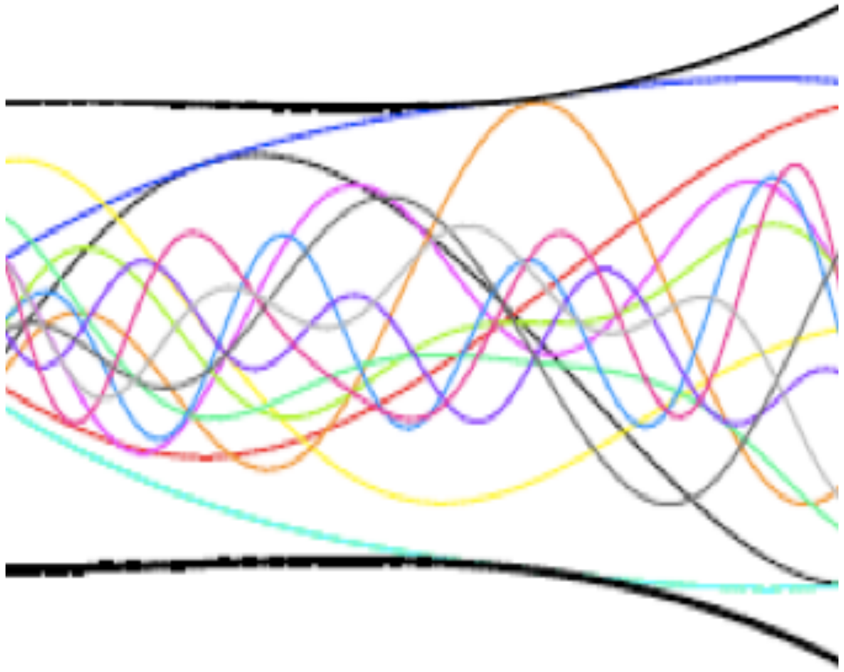
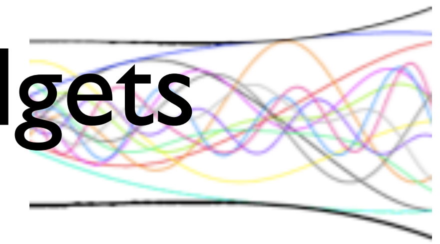
functional marginalization: do not parameterise your ignorance



Tom Kitching
Andy Taylor

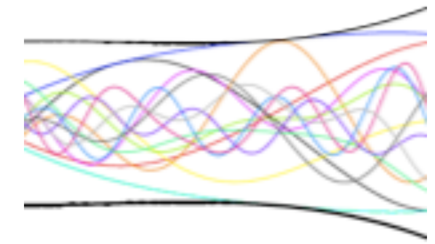
1003.1136
1005.2063
1012.3479

Need correct accounting of *systematic* error budgets



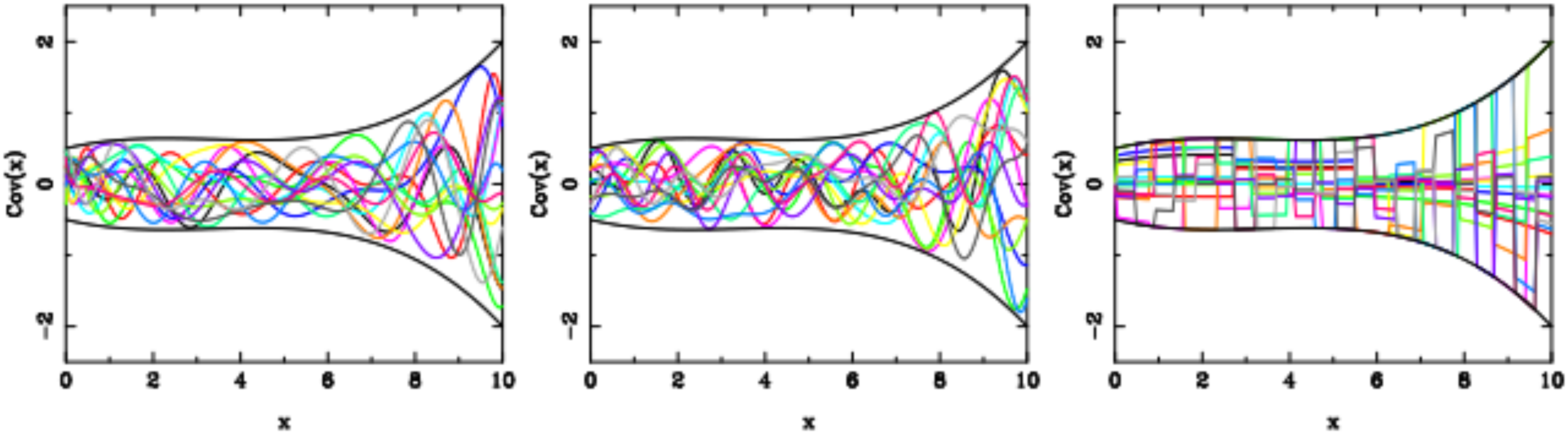
- Standard approach is to use nuisance parameters
  - Truncated specific functional form that encodes uncertainty in understanding
  - Likelihoods calculations can be >~seconds

