Notes about lasers, random processes, stochastic differential equations, power spectral densities

Error of time series

True series x_k, with $k = 1 \dots N$. Estimated series $\widehat{x_k}$. How to quantify the error between them?

Mean absolute error

$$MAE = \frac{1}{N}\sum_{k=1}^{N}|x_k - \widehat{x_k}|$$

Root men square error

$$RMSE = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(x_k - \widehat{x_k})^2}$$

Mean absolute percentage error

$$MAPE = \frac{1}{N}\sum_{k=1}^{N}\left|\left(1 - \frac{\widehat{x_k}}{x_k}\right) \cdot 100\right|$$

Symmetric mean absolute percentage error

$$sMAPE = \frac{1}{N}\sum_{k=1}^{N}\frac{2|x_k - \widehat{x_k}|}{|x_k| + |\widehat{x_k}|}$$

Numerical derivative approximation

Suppose we have a function of multivariable $f(\boldsymbol{x})$ and we want to approximate the gradient. We can use complex differentiation

$$\nabla f_i(\boldsymbol{x_0}) \approx \frac{imag[\,f(\boldsymbol{x_0} + ih\boldsymbol{e}_i)]}{h}$$

Correlation of phase noise traces

Electric field of an OFC

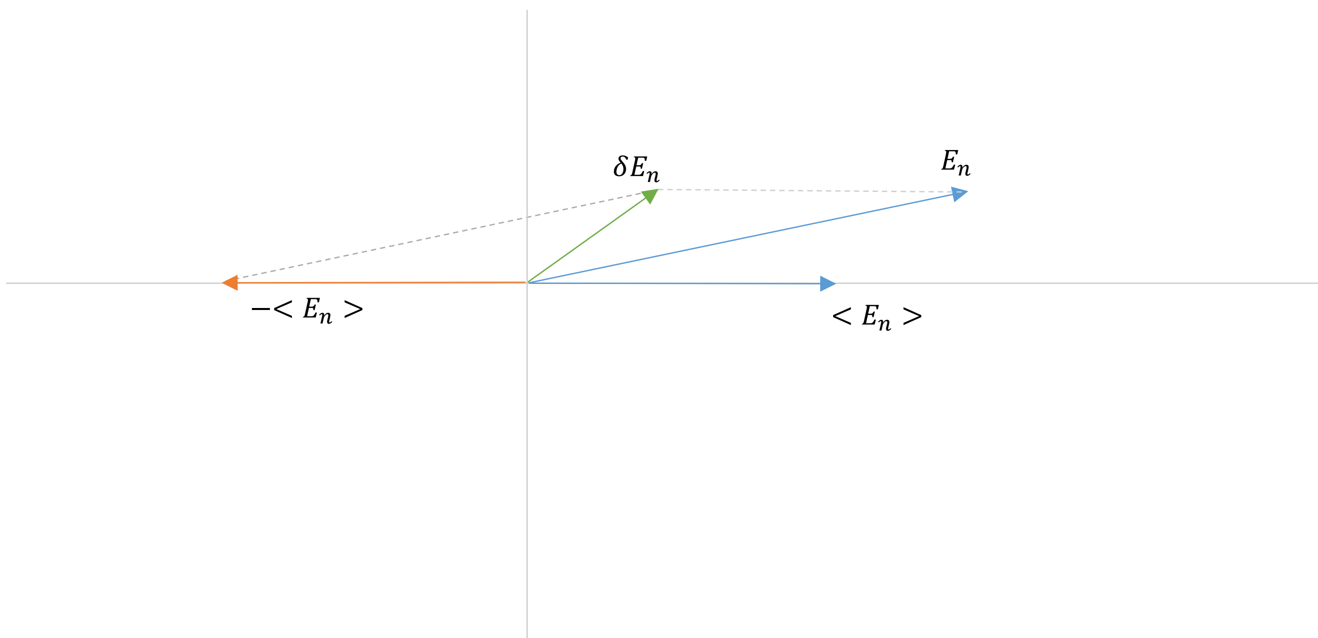
$$e(t) = \sum_n E_n = \sum_n A_n \exp\big(i(\phi_n - \omega_n t)\big)$$

Where $\omega_n = \omega_{CEO} + n\omega_{rep}$.

The noise affects the signal of individual lines, E_n. We can describe the effects of the perturbation as the deviation from the mean value $< E_n >$, using

$$\delta E_n = E_n - < E_n >$$

Using the following vector diagram



We operate in the hypothesis that $< \phi_n > = 0$. This implies a transform-limited pulse (?). But then, the analysis is simplified because

$$E_n - < E_n > = (A_n \cos\phi_n - < A_n >) + iA_n \sin\phi_n$$

Now, if we hypothesize very small angles, i.e. $\phi_n \to \delta\phi_n$ we can say that $\cos\phi_n \to 1$ and $\sin\phi_n \to \delta\phi_n$ [NOT SUPER CLEAR ARGUMENT]

$$\delta E_n = E_n - < E_n > = \ (A_n - < A_n >) + iA_n\delta\phi_n = \ \delta A_n + iA_n\delta\phi_n$$

■

Note: transform-limited pulse means that

Transform Limit

Definition: a limit for the time–bandwidth–product of an optical pulse

Category: light pulses

How to cite the article; suggest additional literature

Ask RP Photonics for advice on the mathematical basics of optical pulses, perhaps in the form of tailored in-house staff training.

In ultrafast optics, the transform limit (or Fourier limit, Fourier transform limit) is usually understood as the lower limit for the **pulse duration** which is possible for a given **optical spectrum** of a **pulse**. A pulse at this limit is called *transform limited*. The condition of being at the transform limit is essentially equivalent to the condition of a frequency-independent spectral phase (which leads to the maximum possible peak power), and basically implies that the **time–bandwidth product** is at its minimum and that there is no **chirp**. The minimum time–bandwidth product depends on the pulse shape, and is e.g. ≈ 0.315 for bandwidth-limited sech^2-shaped pulses and ≈ 0.44 for Gaussian-shaped pulses. (These values hold when a full-width-at-half-maximum criterion is used for the temporal and spectral width.)

For a given pulse duration, transform-limited pulses are those with the minimum possible spectral width. This is important e.g. in optical fiber communications: a transmitter emitting close to transform-limited pulses can minimize the effect of chromatic dispersion during propagation in the transmission fiber, and thus maximize the possible transmission distance.

Many mode-locked lasers, particularly soliton lasers, are able to generate close to transform-limited pulses. During propagation e.g. in transparent media, phenomena such as chromatic dispersion and optical nonlinearities can cause chirp and thus can lead to non-transform-limited pulses. Such pulses may be brought back to the transform limit (and thus temporally compressed) by modifying their spectral phase, e.g. by applying a proper amount of chromatic dispersion. This is called ***dispersion compensation***. For not too broad spectra, compensation of second-order dispersion is often sufficient, whereas very broad spectra may require compensation also of higher-order dispersion in order to approach the transform limit.

Spectral analysis of the correlation

Supposing we have two r.v. X and Y. Let $\mathbf{x}_{0:\mathrm{T}}$ and $\mathbf{y}_{0:\mathrm{T}}$ be the respective realizations, or the sequence of observations over time. We may want to calculate the correlation or covariance. But because we are dealing with sampled random variable, we need to use a sample estimator.

For instance, the cross correlation can be estimated as

$$r_M = \frac{\sum_{k\in M}(x_k - \bar{x})(y_k - \bar{y})}{\sqrt{\sum_{k\in M}(x_k - \bar{x})^2 \sum_{k\in M}(y_k - \bar{y})^2}}$$

Where M is a set of the time indices we want to infer the correlation. Typically:

- M contains contiguous indices, i.e. $M = \{5,6,7, \dots .99,100\}$

- The cardinality of M, denoted as $|M|$, indicates that r_M represent a correlation of length $|M|$samples. If we have the sampling time T_s between samples, then $|M|T_s = \tau_{obs}$ represent the observation time.
- As stated above, and by knowing T_s, the correlation can then be expressed as a function of the observation time τ_{obs}. Now calling $f_{obs} = 1/\tau_{obs}$ the observation frequency, the correlation can also be expressed as a function of frequency. So we have the equivalence $r_M = r(\tau_{obs}) = r(f_{obs})$
- There are several ways to characterize the correlation for a given set of times:
    - One can calculate each correlation by increasing every step the length of the sequence, for instance considering the first index as the time zero and expanding the last index at every iteration. This is the fastest method
    - The slowest but smoothest method: One can try to calculate, for every given observation time, the window length in samples. Then, the requested correlation is given by averaging some values over a sliding window that slide over the time series. The step for which we need to move the window defines how fast this happens. It is similar to windowing techniques for calculating spectrograms.

Of course, spectral correlating wouldn't even make sense if the stochastic signals we are analyzing aren't even stationary.

How to check the hypothesis that the time series is stationary?

Covariance and correlation of a random walk

Introduction

Many phenomena that are slow-varying can be modelled as Gaussian random walk. For example, a random walk phase gives origin to a flat (white) frequency noise spectrum.

We talk about the following 1st order Markov model with additive Gaussian noise. For every k,

$$\phi_k = \phi_{k-1} + q_{k-1}$$

Where $q \sim \mathcal{N}(0, \sigma^2)$ and $\phi_0 = 0$. We have the following, that

$$\phi_T \sim \mathcal{N}(0, T\sigma^2)$$

Since the noise sources q_i are independent.

Multiple phases

Now, to complicate the things, we have multiple phases (like in the case of an optical frequency comb). In vector notation, the dynamic equation becomes

$$\boldsymbol{\phi}_k = \boldsymbol{\phi}_{k-1} + \boldsymbol{q}_{k-1}$$

Where $q \sim \mathcal{N}(0, \boldsymbol{Q})$ and $\boldsymbol{\phi}_0 = \boldsymbol{0}$. We supposed to have n_d total phases, so $\boldsymbol{Q} \in \mathbb{R}^{n_d \times n_d}$ and it is symmetric and positive definite.

Now the fundamental topic. We are interested in calculate the phase correlation: we have two ways:

1. We can generate phases with some length, and calculate the correlation as follow. The Pearson's sample correlation coefficient, using T samples in total, between the phase i and phase j is as follow

$$r(\phi^i_{1:T}, \phi^j_{1:T}) = \frac{\sum_{k=1}^T \left(\phi^i_k - \frac{1}{T}\sum_{k=1}^T \phi^i_k\right)\left(\phi^j_k - \frac{1}{T}\sum_{k=1}^T \phi^j_k\right)}{\sqrt{\sum_{k=1}^T \left(\phi^i_k - \frac{1}{T}\sum_{k=1}^T \phi^i_k\right)^2} \sqrt{\sum_{k=1}^T \left(\phi^j_k - \frac{1}{T}\sum_{k=1}^T \phi^j_k\right)^2}}$$

If we define $\bar{\phi}^i = \frac{1}{T}\sum_{k=1}^T \phi^i_k$ and $\bar{\phi}^j = \frac{1}{T}\sum_{k=1}^T \phi^j_k$, we simplify the notation a little bit

$$r(\phi^i_{1:T}, \phi^j_{1:T}) = \frac{\sum_{k=1}^T (\phi^i_k - \bar{\phi}^i)(\phi^j_k - \bar{\phi}^j)}{\sqrt{\sum_{k=1}^T (\phi^i_k - \bar{\phi}^i)^2} \sqrt{\sum_{k=1}^T (\phi^j_k - \bar{\phi}^j)^2}}$$

2. The sample correlation is good, but it's biased by the discrete set of measurements. For T going to infinity, it is accurate but for finite sample length is not. Therefore, we can compute it from the covariance matrix $\mathbf{\Sigma}_{\boldsymbol{\phi}}$

$$\boldsymbol{\rho} = \left[\sqrt{\mathrm{diag}(\mathbf{\Sigma}_{\boldsymbol{\phi}})}\right]^{-1} \mathbf{\Sigma}_{\boldsymbol{\phi}} \left[\sqrt{\mathrm{diag}(\mathbf{\Sigma}_{\boldsymbol{\phi}})}\right]^{-1}$$

Now, we have to get $\mathbf{\Sigma}_{\boldsymbol{\phi}}$the covariance matrix of the phases (Caution! It is not $\boldsymbol{Q}$). It can be of course, estimated from multiple observation (and it becomes the sample covariance estimator) but we would like it to obtain it as a closed form, independent from the measured data

Covariance matrix of a Multidimensional random walk

To obtain so, we need to think multiple phases as a stochastic signal. So $\boldsymbol{\phi}$ is the vector whose component are the individual phases, i.e. $\boldsymbol{\phi} \coloneqq [\phi^1, \dots, \phi^{n_d}]^\top$. Since from now on, the phases will be treated as random variable (and not a realization), we can effectively compute the theoretical covariance $\mathbf{\Sigma}_{\boldsymbol{\phi}}$. The mean $\overline{\boldsymbol{\phi}}$ it is denoted as expected value here, because It is not a sample average

$$\mathbf{\Sigma}_{\boldsymbol{\phi}} = \mathbb{E}[(\boldsymbol{\phi} - \mathbb{E}[\boldsymbol{\phi}])(\boldsymbol{\phi} - \mathbb{E}[\boldsymbol{\phi}])^\top]$$

Now some simplification. Since we know that no matter what the time index is, from the way it is defined, $\boldsymbol{\phi}$ has always zero mean, i.e. $\mathbb{E}[\boldsymbol{\phi}] \coloneqq \mathbf{0}$. Therefore,

$$\mathbf{\Sigma}_{\boldsymbol{\phi}} = \mathbb{E}[\boldsymbol{\phi}\boldsymbol{\phi}^\top] = \mathbb{E}\begin{bmatrix} [\phi^1]^2 & \phi^1\phi^2 & \cdots & \phi^1\phi^{n_d} \\ \phi^2\phi^1 & [\phi^2]^2 & \cdots & \phi^2\phi^{n_d} \\ \vdots & \vdots & \ddots & \vdots \\ \phi^{n_d}\phi^1 & \phi^{n_d}\phi^2 & \cdots & [\phi^{n_d}]^2 \end{bmatrix} = \begin{bmatrix} \mathbb{E}[[\phi^1]^2] & \mathbb{E}[\phi^1\phi^2] & \cdots & \mathbb{E}[\phi^1\phi^{n_d}] \\ \mathbb{E}[\phi^2\phi^1] & \mathbb{E}[[\phi^2]^2] & \cdots & \mathbb{E}[\phi^2\phi^{n_d}] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}[\phi^{n_d}\phi^1] & \mathbb{E}[\phi^{n_d}\phi^2] & \cdots & \mathbb{E}[[\phi^{n_d}]^2] \end{bmatrix} =$$

$$= \begin{bmatrix} \mathrm{Var}[\phi^1] & \mathbb{E}[\phi^1\phi^2] & \cdots & \mathbb{E}[\phi^1\phi^{n_d}] \\ \mathbb{E}[\phi^2\phi^1] & \mathrm{Var}[\phi^2] & \cdots & \mathbb{E}[\phi^2\phi^{n_d}] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}[\phi^{n_d}\phi^1] & \mathbb{E}[\phi^{n_d}\phi^2] & \cdots & \mathrm{Var}[\phi^{n_d}] \end{bmatrix}$$

Here we have the first problem. $\mathrm{Var}[\phi^i]$ does not have a constant value, because it depends on the sequence length. In fact, the phases $\boldsymbol{\phi}$ are not stationary signals. The random walk model is not stationary. It is just easier to implement in the Extended Kalman Filter. By the way, let's actually try to get some calculation. Be k the time index

$$\boldsymbol{\Sigma}_{\boldsymbol{\phi}}(k) = \begin{bmatrix} k\boldsymbol{Q}_{1,1} & k\boldsymbol{Q}_{1,2} & \cdots & k\boldsymbol{Q}_{1,n_d} \\ k\boldsymbol{Q}_{2,1} & k\boldsymbol{Q}_{2,2} & \cdots & k\boldsymbol{Q}_{2,n_d} \\ \vdots & \vdots & \ddots & \vdots \\ k\boldsymbol{Q}_{n_d,1} & k\boldsymbol{Q}_{n_d,2} & \cdots & k\boldsymbol{Q}_{n_d,n_d} \end{bmatrix} = k\boldsymbol{Q}$$

Correlation matrix of a Multidimensional random walk

So, the true covariance matrix depends on k. But the correlation matrix should not. In fact,

$$\boldsymbol{\rho}(k) = \left[\sqrt{diag\left(\boldsymbol{\Sigma}_{\boldsymbol{\phi}}(k)\right)}\right]^{-1} \boldsymbol{\Sigma}_{\boldsymbol{\phi}}(k) \left[\sqrt{diag\left(\boldsymbol{\Sigma}_{\boldsymbol{\phi}}(k)\right)}\right]^{-1}$$

$$= \begin{bmatrix} \sqrt{k}Std[\phi^1] & 0 & \cdots & 0 \\ 0 & \sqrt{k}Std[\phi^2] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{k}Std[\phi^{n_d}] \end{bmatrix}^{-1} \begin{bmatrix} k\boldsymbol{Q}_{1,1} & k\boldsymbol{Q}_{1,2} & \cdots & k\boldsymbol{Q}_{1,n_d} \\ k\boldsymbol{Q}_{2,1} & k\boldsymbol{Q}_{2,2} & \cdots & k\boldsymbol{Q}_{2,n_d} \\ \vdots & \vdots & \ddots & \vdots \\ k\boldsymbol{Q}_{n_d,1} & k\boldsymbol{Q}_{n_d,2} & \cdots & k\boldsymbol{Q}_{n_d,n_d} \end{bmatrix} \begin{bmatrix} \sqrt{k}Std[\phi^1] & 0 & \cdots & 0 \\ 0 & \sqrt{k}Std[\phi^2] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{k}Std[\phi^{n_d}] \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \boldsymbol{Q}_{1,1}/\mathrm{Var}\left[\phi^1\right] & \boldsymbol{Q}_{1,2}/\mathrm{Std}\left[\phi^1\right]\mathrm{Std}\left[\phi^2\right] & \cdots & \boldsymbol{Q}_{1,n_d}/\mathrm{Std}\left[\phi^1\right]\mathrm{Std}\left[\phi^{n_d}\right] \\ \boldsymbol{Q}_{2,1}/\mathrm{Std}\left[\phi^2\right]\mathrm{Std}\left[\phi^1\right] & \boldsymbol{Q}_{2,2}/\mathrm{Var}\left[\phi^2\right] & \cdots & \boldsymbol{Q}_{2,n_d}/\mathrm{Std}\left[\phi^2\right]\mathrm{Std}\left[\phi^{n_d}\right] \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{Q}_{n_d,1}/\mathrm{Std}\left[\phi^{n_d}\right]\mathrm{Std}\left[\phi^1\right] & \boldsymbol{Q}_{n_d,2}/\mathrm{Std}\left[\phi^{n_d}\right]\mathrm{Std}\left[\phi^2\right] & \cdots & \boldsymbol{Q}_{n_d,n_d}/\mathrm{Var}\left[\phi^{n_d}\right] \end{bmatrix}$$

So to conclude, it is just calculating the correlation matrix on **Q**

$$\boldsymbol{\rho}_{\boldsymbol{\phi}} = \left[\sqrt{diag(\boldsymbol{Q})}\right]^{-1} \boldsymbol{Q} \left[\sqrt{diag(\boldsymbol{Q})}\right]^{-1}$$

■

Converting a correlation matrix into a covariance matrix

It is necessary to know the standard deviations of the variable. Not possible in general

Calculation of spectral covariances

Spectral covariances

We can define a covariance estimation over a sequence of length M on with initial and final indeces k_1 and k_2 as (in a context where the sequence correspond to multiple phases)

$$\boldsymbol{\Sigma}_{M,k_1,k_2} = \frac{1}{M-1} \sum_{\substack{k=k_1 \\ \text{s.t. } k_2-k_1=M}}^{k_2} (\boldsymbol{\phi}_k - \overline{\boldsymbol{\phi}})(\boldsymbol{\phi}_k - \overline{\boldsymbol{\phi}})^\top$$

What we usually do, by hypothesizing the matrix as stationary, take and average over all the possible pairs of k_1, k_2

$$\boldsymbol{\Sigma}_M = \frac{1}{|\mathcal{K}_M|} \sum_{k_1,k_2 \in \mathcal{K}_M} \boldsymbol{\Sigma}_{M,k_1,k_2}$$

Where $\mathcal{K}$ is defined as

$$\mathcal{K}_M \stackrel{\text{def}}{=} \{(k_1,k_2) | k_2 - k_1 = M\}$$

In real life we don't take an average over all the possible pairs, as computing the matrix everytime can be computationally expensive. But, it is possible to reduce the computation time, at the cost to have huge memory. Let us define

$$\widetilde{\mathbf{\Sigma}}_k = (\boldsymbol{\phi}_k - \overline{\boldsymbol{\phi}})(\boldsymbol{\phi}_k - \overline{\boldsymbol{\phi}})^\top$$

The vector outer product, which can be first calculated by subtracting the dataset mean $\overline{\boldsymbol{\phi}}$. We need to have it ready for every value of k. Then $\mathbf{\Sigma}_M$ can be calculated over all $M - spans$ by summing $\widetilde{\mathbf{\Sigma}}_k$, and divide by the number of sums.

