boldsymbol{\mu})^T \, \boldsymbol{S}^{-1} \, (\boldsymbol{X}_\mathrm{o} - \boldsymbol{\mu})$$

Merci :)

–End of Talk–