- New functional approach is needed
  - Assess functions not parameters
  - Start to address the lack of functional analyses in cosmology
  - Note that do not need to find the best fit function (it is a nuisance by definition), only account for its marginalisation

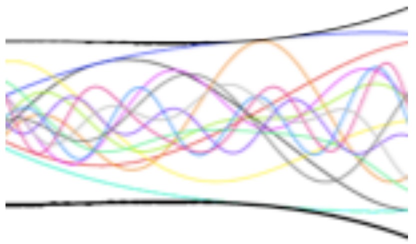


Nuisance function (don't need its value, do want to account for it)

f(x) $\longrightarrow$



- Parameterise
  - Ok if physically motivated (from a theory)
  - Not ok if unknown
  - Could results in a many-parameters problem
    - 100s or 1000s nuisance parameters for 10s of cosmological parameters



f(x) $\longrightarrow$

Can we do something analytic?

- Functional "Path"-Integral Marginalization
- One example (more in papers)

Analytic Marginalization

- Faced with a many parameters problem in cosmology
  - Both systematics and interesting parameters

- Expand log-likelihood about nuisance parameters

$$\mathcal{L} = \mathcal{L}_0 + \delta\psi_\alpha \mathcal{L}_\alpha + \frac{1}{2}\delta\psi_\alpha \delta\psi_\beta \mathcal{L}_{\alpha\beta},$$

- Assume Gaussian likelihood for nuisance parameters

$$\mathcal{L} = \mathcal{L}_0 - \frac{1}{2}\mathcal{L}_\alpha \mathcal{L}^{-1}_{\alpha\beta} \mathcal{L}_\beta + \mathrm{Tr}\ \ln\left(V_\psi^{2/M} \mathcal{L}_{\alpha\beta}\right)$$

Analytic Marginalization

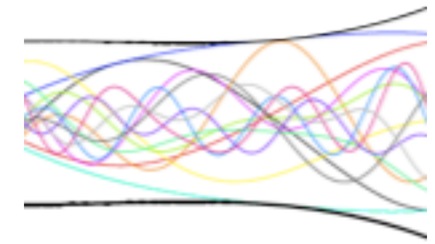
- A very simple expression for the likelihood, marginalized over the nuisance parameters

$$\mathcal{L} = \Delta D C_M^{-1} \Delta D^t + \mathrm{Tr}\, \ln V_\psi^{2/M} F_{\alpha\beta}$$

- Change in the likelihood encapsulated in an increase in the covariance

$$\begin{aligned} C_M &= \left(C^{-1} - C^{-1} \mu_\alpha^t [F_{\alpha\beta} + C_{\alpha\beta}^{-1}]^{-1} \mu_\beta C^{-1}\right)^{-1} \\ &= C + C_{\alpha\beta} \mu_\alpha \mu_\beta^t, \end{aligned}$$

Analytic Marginalization



- Within the same formalism we can also
  - Estimate the maximum likelihood point

$$\delta\Phi_\mu = -\mathcal{L}_\nu \mathcal{L}^{-1}_{\nu\mu}\,.$$

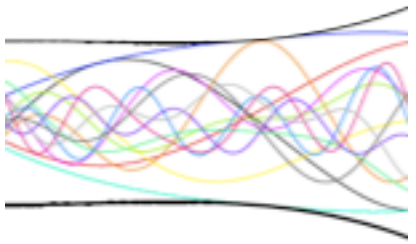
  - Like
    - Newtons Method
    - and Tegmarks "Quadratic Estimator"
  - Write a closed expression for the Bayesian Evidence (!)