How big the memory needs to be? If each element is 4 byte (float 32 bit), we have N_ϕ phases and T total samples, then $S = \ 4TN_\phi^2$ bytes. That is huge. Let's see if we can reduce the complexity further

Let's see in detail. Let's call $\widetilde{\boldsymbol{\phi}}_k = \boldsymbol{\phi}_k - \overline{\boldsymbol{\phi}}$

$$\mathbf{\Sigma}_M = \left(\frac{1}{M}\right)\left(\frac{1}{T-M-1}\right)\sum_{i=1}^{T-M}\sum_{k=i}^{i+M}\widetilde{\mathbf{\Sigma}}_k = \left(\frac{1}{M}\right)\left(\frac{1}{T-M-1}\right)\sum_{i=1}^{T-M}\sum_{k=i}^{i+M}\widetilde{\boldsymbol{\phi}}_k\widetilde{\boldsymbol{\phi}}_k{}^\top$$

A universal framework for linewidth extraction from frequency noise power spectral density

Introduction

We are interested in determines the relationship between the frequency noise of an oscillating signal with the linewidth/lineshape of its characteristic power spectral density of the signal output [1]. Here some highlights on the terms:

- The linewidth: typically the bandwidth of the output signal tone at half the power. Commonly used the definition FWHM (Full width half maximum). For white FN spectrum, this corresponds to the Schawlow Townes linewidth
- The lineshape: referred to the shape assumed by the tone of the power spectral density. Pure tones with no frequency noise have a dirac comb, but typically the shape changes based on the power spectral density form of the frequency noise

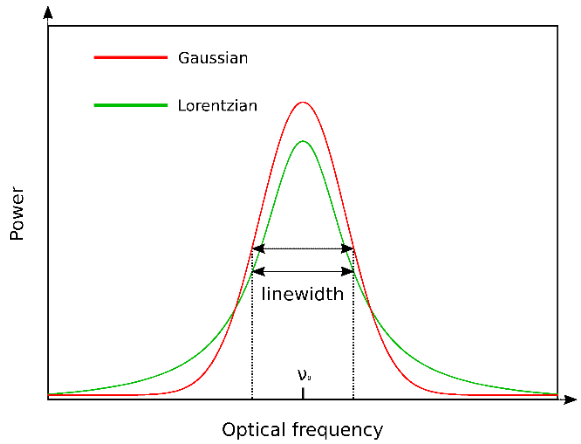


Figure 1 - Linewidth defined as FWHM. Lineshape refers to the general "shape" of the optical spectrum. In the figure different fittings

We want to determine the **linewidth** of a single tone with a custom **FN PSD.** Our signal is the following complex light field

$$E(t) = E_0 \exp\{i[2\pi\nu_0 t + \phi(t)]\} \tag{0.1}$$

The autocorrelation $\Gamma_E(\tau)$ of $E(t)$ is determined by assuming that $E(t)$ is **stationary** and thus is depends only on the lag τ

$$\begin{aligned}
&\Gamma_E(\tau) = \lim_{T\to+\infty} \frac{1}{T}\int_{-T/2}^{T/2} E(t)E^*(t+\tau)dt = \\
&E_0^2 \exp\{i2\phi\nu_0\tau\}\,\mathrm{Re}\left\langle \exp\{i[\phi(t)-\phi(t+\tau)]\}\right\rangle = \\
&E_0^2 \exp\{i2\phi\nu_0\tau\}\exp\left\{-\frac{1}{2} < [\phi(t)-\phi(t+\tau)]^2 >\right\}
\end{aligned} \tag{0.2}$$

Additional assumptions are that $\phi(t)$ is a Gaussian process because $\delta\nu(t)$ - instantaneous frequency deviation – is a Gaussian process as well. Here is the definition of phase oscillation from frequency deviations

$$\phi(t) = 2\pi\int_0^{+\infty} \delta\nu(t')dt' \tag{0.3}$$

And Gaussian process are an invariant class under integrals (a formal proof would be needed). The question still remain, why is $\phi(t)$ a Gaussian process? It means it can be described in terms of its mean $\mu(\phi)$ and covariance $K(\phi_1, \phi_2)$ functions

It turns out that Gaussianity of implies that the odd moments of a Gaussian process vanishes, and even moments are related to the 2nd order moments. Full proof is in [2], which we report here

We outline here the evaluation of the term $\langle \exp\{i[\phi(t)-\phi(t+\tau)]\}\rangle$, leading to Eq. (3). The procedure involves the series expansion of the exponential

$$\langle \exp\{i[\phi(t)-\phi(t+\tau)]\}\rangle = \sum_{n=0}^{\infty} \frac{\langle \{i[\phi(t)-\phi(t+\tau)]\}^n\rangle}{n!} .$$

The Gaussian moment theorem[5] states that the odd moments of a Gaussian process vanish, and that the even moments are related to the second-order moment by the relation $\langle A^{2n}\rangle = [(2n!/2^n n!]\langle A^2\rangle^n$. This leads to

[5]D. Middleton, *An Introduction to Statistical Communication Theory* (McGraw-Hill, New York, 1960).

$$\begin{aligned}\langle \exp\{i[\phi(t)-\phi(t+\tau)]\}\rangle &= \sum_{n=0}^{\infty} \frac{(i)^{2n}\langle [\phi(t)-\phi(t+\tau)]^{2n}\rangle}{(2n)!} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n \frac{(2n)!}{2^n n!}\langle [(\phi(t)-\phi(t+\tau)]^2\}\rangle^n}{(2n)!} .\end{aligned}$$

Rearranging, and contracting the series back into an exponential, we obtain the desired result, which is used in Eq. (3)

$$\begin{aligned}\langle \exp\{i[\phi(t)-\phi(t+\tau)]\}\rangle &= \sum_{n=0}^{\infty} \frac{[-\frac{1}{2}\langle(\phi(t)-\phi(t+\tau)]^2\rangle]^n}{n!} \\ &= \exp\{-\tfrac{1}{2}\langle[\phi(t)-\phi(t+\tau)]^2\rangle\} .\end{aligned}$$

From the last step of (0.2) we can see that $<\left[\phi(t)-\phi(t+\tau)\right]^2>$ is expressing some kind of variance, precisely is the **Mean square phase fluctuation**, as defined in [3] and it is related to the phase noise power spectral density by the following

$$\begin{aligned}
&\sigma_\phi^2(\tau) = \left\langle \left[\phi(t)-\phi(t+\tau)\right]^2 \right\rangle = \\
&(2\pi)^2 \left\langle \left[\int_0^t \delta\nu(t')dt' - \int_0^{t+\tau} \delta\nu(t')dt'\right]^2 \right\rangle = \\
&(2\pi)^2 \left\langle \left[\int_0^\tau \delta\nu(t')dt'\right]^2 \right\rangle = \\
&(2\pi)^2 \int_0^\tau (\tau-t)\Gamma_{\delta\nu}(t)dt = \\
&(2\pi)^2 \int_0^\tau (\tau-t)\int_{-\infty}^{+\infty} S_{\delta\nu}(f)\exp(i2\pi ft)dfdt = \\
&-(2\pi)^2 \int_{-\infty}^{+\infty} S_{\delta\nu}(f)\frac{\exp(i2\pi ft)}{f^2}dfdt = \\
&-2\int_0^{+\infty} S_{\delta\nu}(f)\frac{\cos(2\pi ft)}{f^2}df = \\
&\cancel{-2\int_0^{+\infty} \frac{S_{\delta\nu}(f)}{f^2}df} + 4\int_0^{+\infty} S_{\delta\nu}(f)\frac{\sin(\pi ft)}{f^2}df = \\
&4\int_0^{+\infty} S_{\delta\nu}(f)\frac{\sin(\pi ft)}{f^2}df
\end{aligned} \qquad (0.4)$$

Here we use the Wiener – Khintchine theorem

Cosine duplication formulae

Since the PSD is even, we keep only the real part

This integral converges to zero?

It is not super clear to me how they reach certain passages, but the final result should be

$$\Gamma_E(\tau) = E_0^2 \exp\{i2\phi\nu_0\tau\}\exp\left\{-2\int_0^{+\infty} S_{\delta\nu}(f)\frac{\sin(\pi ft)}{f^2}df\right\} \tag{0.5}$$

Now again the relationship between Signal power spectrum and autocorrelation (Wiener-Khintchine):

$$\begin{aligned} &S_E(f) = 2\int_{-\infty}^{+\infty}\Gamma_E(\tau)\exp(-i2\pi f\tau)d\tau = \\ &2\int_{-\infty}^{+\infty}E_0^2\exp\{i2\phi\nu_0\tau\}\exp\left\{-2\int_0^{+\infty}S_{\delta\nu}(f)\frac{\sin(\pi ft)}{f^2}df\right\}\exp(-i2\pi f\tau)d\tau \end{aligned} \tag{0.6}$$

This was just to show how lineshape/linewidth are connected with the frequency noise PSD.

We show other chunks that maybe help to understand

We are now in a position to apply the Wiener-Khintchine Theorem,[5] which relates the power spectrum $W(f)$, and the autocorrelation function $R(t)$, by the Fourier Transform integral:

$$W(f) = 4\int_0^\infty R(t)\cos\omega t\, dt$$

and

$$R(t) = \int_0^\infty W(f)\cos\omega t\, df\ .$$

This results in a power spectrum of the form

$$W_F(f) = A_0^2\int_0^\infty d\tau[\cos(\omega_0-\omega)\tau + \cos(\omega_0+\omega)\tau] \times\exp\{-\tfrac{1}{2}\langle[\phi(t)-\phi(t+\tau)]^2\rangle\}\ .$$

The rapidly varying $\cos(\omega_0+\omega)\tau$ term can be ignored when $\omega_0+\omega \gg \omega_0-\omega$. Under this condition, the shape and width of the rf power spectrum is independent of the central frequency ω_0. This means that our power spectrum is not changed by shifting ω_0 from 3.3 GHz to 200 MHz, since we restrict the width to values much less than 200 MHz.

The argument of the exponential function must now be determined for our specific application, namely when ϕ is the time integral of a noisy voltage. Since $V(t)$ is a stationary process, the following relation can be shown:

$$\begin{aligned} \langle[\phi(t)-\phi(t+\tau)]^2\rangle &= (2\pi D)^2\left\langle \left[\int_0^t V(t)dt - \int_0^{t+\tau}V(t)\,dt\right]^2\right\rangle \\ &= (2\pi D)^2\left\langle\left[\int_0^\tau V(t)dt\right]^2\right\rangle \\ &= (2\pi D)^2\int_0^\tau\int_0^\tau\langle V(t_1)V(t_2)\rangle\, dt_1dt_2 \\ &= 2(2\pi D)^2\int_0^\tau(\tau-t)R_V(t)\,dt\ , \end{aligned} \tag{4}$$

where $R_V(t)$ is the autocorrelation function of the noise. The final step above is explained in Appendix B.

Applying the Weiner-Khintchine Theorem to the noise power spectrum, we find the rf power spectrum to be

$$W_F(f) = A_0^2\int_0^\infty d\tau\cos(\omega-\omega_0)\tau\exp\left[-(2\pi D)^2\int_0^\infty W_V(f')\int_0^\tau(\tau-t)\cos\omega' t\, dt\, df'\right]\ ,$$

or, reversing the order of integration in the exponential and evaluating the time integral:

$$W_F(\Delta\omega) = \frac{A_0^2}{2\pi}\int_0^\infty d\tau\cos(\Delta\omega\tau)\left[\exp - 2(2\pi D)^2\int_0^\infty W_V(\omega')\left[\frac{\sin\omega'\tau/2}{\omega'}\right]^2 d\omega'\right] \tag{5}$$

where $\Delta\omega = \omega-\omega_0$, and $W(\omega) = W(f)/2\pi$ have been used.

Linewidth from custom PSD

Depending on the behavior of the frequency noise spectrum, there are 2 regions:

- 1st region, where $S_{\delta\nu}(f) > 8\ln 2 f / \pi^2$: 11frequency noise contributes to the **central part** of the lineshape, thus **increasing the linewidth**
- 2nd region, where $S_{\delta\nu}(f) < 8\ln 2 f / \pi^2$:11frequency noise contributes to the wings of the lineshape, thus not contributing to the linewidth

One can obtain a **good approximation** of the linewidth using the following

$$LW = \sqrt{8A\ln 2} \qquad (0.7)$$

Where A is the surface of the high modulation area – the integral of the frequency noise spectra below the frequency defined by the beta-separation line

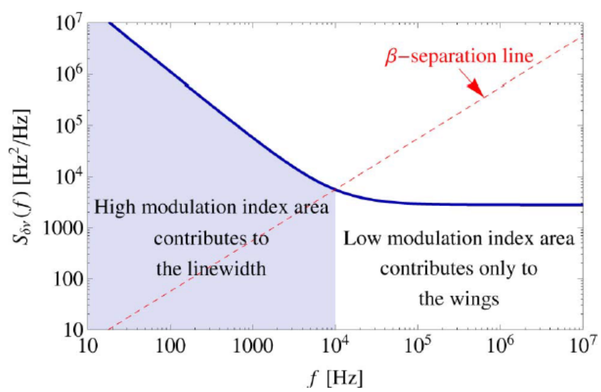


Fig. 3. (Color online) A typical laser frequency noise spectral density composed of flicker noise at low frequencies and white noise at high frequencies. The dashed line given by $S_{\delta\nu}(f) = 8\ln(2)f/\pi^2$ separates the spectrum into two regions whose contributions to the laser line shape is very different: the high modulation index area contributes to the linewidth, whereas the low modulation index area contributes only to the wings of the line shape (see the text for details).

$$A = \int_{1/T_o}^{\infty} H(S_{\delta\nu}(f) - 8\ln(2)f/\pi^2)S_{\delta\nu}(f)\mathrm{d}f,$$

(0.8)

H is the heavside function (the unit step) and T_o is the measurement time that prevent the observation of low frequencies below $1/T_o$.

In other words it is the smallest detectable fourier frequency, being equal to $f_{low} = 1/N/\Delta t$ with N collected samples and Δt sampling time

In other words, we have two methods for evaluating the Linewidth:

1. The approximation given here, as

$$LW = \sqrt{8\ln 2\int_{f_{low}}^{+\infty} H(S_{\delta\nu}(f) - 8\ln 2 f / \pi^2)S_{\delta\nu}(f)df} \quad (0.9)$$

2. The full numerical integration: first we determine the lineshape using (0.6), then we look for the points that are Half Maximum, and from there we have the linewidth

Full Numerical integration

If one wants to try to fully integrate (0.6) to get the lineshape, the procedure is described in [4]:

- First we need to integrate

$$\left(-2\int_0^\infty S_{\delta\nu}(f)\frac{\sin^2(\pi f\tau)}{f^2}\mathrm{d}f\right) \quad (0.10)$$

which requires to use different values of tau. Wrong values lead to numerical errors, and also the interval of integration has to be restricted

I don't know if this method is easier to implement, but the fact that it is necessary to evaluate different values of tau makes me think that the approximate method is maybe the solution.

Numerical integration methods

Theory

We seek methods to evaluate a deterministic integral

$$I = \int_a^b f(x)dx$$

In which we don't have f(x) ready to be evaluated, but instead we have samples of f(x) on a defined interval, that's exactly the case when we estimate the FN spectra and we want to integrate it over a specific interval – like the linewidth evaluations require to.

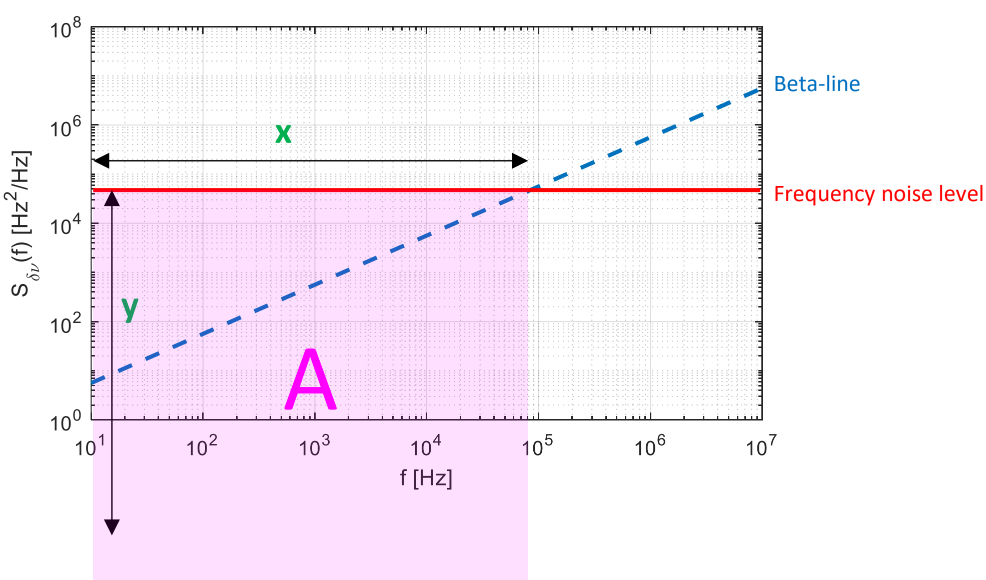
Since the data is discretized in samples, we will use the trapezoidal rule

Practice

Straighforward in Matlab, we just use the trapz function

Troubleshooting and validation

Simple case: Wiener phase noise – White frequency noise



$$\begin{aligned} LW &= \sqrt{A 8 \ln 2} \\ A &= xy \\ x &= y \frac{\pi^2}{8 \ln 2} \end{aligned} \tag{0.11}$$

Therefore we should link the FN level $S_{\delta v}$ to the linewidth LW via the following

$$\begin{aligned} LW &= \sqrt{\frac{{S_{\delta v}}^2 \pi^2}{8 \ln 2} 8 \ln 2} = S_{\delta v} \pi \\ S_{\delta v} &= LW / \pi \end{aligned} \tag{0.12}$$

Problems detected: I have a problem of scaling the power of noise:

Suppose I have certain random Gaussian white noise. Its power should not depend on how many samples are generated, nor on the sampling time. But its power density, integrated on the whole frequency axis, should give me the total power of the noise.

