Christos' notes on power spectra

Statistics of noise in power spectra

In this section, we take a timestream signal that can be decomposed in the true signal + noise. Consider just the discrete noise signal:

$$x[n], \quad n = 0, 1, \ldots, N-1,$$

where each $x[n]$ is an independent and identically distributed (i.i.d.) Gaussian random variable with zero mean and variance σ^2:

$$x[n] \sim \mathcal{N}(0, \sigma^2).$$

When making power spectra we first take the fourier transform and the take the absolute square of the real and imaginary components of the complex Fourier amplitude. The discrete Fourier transform (DFT) of the sequence $x[n]$ at frequency bin k is defined as:

$$X[k] = \sum_{n=0}^{N-1} x[n] \, e^{-j \frac{2\pi}{N} kn}.$$

Since $x[n]$ are Gaussian and the DFT is a linear transform, $X[k]$ is a complex Gaussian random variable:

$$X[k] = X_r[k] + j X_i[k],$$

where

$$X_r[k] = \sum_{n=0}^{N-1} x[n] \cos\left(\frac{2\pi}{N} kn\right), \quad X_i[k] = -\sum_{n=0}^{N-1} x[n] \sin\left(\frac{2\pi}{N} kn\right).$$

Both $X_r[k]$ and $X_i[k]$ are linear combinations of Gaussian variables, which makes them also Gaussian. I am now interested in the statistical properties of the real and imaginary part of the Fourier transformed noise signal.

Because each $x[n]$ has zero mean,

$$\mathbb{E}[X_r[k]] = \mathbb{E}[X_i[k]] = 0.$$

The variances are:

$$\mathrm{Var}[X_r[k]] = \sigma^2 \sum_{n=0}^{N-1} \cos^2\left(\frac{2\pi}{N} kn\right),$$

$$\mathrm{Var}[X_i[k]] = \sigma^2 \sum_{n=0}^{N-1} \sin^2\left(\frac{2\pi}{N}kn\right).$$

Using the trigonometric identity $\cos^2\theta + \sin^2\theta = 1$, we have

$$\sum_{n=0}^{N-1} \cos^2\left(\frac{2\pi}{N}kn\right) + \sum_{n=0}^{N-1} \sin^2\left(\frac{2\pi}{N}kn\right) = N.$$

The orthogonality of the DFT basis vectors implies that the sums are equal:

$$\sum_{n=0}^{N-1} \cos^2\left(\frac{2\pi}{N}kn\right) = \sum_{n=0}^{N-1} \sin^2\left(\frac{2\pi}{N}kn\right) = \frac{N}{2}.$$

Hence,

$$\boxed{\mathrm{Var}[X_r[k]] = \mathrm{Var}[X_i[k]] = \frac{N}{2}\sigma^2.}$$

The power at frequency bin k is the squared magnitude of $X[k]$:

$$P[k] = |X[k]|^2 = X_r[k]^2 + X_i[k]^2.$$

Since $X_r[k]$ and $X_i[k]$ are independent zero-mean Gaussian variables with equal variance, $P[k]$ is the sum of squares of two independent Gaussian random variables.

Recall that if Z is a zero-mean Gaussian with variance σ^2, then Z^2/σ^2 follows a chi-square distribution with 1 degree of freedom:

$$\frac{Z^2}{\sigma^2} \sim \chi_1^2.$$

Therefore,

$$\frac{X_r[k]^2}{\mathrm{Var}[X_r[k]]} + \frac{X_i[k]^2}{\mathrm{Var}[X_i[k]]} \sim \chi_2^2,$$

which is a chi-square distribution with 2 degrees of freedom.
Thus, the power $P[k]$ is distributed as

$$P[k] \sim \text{Scaled } \chi_2^2,$$

with scale parameter $\mathrm{Var}[X_r[k]] = \frac{N}{2}\sigma^2$.

Probability density function (PDF) of $P[k]$: A χ^2 distribution with 2 degrees of freedom is equivalent to an exponential distribution with parameter $\frac{1}{2}$:

$$f_Z(z) = \frac{1}{2}e^{-z/2}, \quad z \geq 0.$$

Therefore, the PDF of $P[k]$ is exponential with mean $\lambda = \frac{N}{2}\sigma^2$:

$$f_{P[k]}(p) = \frac{1}{\lambda}e^{-p/\lambda}, \quad p \geq 0.$$

Substituting in and writin in terms of x to match Colin's notation we get:

$$\boxed{P_\nu(x) = \frac{1}{N\sigma^2}\exp\left(-\frac{x}{N\sigma^2}\right), \quad x \geq 0.}$$

Which is equation (2) in Colin's document, assuming $\sigma = 1$.