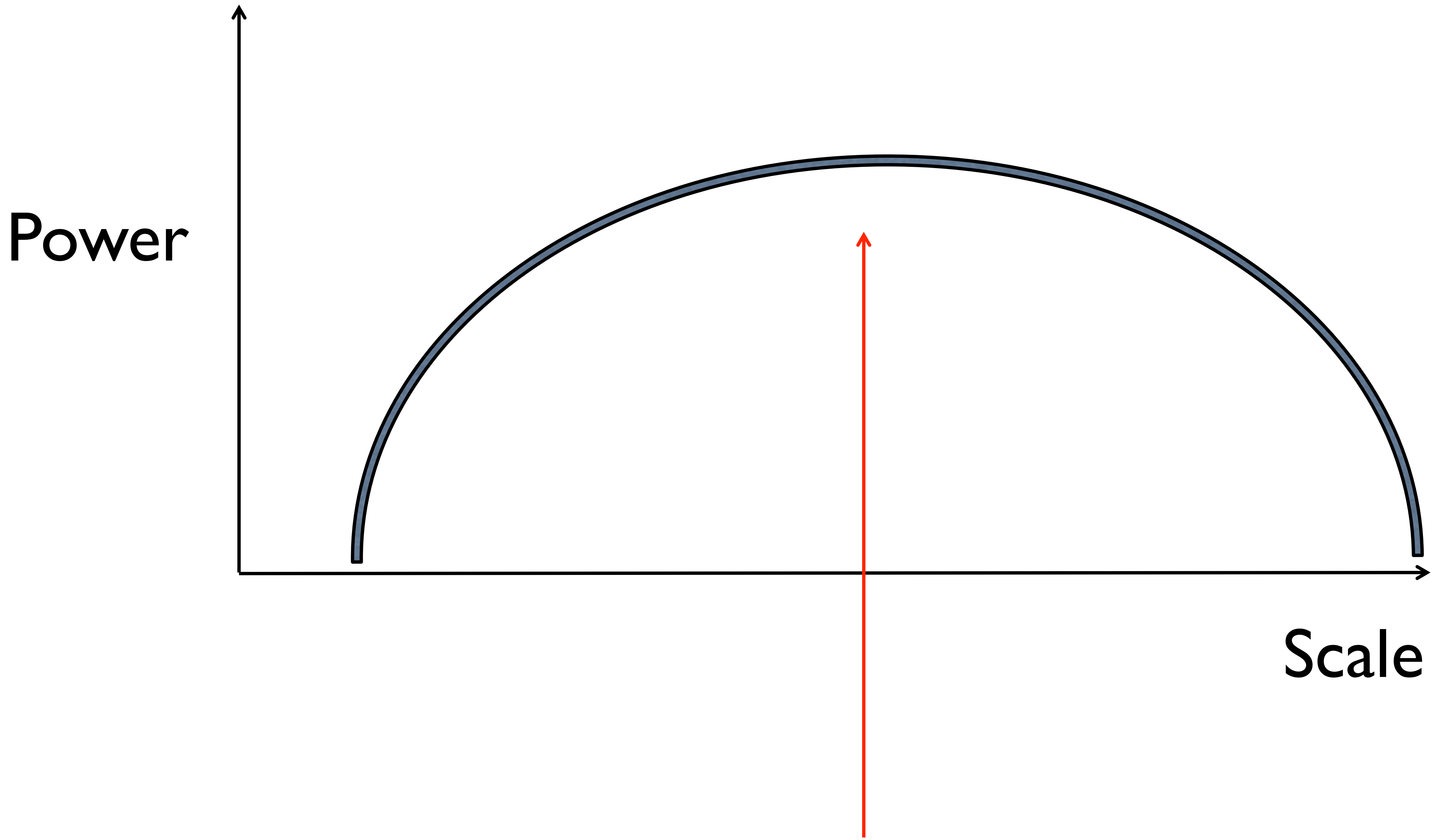
$$\begin{aligned}\mathcal{E}(D|\mathcal{M}) &= \Delta D\left(C^{-1} - C^{-1}\mu_i^t F_{ij}^{-1}\mu_j C^{-1}\right)\Delta D^t \\ &+ \mathrm{Tr}\,\ln C + 2\ln(V_\theta\sqrt{\det F_{ij}}) - N_p \ln 2\pi\,.\end{aligned}$$

