Suggested exercises for stochastic GW background lectures

J.D. Romano

HUST Summer School, July 2022

Abstract

Some suggested exercises accompanying the lectures on searches for stochastic gravitational-wave backgrounds.

Contents

1. *Practical application of Bayes' theorem*

   Suppose on your last visit to the doctor's office you took a test for some rare disease. This type of disease occurs in only 1 out of 10,000 people, as determined by a random sample of the population. The test that you took is rather effective in that it can correctly identify the presence of the disease 95% of the time, but it gives false positives 1% of the time.

   Suppose the test came up positive. What is the probability that you have the disease?

2. *Comparing frequentist and Bayesian analyses for a constant signal in white noise*

   Consider a constant amplitude signal $a = \text{const}$ in N samples of white Gaussian noise with fixed known variance σ^2. The likelihood functions describing the noise-only and signal+noise models are

$$p(d|\mathcal{M}_0) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^N d_i^2\right] \tag{1}$$

   and

$$p(d|a,\mathcal{M}_1) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^N (d_i - a)^2\right] \tag{2}$$

   respectively. Asumme for the Bayesian analyses that the amplitude a of the signal is described by the prior distribution

$$p(a|\mathcal{M}_1) = \frac{1}{a_{\max}} \tag{3}$$

   This exercise asks you to perform both frequentist and Bayesian analyses for the above signal and noise models, calculating various analytic expressions (see below) and doing numerical calculation for simulated data.

   For the analytic calculations:

   (a) show that the ML estimator of a is given by

$$\hat{a} \equiv a_{\mathrm{ML}}(d) = \frac{1}{N}\sum_{i=1}^N d_i \equiv \bar{d} \tag{4}$$

   (b) derive the following identity

$$\sum_{i=1}^N (d_i - a)^2 = \sum_i d_i^2 - N\hat{a}^2 + N(a - \hat{a})^2 = N\left(\mathrm{Var}[d] + (a - \hat{a})^2\right) \tag{5}$$

   (c) show that the likelihood function $p(d|a,\mathcal{M}_1)$ can be rewritten in terms of $\hat{a}$ as

$$p(d|a,\mathcal{M}_1) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \exp\left[-\frac{\mathrm{Var}[d]}{2\sigma_{\hat{a}}^2}\right] \exp\left[-\frac{(a-\hat{a})^2}{2\sigma_{\hat{a}}^2}\right] \tag{6}$$

   (d) show that the evidence is given by

$$p(d|\mathcal{M}_1) = \frac{\exp\left[-\frac{\mathrm{Var}[d]}{2\sigma_{\hat{a}}^2}\right]\left[\mathrm{erf}\left(\frac{a_{\max}-\hat{a}}{\sqrt{2}\sigma_{\hat{a}}}\right) + \mathrm{erf}\left(\frac{\hat{a}}{\sqrt{2}\sigma_{\hat{a}}}\right)\right]}{2a_{\max}\left(\sqrt{2\pi}\sigma\right)^{N-1}\sqrt{N}} \tag{7}$$

   (e) show that the posterior distribution for a is given by

$$p(a|d,\mathcal{M}_1) = \frac{1}{\sqrt{2\pi}\sigma_{\hat{a}}}\exp\left[-\frac{(a-\hat{a})^2}{2\sigma_{\hat{a}}^2}\right] 2\left[\mathrm{erf}\left(\frac{a_{\max}-\hat{a}}{\sqrt{2}\sigma_{\hat{a}}}\right) + \mathrm{erf}\left(\frac{\hat{a}}{\sqrt{2}\sigma_{\hat{a}}}\right)\right]^{-1} \tag{8}$$

(f) show that the Bayes factor is

$$\mathcal{B}_{10}(d) = \exp\left[\frac{\hat{a}^2}{2\sigma_{\hat{a}}^2}\right]\left(\frac{\sqrt{2\pi}\sigma_{\hat{a}}}{a_{\max}}\right)\frac{1}{2}\left[\text{erf}\left(\frac{a_{\max}-\hat{a}}{\sqrt{2}\sigma_{\hat{a}}}\right)+\text{erf}\left(\frac{\hat{a}}{\sqrt{2}\sigma_{\hat{a}}}\right)\right] \tag{9}$$

$$\simeq \exp\left[\frac{\hat{a}^2}{2\sigma_{\hat{a}}^2}\right]\left(\frac{\sqrt{2\pi}\sigma_{\hat{a}}}{a_{\max}}\right) \tag{10}$$

where the last approximate equality uses the Laplace approximation.