I checked, old methods to calculate the FN from the Laser class works. SO we have to debug the following – Why do I get two different power spectral densities?

- *Laser.SSPSD* (single side power spectral density) calculated as: PSD_dv = abs(fft(dv)*Ts).^2/N/Ts
- *Laser.FN* (Frequency noise) calculated as FN = abs(fft(diff(phi)/2/pi/Ts).*Ts).^2/N/Ts

So we have a factor of Ts*2*pi in excess when using Laser.SSPSD. But it turns out that we generate dv as

$$\delta v_t \sim \mathrm{N}\,(0, LW 2\pi) \tag{0.13}$$

We generate it in this way to achieve a specific linewidth on output. And then it is added to the phase noise by using the Euler method

$$\phi_t = \phi_{t-1} + \sqrt{T_S}\,\delta v_{t-1} \tag{0.14}$$

So diving by 2*pi*Ts dv before power spectrum calculation we get a power, that once integrated all should give as LW

Testing the performance

Test 1 : accuracy of the method for different level of white frequency noise

The linewidth acts directly on the phase time series. So no spectral averaging what so ever. But the method is designed to work on frequency noise spectra (but the technique is the same)

Within the same graph, we show also the impact of averaging the white frequency noise in error reduction

Lorentzian lineshape expression as function of the linewidth

References

[1] Di Domenico, G., Schilt, S., Thomann, P.: 'Simple approach to the relation between laser frequency noise and laser line shape', Applied Optics, 2010, 49, (25), pp. 4801-4807.

[2] R. R. a. S. S. D. S. Elliott, "Extracavity laser band-shape and bandwidth modification," *PHYSICAL REVIEW A,* vol. 26, no. 1, pp. 12-18, JULY 1982.

[3] L. B. Mercer, "1 /f Frequency Noise Effects on Self-Heterodyne Linewidth Measurements," *JOURNAL OF LIGHTWAVE TECHNOLOGY,* vol. 9, no. 4, pp. 485-493, April 1991.

[4] *. V. D. C. S. P. T. G. D. D. a. S. S. Nikola Bucalovic, "Experimental validation of a simple approximation to determine the linewidth of a laser from its frequency noise spectrum," *APPLIED OPTICS,* vol. 51, no. 20, pp. 4582-4588, July 2012.

Noise of mode Locked lasers

Based mainly on [1]

The Master Equation and solutions

Differential equation that describes the averaged pulse dynamics of passively mode-locked fiber lasers.

Fiber because the cavity should be an optical fiber.

Assumptions: **small linear and non-linear changes** in the pulse for each cavity round trip

$$T_R \frac{\partial}{\partial T} A(T,t) = \left[-l + g\left(1 + \frac{1}{\Omega_g^2} \frac{\partial^2}{\partial t^2}\right) - jD \frac{\partial^2}{\partial t^2} + j\delta |A|^2 - q(T,t) \right] A,$$

Where:

- T_R is the round trip time in the cavity (pulse bouncing back and forth)
- $A(T,t)$ is the field (pulse) envelope, slow varying in according to the assumptions
- t is the short scale time variable (instantaneous time variable)
- T is the long scale time variable (many T_R)
- l lumped amplitude attenuation
- g saturated amplitude gain
- Ω_g half width half maximum (HWHM) gain bandwidth
- D cavity group delay dispersion (GDD) coefficient
- $q(T,t)$ takes account of self-amplitude modulation (SAM) of saturable absorbers. In fact it is modulating the amplitude A. It is the response of the saturable absorber.
- δ lumped self-phase modulation (SPM) length, given by $\delta = \frac{2\pi}{\lambda_0} \frac{n_2}{A_{eff}} L$ where
  - n_2 non-linear refractive index of the optical fiber
  - A_{eff} effective mode area of the fiber core
  - L length of the fiber
  - λ_0 carrier wavelength

Steady state solution

SS solution : a solution that does not change in time, meaning that $\frac{\partial A}{\partial T} = 0$? if we try to set that we have

$$\underbrace{lA(T,t)}_{\text{Losses, related to the pulse amplitude}} + \underbrace{jD \frac{\partial^2 A(T,t)}{\partial t^2}}_{\text{Group delay dispersions, losses related to amplitude “acceleration”}} + \underbrace{q(T,t)A(T,t)}_{\text{Pulse self amplitude modulation: q is modulating the pulse amplitude}} = \underbrace{gA(T,t)}_{\text{Gains, related to the pulse amplitude}} + \underbrace{\frac{g}{\Omega_g^2} \frac{\partial^2 A(T,t)}{\partial t^2}}_{\text{Gains, but related to amplitude “acceleration”}} + \underbrace{j\delta |A(T,t)|^2 A(T,t)}_{\text{Gains, but related to amplitude “acceleration”}}$$

Assumption for SS solution:

- Ignoring gain dynamics in one round trip. The gain is saturated by many pulses in successions. The lumped gain becomes

$$g(T) = \frac{g_0}{1 + \frac{1}{P_s T_R} \int |A(t,T)|^2 dt}$$

With g_0 being the small-signal gain, P_s the gain saturation power.

- Considering Ideal fast saturable absorber, the response is given

$$q(T,t) = \frac{q_0}{1 + \frac{|A(t,T)|^2}{P_A}}$$

Where q_0 is the non-saturated loss, P_A the saturation power of the saturable absorber

- In addition, considering small change conditions for each round trip (the one stated before). It means that the absorber response can be linearized into

$$q(t,T) \approx q_0 - \frac{q_0}{P_A}|A(t,T)|^2 = q_0 - \gamma|A(t,T)|^2$$

γ takes the name of saturable absorption coefficient

The pulsed solution of the master equation becomes

$$A(T,t) = A_0 \operatorname{sech}^{(1+j\beta)}\left(\frac{t}{\tau}\right) \exp\left[j\psi \frac{T}{T_R} + j\theta\right], \tag{4}$$

where

$$j\psi = g - l_0 + \left(\frac{1}{\Omega_g^2 \tau^2} + j\frac{D}{\tau^2}\right)(1+j\beta)^2, \tag{5}$$

and

$$(2 + 3j\beta - \beta^2)\left(\frac{1}{\Omega_g^2 \tau} + j\frac{D}{\tau}\right) - \frac{1}{2}(\gamma - j\delta)E_p = 0. \tag{6}$$

Where

- This pulse has a duration of (FMHM) $\tau_{FWHM} = 1.76\tau$
- This pulse has an energy of $E_p = 2A_0^2\tau$
- β is the chirp parameter
- θ is the arbitrary phase of the pulse
- ψ is the carrier phase slip respect to the envelope in one round trip time. It correspond to the carrier envelope offset phase

What parameters matters in the pulse?

- τ changes the pulse duration, the width.
- β changes the ripples around the center

PROOF (that this pulse is actually solution) – not sure of the calculus, need to check with wolfram

$$\frac{\partial A(T,t)}{\partial T} = A_0 \operatorname{sech}\left(\frac{t}{\tau}\right)\frac{j\psi}{T_R} \exp\left[j\beta \log\left(\operatorname{sech}\left(\frac{t}{\tau}\right)\right) + j\psi\frac{T}{T_R} + j\theta\right]$$

$$\frac{\partial A(T,t)}{\partial t} = -\frac{A_0}{\tau}[1+j\beta] \tanh\left(\frac{t}{\tau}\right) \operatorname{sech}\left(\frac{t}{\tau}\right) \exp\left[j\beta \log\left(\operatorname{sech}\left(\frac{t}{\tau}\right)\right) + j\psi\frac{T}{T_R} + j\theta\right]$$

$$\frac{\partial^2 A(T,t)}{\partial t^2} = -\frac{A_0}{\tau^2}(1+j\beta)\left[1 - \tanh^2\left(\frac{t}{\tau}\right)(2+j\beta)\right] \exp\left[j\beta \log\left(\operatorname{sech}\left(\frac{t}{\tau}\right)\right) + j\psi\frac{T}{T_R} + j\theta\right]$$

Homework: complete the proof.

- The pulse intracavity GDD (parameter D) can adjust the pulse dynamics. Shortest pulse duration can be obtained in the lowest possible dispersion.

- In the case of slow saturable absorber

$$\frac{\partial q(T,t)}{\partial t} = -\frac{q(T,t) - q_0}{\tau_A} - q(T,t)\frac{|A(T,t)|^2}{E_A}$$

  - τ_A recovery time
  - E_A saturation energy of the absorber

The solution of the master equation can be found as well, but is not reported here (in case of slow absorber)

- In the case of a non-trivial gain dynamics

$$\frac{\partial g(T,t)}{\partial t} = -\frac{g(T,t) - g_0}{\tau_L} - \frac{P}{E_{sat,L}} g(T,t)$$

  - τ_L finite gain relaxation time
  - $E_{sat,L}$ saturation energy
  - P average laser power

A solution to master equation exists but it not reported here

Pulse shaping

How to achieve pulse shaping: combination of two effects:

- Linear effects (i.e. dispersion)
- Non Linear effects (i.e. SPM and SAM)

Noise in mode-locked lasers

Electric field of the optical pulse train with noise, is described in the time domain as

$$A(t) = [A_0 + \Delta A_0(t)] \sum_{m=-\infty}^{+\infty} a(t - mT_R + \Delta T_R(t)) \exp[j\{2\pi\nu_c t + m\Phi_{\mathrm{CE}} + \Delta\theta(t)\}], \tag{12}$$

Where

- $A(t)$ is the electric field
- A_0 is the amplitude
- $\Delta A_0(t)$ is the **amplitude noise**
- $a(t)$ is the pulse envelope function (can be sech() for soliton pulses)
- $\Delta T_R(t)$ is the pulse **timing jitter**
- T_R is the repetition period

- ν_c is the carrier frequency
- ϕ_{CE} is the carrier-envelope phase shift from pulse to pulse
- $\Delta\theta(t)$ is the phase fluctuations of the field, or the common **phase noise**

There are 3 kind of noise we are interested in:

Intensity noise	**Timing jitter**	**Comb-line frequency noise**
$\Delta A_0(t)$	$\Delta T_R(t)$	$\Delta\theta(t)$
Fluctuations in the average power	Equivalent to the repetition rate phase noise in frequency domain	The residual ceo phase noise, including the ceo frequency fluctuation f_{ceo}

Intensity noise

Represents average power stability of the optical pulse train.

It is a wide sense STATIONARY PROCESS.

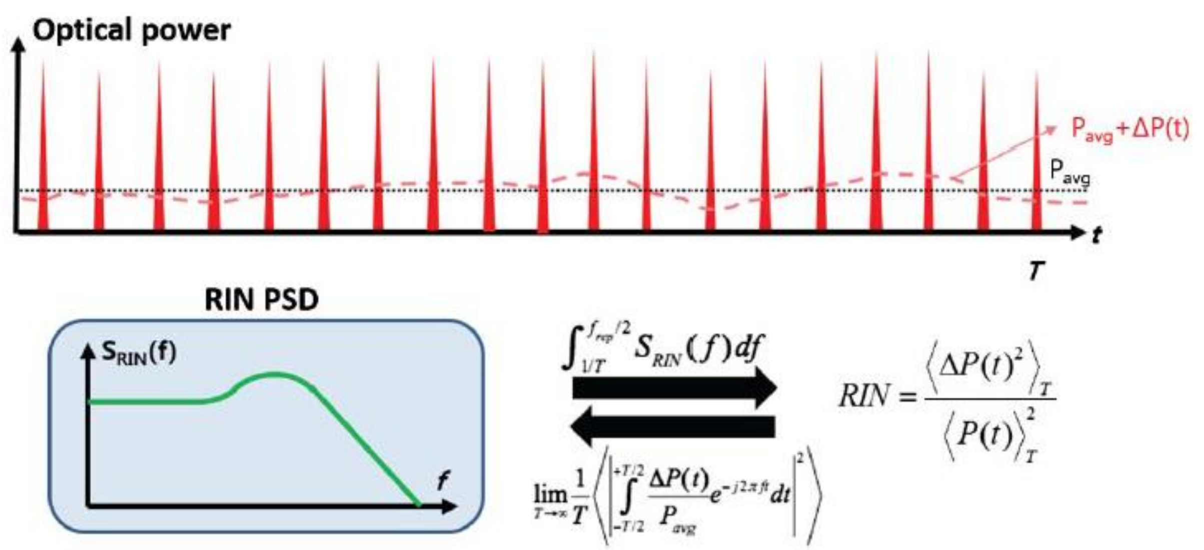
Quantified by the relative intensity noise RIN, over a measurement time T

$$RIN = \frac{\langle \Delta P(t)^2 \rangle_T}{\langle P(t) \rangle_T^2}$$

Mean square optical power fluctuation

Pulse train average optical power

The RIN is different from the PSD of the RIN power. The RIN can be obtained by integrating the PSD of the RIN power, and it is a single parameter.



The integration is from $\frac{1}{T}$, which is the lowest possible frequency (the inverse of the total signal duration) till $\frac{f_{rep}}{2}$ which is the Nyquist frequency of the pulse train

WHERE INTENSITY NOISE MATTERS:

- sampling and photonic **analog-to-digital converters**
- arbitrary optical waveform generation

- Seeding of optical amplifiers
- optical communication systems
- laser-based materials processing¨

IT IS CAUSED BY: quantum sources (such as ASE noise and vacuum fluctuation noise) and external technical noise sources (such as pump laser intensity noise, including power supply noise coupled to the Pump laser driver).

NOISE RESPONSE: Overshoot with relaxation oscillation frequency, then roll off like a lowpass, because noise contributes only till the peak. The technical noise may dominate over the ASE noise before roll off, so the peak may not be visible. Beyond relaxation, the RIN approaches the quantum limit.

The quantum limit (or shot noise limit), set by the vacuum fluctuation of the RIN is

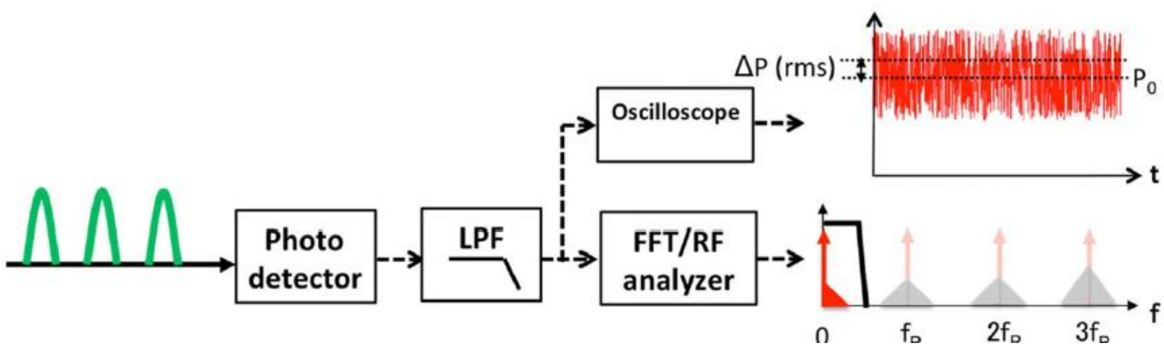
$$S_{\mathrm{RIN}}^{\mathrm{shot\ noise}}(f) = \frac{2h\nu_c}{P_{\mathrm{avg}}},$$

h is the Planck's constant, ν_c the optical frequency and P_{avg} is the average output power from the output coupler.

Measuring intensity noise

WE don't need to measure the correlation by definition. Just low bandwidth photodetection of the pulse trains. Only baseband noise spectrum is necessary for RIN measurement. The RIN becomes the ration between the voltage noise PSD and the DC voltage level of the signal,

$$S_{RIN}(f) = \frac{S_V(f)}{V_0^2}$$



This way of measurement is limited by the shot noise of the photodetection, which is higher than the vacuum fluctuation of the RIN

Timing jitter

Pulse envelope temporal deviation from the perfect periodicity.

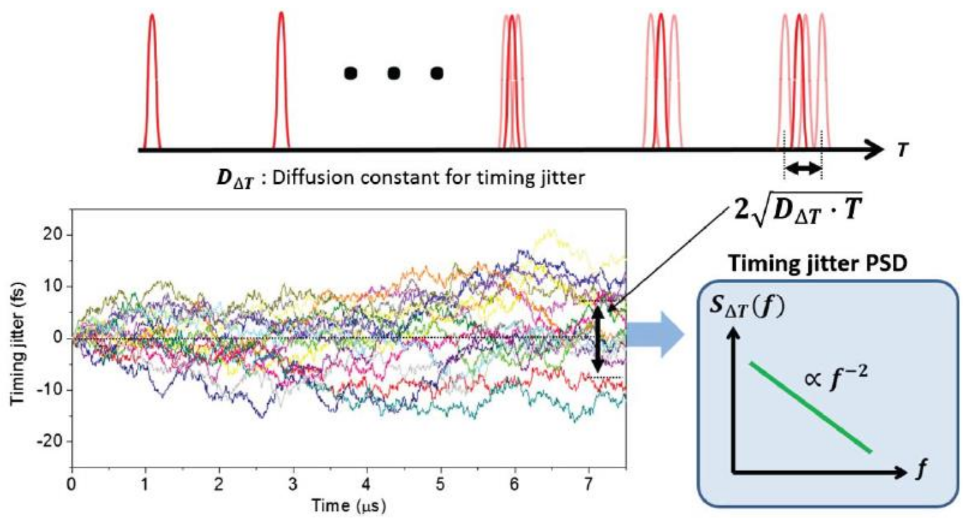
The timing jitter PSD is the SAME as the phase noise PSD of the repetition rate and its harmonics (if no excess noise is introduces in the electro-opto conversion.

It is not bounded, diverges without bound and undergoes A RANDOM WALK (in free running passively mode-locked lasers).

Dispersion influence the coupling between timing jitter and center frequency noise. When there's no coupling, the variance in timing is PROPORTIONAL to the observation time (like in any random walk)

$$\langle \Delta T_R^2(T) \rangle \cong D_{\Delta T} \cdot T$$

Where $D_{\Delta T}$ is the diffusion coefficient of the random walk.