Analytic Marginalization

- Now have a self consistent way of treating likelihoods in an analytic way
  - Finding maximum likelihoods
  - Marginalizing over nuisance parameters
  - Calculating Bayesian Evidence

Unknown (nuisance) Amplitude

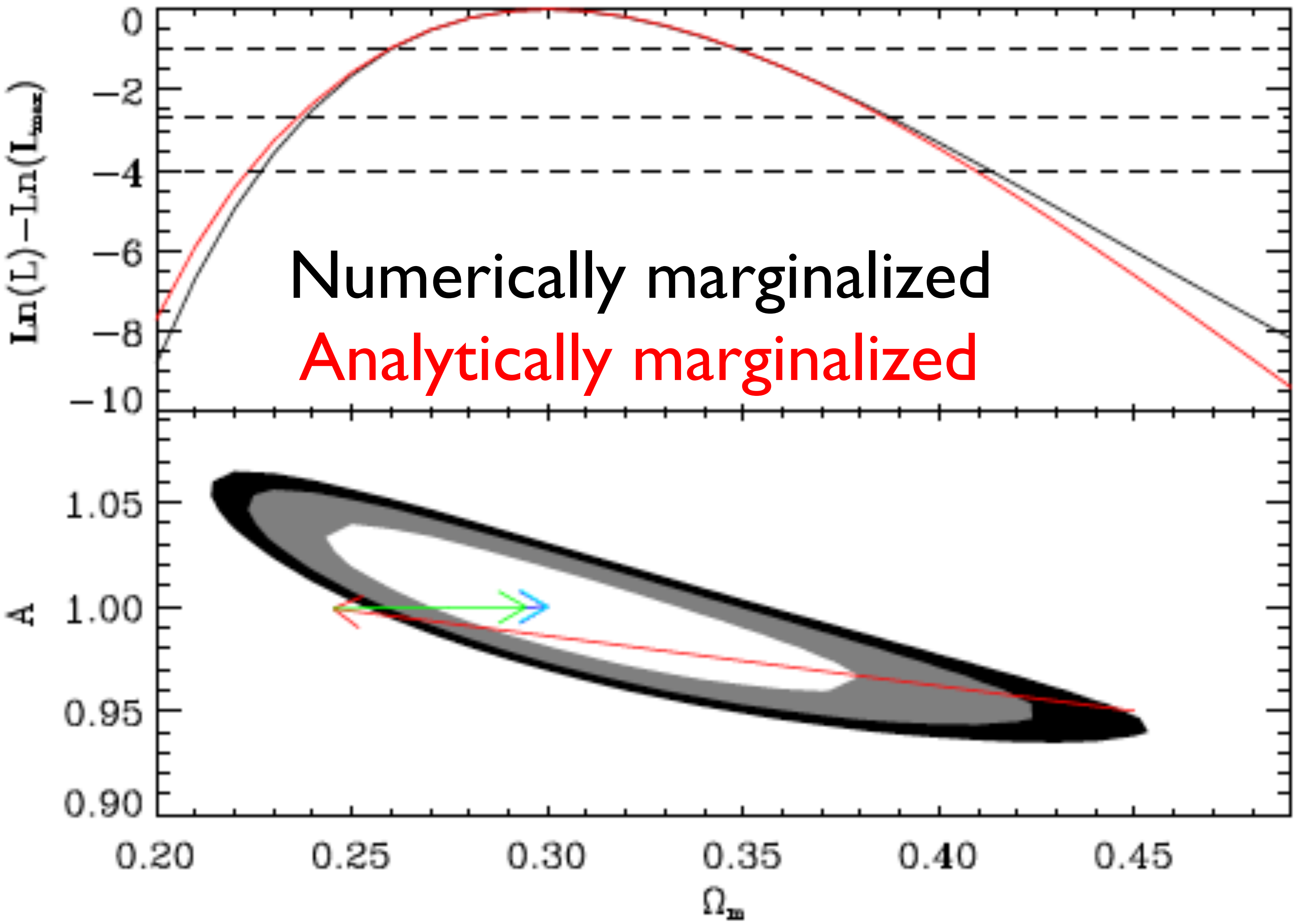




Standard Practice (camb/cosmomc Lewis; Bridle et al. 2002)

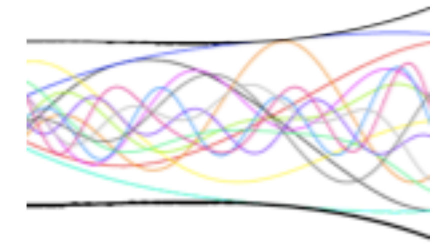
Example : Nuisance parameters

- Redshift survey power spectrum measurement
- Assume max k=0.2 $hMpc^{-1}$ and volume 19 $h^{-3}Gpc^3$
- Marginalize over an unknown amplitude



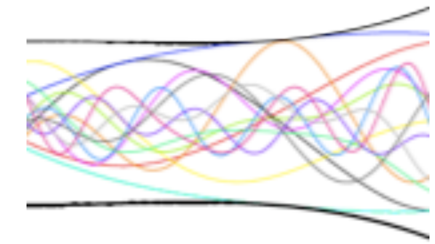
Single amplitude, now extend formalism to an infinite set of amplitudes !

Functional Integration



- Q: What do we do when we don't know the systematic functional form?
- ~~A: Parameterize, estimate and then marginalize~~ X
- A: Marginalize over all functional behavior ✓

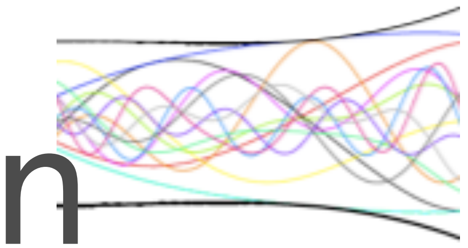
- Can use functional-integrals "path-integrals"
  - Very common in physics
    - Feynman Sum over Histories
    - Action principles