(g) show that the maximum-likelihood ratio statistic is

$$\Lambda_{\rm ML}(d) = \exp\left(\frac{\hat{a}^2}{2\sigma_{\hat{a}}^2}\right) \tag{11}$$

(h) show that frequentist test statistic $\Lambda(d)$ constructed from $\Lambda_{\rm ML}(d)$ is

$$\Lambda(d) \equiv 2\ln\Lambda_{\rm ML}(d) = \frac{\hat{a}^2}{\sigma_{\hat{a}}^2} = \left(\frac{\sqrt{N}\bar{d}}{\sigma}\right)^2 \equiv \rho^2 \tag{12}$$

which is just the squared SNR of the data.

(i) show that the sampling distributions of the test statistic are

$$p(\Lambda|\mathcal{M}_0) = \frac{1}{\sqrt{2\pi\Lambda}}e^{-\Lambda/2} \tag{13}$$

$$p(\Lambda|a,\mathcal{M}_1) = \frac{1}{\sqrt{2\pi\Lambda}}\frac{1}{2}\left[e^{-\frac{1}{2}(\sqrt{\Lambda}-\sqrt{\lambda})^2}+e^{-\frac{1}{2}(\sqrt{\Lambda}+\sqrt{\lambda})^2}\right] \tag{14}$$

where

$$\lambda = \langle\rho\rangle^2 = \frac{Na^2}{\sigma^2} \tag{15}$$

is the non-centrality parameter for a non-central chi-square distribution with one degree of freedom.

For the numerical simulations, take e.g.,

$$N = 100\,,\quad \sigma = 1\,,\quad 0 \le a \le a_{\max} = 1\,,\quad a_0 = 0.1 = \text{true value}\,, \tag{16}$$

but these values can be changed to consider stronger (or weaker) injections, etc.

For a realization of simulated data, calculate the following quantities, making plots when possible:

(a) the value of the threshold Λ_* corresponding to a false alarm probability $\alpha = 0.1$.

(b) the observed value $\Lambda_{\rm obs}$ of the test statistic $\Lambda(d)$ and its corresponding p value.

(c) the frequentist detection probability curve, and the value $a^{90\%,\rm DP}$ of a needed for 90% detection probability.

(d) the frequentist 95% confidence interval $[\hat{a}-\sigma_{\hat{a}}, \hat{a}+\sigma_{\hat{a}}]$.

(e) the frequentist 90% confidence level upper limit $a^{90\%,\rm UL}$.

(f) the Bayesian 95% credible interval centered on the mode of $p(a|d,\mathcal{M}_1)$.

(g) the Bayesian 90% credible upper limit $a^{90\%,\rm UL}$.

(h) the Bayes factor $\mathcal{B}_{10}(d)$, twice the log of the Bayes factor $2\ln\mathcal{B}_{10}(d)$, and the Laplace approximation of that quantity, see Eq. (10).

3. *Rate estimate of stellar-mass binary black hole mergers:*

Estimate the total rate (number of events per time) of stellar-mass binary black hole mergers throughout the universe by multiplying LIGO's local rate estimate $R_0 \sim 10$ - $200\ \text{Gpc}^{-3}\,\text{yr}^{-1}$ by the comoving volume out to some large redshift, e.g., $z = 10$. (For this calculation you can ignore any dependence of the rate density with redshift.) You should find a merger rate of $\sim$1 per minute to a few per hour.

Hint: You will need to do numerically evaluate the following integral for proper distance today as a function of source redshift:

$$d_0(z) = \frac{c}{H_0}\int_0^z \frac{dz'}{E(z')}\,, \qquad E(z) \equiv \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}\,, \tag{17}$$

with

$$\Omega_m = 0.31\,, \qquad \Omega_\Lambda = 0.69\,, \qquad H_0 = 68\ \mathrm{km\,s^{-1}\,Mpc^{-1}}\,. \tag{18}$$

Doing that integral, you should find what's shown in Figure 1, which you can then evaluate at $z = 10$ to convert R_0 (number of events per comoving volume per time) to total rate (number of events per time) for sources out to redshift $z = 10$.

