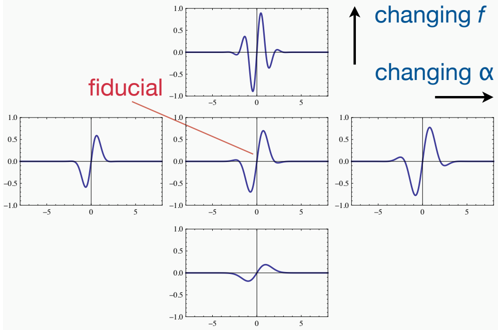
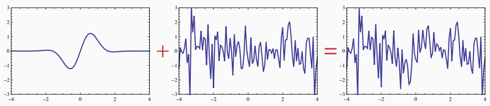
$$h(t; A, \alpha, f) = A\, e^{-\frac{t^2}{2\alpha^2}} \sin(2\pi f t)$$



1 *The physics that can be extracted from GW signals is encoded in the functional dependence of the waveforms on the source parameters. Here we consider a simple sine-Gaussian.*

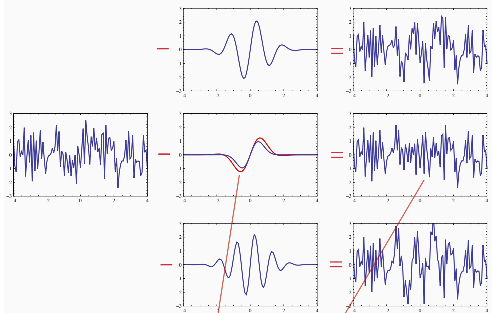
signal + noise = data



2 *Of course, all detected signals will be embedded in detector noise!*

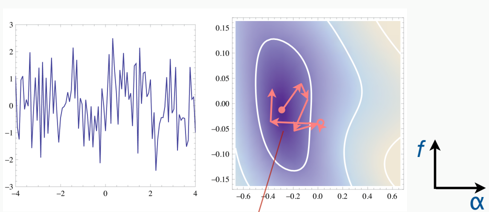
data – signal = noise

hence p(signal pars) = p(noise residual)



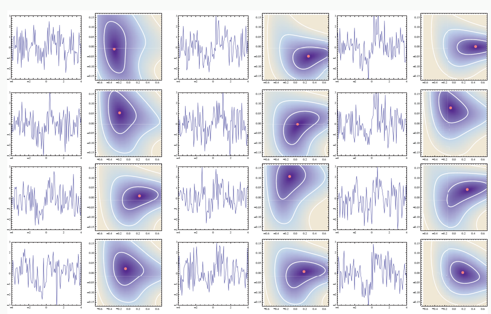
3 *Thus, parameter estimation becomes an exercise in hypothetical subtraction: the maximally likely waveform is the one that, when taken out of the data, yields the least noisy residual.*

noise realization → likelihood map



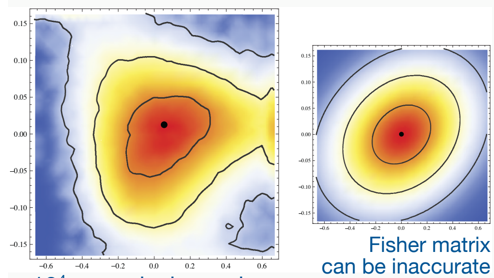
4 *For a single experiment (i.e., a given noise realization) we can then map the likelihood of source parameters. Markov Chain Monte Carlos are a good way to explore these maps.*

many noises → many maps



5 *For each possible noise, we would get a different map, and therefore different parameter-estimation uncertainties. To study measurement prospects, we need to characterize these variations.*

goal: map the distribution of the maximum-likelihood estimator



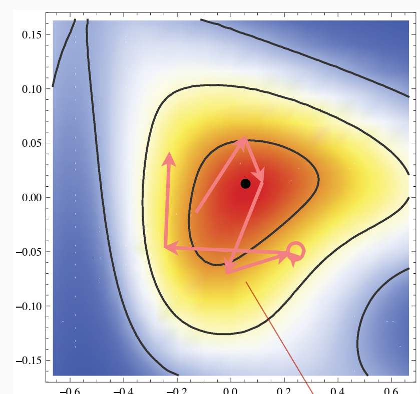
6 *Unfortunately, computing maps for all noises is unfeasible (an exaproblem!). Instead we derive the exact distribution of the maximum-likelihood estimator (the pink dot) over all noise realizations.*