- Replace the fundamental variables (parameters) with parameters
- Nuisance parameters becomes Nuisance functions



$$\int d^n x \, e^{-\frac{1}{2} x_i C^{-1}_{ij} x_j} ;$$

$$L(\boldsymbol{D}|\boldsymbol{\theta}) = \int d[\delta\psi] \, \frac{e^{-1/2(\Delta \boldsymbol{D} - \delta\psi_i \, \partial_{\psi_i}\boldsymbol{\mu}) \boldsymbol{C}^{-1} (\Delta \boldsymbol{D} - \delta\psi_j \, \partial_{\psi_j}\boldsymbol{\mu})^t}}{\sqrt{(2\pi)^N \det \boldsymbol{C}}}$$

Functional Marginalization

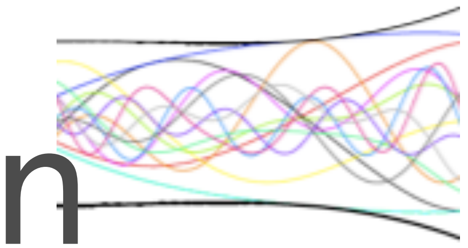


- Expand likelihood about ~~parameters~~ nuisance functions

$$\mathcal{L} = \mathcal{L}_0 + \delta\psi_\alpha \mathcal{L}_\alpha + \frac{1}{2}\delta\psi_\alpha \delta\psi_\beta \mathcal{L}_{\alpha\beta},$$

$$\begin{aligned}\mathcal{L} &= \mathcal{L}_0 + \int dx'\, \delta\psi_\alpha(x') \frac{\delta \mathcal{L}_0}{\delta\psi_\alpha(x')} \\ &+ \frac{1}{2}\int dx' dx''\, \delta\psi_\alpha(x') \left[\frac{\delta^2 \mathcal{L}_0}{\delta\psi_\alpha(x')\delta\psi_\beta(x'')}\right] \delta\psi_\beta(x''),\end{aligned}$$