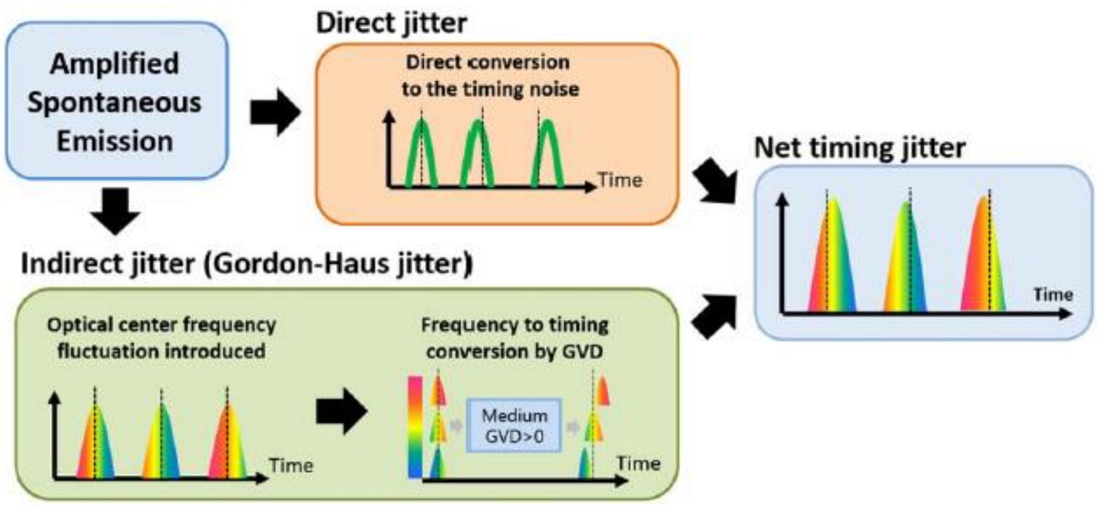


NOISE RESPONSE: Typical of a random walk, when there's no coupling with the center frequency it rolls off like $1/f^2$

$$S_{\Delta T}(f) = \int_{-\infty}^{+\infty} \langle \Delta T_R(t) \rangle \langle \Delta T_R(t+\tau) \rangle e^{-j2\pi f \tau} d\tau = \frac{D_{\Delta T}}{f^2}$$

However can happen that when the center frequency noise is coupled it rolls off with different power, even $1/f^4$. Measurement of timing jitter psd can reveal different noise influence at different time scales

ORIGIN OF TIMING JITTER



Contribution from ASE noise directly coupled into the timing jitter and indirect coupling of the central frequency fluctuation (Gordon-Haus jitter, coupling via dispersion and pulse chirp).

Other technical noise sources (acoustic and intensity noise) can couple into the timing jitter.

MEASURING TIMING JITTER

RF spectrum analysis: analyze the RF spectrum of the photodetected signal of mode-locked laser output. Such power spectrum can be

$$P_{\mathrm{RF}}(f) \propto \sum_{k=0}[\delta(f-kf_{\mathrm{rep}})+P_{\Delta A}(f-kf_{\mathrm{rep}})+(2\pi k)^2 P_{\Delta T}(f-kf_{\mathrm{rep}})],$$

Where $P_{\Delta A}$ is the amplitude noise power, $P_{\Delta T}$ is the timing noise power. RIN spectrum can be measured at baseband (k=0) and timing jitter spectrum can be measured at higher harmonic (k >>0). The latter can be estimated by measuring the Lineshape and linewidth of the RF components of higher harmonics.

There are other methods, but here we focus on direct detection.

Comb-line frequency noise

Optical power spectrum of pulse train can be seen as the periodic repetition of Lorentzian line at the comb-line positions

$$\tilde{A}(\nu) = \frac{|\tilde{a}(\nu-\nu_c)|^2}{2\pi T_R^2}\sum_n \frac{2\Delta\nu_n}{(\nu-\nu_n)^2+\Delta\nu_n^2},$$

$\tilde{a}(\nu)$ is the Fourier transform of the pulse envelope $a(t)$. ν_n is the comb line position and $\Delta\nu_n$ is the linewidth of the line n.

THE ELASTIC TAPE MODEL:

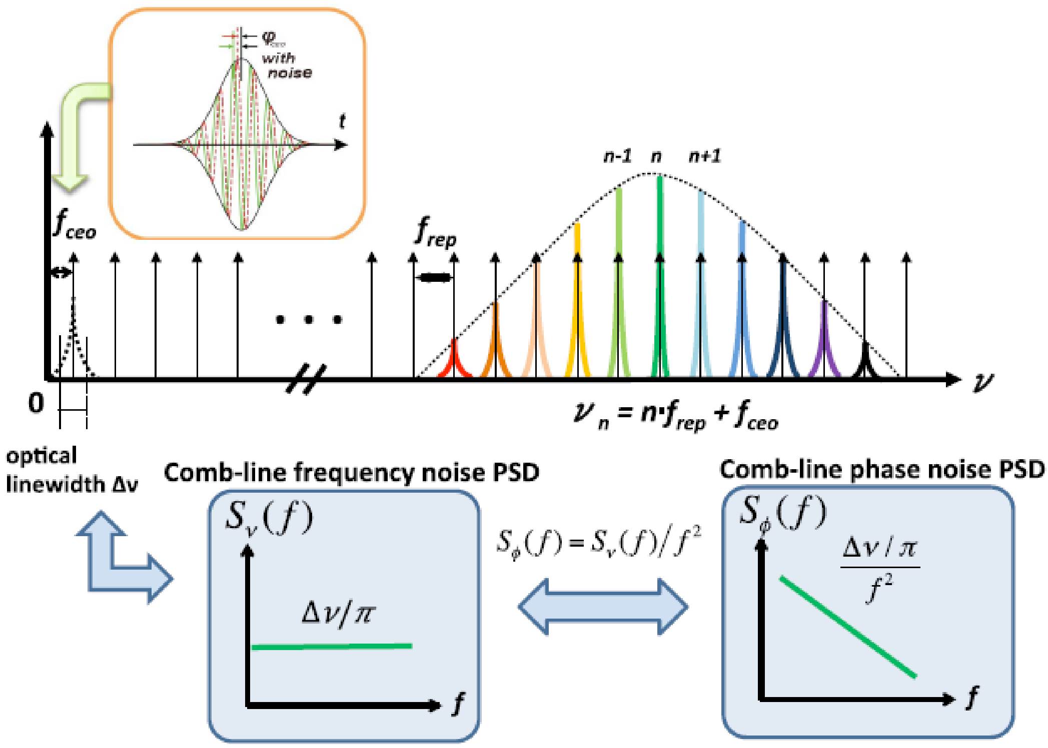
$\nu_n(t) = f_{ceo}(t) + nf_{rep}(t)$ relationship for the frequencies

$\varphi_n(t) = \varphi_{ceo}(t) + n\varphi_{rep}(t) + \varphi_n$ relationship for the phases

When timing ΔT_R and phase fluctuations $\Delta\theta$ are uncorrelated, the linewidth can be expressed as

$$\Delta\nu_n = \Delta\nu_{\Delta\theta} + [2\pi\tau(\nu_n - \nu_c)^2]\Delta\nu_{\Delta T}$$

With $\Delta\nu_{\Delta\theta}$ and $\Delta\nu_{\Delta T}$ diffusion constants for $\Delta\theta$ and ΔT_R and τ is the pulse width (temporal). The linewidth is minimum at the central frequency, and then it increases in the wings.



Conclusions

When measuring the optical train with a photodiode, only relative intensity noise and timing jitter can be inferred

$$P_{RF}(f) \propto \sum_{k=0}[\delta(f - kf_{rep}) + P_{\Delta A}(f - kf_{rep}) + (2\pi k)^2 P_{\Delta T}(f - kf_{rep})],$$

- using k = 0, the RIN can be extracted
- usgin k >> 1, the term regarding timing jitter becomes more dominant. But of course the photodetector bandwidth has to allow that. However, the timing jitter spectrum (or repetition rate phase noise, is the same) will suffer of poor characterization in low frequency.

Can we apply a Kalman like equation to extract things in real time?

The only suitable equation would be:

$$A(t) = [A_0 + \Delta A_0(t)] \sum_{m=-\infty}^{+\infty} a(t - mT_R + \Delta T_R(t)) \exp[j\{2\pi\nu_c t + m\Phi_{CE} + \Delta\theta(t)\}], \tag{12}$$

Of course, by removing the high frequency part, aka the complex exponential

$$A(t) = [A_0 + \Delta A_0(t)] \sum_{m=-\infty}^{+\infty} a(t - mT_R + \Delta T_R(t))$$

The only problem is to determine now the form of the envelope a(t). It is not trivial.

Noise of mode-locked lasers [2]

All noise affecting these lasers are coupled each other. We are going to analyze the noise in several parameter of the pulses generated by a mode locked laser. Parameters such as the pulse energy E_p, the pulse duration FWHM τ_p, deviation of central frequency from reference $\Delta f_c = f_c - f_{ref}$, the optical phase of the pulse φ_{opt}

Timing jitter

Δt is the timing error (or: temporal displacement of the pulse wrt a reference. To check the displacement wrt to a reference, we can set a particular point of the pulse, but instead to be the peak, it is the center of gravity: the temporal position of the pulse

$$t_p = \frac{\int tP(t)dt}{\int P(t)dt}$$

The timing error can be transformed in a phase error: $\Delta\varphi = 2\pi f_{rep}\Delta t$, with f_{rep} the pulse repetition rate. This phase noise correspond to the lowest harmonic of the output of a photodiode detecting the pulse train (centered in f_{rep} of course, as it is the first harmonic).

Quantifying the noise

We use double or single side power spectral densities. Attention: this way of describing stochastic processes has the implicit assumption that those processes are stationary. Other ways to characterize them is to calculate the variance.

For example, by considering the phase noise of the repetition frequency (or the timing jitter spectra, that is the same):

$$S_\varphi(f) = \lim_{T\to+\infty} \frac{1}{T} \left| \int_{-\frac{T}{2}}^{+\frac{T}{2}} \Delta\varphi(t) e^{j2\pi f t} dt \right|^2 = \int_{+\infty}^{-\infty} \langle \Delta\varphi(t)\Delta\varphi(t+\tau)\rangle e^{j2\pi f \tau} d\tau$$

$$S_\varphi(f) = S_\varphi(-f)$$

$$S_{\varphi}(f) = \frac{1}{f^2} S_{f_{rep}}(f)$$

$$\sigma_{\Delta\varphi}^2 = \langle|\Delta\varphi|^2\rangle = \int_{-\infty}^{+\infty} S_{\varphi}(f)df$$

Spectrum detected with direct detection

When a photodiode detects a pulse train: there is significant frequency content around f_{rep}, if the photodiode is fast enough can detect single light pulses. This spectrum is related to the timing error spectra. However, there is also the intensity noise spectra affecting the detected spectra, so it is not so trivial to separate the two.

All noise spectra

Phase noise (timing jitter)	Intensity noise (pulse energy divided by its average value)	Normalized pulse duration noise	Normalized center frequency noise
$S_{\varphi}(f)$	$S_I(f)$	$S_{\tau}(f)$	$S_{\Delta f_c}(f)$

Pulse propagation in the cavity

The state of the pulse is described by the complex envelope in time domain $A(t)$ or in frequency domain $A(f)$. The relation among the two forms is straightforward

$$A(f) = \int_{-\infty}^{+\infty} A(t)\exp(j2\pi ft)dt \qquad A(t) = \int_{-\infty}^{+\infty} A(f)\exp(-j2\pi ft)df$$

The power is the squared envelope $P(t) = |A(t)|^2$.

The pulse energy is $E = \int_{-\infty}^{+\infty} P(t)dt = \int_{-\infty}^{+\infty} A(f)^2 df$

The electric field is $E(t) = \Re\mathrm{e}\big[A(t)\exp\big(-j2\pi f_{ref} t\big)\big]$; for the real quadrature. We removed any fast oscillation in $A(t)$,by shifting with the complex exponential the frequency f_{ref}. In this way, the envelope is just slow varying.

Gain

We can have different situations, depending on how the gain is modelled. G is the power amplification factor, while g is the gain coefficient

- Constant gain G. This is just a factor that multiplies the envelope $A(t)$
- Wavelength-dependent gain: $G(f)$ Is applied in the frequency domain, to $A(f)$

It can have some phase profile associated, which has to be accounted for additional dispersion

$$\Delta\varphi(f) = -\frac{f}{\pi}\wp\int_0^{+\infty}\frac{\ln G(f')}{f'^2 - f^2}df'$$

- Fluctuating gain
- Saturated gain

Steady state small signal gain

$$\frac{\partial g(t)}{\partial t} = -\frac{g(t) - g_{ss}}{\tau_g} - \frac{g(t)P}{E_{sat,g}}$$

Intracavity laser power

Gain recovery lifetime

Saturation energy of the gain

The power amplification factor in this case is $G = \exp(g(t))$

- Influence by pump noise: within one pulse period, the pump contributes to an average value to the gain equals to $\Delta g = \frac{T_{rt}}{\tau_g} g_{ss}$, where $T_{rt} = 1/f_{rep}$ is the round trip time.

The change of gain due to pump noise increments the carrier population

$$\Delta N = \frac{\Delta g E_{sat,g}}{h\nu}$$

Saturable absorber

It makes the loss dependent on the optical power. For a fast saturable absorber

$$l(t) = \frac{l_0}{1 + P(t)/P_{sat}}$$

Unsaturated loss

Saturation power

Numerical model [3]

The master equation for pulse, with envelope $a(T,t)$ propagation:

$$T_R \frac{\partial a(T,t)}{\partial T} = \left[-l(t) + g(t)\left(1 - \frac{1}{\Omega_g}\frac{\partial}{\partial t} + \frac{1}{\Omega_g^2}\frac{\partial^2}{\partial t^2}\right) + jD\frac{\partial^2}{\partial t^2} + (\gamma - j\delta)|a(T,t)|^2\right] a(T,t) + T_R S(T,t)$$

Round trip time

Incremental loss

Incremental gain

Gain bandwidth

Effective saturable absorber action

Group velocity dispersion coeff

Kerr phase modulation coefficient

Noise source, perturbation

References

[1] "Jungwon Kim and Youjian Song, "Ultralow-noise mode-locked fiber lasers and frequency combs: principles, status, and applications," Adv. Opt. Photon. 8, 465-540 (2016)".

[2] "Rüdiger Paschotta, Noise of mode-locked lasers, Applied Physics B 79".

[3] "H. A. Haus and A. Mecozzi, "Noise of mode-locked lasers," in IEEE Journal of Quantum Electronics, vol. 29, no. 3, pp. 983-996, March 1993.".

Meaning of state covariance matrix

I read somewhere that the meaning of it is the distribution of the state estimation error. Here (Ian Reid, Estimation II)

1 Discrete-time Kalman filter

We ended the first part of this course deriving the Discrete-Time Kalman Filter as a recursive Bayes' estimator. In this lecture we will go into the filter in more detail, and provide a new derivation for the Kalman filter, this time based on the idea of **Linear Minimum Variance (LMV) estimation of discrete-time systems**.

1.1 Background

The problem we are seeking to solve is the continual estimation of a set of parameters whose values change over time. Updating is achieved by combining a set of observations or measurements $\mathbf{z}(t)$ which contain information about the signal of interest $\mathbf{x}(t)$. The role of the estimator is to provide an estimate $\hat{\mathbf{x}}(t+\tau)$ at some time $t+\tau$. If $\tau > 0$ we have a **prediction** filter, if $\tau < 0$ a **smoothing** filter and if $\tau = 0$ the operation is simply called **filtering**.

Recall that an estimator is said to be **unbiased** if the expectation of its output is the expectation of the quantity being estimated, $E[\hat{\mathbf{x}}] = E[\mathbf{x}]$.

Also recall that a **minimum variance unbiased estimator (MVUE)** is an estimator which is unbiased and minimises the mean square error:

$$\hat{\mathbf{x}} = \arg\min_{\hat{\mathbf{x}}} E[||\hat{\mathbf{x}} - \mathbf{x}||^2|\mathbf{z}] = E[\mathbf{x}|\mathbf{z}]$$

The term $E[||\mathbf{x} - \hat{\mathbf{x}}||^2]$, the so-called **variance of error**, is closely related to the **error covariance matrix**, $E[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T]$. Specifically, the variance of error of an estimator is equal to the trace of the error covariance matrix,

$$E[||\mathbf{x} - \hat{\mathbf{x}}||^2] = \text{trace} E[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T].$$

The Kalman filter is a **linear** minimum variance of error filter (i.e. it is the best linear filter over the class of all linear filters) over time-varying and time-invariant filters. In the case of the state vector $\mathbf{x}$ and the observations $\mathbf{z}$ being jointly Gaussian distributed, the MVUE estimator is a linear function of the measurement set $\mathbf{z}$ and thus the MVUE (sometimes written MVE for Minimum Variance of Error estimator) is also a LMV estimator, as we saw in the first part of the course.

1.2 System and observation model

We now begin the analysis of the Kalman filter. Refer to figure 1. We assume that the system can be modelled by the state transition equation,

$$\mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{G}_k \mathbf{u}_k + \mathbf{w}_k \tag{1}$$

where $\mathbf{x}_k$ is the state at time k, $\mathbf{u}_k$ is an input control vector, $\mathbf{w}_k$ is additive system or process noise, $\mathbf{G}_k$ is the input transition matrix and $\mathbf{F}_k$ is the state transition matrix.

We further assume that the observations of the state are made through a measurement system which can be represented by a linear equation of the form,

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k, \tag{2}$$

where $\mathbf{z}_k$ is the observation or measurement made at time k, $\mathbf{x}_k$ is the state at time k, $\mathbf{H}_k$ is the observation matrix and $\mathbf{v}_k$ is additive measurement noise.

1.3 Assumptions

We make the following assumptions;

- The process and measurement noise random processes $\mathbf{w}_k$ and $\mathbf{v}_k$ are uncorrelated, zero-mean white-noise processes with known covariance matrices. Then,

$$E[\mathbf{w}_k \mathbf{w}_l^T] = \begin{cases} \mathbf{Q}_k & k = l, \\ \mathbf{0} & \text{otherwise}, \end{cases} \tag{3}$$

$$E[\mathbf{v}_k \mathbf{v}_l^T] = \begin{cases} \mathbf{R}_k & k = l, \\ \mathbf{0} & \text{otherwise}, \end{cases} \tag{4}$$

$$E[\mathbf{w}_k \mathbf{v}_l^T] = 0 \quad \text{for all } k, l \tag{5}$$

where $\mathbf{Q}_k$ and $\mathbf{R}_k$ are symmetric positive semi-definite matrices.

- The initial system state, $\mathbf{x}_0$ is a random vector that is uncorrelated to both the system and measurement noise processes.