- θ^i ≡ *source parameters* ($i = 1,\dots,d$);
  $h(\theta^i)$ ≡ waveform (of length $N \gg d$);
  (x,y) ≡ noise-weighted correlation product such that $p(n) = \exp[-(n,n)/2]$
- log likelihood:
  $$\log p(\theta) = -\big(n + h(\theta_{\rm true}) - h(\theta), n + h(\theta_{\rm true}) - h(\theta)\big)/2$$
- max-likelihood (ML) equation:
  $$\underbrace{\big(\partial_i h(\theta), n + h(\theta_{\rm true}) - h(\theta)\big)}_{\mathrm{ML}_i(\theta;\, n,\, \theta_{\rm true})} = 0$$
- formal solution of ML equation: $\theta_{\rm ML}(n, \theta_{\rm true})$
- formal distribution of ML estimator:
  $$p(\theta) = \int \delta\big(\theta_{\rm ML}(n, \theta_{\rm true}) - \theta\big)\, p(n)\, dn$$
- but it's much easier to solve the ML equation for n than for θ:
  $$\frac{\delta(\theta_{\rm ML}(n, \theta_{\rm true}) - \theta)}{|\partial \mathrm{ML}_i / \partial\theta_j|} = \delta(\mathrm{ML}_i(\theta; n, \theta_{\rm true}))$$
- thus we get the integral
  $$\boxed{p(\theta) = \mathcal{N} \int \delta(\mathrm{ML}_1(n)) \cdots \delta(\mathrm{ML}_d(n))\, |\partial \mathrm{ML}_i/\partial\theta_j|\, e^{-(n,n)/2} dn}$$
  performed over ~d^2 dimensions, not N!

- why? The noise can be decomposed in a basis that orthonormalizes the derivatives of the signal
  $$\Delta h = h(\theta) - h(\theta_{\rm true})$$
  $$n = N^k \hat{n}_k,$$
  $$h_i = C_i^k \hat{n}_k, \quad k = 1 \text{ to } d \qquad \mathrm{ML}_i = C_i^k N_k - C_i^k(\Delta h, \hat{n}_k),$$
  $$h_{i,j} = C_{ij}^k \hat{n}_k, \; k = 1 \text{ to } d{+}d(d{+}1)/2 \qquad \mathrm{ML}_{i,j} = C_{ij}^m N_m - C_{ij}^m(\Delta h, \hat{n}_m) - C_i^k C_{jk}$$
- now, all the N_k integrate out except for the projections of n over h_i and h_{ij}; in addition, the δs satisfy the integration over the projections on the h_i
  $$p(\theta) = \frac{e^{-\Sigma_{k=1}^d (\Delta h, \hat{n}_k)^2/2}}{(2\pi)^{d(d+3)/4} \prod_{k=1}^d C_k^k} \int \left| C_{ij}^m N_m - C_{ij}^m(\Delta h, \hat{n}_m) - C_i^k C_{jk} \right| e^{-\Sigma_m (N_m)^2/2} dN_m$$
  $$\boxed{p(\theta) = \frac{e^{-(\Delta h, h_i)(F^{-1})^{ij}(\Delta h, h_j)/2}}{\sqrt{(2\pi)^d |F_{ij}|}} \times \frac{1}{\sqrt{(2\pi)^{d(d-1)/2} |D_{\mu\nu}|}} \int \left| F_{ij} + (\Delta h, h_{ij}) - M_{(ij)} \right| e^{-M_\mu (D^{-1})^{\mu\nu} M_\nu/2} dM_\mu}$$
- ...where the $M_\mu \equiv M_{ij}$ are normal random variables with covariance matrix given by $D_{\mu\nu} \equiv (Ph_{ij}, Ph_{kl})$, where P denotes projection orthogonal to the h_k
- this $d(d+1)/2$-dimensional integral is trivial numerically. All that's needed are the F_{ij} and $D_{\mu\nu}$, which require ~$d^4/8$ inner products (each an N-point FFT)

7 *Our purpose is to map the distribution of the maximally likely source parameters $\theta_{\rm ml}$ over all noise realizations n. We can do this by enumerating the n, figuring out the $\theta_{\rm ml}$ corresponding to each, and accumulating the resulting distribution. However, it is much more efficient to enumerate the $\theta_{\rm ml}$, and compute the total probability weight of the n that are compatible with each. Surprisingly, this involves only low-dimensional integrals.*

result: we obtain the ML distribution with 1000x less computation



8 *This technique generates the exact frequentist error of the maximum-likelihood estimator for any SNR, and can be used to seed Bayesian-inference Markov-chain Monte Carlos.*