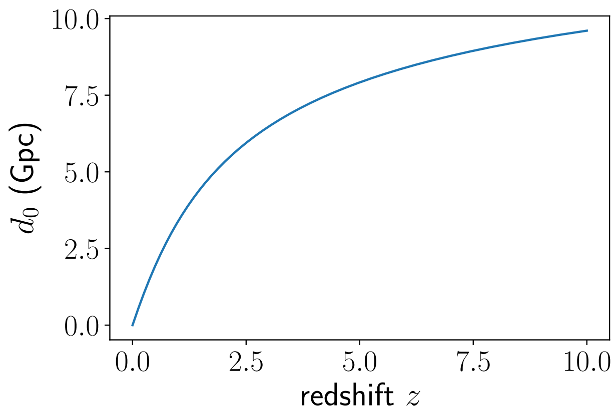

Figure 1:

4. *Relating $S_h(f)$ and $\Omega_{gw}(f)$:*
   Derive the relationship

$$S_h(f) = \frac{3H_0^2}{2\pi^2}\frac{\Omega_{gw}(f)}{f^3} \tag{19}$$

   between the strain power spectral density $S_h(f)$ and the dimensionless fractional energy density spectrum $\Omega_{gw}(f)$. (*Hint*: You will need to use the various definitions of these quantities and also

$$\rho_{gw} = \frac{c^2}{32\pi G}\langle \dot{h}_{ab}(t,\vec{x})\dot{h}^{ab}(t,\vec{x})\rangle\,, \tag{20}$$

   which expresses the energy-density in gravitational-waves to the metric perturbations $h_{ab}(t,\vec{x})$.)

5. *Cosmology and the "Phinney formula" for astrophysical backgrounds:*
   (a) Using the Friedmann equation

$$\left(\frac{\dot{a}}{a}\right)^2 = H_0^2\left(\frac{\Omega_m}{a^3} + \Omega_\Lambda\right) \tag{21}$$

   for a spatially-flat FRW spacetime with matter and cosmological constant, and the relationship

$$1 + z = \frac{1}{a(t)}\,, \qquad a(t_0) \equiv 1 \quad (t_0 \equiv \text{today})\,, \tag{22}$$

between redshift z and scale factor $a(t)$, derive

$$\frac{\mathrm{d}t}{\mathrm{d}z} = -\frac{1}{(1+z)H_0 E(z)}\,, \qquad E(z) = \sqrt{\Omega_{\rm m}(1+z)^3 + \Omega_\Lambda}\,. \tag{23}$$

(b) Using this result for $\mathrm{d}t/\mathrm{d}z$, show that

$$\Omega_{\rm gw}(f) = \frac{f}{\rho_{\rm c} H_0}\int_0^\infty \mathrm{d}z\, R(z)\,\frac{1}{(1+z)E(z)}\left(\frac{\mathrm{d}E_{\rm gw}}{\mathrm{d}f_{\rm s}}\right)\Bigg|_{f_{\rm s}=f(1+z)} \tag{24}$$

in terms of the rate density $R(z)$ as measured in the source frame (number of events per comoving volume per time interval in the source frame). (*Hint*: The expression for $\mathrm{d}t/\mathrm{d}z$ from part (a) will allow you to go from the "Phinney formula" for $\Omega_{\rm gw}(f)$ written in terms of the number density $n(z)$,

$$\Omega_{\rm gw}(f) = \frac{1}{\rho_c}\int_0^\infty \mathrm{d}z\, n(z)\,\frac{1}{1+z}\left(f_{\rm s}\,\frac{\mathrm{d}E_{\rm gw}}{\mathrm{d}f_{\rm s}}\right)\Bigg|_{f_{\rm s}=f(1+z)}\,, \tag{25}$$

to one in terms of the rate density $R(z)$, where $n(z)\,\mathrm{d}z = R(z)\,|\mathrm{d}t|_{t=t(z)}$. Note: Both of the above expressions for $\Omega_{\rm gw}(f)$ assume that there is only one type of source, described by some set of average source parameters. If there is more than one type of source, one must sum the contributions of each source to $\Omega_{\rm gw}(f)$.)