Functional Marginalization



- Marginalize nuisance ~~parameters~~ functions
  - Assume Gaussian Prior in Function-Space

$$\mathcal{L} = \Delta D C_M^{-1} \Delta D^t + \mathrm{Tr}\,\ln V_\psi^{2/M} F_{\alpha\beta}$$

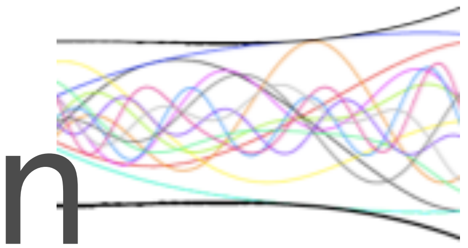
$$\begin{aligned} C_M &= \left(C^{-1} - C^{-1}\mu_\alpha^t [F_{\alpha\beta} + C_{\alpha\beta}^{-1}]^{-1}\mu_\beta C^{-1}\right)^{-1} \\ &= C + C_{\alpha\beta}\mu_\alpha\mu_\beta^t, \end{aligned}$$

↓

$$\mathcal{L} = \Delta D C_M^{-1} \Delta D^t + \mathrm{Tr}\,\ln V_\psi^{2/M} F_{\alpha\beta}$$

$$C_M = C + \int dx'\,dx''\,C_{\alpha\beta}(x', x'')\frac{\delta\mu[\psi(x)]}{\delta\psi_\alpha(x')}\frac{\delta\mu^t[\psi(x)]}{\delta\psi_\beta(x'')}$$

Functional Marginalization



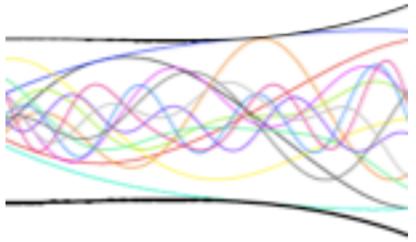
- Simple change to the covariance
- Can now marginalize over <u>all</u> possible functional behavior (as prior variance → infinity)