Note on χ^2 distributions

Let

$$Z_1, Z_2, \ldots, Z_k \sim \mathcal{N}(0,1)$$

be independent and identically distributed standard normal random variables. Define the random variable:

$$X = \sum_{i=1}^{k} Z_i^2.$$

The goal is to find the probability density function (PDF) $f_X(x)$ of X, where x is the actual value that the random variable X takes.

Express X as squared norm of a Gaussian vector
Consider the vector

$$\mathbf{Z} = (Z_1, Z_2, \ldots, Z_k) \in \mathbb{R}^k.$$

Since each $Z_i \sim \mathcal{N}(0,1)$ independently, the joint PDF of $\mathbf{Z}$ is

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{1}{(2\pi)^{k/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{k} z_i^2\right) = \frac{1}{(2\pi)^{k/2}} e^{-\frac{r^2}{2}},$$

where

$$r = \|\mathbf{z}\| = \sqrt{z_1^2 + z_2^2 + \cdots + z_k^2}.$$

Note that:

$$X = r^2.$$

This might seem a little strange, but essentially what we have done by defining the vector Z is, to define a k-dimensional space where Z lives. Since $(Z_1, Z_2, \ldots, Z_k)$ are i.i.d. we have k indipendent dimensions to this space, and r is the radial coordinate of this k-dimensional space.

Spherical coordinates
In k-dimensional Euclidean space, the volume element in spherical coordinates is given by

$$dV = r^{k-1} dr\, d\Omega,$$

where

- $r \geq 0$ is the radius,
- $d\Omega$ is the differential solid angle on the unit $(k-1)$-sphere,
- The surface area of the unit sphere in $\mathbb{R}^k$ is

$$S_k = \frac{2\pi^{k/2}}{\Gamma(k/2)}.$$

Cumulative distribution function (CDF) of X

The CDF of X is

$$F_X(x) = P(X \le x) = P(r^2 \le x) = P(r \le \sqrt{x}).$$

By integrating over the spherical shell,

$$F_X(x) = \int_{\|\mathbf{z}\| \le \sqrt{x}} f_{\mathbf{Z}}(\mathbf{z})\, dV = \int_0^{\sqrt{x}} \int_\Omega \frac{1}{(2\pi)^{k/2}} e^{-r^2/2} r^{k-1}\, d\Omega\, dr.$$

Because the density is radially symmetric, the integral over the angles is just the surface area S_k:

$$F_X(x) = \frac{S_k}{(2\pi)^{k/2}} \int_0^{\sqrt{x}} e^{-r^2/2} r^{k-1}\, dr.$$

Get the PDF of X

Differentiate $F_X(x)$ with respect to x using the Leibniz rule to account for the change in boundaries since the integration limits depend on the parameter we are differentiating:

$$f_X(x) = \frac{d}{dx} F_X(x) = \frac{d}{dx} \left[\frac{S_k}{(2\pi)^{k/2}} \int_0^{\sqrt{x}} e^{-r^2/2} r^{k-1}\, dr \right].$$

Since the upper limit depends on x, we have

$$f_X(x) = \frac{S_k}{(2\pi)^{k/2}} e^{-x/2} (\sqrt{x})^{k-1} \frac{d}{dx} \sqrt{x}.$$

Calculate the derivative:

$$\frac{d}{dx} \sqrt{x} = \frac{1}{2\sqrt{x}}.$$

Therefore,

$$f_X(x) = \frac{S_k}{(2\pi)^{k/2}} e^{-x/2} x^{\frac{k-1}{2}} \cdot \frac{1}{2\sqrt{x}} = \frac{S_k}{(2\pi)^{k/2}} e^{-x/2} \frac{x^{\frac{k}{2}-1}}{2}.$$

Substituting into the expression for $f_X(x)$:

$$f_X(x) = \frac{1}{(2\pi)^{k/2}} \cdot \frac{2\pi^{k/2}}{\Gamma(k/2)} \cdot e^{-x/2} \cdot \frac{x^{\frac{k}{2}-1}}{2} = \frac{1}{2^{k/2}\Gamma(k/2)} x^{\frac{k}{2}-1} e^{-x/2}.$$

Finally, the PDF of the chi-square distribution with k degrees of freedom is:

$$\boxed{f_X(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{\frac{k}{2}-1} e^{-x/2}, \quad x \ge 0.}$$

which is what we used in the case of 2 degrees of freedom in the previous section.

Practical example: Noise in ground maps

Take simulated noise timestream data as in figure 1 and take the power spectrum. Then use the likelihood approach described in Colin's document to find the parameters that fit the model best.

Do this by minimizing the $-\log$ of the likelihood, which is mathematically equivalent to maximizinf the $-\log$ of the likelihood.