6. *Optimal filtering for the cross-correlation statistic:*
Verify the form

$$\tilde{Q}(f) \propto \frac{\Gamma_{12}(f)H(f)}{P_1(f)P_2(f)}\,, \tag{26}$$

of the optimal filter function in the weak-signal limit, where $H(f)$ is the assumed spectral shape of the gravitational-wave background, $\Gamma_{12}(f)$ is the overlap function, and $P_1(f)$, $P_2(f)$ are the power spectral densities of the outputs of the two detectors (which are approximately equal to $P_{n_1}(f)$, $P_{n_2}(f)$, respectively). Recall that the optimal filter $\tilde{Q}(f)$ maximizes the signal-to-noise ratio of the cross-correlation statistic. (*Hint*: Introduce an inner product on the space of functions of frequency $A(f)$, $B(f)$:

$$(A, B) \equiv \int df\, A(f)B^*(f)P_1(f)P_2(f)\,. \tag{27}$$

This inner product has all of the properties of the familiar dot product of vectors in 3-dimensional space. The signal-to-noise ratio of the cross-correlation statistic can be written in terms of this inner product.)

7. *Maximum-likelihood estimators for single and multiple parameters:*
(a) Show that the maximum-likelihood estimator $\hat{a}$ of the single parameter a in the likelihood function

$$p(d|a,\sigma) \propto \exp\left[-\frac{1}{2}\sum_{i=1}^{N}\frac{(d_i-a)^2}{\sigma_i^2}\right] \tag{28}$$

is given by the noise-weighted average

$$\hat{a} = \sum_i \frac{d_i}{\sigma_i^2}\Bigg/\sum_j \frac{1}{\sigma_j^2}\,. \tag{29}$$

(b) Extend the previous calculation to the likelihood

$$p(d|A,C) \propto \exp\left[-\frac{1}{2}(d-MA)^\dagger C^{-1}(d-MA)\right]\,, \tag{30}$$

where $A \equiv A_\alpha$ is a vector of parameters, $C \equiv C_{ij}$ is the noise covariance matrix, and $M \equiv M_{i\alpha}$ is the response matrix mapping A_α to data samples, $MA \equiv \sum_\alpha M_{i\alpha} A_\alpha$. For this more general case you should find:

$$\hat{A} = F^{-1} X \,, \tag{31}$$

where

$$F \equiv M^\dagger C^{-1} M \,, \qquad X \equiv M^\dagger C^{-1} d \,. \tag{32}$$

In general, the matrix F (called the *Fisher* matrix) is not invertible, so some sort of regularization is needed to do the matrix inversion.

8. *Timing-residual response for a 1-arm, 1-way detector:*
   Derive the timing residual reponse function

$$R^A(f, \hat{k}) = \frac{1}{2} u^a u^b e^A_{ab}(\hat{k}) \frac{1}{i2\pi f} \frac{1}{1 - \hat{k}\cdot\hat{u}} \left[1 - e^{-\frac{i2\pi f L}{c}(1 - \hat{k}\cdot\hat{u})}\right] \tag{33}$$

   for a single-link (i.e., a one-arm, one-way detector like that for pulsar timing). Here $\hat{u}$ is the direction of propagation of the electromagnetic pulse (the direction to the pulsar $\hat{p} = -\hat{u}$), and $\hat{k}$ is the direction of propagation of the GW. The origin of coordinates is taken to be at the position of the detector.

9. *Overlap function for colocated electric dipole antennae:*
   Show that the overlap function for a pair of (short) colocated electric dipole antennae pointing in directions $\hat{u}_1$ and $\hat{u}_2$ is given by

$$\Gamma_{12} \propto \hat{u}_1 \cdot \hat{u}_2 \equiv \cos\zeta \tag{34}$$

   for the case of an unpolarized, isotropic electromagnetic field. (*Hint*: "short" means that the phase of the electric field can be taken to be constant over of the lengths of the dipole antennae, so that the reponse of antenna $I = 1, 2$ to the field is given by $r_I(t) = \hat{u}_I \cdot \vec{E}(t, \vec{x}_0)$, where $\vec{x}_0$ is the common location of the two antenna.)