- The initial system state has a known mean and covariance matrix

$$\hat{\mathbf{x}}_{0|0} = E[\mathbf{x}_0] \quad \text{and} \quad \mathbf{P}_{0|0} = E[(\hat{\mathbf{x}}_{0|0} - \mathbf{x}_0)(\hat{\mathbf{x}}_{0|0} - \mathbf{x}_0)^T] \tag{6}$$

Given the above assumptions the task is to determine, given a set of observations $\mathbf{z}_1, \ldots, \mathbf{z}_{k+1}$, the estimation filter that at the $k+1$th instance in time generates an optimal estimate of the state $\mathbf{x}_{k+1}$, which we denote by $\hat{\mathbf{x}}_{k+1}$, that minimises the expectation of the squared-error loss function,

$$E[||\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1}||^2] = E[(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1})^T(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1})] \tag{7}$$

1.4 Derivation

Consider the estimation of state $\hat{\mathbf{x}}_{k+1}$ based on the observations up to time k, $\mathbf{z}_1 \ldots, \mathbf{z}_k$, namely $\hat{\mathbf{x}}_{k+1|\mathbf{Z}^k}$. This is called a one-step-ahead prediction or simply a **prediction**. Now, the solution to the minimisation of Equation 7 is the expectation of the state at time $k+1$ conditioned on the observations up to time k. Thus,

$$\hat{\mathbf{x}}_{k+1|k} = E[\mathbf{x}_{k+1}|\mathbf{z}_1, \ldots, \mathbf{z}_k] = E[\mathbf{x}_{k+1}|\mathbf{Z}^k] \tag{8}$$

Then the predicted state is given by

$$\begin{aligned} \hat{\mathbf{x}}_{k+1|k} &= E[\mathbf{x}_{k+1}|\mathbf{Z}^k] \\ &= E[\mathbf{F}_k \mathbf{x}_k + \mathbf{G}_k \mathbf{u}_k + \mathbf{w}_k|\mathbf{Z}^k] \\ &= \mathbf{F}_k E[\mathbf{x}_k|\mathbf{Z}^k] + \mathbf{G}_k \mathbf{u}_k + E[\mathbf{w}_k|\mathbf{Z}^k] \\ &= \mathbf{F}_k \hat{\mathbf{x}}_{k|k} + \mathbf{G}_k \mathbf{u}_k \end{aligned} \tag{9}$$

where we have used the fact that the process noise has zero mean value and $\mathbf{u}_k$ is known precisely.

The estimate variance $\mathbf{P}_{k+1|k}$ is the mean squared error in the estimate $\hat{\mathbf{x}}_{k+1|k}$.

Thus, using the facts that $\mathbf{w}_k$ and $\hat{\mathbf{x}}_{k|k}$ are uncorrelated:

$$\begin{aligned} \mathbf{P}_{k+1|k} &= E[(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k})(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k})^T|\mathbf{Z}^k] \\ &= \mathbf{F}_k E[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})^T|\mathbf{Z}^k]\mathbf{F}_k^T + E[\mathbf{w}_k \mathbf{w}_k^T] \\ &= \mathbf{F}_k \mathbf{P}_{k|k} \mathbf{F}_k^T + \mathbf{Q}_k \end{aligned} \tag{10}$$

Having obtained a predictive estimate $\hat{\mathbf{x}}_{k+1|k}$ suppose that we now take another observation $\mathbf{z}_{k+1}$. How can we use this information to update the prediction, ie. find $\hat{\mathbf{x}}_{k+1|k+1}$? We assume that the estimate is a linear weighted sum of the prediction and the new observation and can be described by the equation,

$$\hat{\mathbf{x}}_{k+1|k+1} = \mathbf{K}'_{k+1} \hat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1} \mathbf{z}_{k+1} \tag{11}$$

where $\mathbf{K}'_{k+1}$ and $\mathbf{K}_{k+1}$ are weighting or **gain** matrices (of different sizes). Our problem now is to reduced to finding the $\mathbf{K}_{k+1}$ and $\mathbf{K}'_{k+1}$ that minimise the conditional mean squared estimation error where of course the estimation error is given by:

$$\tilde{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k+1} - \mathbf{x}_{k+1} \tag{12}$$

It can be verified that the state error distribution has the same covariance matrix of the Kalman covariance matrix. What is then the relation with the mean square error?

Central and Noncentral χ^2 Distributions

The χ^2 distribution arises from sums of squared, normally distributed, random variables — if $x_i \sim \mathrm{N}(0,1)$, then $u = \sum_{i=1}^{n} x_i^2 \sim \chi_n^2$, a **central** χ^2 distribution with n degrees of freedom. It follows that the sum of two χ^2 random variables is also χ^2 distributed, so that if $u \sim \chi_n^2$ and $v \sim \chi_m^2$, then

$$u + v \sim \chi^2_{(n+m)} \qquad \text{(A5.14a)}$$

Two other useful results are that if $x_i \sim \mathrm{N}(0, \sigma^2)$, then

$$\sum_{i=1}^{n} x_i^2 \sim \sigma^2 \cdot \chi_n^2 \qquad \text{(A5.14b)}$$

and for $\overline{x} = n^{-1} \sum_{i=1}^{n} x_i$,

$$\sum_{i=1}^{n} (x_i - \overline{x})^2 \sim \sigma^2 \cdot \chi^2_{(n-1)} \qquad \text{(A5.14c)}$$

In this last case, subtraction of the mean causes the loss of one degree of freedom.

Gamma, exponential, and related distributions [edit]

The chi-square distribution $X \sim \chi_k^2$ is a special case of the gamma distribution, in that $X \sim \Gamma\left(\frac{k}{2}, \frac{1}{2}\right)$ using the rate parameterization of the gamma distribution (or $X \sim \Gamma\left(\frac{k}{2}, 2\right)$ using the scale parameterization of the gamma distribution) where k is an integer.

The gamma distribution can be parameterized in terms of a shape parameter $\alpha = k$ and an inverse scale parameter $\beta = 1/\theta$, called a rate parameter. A random variable X that is gamma-distributed with shape α and rate β is denoted

$$X \sim \Gamma(\alpha, \beta) \equiv \mathrm{Gamma}(\alpha, \beta)$$

The corresponding probability density function in the shape-rate parametrization is

$$f(x; \alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \quad \text{for } x > 0 \quad \alpha, \beta > 0,$$

$$f(x; \alpha, \theta) = \frac{x^{\alpha-1} e^{-x/\theta}}{\theta^\alpha \Gamma(\alpha)} \quad \text{for } x > 0, \quad \alpha, \theta > 0.$$

where $\Gamma(\alpha)$ is the gamma function.

Both parametrizations are common because either can be more convenient depending on the situation.

Scaling [edit]

If

$$X \sim \mathrm{Gamma}(k, \theta),$$

then, for any *c > 0*,

$cX \sim \mathrm{Gamma}(k, c\,\theta)$, by moment generating functions,

or equivalently

$$cX \sim \mathrm{Gamma}\left(\alpha, \frac{\beta}{c}\right),$$

Machine learning framework for OFC noise estimation

Bayesian filtering state and measurement equation

$$\delta \boldsymbol{a}_k = \delta \boldsymbol{a}_{k-1} + \boldsymbol{q}_{k-1}^{a}$$

$$\boldsymbol{\phi}_k = \boldsymbol{\phi}_{k-1} + \boldsymbol{q}_{k-1}^{\phi}$$

$$\boldsymbol{q}_k^{a} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}^{a})$$

$$\boldsymbol{q}_k^{\phi} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}^{\phi})$$

Joint hidden state: amplitude and phase noise $\mathbf{x}_k = [\delta \boldsymbol{a}_k, \boldsymbol{\phi}_k]^{\top}$

$$y_k = \sum_m [\bar{a}_k^m + \delta a_k^m] \cos(2kT_S \nu_m + \phi_k^m) + n_k$$

$$n_k \sim \mathcal{N}(0, \sigma_n^2)$$

Optimized EKF-EKS-EM

Prefiltering

Necessary to adapt the signal to be “Optimal” for the assumed model. In an ideal world, everything that is not part of the comb should be white noise. In practice, it is not always the case.

Brickwall filter

Removing any frequency content that is outside the desired range. This can fool EM to estimate the current level of measurement noise. Which would not be white anymore in this case.

<u>Advantages</u>: fast, we make sure we remove any nasty stuff

<u>Disadvantages</u>: not a real filter, it is not used in signal processing, so we may are making some mistakes

Brickwall filter + whitening

This idea comes from the fact that we make everything white whatever is outside our frequency of interests. But is it a good one?

<u>Challenges:</u> How do we generate white noise samples in frequency domain? Shall we just put a constant power spectrum? Or should we randomize also a power spectrum? Or we generate in time domain and then transform?

<u>Proposals:</u>

- Generate white samples in time domain. Use brickwall to mask the signal and insert these white samples in between. We need a good value for the noise variance

No prefiltering

Important: it may be suspected that it can cause issues to the signal statistics. It would be better to try out if we can increase the number of combs and eventually treat all the rest of the signal as white noise. With the proper parameter tuning is maybe possible

Phase-Amplitude noise decoupling

This technique consists in neglecting any possible correlation between phase and amplitude noise. Can be used if we know there’s no correlation. It consists in consider the state covariance matrix and noise covariance matrix as composed by 2 macroblocks:

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_a & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_\phi \end{bmatrix} \qquad \mathbf{Q} = \begin{bmatrix} \mathbf{Q}_a & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_\phi \end{bmatrix}$$

The $\mathbf{P}$ matrix is intented to be decoupled for every time step, and for both filtering and smoothing.

The question is whether this can improve or not the algorithm.

Spectra smoothing

The psd estimate of the phase traces is quite noisy.

The lazy way to obtain nice results is to use a movmean filter in frequency domain, to get the main envelope of the PSD nosie shape.

The harder but most correct way is to renounce on some frequency resolution, and actually calculate the psd on a sliding window, and average the results.

Parameter initialization

It is quite stable to use the following

$$\mathbf{P_0} = \alpha \boldsymbol{Q} \text{ , } \mathbf{m_0} = [\mathbf{0}, \mathbf{0}]^{\top}$$

Where α is a scaling parameter. This works because we indeed have a rough idea of phase and amplitude estimate, and it is given by the conventional way

Parameter estimation

Measurement noise level

Supposing our detected photocurrent is affected by measurement noise, which is white and characterized by a certain standard deviation σ_n. If we compute the power spectral density, what level of power density will correspond?

For a single side PSD,

$$SSPSD_n(f) = 2T_s\sigma_n^2$$

This is of course according to the MATLAB tools to calculate the PSD of a signal. Of course, this is an approximation because we work with discrete time signals / vectors, with sampling time T_s

$$X(f) = \text{fft}\big(x(t)\big)$$

$$PSD_x(f) = |X(f)T_s|^2/(N_T T_s)$$

In case of single side PSD, we retain just the first half of the signal and double its power (in all cases that we treat real signals, the double side psd is symmetric around the Nyquist frequency)

We CAN estimate the noise level by INSPECTION of the signal power spectral density. For a given level PSD_n outside the desired range, the extracted variance would be

$$\sigma_n^2 = \frac{PSD_n}{T_s} = \frac{SSPSD_n}{2T_s}$$

Now it is better to do this estimation with all the signal trace. Like all filters are applied to the whole trace

Performance monitoring

We should monitor:

- The Kalman gain, should converge
- The Covariance matrix, should converge
- The innovation (prediction error) sequence
    - Should be stationary Gaussian distributed according to its covariance value S
    - Normalized innovation squared must be χ^2 distributed with mean 1
    - Autocorrelation should be as white as possible

Basing on the results of these testing, we may have some hints on how to adjust the parameters

Parameter refinement: BER on modulated signal

Phase and amplitude noise

Be b_k a sequence of +1-1 symbols. We can then apply it to the modulated signal (single frequency)

$$y_k^{mod} = b_k y_k = b_k\{[\bar{a} + \delta a_k] \exp[j(2\pi k T_S \nu_{beat} + \phi_k)] + n_k\}$$

What would happen if we try to demodulate the signal? – digitally by applying a frequency shift on the other direction, i.e. $\exp[-j(2\pi k T_S \nu_{beat})]$. But because we have also a phase noise estimate, $\hat{\phi}_k$, we can try to correct for this term by actual adding it to the demodulation term, i.e. $\exp\left[-j\left(2\pi k T_S \nu_{beat} + \hat{\phi}_k\right)\right]$

$$y_k^{demod} = [b_k y_k] \exp\left[-j\left(2\pi k T_S \nu_{beat} + \hat{\phi}_k\right)\right] =$$

$$= b_k[\bar{a} + \delta a_k] \exp\left[j(\phi_k - \hat{\phi}_k)\right] + b_k n_k \exp\left[-j\left(2\pi k T_S \nu_{beat} + \hat{\phi}_k\right)\right]$$

As you can notice, the term $(\phi_k - \hat{\phi}_k)$ is the error phase estimate. We now can retain only the real part of the signal, as following

$$y_k^{demod,i} = \Re\mathrm{e}\left[b_k[\bar{a} + \delta a_k] \exp\left[j(\phi_k - \hat{\phi}_k)\right] + b_k n_k \exp\left[-j\left(2\pi k T_S \nu_{beat} + \hat{\phi}_k\right)\right]\right] =$$

$$= b_k[\bar{a} + \delta a_k] \cos(\phi_k - \hat{\phi}_k) + \underbrace{b_k\left\{\Re\mathrm{e}(n_k) \cos\left(2\pi k T_S \nu_{beat} + \hat{\phi}_k\right) - \Im\mathrm{m}(n_k) \sin\left(-2\pi k T_S \nu_{beat} + \hat{\phi}_k\right)\right\}}_{\textbf{Noise terms}}$$

The second term is just noise, and let's call it n'_k

$$y_k^{demod,i} = b_k[\bar{a} + \delta a_k] \underbrace{\cos(\phi_k - \hat{\phi}_k)}_{\textbf{This term >0}} + n'_k$$

If our estimator is good, then $(\phi_k - \hat{\phi}_k) \approx 0$ and $\cos(\phi_k - \hat{\phi}_k) \approx 1$.

In the hypothesis of good estimations,

$$y_k^{demod,i} \approx b_k[\bar{a} + \delta a_k] + n'_k$$

And to recover the symbol sequence

$$\hat{b}_k = \mathrm{sign}(b_k[\bar{a} + \delta a_k] + n'_k)$$

This procedure can be evaluated to determine how good is our phase noise estimation

Phase noise only

Be b_k a sequence of +1-1 symbols. We can then apply it to the modulated signal (single frequency)

$$y_k^{mod} = b_k y_k = b_k\{a \exp[j(2\pi k T_S \nu_{beat} + \phi_k)] + n_k\}$$

What would happen if we try to demodulate the signal? – digitally by applying a frequency shift on the other direction, i.e. $\exp[-j(2\pi k T_S \nu_{beat})]$. But because we have also a phase noise estimate, $\hat{\phi}_k$, we can try to correct for this term by actual adding it to the demodulation term, i.e. $\exp\left[-j\left(2\pi k T_S \nu_{beat} + \hat{\phi}_k\right)\right]$

$y_k^{demod} = [b_k y_k] \exp\left[-j\left(2\pi k T_S \nu_{beat} + \hat{\phi}_k\right)\right] \quad =$

$= b_k a \exp\left[j(\phi_k - \hat{\phi}_k)\right] + b_k n_k \exp\left[-j\left(2\pi k T_S \nu_{beat} + \hat{\phi}_k\right)\right]$

As you can notice, the term $(\phi_k - \hat{\phi}_k)$ is the error phase estimate. We now can retain only the real part of the signal, as following

$y_k^{demod,i} = \Re\mathrm{e}\left[b_k a \exp\left[j(\phi_k - \hat{\phi}_k)\right] + b_k n_k \exp\left[-j\left(2\pi k T_S \nu_{beat} + \hat{\phi}_k\right)\right]\right] =$

$= b_k a \cos(\phi_k - \hat{\phi}_k) + \underbrace{b_k\left\{\Re\mathrm{e}(n_k) \cos\left(2\pi k T_S \nu_{beat} + \hat{\phi}_k\right) - \Im\mathrm{m}(n_k) \sin\left(-2\pi k T_S \nu_{beat} + \hat{\phi}_k\right)\right\}}_{\textbf{Noise terms}}$

The second term is just noise, and let's call it n'_k

$y_k^{demod,i} = b_k a \underbrace{\cos(\phi_k - \hat{\phi}_k)}_{\textbf{This term >0}} + n'_k$

If our estimator is good, then $(\phi_k - \widehat{\phi}_k) \approx 0$ and $\cos(\phi_k - \hat{\phi}_k) \approx 1$.

In the hypothesis of good estimations,

$y_k^{demod,i} \approx b_k a + n'_k$

And to recover the symbol sequence

$\hat{b}_k = \mathrm{sign}(b_k a + n'_k)$

This procedure can be evaluated to determine how good is our phase noise estimation, and how good is how noise removal algorithm. But we are limited if the measurement noise is way too high.

But, is this a good argument? I mean, can we also compare the signal without modulation? Let's try

Feedback without data modulation

$$y_k = a \exp[j(2\pi k T_S \nu_{beat} + \phi_k)] + n_k$$

Suppose we demodulate, with the current phase estimate

$$y_k^{demod} = y_k \exp\left[-j\left(2\pi k T_S \nu_{beat} + \hat{\phi}_k\right)\right]$$

$y_k^{demod} = \{a \exp[j(2\pi k T_S \nu_{beat} + \phi_k)] + n_k\} \exp\left[-j\left(2\pi k T_S \nu_{beat} + \hat{\phi}_k\right)\right] =$
$= a \exp\left[j\left(\phi_k - \hat{\phi}_k\right)\right] + \underbrace{n_k \exp\left[-j\left(2\pi k T_S \nu_{beat} + \hat{\phi}_k\right)\right]}_{still\ noise, but\ different\ distribution} = a \exp\left[j\left(\phi_k - \hat{\phi}_k\right)\right] + n_k'$

If our estimator is good, then $(\phi_k - \widehat{\phi}_k) \approx 0$ and

$$y_k^{demod} = a + n_k'$$

The results should be closer to a as much as possible, but we cannot neglect the measurement noise! The previous BER criteria can actually work as it has some additional robusteness against measurement noise

Including measurement noise estimation

$$y_k = a \exp[j(2\pi k T_S \nu_{beat} + \phi_k)] + n_k$$

Our EKF filter van give us the clean signal, namely

$$\hat{y}_k = a \exp\left[j\left(2\pi k T_S \nu_{beat} + \hat{\phi}_k\right)\right]$$

And we can get an estimation of the measurement noise by applying

$$\hat{n}_k = y_k - \hat{y}_k$$

We can use this to improve the estimate, before data modulation. In fact we have

$$y_k^{mod} = b_k y_k = b_k\{a \exp[j(2\pi k T_S \nu_{beat} + \phi_k)] + (n_k - \hat{n}_k)\}$$

All the steps are the same and we end up with

$$y_k^{demod} = b_k a \exp\left[j(\phi_k - \hat{\phi}_k)\right] + b_k(n_k - \hat{n}_k) \exp\left[-j\left(2\pi k T_S \nu_{beat} + \hat{\phi}_k\right)\right]$$

When phase and measurement noise estimators are good, i.e. $(\phi_k - \hat{\phi}_k) \approx 0$ and $(n_k - \hat{n}_k) \approx 0$, we get back

$$y_k^{demod} \approx b_k a$$

Which would lead to total error free signal.