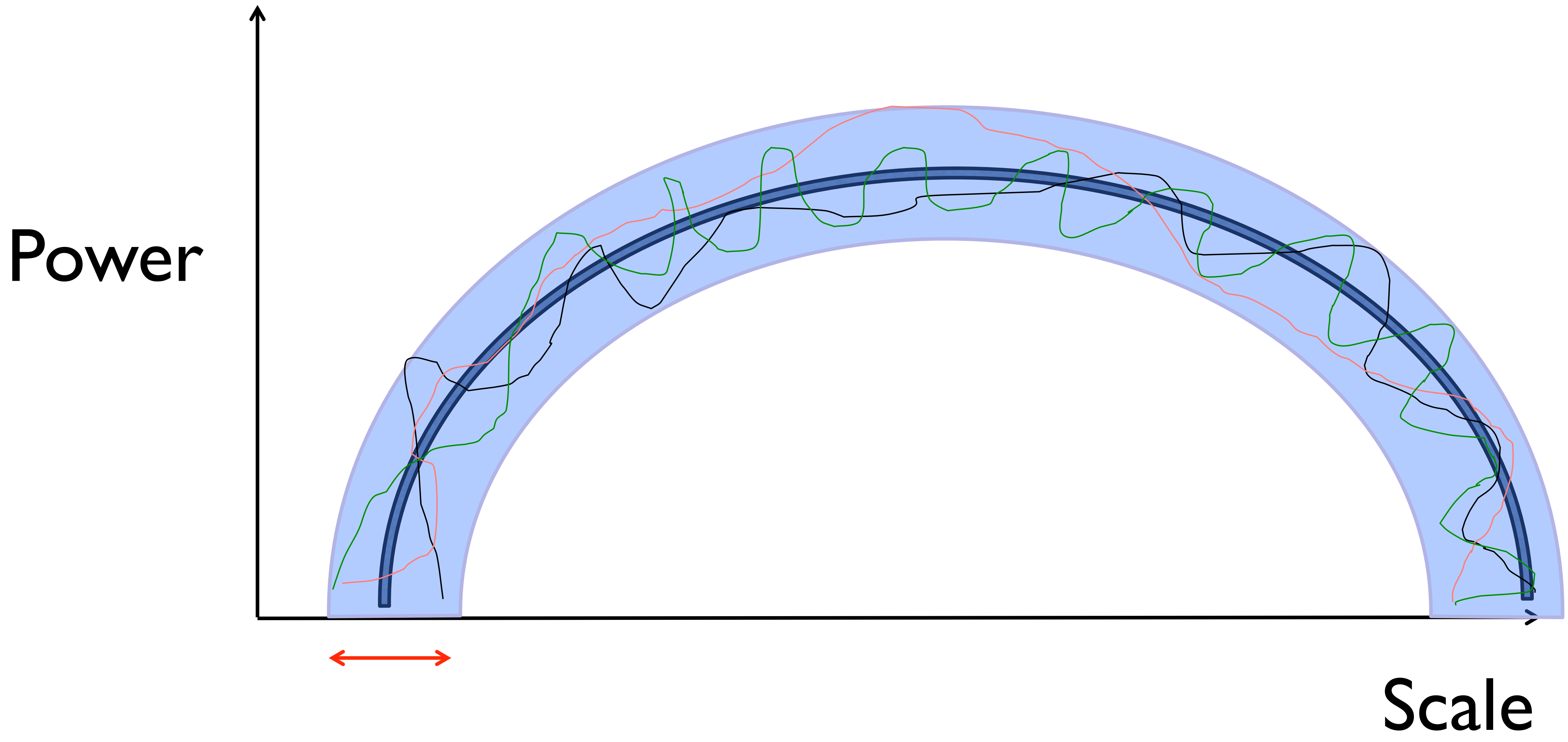
$$C_M = C + \int dx' dx'' \, C_{\alpha\beta}(x', x'') \frac{\delta \mu[\psi(x)]}{\delta \psi_\alpha(x')} \frac{\delta \mu^t[\psi(x)]}{\delta \psi_\beta(x'')}$$