10. *Maximum-likelihood estimators for the standard cross-correlation statistic:*
   Verify that

$$\hat{C}_{11} \equiv \frac{1}{N} \sum_{i=1}^N d_{1i}^2 \,, \qquad \hat{C}_{22} \equiv \frac{1}{N} \sum_{i=1}^N d_{2i}^2 \,, \qquad \hat{C}_{12} \equiv \frac{1}{N} \sum_{i=1}^N d_{1i} d_{2i} \tag{35}$$

   are maximum-likelihood estimators of

$$S_1 \equiv S_{n_1} + S_h \,, \quad S_2 \equiv S_{n_2} + S_h \,, \quad S_h \,, \tag{36}$$

   for the case of N samples of a white GWB in uncorrelated white detector noise, for a pair of colocated and coaligned detectors. Recall that the likelihood function is

$$p(d|S_{n_1}, S_{n_2}, S_h) = \frac{1}{\sqrt{\det(2\pi C)}} \exp\left[-\frac{1}{2} d^T C^{-1} d\right] \,, \tag{37}$$

   where

$$C = \begin{bmatrix} (S_{n_1} + S_h)\, 1_{N\times N} & S_h\, 1_{N\times N} \\ S_h\, 1_{N\times N} & (S_{n_2} + S_h)\, 1_{N\times N} \end{bmatrix} \tag{38}$$

   and

$$d^T C^{-1} d \equiv \sum_{I,J=1}^2 \sum_{i,j=1}^N d_{Ii} \left(C^{-1}\right)_{Ii,Jj} d_{Jj} \,. \tag{39}$$

11. *Derivation of the maximum-likelihood ratio detection statistic:*
Verify that twice the log of the maximum-likelihood ratio for the standard stochastic likelihood function goes like the square of the (power) signal-to-noise ratio,

$$2\ln\Lambda_{\rm ML}(d) \simeq \frac{\hat{C}_{12}^2}{\hat{C}_{11}\hat{C}_{22}/N}\,, \tag{40}$$

in the weak-signal approximation. (*Hint:* For simplicity, do the calculation in the context of N samples of a white GWB in uncorrelated white detector noise, for a pair of colocated and coaligned detectors, using the results of Exercise 10.)

12. *Standard cross-correlation likelihood by marginalizing over stochastic signal prior:*
Derive the standard form of the likelihood function for stochastic background searches

$$p(d|S_{n_1},S_{n_2},S_h) = \frac{1}{\sqrt{\det(2\pi C)}}\exp\left[-\frac{1}{2}\sum_{I,J=1}^{2} d_I\left(C^{-1}\right)_{IJ}d_J\right]\,, \tag{41}$$

where

$$C \equiv \left[\begin{array}{cc} S_{n_1}+S_h & S_h \\ S_h & S_{n_2}+S_h\end{array}\right]\,, \tag{42}$$

by marginalizing

$$p_n(d-h|S_{n_1},S_{n_2}) = \frac{1}{2\pi\sqrt{S_{n_1}S_{n_2}}}\exp\left[-\frac{1}{2}\left\{\frac{(d_1-h)^2}{S_{n_1}}+\frac{(d_2-h)^2}{S_{n_2}}\right\}\right] \tag{43}$$

over the signal samples h for the *stochastic* signal prior

$$p(h|S_h) = \frac{1}{\sqrt{2\pi S_h}}\exp\left[-\frac{1}{2}\frac{h^2}{S_h}\right]\,. \tag{44}$$

In other words, show that

$$p(d|S_{n_1},S_{n_2},S_h) = \int_{-\infty}^{\infty}{\rm d}h\; p_n(d-h|S_{n_1},S_{n_2})p(h|S_h)\,. \tag{45}$$

(*Hint*: You'll have to complete the square in the argument of the exponential in the marginalization integral.)