Calculate the BER given an estimate of the phase noise

```
% X is the input data. Raw signal from the photodiode
% a_k is a sequence of random symbols +1 -1
% k is the integer time index
% Ts is the sampling time,
% freq is the beating angular frequency
% Y is the modulated data, after coverting the signal into an analytic signal
using the hilbert transform
% Y_demod is the demodulated signal, applying the frequency and phase noise
correction, using the estimated phase noise


a_k = sign(randn(1,param.L));
k = 1 : param.L;
Y = a_k.*hilbert(movmean(X',1));
Y_demod = Y.*exp(-1i*(freq*param.Ts*k + Xest_EKF(1:end)'));
a_k_det = sign(real(Y_demod));
errors = sum(a_k(10e3:end) ~= a_k_det(10e3:end));
BER_num = errors/(param.L-10e3)
```

Testing

Simulating noise correlation

- Check how the correlation matrix vs observation time looks like for simulated data. Also look at the eigenvalue analysis for simulated data

Should another type of prefiltering should be done?

- Test no prefiltering + increase the number of tracked lines

Should amplitude and phase noise have the possibility to be correlated?

Should R be optimized?

What is a good number of EM iterations? How long should the sequence be?

What is a good numerical value for the starting parameters?

Results on single line

Single tone lasers

Pre-Filtered OFC

Eigenvalue decomposition and mode analysis

We obtain our matrices representing the phase noise time traces as

$$\boldsymbol{\phi} = \begin{array}{cc} \begin{bmatrix} \phi_0^{-M} & \phi_1^{-M} & \cdots & \phi_T^{-M} \\ \phi_0^{-M+1} & \phi_0^{-M+1} & & \phi_T^{-M+1} \\ \vdots & & \ddots & \vdots \\ \phi_0^{M} & \phi_1^{M} & \cdots & \phi_T^{M} \end{bmatrix} & \Big\downarrow \ \text{Line index} \\ \xrightarrow[\text{Time}]{\qquad\qquad\qquad} & \end{array}$$

For the sake of notation, we have 2M+1 lines indexed as $-M, \ldots M$

For the correlation, we use the Pearson sample correlation. In our case, the time index of the phase noise defines a sample of the corresponded stochastic variable

$$r_{ij} = \frac{\sum_{k=0}^{T}\left(\phi_k^i - \bar{\phi}^i\right)\left(\phi_k^i - \bar{\phi}^j\right)}{\sqrt{\sum_{k=0}^{T}\left(\phi_k^i - \bar{\phi}^i\right)^2 \sum_{k=0}^{T}\left(\phi_k^j - \bar{\phi}^j\right)^2}}$$

Where we have denoted $\bar{\phi}^i$ as the mean value (average in time) of the phase noise of the line i.

The covariance matrix differs from the correlation matrix from the fact that we do not scale it with the standard deviation, i.e.

$$\operatorname{cov}\left(\phi^i\phi^j\right) = \frac{1}{T}\sum_{k=0}^{T}\left(\phi_k^i - \bar{\phi}^i\right)\left(\phi_k^i - \bar{\phi}^j\right)$$

We know that the covariance matrix of the phases should follow the simple scheme (2 degrees of freedom), given that it is an EO comb

$$\boldsymbol{\Sigma}_{\mathrm{ij}} = \operatorname{cov}\left(\phi^i\phi^j\right) = \sigma_L^2 + \left(\boldsymbol{k}\cdot\boldsymbol{k}^{\top}\right)_{ij}\sigma_{RF}^2$$

where $\boldsymbol{k} = [-M, \ldots M\ ]^{\top}$is the vector of the line indeces – negative and positive indeces. This because, form the EO comb theory $\phi^i = \phi^L + i\phi^{RF}$. Then we know that

$$\mathrm{cov}(\phi^i\phi^j) = E\left[\phi^i{\phi^j}^\top\right] - E\left[\phi^i\right]E\left[{\phi^j}^\top\right] = E[(\phi^L + i\phi^{RF})(\phi^L + j\phi^{RF})^\top] =$$

$$= E\left[\phi^L{\phi^L}^\top\right] + ijE\left[\phi^{RF}{\phi^{RF}}^\top\right] + iE\left[\phi^L{\phi^{RF}}^\top\right] + jE\left[\phi^{RF}{\phi^L}^\top\right] =$$

$$= E\left[\phi^L{\phi^L}^\top\right] + ijE\left[\phi^{RF}{\phi^{RF}}^\top\right] = \sigma_L^2 + ij\sigma_{RF}^2$$

Let $(\boldsymbol{k}\cdot\boldsymbol{k}^\top) = \mathbf{K}$, and $\mathbf{1}$ a matrix formed by all ones,i.e.

$$\mathbf{K} = \begin{bmatrix} M^2 & M(M-1) & \cdots & M \\ M(M-1) & (M-1)^2 & & (M-1) \\ \vdots\ \vdots & & \ddots & \vdots \\ M \quad (M-1) & & \cdots & M^2 \end{bmatrix}$$

$$\mathbf{1} = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}$$

We have indeed the following

$$\boldsymbol{\Sigma} = \mathbf{1}\sigma_L^2 + \mathbf{K}\sigma_{RF}^2$$

It means that the covariance matrix can be expressed as contribution of its noise modes. [need further investigation]

Now if we try to perform eigenvalue decomposition on , basing on the spectral theorem, the eigenvalues are all real and positive, thus

$$\boldsymbol{\Sigma} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^\top$$

Now in this case the inverse of $\mathbf{V}$ is $\mathbf{V}^\top$ because $\boldsymbol{\Sigma}$ is symmetric. Be $\lambda_1 \ldots \lambda_{2M+1}$ the resulting eigenvalues in order of magnitude. Let $\mathbf{V} = [\boldsymbol{v}^1, \ldots\ \boldsymbol{v}^{2M+1}]$ formed by columns of eigenvectors. We have

$$\boldsymbol{\Sigma} = \boldsymbol{v}^1\lambda_1{\boldsymbol{v}^1}^\top + \boldsymbol{v}^2\lambda_2{\boldsymbol{v}^2}^\top + \cdots\ \boldsymbol{v}^{2M+1}\lambda_{2M+1}{\boldsymbol{v}^{2M+1}}^\top = \sum_{i=1}^{2M+1}\boldsymbol{v}^i\lambda_i{\boldsymbol{v}^i}^\top$$

If we neglect high order terms and keep only the main eigenvalues, they have necessarily to match. Other high order terms may come from numerical imprecisions of the calculations.

$$\mathbf{1}\sigma_L^2 + \mathbf{K}\sigma_{RF}^2 = \boldsymbol{v}^1\lambda_1{\boldsymbol{v}^1}^\top + \boldsymbol{v}^2\lambda_2{\boldsymbol{v}^2}^\top$$

Therefore we must have the eigenvector matching the noise modes and the eigenvalues matching the noise power. $\boldsymbol{v}^i\lambda_i{\boldsymbol{v}^i}^\top$ shall have the same measurement unit as the covariance matrix, in this case [rad^2]

However, the eigenvectors are orthonormal, so if we want to find a 1-to-1 comparison, we should normalize the vectors on the left hand side. The normalization constant can then be absorbed into the multiplicative terms, and then we can compare the eigenvalues with the noise sources.

To simplify the terms, we can write the equation as follow:, let $\boldsymbol{c} = \Big[\underbrace{1,1,\dots 1}_{M\ times}\Big]^\top$ then $\mathbf{1} = \Big[\underbrace{\boldsymbol{c},\boldsymbol{c},\dots \boldsymbol{c}}_{\boldsymbol{M\ times}}\Big] = \boldsymbol{c}\boldsymbol{c}^\top$

$$\mathbf{1}\sigma_L^2 + \mathbf{K}\sigma_{RF}^2 = \boldsymbol{v}^1\lambda_1{\boldsymbol{v}^1}^\top + \boldsymbol{v}^2\lambda_2{\boldsymbol{v}^2}^\top$$

$$\boldsymbol{c}\boldsymbol{c}^\top\sigma_L^2 + \boldsymbol{k}\boldsymbol{k}^\top\sigma_{RF}^2 = \boldsymbol{v}^1\lambda_1{\boldsymbol{v}^1}^\top + \boldsymbol{v}^2\lambda_2{\boldsymbol{v}^2}^\top$$

The vectors $\boldsymbol{c}$ and $\boldsymbol{k}$ **are orthogonal. And so** $\boldsymbol{v}^i, \boldsymbol{v}^j$ $\forall i \neq j$ by definition of orthonormal eigenvector.

$\boldsymbol{c}\boldsymbol{c}^\top\sigma_L^2 + \boldsymbol{k}\boldsymbol{k}^\top\sigma_{RF}^2 = (\boldsymbol{c}\sigma_L^2 + \boldsymbol{k}\sigma_{RF}^2)(\boldsymbol{c}^\top + \boldsymbol{k}^\top)$

$\boldsymbol{v}^1\lambda_1{\boldsymbol{v}^1}^\top + \boldsymbol{v}^2\lambda_2{\boldsymbol{v}^2}^\top = (\boldsymbol{v}^1\lambda_1 + \boldsymbol{v}^2\lambda_2)\left({\boldsymbol{v}^1}^\top + {\boldsymbol{v}^2}^\top\right)$

Now we can normalize our vectors. Let $\hat{\boldsymbol{c}}$ and $\widehat{\boldsymbol{k}}$ be the normalized version. They correspond to

$\hat{\boldsymbol{c}} = \boldsymbol{c}/\sqrt{2M+1}$

$$\widehat{\boldsymbol{k}} = \frac{[-M,\dots M\,]^\top}{\sqrt{(-M)^2+\dots+M^2}} = \frac{[-M,\dots M\,]^\top}{\sqrt{2\sum_{j=1}^{M} j^2}} = \sqrt{\frac{3}{M(M+1)(2M+1)}}\ \boldsymbol{k}$$

Let $\|\boldsymbol{c}\| = \sqrt{2M+1}$ and $\|\boldsymbol{k}\| = \sqrt{\frac{3}{M(M+1)(2M+1)}}$

We should have the following

$$\hat{\boldsymbol{c}}\sigma_L^2\|\boldsymbol{c}\| + \widehat{\boldsymbol{k}}\sigma_{RF}^2\|\boldsymbol{k}\| = \boldsymbol{v}^1\lambda_1 + \boldsymbol{v}^2\lambda_2$$

With the equivalence

$\sigma_L^2\|\boldsymbol{c}\| = \lambda_1, \quad \sigma_{RF}^2\|\boldsymbol{k}\| = \lambda_2, \quad \hat{\boldsymbol{c}} = \boldsymbol{v}^1, \qquad \widehat{\boldsymbol{k}} = \boldsymbol{v}^2$

For the particular way the signal is generated in a simulated EO comb, the variance increase linearly with the observation time so

$$\sigma_L^2(t) = 2\pi\cdot\mathrm{LW_L}\cdot t$$

$$\sigma_{RF}^2(t) = 2\pi\cdot\mathrm{LW_{RF}}\cdot t$$

Therefore the theoretical values should be

$\lambda_1(t) = 2\pi\cdot\mathrm{LW_L}\cdot t\cdot\|\boldsymbol{c}\|$

$\lambda_2(t) = 2\pi\cdot\mathrm{LW_{RF}}\cdot t\cdot\|\boldsymbol{k}\|$

By projecting the data onto the eigenmodes, we can literally extract the noise source corresponding to the mode. Say $\boldsymbol{\Phi}$ is the matrix containing the phase time series

$$\boldsymbol{\Phi} = \begin{bmatrix} \phi_{\boldsymbol{1}}^{-\boldsymbol{M}} & \phi_{\boldsymbol{1}}^{-\boldsymbol{M}+\boldsymbol{1}} & \dots & \phi_{\boldsymbol{1}}^{\boldsymbol{M}} \\ \phi_{\boldsymbol{2}}^{-\boldsymbol{M}} & \phi_{\boldsymbol{2}}^{-\boldsymbol{M}+\boldsymbol{1}} & \dots & \phi_{\boldsymbol{2}}^{\boldsymbol{M}} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{\boldsymbol{T}}^{-\boldsymbol{M}} & \phi_{\boldsymbol{T}}^{-\boldsymbol{M}+\boldsymbol{1}} & \dots & \phi_{\boldsymbol{T}}^{\boldsymbol{M}} \end{bmatrix}$$

With 2M+1 columns and T rows. Then, you can multiply this matrix by the eigenmodes to get an estimate of the noise

$$\phi_{RF} = \boldsymbol{\Phi} \cdot \widehat{\boldsymbol{k}}$$

$$\phi_S = \boldsymbol{\Phi} \cdot \hat{\boldsymbol{c}}$$

Using the correlation matrix

The difference is that I have used the correlation instead of the covariance matrix. In the specific

$$\mathbf{R} = \left(\sqrt{\operatorname{diag}(\boldsymbol{\Sigma})}\right)^{-\mathbf{1}} \boldsymbol{\Sigma} \left(\sqrt{\operatorname{diag}(\boldsymbol{\Sigma})}\right)^{-\mathbf{1}}$$

We can use the previous results to include the identity $\boldsymbol{\Sigma} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\top}$ s.t.

$$\mathbf{R} = \left(\sqrt{\operatorname{diag}(\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\top})}\right)^{-\mathbf{1}} \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\top} \left(\sqrt{\operatorname{diag}(\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\top})}\right)^{-\mathbf{1}}$$

$$\mathbf{R} = \left(\sqrt{\operatorname{diag}\left(\sum_{i=1}^{2M+1} \boldsymbol{v}^i \lambda_i \boldsymbol{v}^{\boldsymbol{i}\top}\right)}\right)^{-\mathbf{1}} \left[\sum_{i=1}^{2M+1} \boldsymbol{v}^i \lambda_i \boldsymbol{v}^{\boldsymbol{i}\top}\right] \left(\sqrt{\sum_{i=1}^{2M+1} \boldsymbol{v}^i \lambda_i \boldsymbol{v}^{\boldsymbol{i}\top}}\right)^{-\mathbf{1}}$$

The diag() of a sum is the sum of the diag()

$$\mathbf{R} = \left(\sqrt{\sum_{i=1}^{2M+1} \operatorname{diag}\left(\boldsymbol{v}^i \lambda_i \boldsymbol{v}^{\boldsymbol{i}\top}\right)}\right)^{-\mathbf{1}} \left[\sum_{i=1}^{2M+1} \boldsymbol{v}^i \lambda_i \boldsymbol{v}^{\boldsymbol{i}\top}\right] \left(\sqrt{\sum_{i=1}^{2M+1} \operatorname{diag}\left(\boldsymbol{v}^i \lambda_i \boldsymbol{v}^{\boldsymbol{i}\top}\right)}\right)^{-\mathbf{1}} =$$

$$= \left(\sqrt{\sum_{i=1}^{2M+1} \begin{bmatrix} {v_1^i}^2 & 0 & \cdots & 0 \\ 0 & {v_2^i}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & {v_{2M+1}^i}^2 \end{bmatrix} \lambda_i}\right)^{-\mathbf{1}} \left[\sum_{i=1}^{2M+1} \boldsymbol{v}^i \lambda_i \boldsymbol{v}^{\boldsymbol{i}\top}\right] \left(\sqrt{\sum_{i=1}^{2M+1} \begin{bmatrix} {v_1^i}^2 & 0 & \cdots & 0 \\ 0 & {v_2^i}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & {v_{2M+1}^i}^2 \end{bmatrix} \lambda_i}\right)^{-\mathbf{1}} =$$

$$= \left(\sqrt{\begin{bmatrix} \sum {v_1^i}^2 \lambda_i & 0 & \cdots & 0 \\ 0 & \sum {v_2^i}^2 \lambda_i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum {v_{2M+1}^i}^2 \lambda_i \end{bmatrix}}\right)^{-\mathbf{1}} \left[\sum_{i=1}^{2M+1} \boldsymbol{v}^i \lambda_i \boldsymbol{v}^{\boldsymbol{i}\top}\right] \left(\sqrt{\begin{bmatrix} \sum {v_1^i}^2 \lambda_i & 0 & \cdots & 0 \\ 0 & \sum {v_2^i}^2 \lambda_i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum {v_{2M+1}^i}^2 \lambda_i \end{bmatrix}}\right)^{-\mathbf{1}} =$$

$$
= \begin{bmatrix} \sqrt{\sum {v_1^i}^2 \lambda_i} & 0 & \cdots & 0 \\ 0 & \sqrt{\sum {v_2^i}^2 \lambda_i} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\sum {v_{2M+1}^i}^2 \lambda_i} \end{bmatrix}^{-1} \left[\sum_{i=1}^{2M+1} \boldsymbol{v}^i \lambda_i \boldsymbol{v}^{i\top}\right] \begin{bmatrix} \sqrt{\sum {v_1^i}^2 \lambda_i} & 0 & \cdots & 0 \\ 0 & \sqrt{\sum {v_2^i}^2 \lambda_i} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\sum {v_{2M+1}^i}^2 \lambda_i} \end{bmatrix}^{-1} =
$$

$$
= \begin{bmatrix} \left(\sqrt{\sum {v_1^i}^2 \lambda_i}\right)^{-1} & 0 & \cdots & 0 \\ 0 & \left(\sqrt{\sum {v_2^i}^2 \lambda_i}\right)^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left(\sqrt{\sum {v_{2M+1}^i}^2 \lambda_i}\right)^{-1} \end{bmatrix} \left[\sum_{i=1}^{2M+1} \boldsymbol{v}^i \lambda_i \boldsymbol{v}^{i\top}\right] \begin{bmatrix} \left(\sqrt{\sum {v_1^i}^2 \lambda_i}\right)^{-1} & 0 & \cdots & 0 \\ 0 & \left(\sqrt{\sum {v_2^i}^2 \lambda_i}\right)^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left(\sqrt{\sum {v_{2M+1}^i}^2 \lambda_i}\right)^{-1} \end{bmatrix} =
$$

$$
= \begin{bmatrix} \left(\sum {v_1^i}^2 \lambda_i\right)^{-1} & 0 & \cdots & 0 \\ 0 & \left(\sum {v_2^i}^2 \lambda_i\right)^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left(\sqrt{\sum {v_{2M+1}^i}^2 \lambda_i}\right)^{-1} \end{bmatrix}
$$

It is difficult to relate the eigenvalues. But for an EO comb

$$
\boldsymbol{\Sigma} = \mathbf{c}\sigma_{\mathrm{L}}^2 \boldsymbol{c}^\top + \mathbf{k}\sigma_{\mathrm{RF}}^2 \mathbf{k}^\top = 2\pi T_s T[\Delta\nu_L \mathbf{c}\mathbf{c}^\top + \Delta\nu_{RF} \mathbf{k}\mathbf{k}^\top]
$$