- Fisher matrix, marginalized over all possible nuisance functions

$$F_{ab} = \frac{1}{2} \mathrm{Tr}\, [\mu_a \mu_b^t + \mu_a \mu_b^t] C_M^{-1}$$

Unknown (nuisance) Amplitude





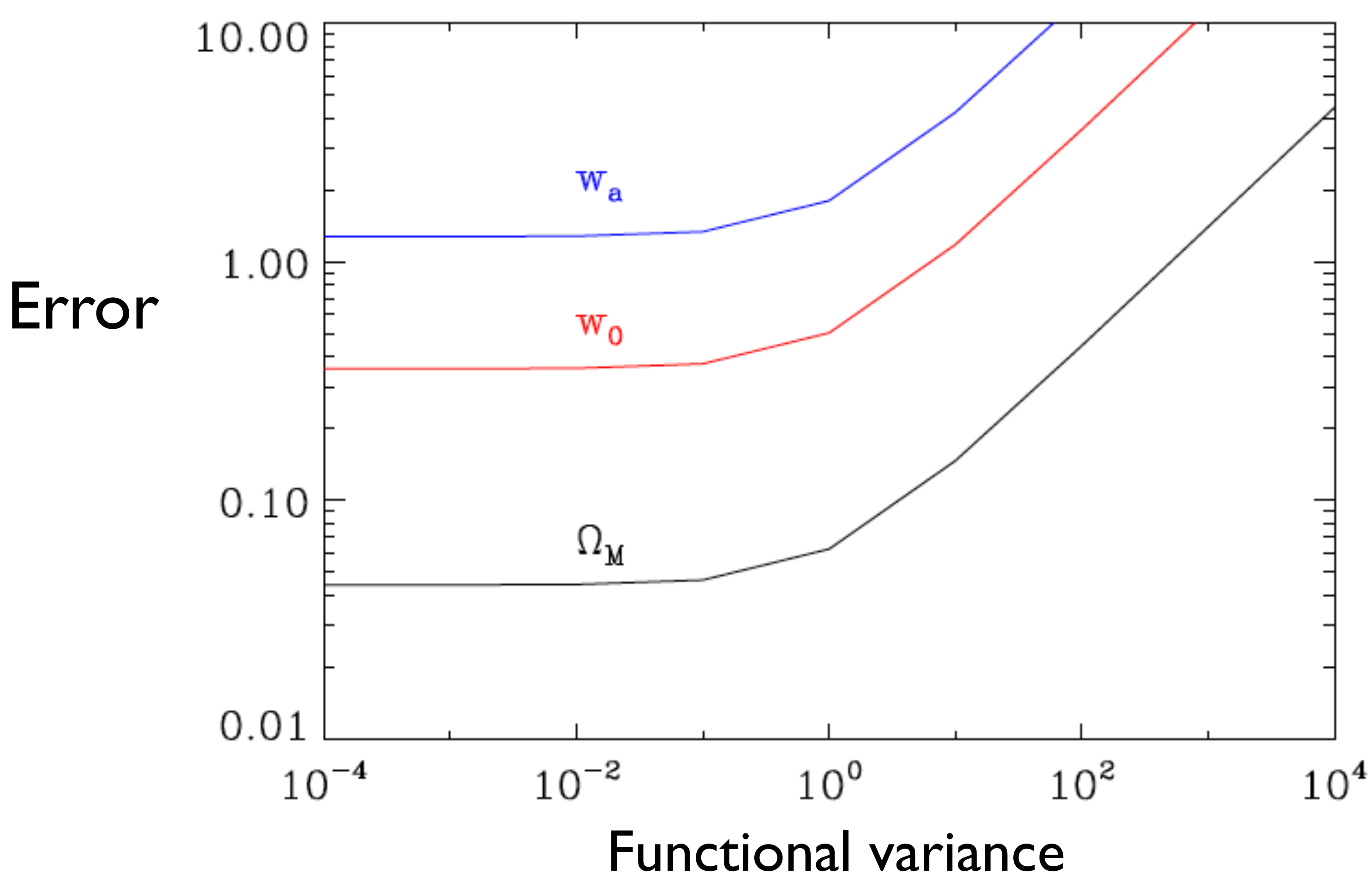
For CMB Standard Practice (Bridle et al. 2002)

Example : Scale dependent bias

$$V_{\rm eff}(k) = \int d^3r \left(\frac{[P^s_{gg}(k)]^2}{2[P^s_{gg}(k) + N(r)]^2 + B(k)} \right)$$

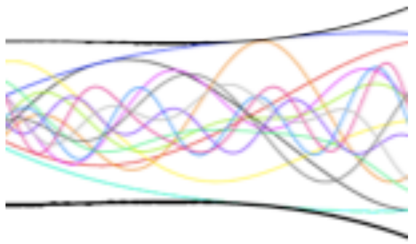
$$B(k) = 4 \left(\frac{P_b(k)}{b^2(k)} \right) |P^s_{gg}(k)|^2$$

- Analytic expression that account for unknown functional behavior in the bias



Need to know functional variance of scale dependent bias b(k) to approx 1% for Euclid-scale experiment

- Q: What do we do when we don't know the systematic functional form?



- ~~A: Parameterize, estimate and then marginalize~~ X
- A: Marginalize over all functional behavior ✓

- Next
  - Many more applications, pick your favorite function
    - H(z), D(z), P(k), w(z)…
  - Address assumptions (e.g. equally weighting all functions)

- Papers (and more examples) on arxiv
  - 1003.1136 (analytical marginalization)
  - 1005.2063 (instrinsic alignments)
  - 1012.3479 (non-linear power)