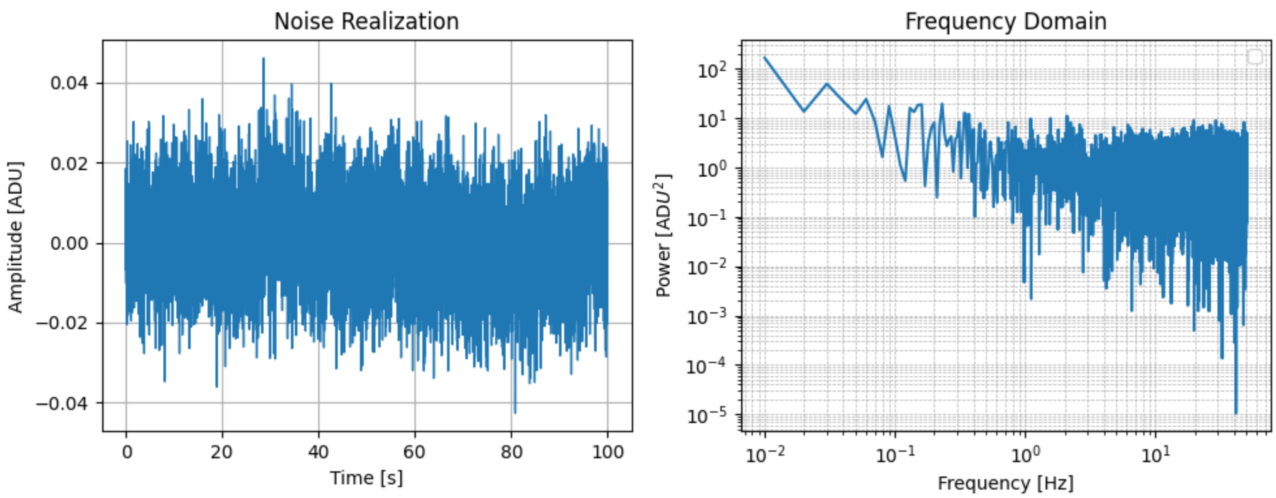

Figure 1: Fit of the noise model to simulated noise spectrum data.

See code snippet below for the likelihood function to be minimized:

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import minimize

# -----------------------------
# Define the Noise Model
# -----------------------------
def noise_model(nu, A, nu_star, alpha):
    return A * (1 + (nu / nu_star)**alpha)
# -----------------------------
# Log-Likelihood Function
# -----------------------------
def log_likelihood(params, nu, x_nu):
    A, nu_star, alpha = params
    if A <= 0 or nu_star <= 0 or alpha >=0:
        return -np.inf  # penalize invalid parameters

    N = noise_model(nu, A, nu_star, alpha)
    if np.any(N <= 0) or np.any(np.isnan(N)) or np.any(np.isinf(N)):
        return -np.inf

```

```python
    return -np.sum(np.log(N) + x_nu / N)
# ------------------------------
# Generate Simulated Data
# ------------------------------
np.random.seed(42)

# True parameters for simulation
A_true = 1.0
nu_star_true = 1.0
alpha_true = -1.0

# # Frequency bins
# nu = np.linspace(0.001, 10, 1000)

# # Generate simulated data
# N_true = noise_model(nu, A_true, nu_star_true, alpha_true)

# # Add exponential noise (same as the assumed likelihood)
# x_nu = np.random.exponential(N_true)

# Use above simulated data of figure 1
nu = freqs[1:]
x_nu = power_spec[1:]
nu_star_true = f_knee
N_true = noise_model(nu, A_true, nu_star_true, alpha_true)
# ------------------------------
# Fit the Model
# ------------------------------
def neg_log_likelihood(params):
    return -log_likelihood(params, nu, x_nu)

initial_guess = [2.0, 2, -1.5]  # [A, nu_star, alpha]

result = minimize(neg_log_likelihood, initial_guess,
                  bounds=[(1e-6, None), (1e-6, None), (None, None)],
                  method='L-BFGS-B')

A_fit, nu_star_fit, alpha_fit = result.x

print("Best-fit parameters:")
print(f"A      = {A_fit:.4f}")
print(f"nu*    = {nu_star_fit:.4f}")
print(f"alpha  = {alpha_fit:.4f}")
print(result.message)
# ------------------------------
# Plot Results
# ------------------------------
plt.figure(figsize=(8, 8))
plt.plot(nu, x_nu, label='Simulated Data')
```

```python
plt.plot(nu, N_true, '--', label='True Model')
plt.plot(nu, noise_model(nu, A_fit, nu_star_fit, alpha_fit), '-',
    label='Fitted Model')
plt.xlabel('Frequency (v)')
plt.ylabel('Power [$ADU^2$]')
plt.xscale('log')
plt.yscale('log')
plt.legend()
plt.grid(True)
plt.title('Power Spectrum Model Fit -- Simulated Data')
plt.tight_layout()
plt.show()
```

Listing 1: Minimize negative log likelihood in Python

The parameters estimated using the maximum likelihood are:

```
Best-fit parameters:
A      = 1.0235
nu*    = 0.8522
alpha  = -1.0438
CONVERGENCE: RELATIVE REDUCTION OF F <= FACTR*EPSMCH
```

It seems to get the parameters very well with the only one being slightly off being the $\nu_\star$.Ploting the simulated data, true model, and estimated model looks like:

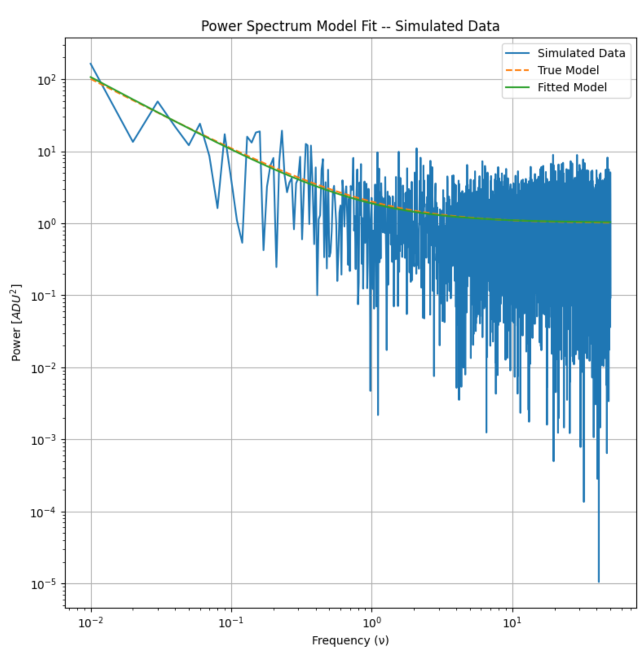

Figure 2: Fit of the noise model to simulated noise spectrum data.

Exercise for the reader

Given the probability density function

$$P_\nu(x) = \frac{1}{N_\nu} e^{-\frac{x}{N_\nu}}, \quad x \geq 0,$$

Calculate the expectation value of x

$$\langle x \rangle = \int_0^\infty x P_\nu(x)\, dx = \int_0^\infty x \frac{1}{N_\nu} e^{-\frac{x}{N_\nu}}\, dx.$$

Let

$$a = \frac{1}{N_\nu}.$$

Then,

$$\langle x \rangle = \int_0^\infty x a e^{-ax}\, dx.$$

Use integration by parts, since using integral tables is lame:

$$u = x \quad \Rightarrow \quad du = dx,$$

$$dv = a e^{-ax} dx \quad \Rightarrow \quad v = -e^{-ax}.$$

$$\begin{aligned}\langle x \rangle &= \int_0^\infty x a e^{-ax}\, dx \\ &= -x e^{-ax}\Big|_0^\infty + \int_0^\infty e^{-ax}\, dx.\end{aligned}$$

first term vanishes:

$$\lim_{x \to \infty} x e^{-ax} = 0, \quad \text{and} \quad x = 0 \implies 0,$$

$$-x e^{-ax}\Big|_0^\infty = 0 - 0 = 0.$$

The remaining integral is:

$$\int_0^\infty e^{-ax}\, dx = \frac{-1}{a} e^{-ax}\Big|_0^\infty = \frac{1}{a}.$$

Therefore:

$$\langle x \rangle = \frac{1}{a} = N_\nu.$$

Moment generating function

Given the probability density function (PDF):

$$P_\nu(x) = \frac{1}{N_\nu} \exp\left(-\frac{x}{N_\nu}\right), \quad x \geq 0,$$

where N_ν is a parameter depending on ν.

Calculate the k-th moment $\langle x^k \rangle$:

$$\langle x^k \rangle = \int_0^\infty x^k P_\nu(x)\, dx = \int_0^\infty x^k \frac{1}{N_\nu} \exp\left(-\frac{x}{N_\nu}\right) dx.$$

Notice how similar this integral looks to the integral definition of the gamma function. Let:

$$t = \frac{x}{N_\nu} \implies x = N_\nu t, \quad dx = N_\nu dt.$$

Then the integral becomes:

$$\langle x^k \rangle = \int_0^\infty (N_\nu t)^k \frac{1}{N_\nu} e^{-t} N_\nu dt = N_\nu^k \int_0^\infty t^k e^{-t} dt.$$

Gamma function definition:

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt.$$

Therefore:

$$\int_0^\infty t^k e^{-t} dt = \Gamma(k+1).$$

The k-th moment is

$$\boxed{\langle x^k \rangle = N_\nu^k \Gamma(k+1).}$$

Var(x) is:

$$\mathrm{Var}(x) = N_\nu^2$$