The correlation is

$$
\begin{aligned}
\boldsymbol{\rho} &= \mathrm{diag}(\boldsymbol{\Sigma})^{-0.5} \boldsymbol{\Sigma} \mathrm{diag}(\boldsymbol{\Sigma})^{-0.5} \\
&= \begin{bmatrix} \frac{\Delta\nu_L + M^2 \Delta\nu_{RF}}{\Delta\nu_L + M^2 \Delta\nu_{RF}} & \frac{\Delta\nu_L + M(M-1)\Delta\nu_{RF}}{\sqrt{\Delta\nu_L + M^2\Delta\nu_{RF}}\sqrt{\Delta\nu_L + (M-1)^2\Delta\nu_{RF}}} & \cdots & \frac{\Delta\nu_L - M^2\Delta\nu_{RF}}{\Delta\nu_L + M^2\Delta\nu_{RF}} \\ \frac{\Delta\nu_L + M(M-1)\Delta\nu_{RF}}{\sqrt{\Delta\nu_L + M^2\Delta\nu_{RF}}\sqrt{\Delta\nu_L + (M-1)^2\Delta\nu_{RF}}} & \frac{\Delta\nu_L + (M-1)^2\Delta\nu_{RF}}{\Delta\nu_L + (M-1)^2\Delta\nu_{RF}} & \cdots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \cdots & \end{bmatrix}
\end{aligned}
$$

So the max correlation is 1 (of course) and the minimum is given by $\frac{\Delta\nu_L - M^2\Delta\nu_{RF}}{\Delta\nu_L + M^2\Delta\nu_{RF}}$

PSD of phase noise

If we want to calculate the PSD of a signal, and we want to returns its value in dBc per Hz, we have to

1. Estimate the noise time series $\varphi(t)$

2. Build a oscillatory signal with that phase noise, i.e. $y(t) = \exp(j\varphi(t))$
3. Calculate the PSD of this signal, PSD_y
4. Return the values in dBc by computing $PSD_y[dBc] = 10\log_{10} PSD_y$

Rate-Equation laser based filtering and parameter extraction

Rate equations

We finally reach a conclusive form for the rate equations

Gain

$$g(N(t), N_p(t)) = \frac{g_0}{1+\epsilon N_p(t)} \ln\left(\frac{N(t)+N_s}{N_{tr}+N_s}\right)$$

And the gain derivative a is

$$a = \frac{\partial g}{\partial N} = \frac{g_0}{(N(t)+N_s)\left(1+\epsilon N_p(t)\right)}$$

$$a_p = \frac{\partial g}{\partial N_p} = -\frac{\epsilon}{\left(1+\epsilon N_p(t)\right)} g = -\frac{\epsilon g_0}{\left(1+\epsilon N_p(t)\right)^2} \ln\left(\frac{N(t)+N_s}{N_{tr}+N_s}\right)$$

$$\frac{\partial^2 g}{\partial N^2} = -\frac{g_0}{(N(t)+N_s)^2\left(1+\epsilon N_p(t)\right)} = -\frac{1}{N(t)+N_s} g$$

$$\frac{\partial^2 g}{\partial N \partial N_p} = -\frac{\epsilon g_0}{(N(t)+N_s)\left(1+\epsilon N_p(t)\right)^2} = -\frac{\epsilon}{\left(1+\epsilon N_p(t)\right)} g$$

Spontaneous emission rate

$$R'_{sp} = \frac{\Gamma v_g n_{sp} g}{V}$$

Volume

$$V_p = V\Gamma^{-1}$$

Threshold current

$$I_{th} = I - \frac{q}{h\nu}\frac{P_{0_{ss}}}{\eta_0}$$

Deterministic

$$\frac{dN(t)}{dt} = \frac{\eta_i I}{qV} - \frac{N(t)}{\tau} - v_g N_p(t) g$$

$$\frac{dN_p(t)}{dt} = \left(\Gamma v_g g - \frac{1}{\tau_p}\right) N_p(t) + \Gamma R'_{sp} = \left(\Gamma v_g g - \frac{1}{\tau_p}\right) N_p(t) + \frac{\Gamma^2 v_g n_{sp}}{V} g$$

Substituting the gain expression, the spontaneous emission rate they becomes

$$\boxed{\begin{aligned} \frac{dN(t)}{dt} &= \frac{\eta_i I}{qV} - \frac{N(t)}{\tau} - \frac{v_g g_0 N_p(t)}{1+\epsilon N_p(t)} \ln\left(\frac{N(t)+N_s}{N_{tr}+N_s}\right) \\ \frac{dN_p(t)}{dt} &= \left(\frac{\Gamma v_g g_0}{1+\epsilon N_p(t)} \ln\left(\frac{N(t)+N_s}{N_{tr}+N_s}\right) - \frac{1}{\tau_p}\right) N_p(t) + \frac{\Gamma^2 v_g V^{-1} g_0 n_{sp}}{1+\epsilon N_p(t)} \ln\left(\frac{N(t)+N_s}{N_{tr}+N_s}\right) \end{aligned}}$$

Steady states

We need to solve the following system

$\begin{cases} \frac{dN(t)}{dt} = 0 \\ \frac{dN_p(t)}{dt} = 0 \end{cases}$	(1)	The solution of this system should be denoted as N_{ss}, N_{pss}. These are the steady states.

The best way to get a solution is to use a numerical method

The second equation gives us

$$N_{ss} = f_1(N_{pss}) = (N_{tr} + N_s) \exp\left(\frac{N_{pss}}{\tau_p}\left[\frac{1+\epsilon N_{pss}}{\Gamma v_g g_0 [N_{pss} + \Gamma V^{-1} n_{sp}]}\right]\right) - N_s$$

And the first equation

$$N_{ss} = f_2(N_{pss}) = \frac{\tau \eta_i I}{qV} - \left[\frac{\tau {N_{pss}}^2}{\tau_p \Gamma [N_{pss} + \Gamma V^{-1} n_{sp}]}\right]$$

Now by using a matlab solver for $f_Z(N_{pss}) = f_1(N_{pss}) - f_2(N_{pss}) = 0$, we will find a solution for N_{pss} which will be used in $f_1(N_{pss})$ or $f_2(N_{pss})$ to find the respective N_{ss}. We can then choose the one that makes (1) as close as possible to zero (It will never be zero due to the limited precision of numerical techniques for root finding).

To improve the method, we should input an interval for finding the root. Well, N_{pss} has to be strictly positive, but an interval [0,+inf) does not suit an optimization method. Instead, we will use a very small value and for the lower bound (Matlab: realmin) and to compute the upper bound, we will chose

$$f_Z(N_p^{max}) = f_1(N_p^{max}) - f_2(N_p^{max}) < realmax$$

Where realmax is the maximum number (positive) representable in Matlab with double precision.

Now the first function will of course dominate the second (because is an exponential vs a polynomial) one so we just solve

$$(N_{tr} + N_s)\, exp\left(\frac{N_p^{max}}{\tau_p}\left[\frac{1+\epsilon N_p^{max}}{\Gamma v_g g_0 [N_p^{max} + \Gamma V^{-1} n_{sp}]}\right]\right) - N_s < realmax$$

$$\frac{N_p^{max}(1+\epsilon N_p^{max})}{\tau_p \Gamma v_g g_0 [N_p^{max} + \Gamma V^{-1} n_{sp}]} < \log\left(\frac{realmax + N_s}{(N_{tr} + N_s)}\right)$$

$$\frac{N_p^{max} + \epsilon N_p^{max\,2}}{\tau_p \Gamma v_g g_0} < \log\left(\frac{realmax + N_s}{(N_{tr} + N_s)}\right)\left[N_p^{max} + \ \Gamma V^{-1} n_{sp}\right]$$

$$N_p^{max\,2} \frac{\epsilon}{\tau_p \Gamma v_g g_0} + N_p^{max}\left(\frac{1}{\tau_p \Gamma v_g g_0} - \log\left(\frac{realmax + N_s}{(N_{tr} + N_s)}\right)\right) - \Gamma V^{-1} n_{sp} \log\left(\frac{realmax + N_s}{(N_{tr} + N_s)}\right) < 0$$

$$N_p^{max\,2} + N_p^{max}\left(\frac{1}{\epsilon} - \frac{\tau_p \Gamma v_g g_0}{\epsilon} \log\left(\frac{realmax + N_s}{(N_{tr} + N_s)}\right)\right) - \frac{\tau_p \Gamma^2 v_g g_0 V^{-1} n_{sp}}{\epsilon} \log\left(\frac{realmax + N_s}{(N_{tr} + N_s)}\right) < 0$$

The solution is just below the highest root (still negative)

$$N_p^{max} < \frac{1}{2}\left\{-\left(\frac{1}{\epsilon} - \frac{\tau_p \Gamma v_g g_0}{\epsilon} log\left(\frac{realmax + N_s}{(N_{tr} + N_s)}\right)\right) + \sqrt{\left(\frac{1}{\epsilon} - \frac{\tau_p \Gamma v_g g_0}{\epsilon} log\left(\frac{realmax + N_s}{(N_{tr} + N_s)}\right)\right)^2 + 4\frac{\tau_p \Gamma v_g g_0 \Gamma V^{-1} n_{sp}}{\epsilon} \log\left(\frac{realmax + N_s}{(N_{tr} + N_s)}\right)}\right\}$$

Stochastic

Phase

Starting from the following phase definition $\frac{d\phi}{dt} = 2\pi\Delta\nu$, we have the frequency deviations maps to carrier deviations

$$\Delta\nu = \frac{\alpha}{4\pi} \Gamma v_g a \Delta N$$

We can use the derivative to approximate the carrier deviation $\Delta N = \frac{dN}{dt}\Delta t$. But for our implementation, we just use the deviation as the neat difference from the steady state value $\Delta N = N(t) - N_{ss}$. This reason is for simplifying analytical calculations later

$$\begin{aligned}
\frac{dN(t)}{dt} &= \frac{\eta_i I}{qV} - \frac{N(t)}{\tau} - v_g N_p(t) g + F_N(t) \\
\frac{dN_p(t)}{dt} &= \left(\Gamma v_g g - \frac{1}{\tau_p}\right) N_p(t) + \Gamma R'_{sp} + F_P(t) \\
\frac{d\phi(t)}{dt} &= \frac{\alpha}{2}\ \Gamma v_g a [N(t) - N_{ss}] + F_\phi(t) \\
P_0(t) &= \eta_0 h\nu V_p \tau_p^{-1} N_p(t) + F_0(t)
\end{aligned}$$

Substituting the gain expression, the spontaneous emission rate, we obtain

$$\begin{aligned}
\frac{dN(t)}{dt} &= \frac{\eta_i I}{qV} - \frac{N(t)}{\tau} - \frac{v_g N_p(t) g_0}{1 + \epsilon N_p(t)} \ln\left(\frac{N(t) + N_s}{N_{tr} + N_s}\right) + F_N(t) \\
\frac{dN_p(t)}{dt} &= \frac{\Gamma v_g g_0}{1 + \epsilon N_p(t)} \ln\left(\frac{N(t) + N_s}{N_{tr} + N_s}\right)\left(N_p(t) + \Gamma V^{-1} n_{sp}\right) - \frac{N_p(t)}{\tau_p} + F_P(t) \\
\frac{d\phi(t)}{dt} &= \frac{\alpha}{2}\ \Gamma v_g a \Delta t [N(t) - N_{ss}] + F_\phi(t) \\
P_0(t) &= \eta_0 h\nu V \Gamma^{-1} \tau_p^{-1} N_p(t) + F_0(t)
\end{aligned}$$

The optical power does not have a proper dynamic but it has a Langevin source associated to it

The Langevin noise strengths are defined as follow

$$\langle F_N F_N \rangle = 2R'_{sp}\Gamma^{-1} N_p(t)\left[1 + \left(2N_p(t)V_p\right)^{-1}\right] - v_g g N_p(t) V^{-1} + \eta_i \left(2I\ - \frac{q}{h\nu}\frac{P_{0ss}}{\eta_0}\right)(qV^2)^{-1}$$

$$\langle F_P F_P \rangle = 2\Gamma R'_{sp} N_p(t) \left[1 + \left(2N_p(t)V_p\right)^{-1}\right]$$

$$\langle F_P F_N \rangle = -2R'_{sp} N_p(t) \left[1 + \left(2N_p(t)V_p\right)^{-1}\right] + v_g g N_p / V_p$$

$$\langle F_\phi F_\phi \rangle = \Gamma R'_{sp} \left(2N_p(t)\right)^{-1}$$

$$\langle F_P F_\phi \rangle = \langle F_N F_\phi \rangle = 0$$

$$\langle F_P F_0 \rangle = -P_0(t) V_p^{-1}$$

$$\langle F_0 F_0 \rangle = h\nu P_0(t)$$

$$\langle F_N F_0 \rangle = \langle F_\phi F_0 \rangle = 0$$

Now replacing R'_{sp} , $V_p = V/\Gamma$, and the expression for the gain, and the stationary output power

$$P_{0_{ss}} = \eta_0 h\nu V_p \tau_p^{-1} N_{pss}$$

$$\langle F_N F_N \rangle = \frac{v_g g_0 N_p(t) V^{-1}}{\left(1+\epsilon N_p(t)\right)} \ln\left(\frac{N(t)+N_s}{N_{tr}+N_s}\right) \left\{2n_{sp}\left[1 + \Gamma\left(2N_p(t)V\right)^{-1}\right] - 1\right\} + \eta_i \left(2I - \frac{q}{h\nu} \frac{\eta_0 h\nu V_p \tau_p^{-1} N_{pss}}{\eta_0}\right) (qV^2)^{-1}$$
$$\langle F_P F_P \rangle = \frac{2\Gamma^2 v_g n_{sp} g_0 N_p(t) V^{-1}}{\left(1+\epsilon N_p(t)\right)} \ln\left(\frac{N(t)+N_s}{N_{tr}+N_s}\right) \left[1 + \Gamma\left(2N_p(t)V\right)^{-1}\right]$$
$$\langle F_P F_N \rangle = \frac{\Gamma v_g g_0 N_p(t) V^{-1}}{\left(1+\epsilon N_p(t)\right)} \ln\left(\frac{N(t)+N_s}{N_{tr}+N_s}\right) \left\{1 - 2n_{sp}\left[1 + \Gamma\left(2N_p(t)V\right)^{-1}\right]\right\}$$
$$\langle F_\phi F_\phi \rangle = \frac{\Gamma^2 v_g n_{sp} g_0 V^{-1}}{2N_p(t)\left(1+\epsilon N_p(t)\right)} \ln\left(\frac{N(t)+N_s}{N_{tr}+N_s}\right)$$
$$\langle F_P F_\phi \rangle = \langle F_N F_\phi \rangle = 0$$
$$\langle F_P F_0 \rangle = -P_0(t) V^{-1} \Gamma$$
$$\langle F_0 F_0 \rangle = h\nu P_0(t)$$
$$\langle F_N F_0 \rangle = \langle F_\phi F_0 \rangle = 0$$

Euler solutions (with parameters)

The time step used is Δt

We use different colors for different parameters

$$N(t+1) = N(t) + \Delta t \left\{\frac{\eta_i I V^{-1}}{q} - \frac{N(t)}{\tau} - \frac{v_g g_0 N_p(t)}{1+\epsilon N_p(t)} \ln\left(\frac{N(t)+N_s}{N_{tr}+N_s}\right)\right\} + \sqrt{\Delta t} F_N(t)$$
$$N_p(t+1) = N_p(t) + \Delta t \left\{\frac{\Gamma v_g g_0}{1+\epsilon N_p(t)} \ln\left(\frac{N(t)+N_s}{N_{tr}+N_s}\right)\left(N_p(t) + \Gamma V^{-1} n_{sp}\right) - \frac{N_p(t)}{\tau_p}\right\} + \sqrt{\Delta t} F_P(t)$$
$$\phi(t+1) = \phi(t) + \Delta t \left\{\frac{\alpha}{2} \frac{\Gamma v_g g_0 [N(t) - N_{ss}]}{(N(t)+N_s)\left(1+\epsilon N_p(t)\right)}\right\} + \sqrt{\Delta t} F_\phi(t)$$
$$P_0(t) = h\eta_0 \nu V \Gamma^{-1} \tau_p^{-1} N_p(t) + \sqrt{\Delta t} F_0(t)$$

$$\langle F_N F_N \rangle = \frac{v_g g_0 N_p(t) V^{-1}}{\left(1 + \epsilon N_p(t)\right)} \ln\left(\frac{N(t) + N_s}{N_{tr} + N_s}\right) \left\{2 n_{sp} \left[1 + \Gamma\left(2 N_p(t) V\right)^{-1}\right] - 1\right\} + \eta_i V^{-1} \left(2 q^{-1} I V^{-1} - N_{pss} \Gamma^{-1} \tau_p^{-1}\right)$$

$$\langle F_P F_P \rangle = \frac{2 \Gamma^2 n_{sp} v_g g_0 N_p(t) V^{-1}}{\left(1 + \epsilon N_p(t)\right)} \ln\left(\frac{N(t) + N_s}{N_{tr} + N_s}\right) \left[1 + \Gamma\left(2 N_p(t) V\right)^{-1}\right]$$

$$\langle F_P F_N \rangle = \frac{\Gamma v_g g_0 N_p(t) V^{-1}}{\left(1 + \epsilon N_p(t)\right)} \ln\left(\frac{N(t) + N_s}{N_{tr} + N_s}\right) \left\{1 - 2 n_{sp} \left[1 + \Gamma\left(2 N_p(t) V\right)^{-1}\right]\right\}$$

$$\langle F_\phi F_\phi \rangle = \frac{\Gamma^2 n_{sp} v_g g_0 V^{-1}}{2 N_p(t) \left(1 + \epsilon N_p(t)\right)} \ln\left(\frac{N(t) + N_s}{N_{tr} + N_s}\right)$$

$$\langle F_P F_\phi \rangle = \langle F_N F_\phi \rangle = 0$$

$$\langle F_P F_0 \rangle = -h \eta_0 \nu \tau_p^{-1} N_p(t)$$

$$\langle F_0 F_0 \rangle = h^2 \eta_0 V \Gamma^{-1} \nu^2 \tau_p^{-1} N_p(t)$$

$$\langle F_N F_0 \rangle = \langle F_\phi F_0 \rangle = 0$$

The measurement equation is $y(t) = E(t) \cos\left(\Delta\omega t T_s + \Delta\phi(t)\right) + r(t)$

According to the LO and LUT contributions

$E(t) \propto 2R\sqrt{P_0(t) P_{LO}(t)}$	$\Delta\phi(t) \propto \phi(t) - \phi_{LO}(t)$

Assuming that we don't care about LO, i.e. constant power and phase noiseless

$$y(t) = 2R\sqrt{P_{LO} P_0(t)} \cos\left(\Delta\omega t T_s + \phi(t)\right) + r(t) = R' \sqrt{P_0(t)} \cos\left(\Delta\omega t T_s + \phi(t)\right) + r(t)$$

R' is a proportionality constant that groups everything (LO power and responsivity)

$$R' = 2R\sqrt{P_{LO}}$$

Static parameter list (complete)

Symbol	Full Name (from Coldren's book)	Order in the parameter list	Starting value	Strict bounds	Loose bounds
τ	Carrier lifetime	1	2.71e-9	[0,+∞]	[1e-9,3e-9]
$v_g g_0$	Group velocity times gain	2	c/3.8*3e5	[0,+∞]	[c/4*1e5,c/3*5e5]
ϵ	gain compression factor	3	1.5e-23	[0,+∞]	[1e-26, 1e-20]
N_s	Gain correction factor	4	1.1e24	[-∞,+∞]	[-1e24, 8e24]
Γ	Confinement factor	5	5e-2	[0,1]	[0,0.5]
N_{tr}	Carriers at transparency	6	2.6e24	[0,+∞]	[1e24, 4e24]
I	injected current	7	0.3	[0,+∞]	[0.2,0.4]
η_i	Injection efficiency	8	0.8	[0,1]	[0,1]
τ_p	Photons lifetime	9	2.77e-12	[0,+∞]	[1e-12,4e-12]

V	Carrier volume	10	2.7e-16	[0,+∞]	[1e-20,1e-15]
n_{sp}	population inversion factor	11	1.5	[0,+∞]	[1,2]
α	Linewidth enhancement factor	12	4	[0,+∞]	[1,7]
η_0	optical efficiency	14	0.45	[0,1]	[0,1]
R'	Effective responsivity	-	from avg_power		
σ_r^2	Measurement noise power	-	from EM		
$\Delta\omega$	beating frequency	-	from conv analysis (polyfit)		
N_{ss}, N_{pss}	Carriers and photons steady state number	-	from steady state function		
f_R	Relaxation frequency	-	There a direct expression from that		order of GHz
ν	Lasing frequency	13	c/1.575e-6		

The lasing frequency has been removed from the parameters to be estimate, because tiny variations of it in the desired range do not affect the final result. Instead we kept eta0 as parameter

UKF implementation by normalization (revise)

It is necessary to normalize the states, by subtracting the steady states from the state itself (so only the oscillation around the state are tracked) and also the state should be divided by the magnitude of the excursion (so maximum excursion should be around unity).

How to transform the state space model:

1. Subtract the steady state
2. Divide by standard deviation of Langevin noise

This for photons and carriers. I would not touch the phase. The power is normalized only by subtracting the steady state

$n(t) = \frac{N(t) - N^*}{\sqrt{L_N^*}}$	$n_p(t) = \frac{N_p(t) - N_p^*}{\sqrt{L_P^*}}$	$n(t)\sqrt{L_N^*} + N^* = N(t)$	$n_p(t)\sqrt{L_P^*} + N_p^* = N_p(t)$	$p_0(t) = P_0(t) - P_0^*$	$p_0(t) + P_0^* = P_0(t)$

Where we define the Langevin strength at the steady state (N^*, N_p^*) as

$$L_N^* = \frac{v_g g_0 N_p^* V^{-1}}{(1 + \epsilon N_p^*)} \ln\left(\frac{N^* + N_s}{N_{tr} + N_s}\right)\left\{2n_{sp}\left[1 + \Gamma\left(2N_p^* V\right)^{-1}\right] - 1\right\} + \eta_i(I + I_{th})(qV^2)^{-1}$$

$$L_P^* = \frac{2\Gamma^2 v_g n_{sp} g_0 N_p^* V^{-1}}{\left(1+\epsilon N_p^*\right)} \ln\left(\frac{N^*+N_s}{N_{tr}+N_s}\right)\left[1+\Gamma\left(2N_p^* V\right)^{-1}\right]$$

$$n(t+1) = n(t) + \frac{\Delta t}{\sqrt{L_N^*}}\left\{\frac{\eta_i I}{qV} - \frac{n(t)\sqrt{L_N^*}+N^*}{\tau} - \frac{v_g\left(n_p(t)\sqrt{L_P^*}+N_p^*\right)g_0}{1+\epsilon\left(n_p(t)\sqrt{L_P^*}+N_p^*\right)}\ln\left(\frac{n(t)\sqrt{L_N^*}+N^*+N_s}{N_{tr}+N_s}\right)\right\} + \frac{\sqrt{\Delta t}}{\sqrt{L_N^*}}F_N(t)$$

$$n_p(t+1) = n_p(t) + \frac{\Delta t}{\sqrt{L_P^*}}\left\{\frac{\Gamma v_g g_0}{1+\epsilon\left(n_p(t)\sqrt{L_P^*}+N_p^*\right)}\ln\left(\frac{n(t)\sqrt{L_N^*}+N^*+N_s}{N_{tr}+N_s}\right)\left(n_p(t)\sqrt{L_P^*}+N_p^*+\Gamma V^{-1}n_{sp}\right) - \frac{n_p(t)\sqrt{L_P^*}+N_p^*}{\tau_p}\right\} + \frac{\sqrt{\Delta t}}{\sqrt{L_P^*}}F_P(t)$$

$$\phi(t+1) = \phi(t) + \Delta t\left\{\frac{\alpha}{2}\frac{\Gamma v_g g_0}{\left(n(t)\sqrt{L_N^*}+N^*+N_s\right)\left(1+\epsilon\left(n_p(t)\sqrt{L_P^*}+N_p^*\right)\right)}\Delta t\left[\frac{\eta_i I}{qV} - \frac{n(t)\sqrt{L_N^*}+N^*}{\tau} - \frac{v_g\left(n_p(t)\sqrt{L_P^*}+N_p^*\right)g_0}{1+\epsilon\left(n_p(t)\sqrt{L_P^*}+N_p^*\right)}\ln\left(\frac{n(t)\sqrt{L_N^*}+N^*+N_s}{N_{tr}+N_s}\right)\right]\right\} + \sqrt{\Delta t}F_\phi(t)$$

$$p_0(t) = \sqrt{L_P^*}\eta_0 h\nu V\Gamma^{-1}\tau_p^{-1}n_p(t) + \sqrt{\Delta t}F_0(t)$$

As alternative, the phase equation can be written as

$$\phi(t+1) = \phi(t) + \Delta t\left\{\frac{\alpha}{2}\frac{\Gamma v_g g_0}{\left(n(t)\sqrt{L_N^*}+N^*+N_s\right)\left(1+\epsilon\left(n_p(t)\sqrt{L_P^*}+N_p^*\right)\right)}n(t)\sqrt{L_N^*}\right\} + \sqrt{\Delta t}F_\phi(t)$$

Then the starting vector can easily be 0

$m_0 = [0,0,0,0]$, $P_0 = I_4$

The Langevin strengths with normalized states are so defined

$$\langle F_N F_N\rangle = \frac{v_g g_0\left(n_p(t)\sqrt{L_P^*}+N_p^*\right)V^{-1}}{1+\epsilon\left(n_p(t)\sqrt{L_P^*}+N_p^*\right)}\ln\left(\frac{n(t)\sqrt{L_N^*}+N^*+N_s}{N_{tr}+N_s}\right)\left\{2n_{sp}\left[1+\Gamma\left(2\left[n_p(t)\sqrt{L_P^*}+N_p^*\right]V\right)^{-1}\right]-1\right\} + \eta_i(I+I_{th})(qV^2)^{-1}$$

$$\langle F_P F_P\rangle = \frac{2\Gamma^2 v_g n_{sp} g_0\left(n_p(t)\sqrt{L_P^*}+N_p^*\right)V^{-1}}{1+\epsilon\left(n_p(t)\sqrt{L_P^*}+N_p^*\right)}\ln\left(\frac{n(t)\sqrt{L_N^*}+N^*+N_s}{N_{tr}+N_s}\right)\left[1+\Gamma\left(2\left(n_p(t)\sqrt{L_P^*}+N_p^*\right)V\right)^{-1}\right]$$

$$\langle F_P F_N\rangle = \frac{\Gamma v_g g_0\left(n_p(t)\sqrt{L_P^*}+N_p^*\right)V^{-1}}{1+\epsilon\left(n_p(t)\sqrt{L_P^*}+N_p^*\right)}\ln\left(\frac{n(t)\sqrt{L_N^*}+N^*+N_s}{N_{tr}+N_s}\right)\left\{1-2n_{sp}\left[1+\Gamma\left(2\left(n_p(t)\sqrt{L_P^*}+N_p^*\right)V\right)^{-1}\right]\right\}$$

$$\langle F_\phi F_\phi\rangle = \frac{\Gamma^2 v_g n_{sp} g_0 V^{-1}}{2\left(n_p(t)\sqrt{L_P^*}+N_p^*\right)\left(1+\epsilon\left(n_p(t)\sqrt{L_P^*}+N_p^*\right)\right)}\ln\left(\frac{n(t)\sqrt{L_N^*}+N^*+N_s}{N_{tr}+N_s}\right)$$

$$\langle F_P F_\phi\rangle = \langle F_N F_\phi\rangle = 0$$

$$\langle F_P F_0\rangle = -\left(p_0(t)+P_0^*\right)V^{-1}\Gamma$$

$$\langle F_0 F_0\rangle = h\nu\left(p_0(t)+P_0^*\right)$$

$$\langle F_N F_0\rangle = \langle F_\phi F_0\rangle = 0$$

Cholesky Decomposition of the Q matrix

Sometimes it is necessary to explicitly use a white normal Gaussian noise and write its dependence on the states using the Cholesky factor

$$\Sigma = \begin{pmatrix} \langle F_N F_N \rangle & \langle F_P F_N \rangle & 0 & 0 \\ \langle F_P F_N \rangle & \langle F_P F_P \rangle & 0 & \langle F_P F_0 \rangle \\ 0 & 0 & \langle F_\phi F_\phi \rangle & 0 \\ 0 & \langle F_P F_0 \rangle & 0 & \langle F_0 F_0 \rangle \end{pmatrix} = LL^\top = \begin{pmatrix} L_{11} & 0 & 0 & 0 \\ L_{21} & L_{22} & 0 & 0 \\ L_{31} & L_{32} & L_{33} & 0 \\ L_{41} & L_{42} & L_{43} & L_{44} \end{pmatrix} \begin{pmatrix} L_{11} & L_{21} & L_{31} & L_{41} \\ 0 & L_{22} & L_{32} & L_{42} \\ 0 & 0 & L_{33} & L_{44} \\ 0 & 0 & 0 & L_{44} \end{pmatrix} = \cdots$$

$$= \begin{pmatrix} L_{11}^2 & 0 & 0 & 0 \\ L_{11}L_{21} & L_{21}^2 + L_{22}^2 & 0 & 0 \\ L_{11}L_{31} & L_{21}L_{31} + L_{22}L_{32} & L_{31}^2 + L_{31}^2 + L_{33}^2 & 0 \\ L_{11}L_{41} & L_{21}L_{41} + L_{22}L_{42} & L_{31}L_{41} + L_{31}L_{42} + L_{33}L_{43} & L_{41}^2 + L_{42}^2 + L_{43}^2 + L_{44}^2 \end{pmatrix}$$

Now by equating the terms we can find that

$$L_{11} = \sqrt{\langle F_N F_N \rangle}$$
$$L_{21} = \frac{\langle F_P F_N \rangle}{\sqrt{\langle F_N F_N \rangle}}$$
$$L_{22} = \sqrt{\langle F_P F_P \rangle - \frac{\langle F_P F_N \rangle^2}{\langle F_N F_N \rangle}}$$
$$L_{31} = 0$$
$$L_{32} = 0$$
$$L_{33} = \sqrt{\langle F_\phi F_\phi \rangle}$$
$$L_{41} = 0$$
$$L_{42} = \frac{\langle F_P F_0 \rangle}{\sqrt{\langle F_P F_P \rangle - \frac{\langle F_P F_N \rangle^2}{\langle F_N F_N \rangle}}}$$
$$L_{43} = 0$$
$$L_{44} = \sqrt{\langle F_0 F_0 \rangle - \frac{\langle F_P F_0 \rangle^2}{\langle F_P F_P \rangle - \frac{\langle F_P F_N \rangle^2}{\langle F_N F_N \rangle}}}$$

Stability of the implementation

We need to check the stability in different parts. The stability of the numerical method, since we use Euler-Mayurama, is included in the stability of discrete time systems.

Stability of the deterministic discrete-time rate equation system (eigenvalues)

According to the 1st method of Lyapunov for discrete-time systems, $\boldsymbol{x}(t+1) = \boldsymbol{f}\big(\boldsymbol{x}(t)\big)$ is asymptotically stable around the equilibrium point $\boldsymbol{x}^*$ if the eigenvalues of the Jacobian evaluated at $\boldsymbol{x}^*$ are strictly less than 1 in absolute value.

$$N(t+1) = N(t) + \Delta t \left\{ \frac{\eta_i I V^{-1}}{q} - \frac{N(t)}{\tau} - \frac{v_g g_0 N_p(t)}{1 + \epsilon N_p(t)} \ln\left(\frac{N(t) + N_s}{N_{tr} + N_s} \right) \right\}$$

$$N_p(t+1) = N_p(t) + \Delta t \left\{ \frac{\Gamma v_g g_0}{1 + \epsilon N_p(t)} \ln\left(\frac{N(t) + N_s}{N_{tr} + N_s} \right) \left(N_p(t) + \Gamma V^{-1} n_{sp} \right) - \frac{N_p(t)}{\tau_p} \right\}$$

With $\boldsymbol{x}(t) = \begin{bmatrix} N(t) \\ N_p(t) \end{bmatrix}$, we can identify the functional form $\boldsymbol{f}$ as follow

$$\boldsymbol{f}\big(\boldsymbol{x}(t)\big) := \mathbf{I}\boldsymbol{x}(t) + \Delta t\boldsymbol{d}\big(\boldsymbol{x}(t)\big)$$

Where $\mathbf{I}$ is the identity matrix, and $\boldsymbol{d}$ is the deterministic part of the rate equation. We need first to find the Jacobian. Now we try to diagonalize it and find the eigenvalues as function of the parameters (and state). We will then replace the state with the steady state and set the eigenvalues conditions, which they would be conditions for the parameters (and the time step, hopefully).

The system Jacobian is

$$\boldsymbol{J}\big(N, N_p\big) = \boldsymbol{I} + \Delta t \begin{bmatrix} -\dfrac{1}{\tau} - v_g N_p \dfrac{dg}{dN} & -v_g\left(g + N_p \dfrac{dg}{dN_p}\right) \\ \Gamma v_g \dfrac{dg}{dN}\big(N_p + \Gamma n_{sp} V^{-1}\big) & -\dfrac{1}{\tau_p} + \Gamma v_g\left(g + \dfrac{dg}{dN_p}\big[N_p + \Gamma n_{sp} V^{-1}\big]\right) \end{bmatrix}$$

The eigenvalues solve $|\boldsymbol{J} - \lambda\boldsymbol{I}| = 0$

$$\begin{vmatrix} 1 + \Delta t\left[-\dfrac{1}{\tau} - v_g N_p \dfrac{dg}{dN}\right] - \lambda & \Delta t\left[-v_g\left(g + N_p \dfrac{dg}{dN_p}\right)\right] \\ \Delta t\left[\Gamma v_g \dfrac{dg}{dN}\big(N_p + \Gamma n_{sp} V^{-1}\big)\right] & 1 + \Delta t\left[-\dfrac{1}{\tau_p} + \Gamma v_g\left(g + \dfrac{dg}{dN_p}\big[N_p + \Gamma n_{sp} V^{-1}\big]\right)\right] - \lambda \end{vmatrix} = 0$$

$$\left\{1 + \Delta t\left[-\frac{1}{\tau} - v_g N_p \frac{dg}{dN}\right] - \lambda\right\}\left\{1 + \Delta t\left[-\frac{1}{\tau_p} + \Gamma v_g\left(g + \frac{dg}{dN_p}\big[N_p + \Gamma n_{sp} V^{-1}\big]\right)\right] - \lambda\right\} - \Delta t^2\left[-v_g\left(g + N_p \frac{dg}{dN_p}\right)\right]\left[\Gamma v_g \frac{dg}{dN}\big(N_p + \Gamma n_{sp} V^{-1}\big)\right] = 0$$

$$\lambda^2 - \lambda\left\{2 + \Delta t\left[-\frac{1}{\tau_p} + \Gamma v_g\left(g + \frac{dg}{dN_p}\big[N_p + \Gamma n_{sp} V^{-1}\big]\right) - \frac{1}{\tau} - v_g N_p \frac{dg}{dN}\right]\right\} + \Delta t^2\left[v_g\left(g + N_p \frac{dg}{dN_p}\right)\right]\left[\Gamma v_g \frac{dg}{dN}\big(N_p + \Gamma n_{sp} V^{-1}\big)\right] = 0$$

$$\lambda^2 - \lambda\left\{2 + \Delta t\left[-\frac{1}{\tau_p} + \Gamma v_g g\left(1 - \frac{\epsilon\big[N_p + \Gamma n_{sp} V^{-1}\big]}{1 + \epsilon N_p}\right) - \frac{1}{\tau} - \frac{v_g N_p g_0}{(1 + \epsilon N_p)(N + N_s)}\right]\right\} + \Delta t^2\left[v_g g\left(1 + \frac{N_p \epsilon}{1 + \epsilon N_p}\right)\right]\left[\frac{\Gamma v_g\big(N_p + \Gamma n_{sp} V^{-1}\big) g_0}{(1 + \epsilon N_p)(N + N_s)}\right] = 0$$

$$\lambda^{(1,2)} = \frac{1}{2}\left\{2 + \Delta t\left[-\frac{1}{\tau_p} + \Gamma v_g\left(g + \frac{dg}{dN_p}\big[N_p + \Gamma n_{sp} V^{-1}\big]\right) - \frac{1}{\tau} - v_g N_p \frac{dg}{dN}\right] \pm \sqrt{\left\{2 + \Delta t\left[-\frac{1}{\tau_p} + \Gamma v_g\left(g + \frac{dg}{dN_p}\big[N_p + \Gamma n_{sp} V^{-1}\big]\right) - \frac{1}{\tau} - v_g N_p \frac{dg}{dN}\right]\right\}^2 - 4\Delta t^2\left[v_g\left(g + N_p \frac{dg}{dN_p}\right)\right]\left[\Gamma v_g \frac{dg}{dN}\big(N_p + \Gamma n_{sp} V^{-1}\big)\right]}\right\}$$

We can try to write down the conditions for the stability

$\left|\lambda^{(1)}\right| < 1, \left|\lambda^{(2)}\right| < 1$

Which translate into

$\left|\lambda^{(1)}(\Delta t)\right| < 1, \left|\lambda^{(2)}(\Delta t)\right| < 1$

This condition can be transformed in a condition for the timestep, which can be chosen small enough to guarantee the stability;

<u>A quick solution to this problem:</u>

The Jacobian is going to be evaluated at the steady states

$$\boldsymbol{J} = \boldsymbol{I} + \Delta t\boldsymbol{A}$$

If λ is an eigenvalue of $\boldsymbol{A}$**,** then $\mu = 1 + \Delta t\lambda$ is an eigenvalue of $\boldsymbol{J}$**.**

Proof:

$$\boldsymbol{Aa} = \lambda \boldsymbol{a}, \qquad \boldsymbol{Ja} = [\boldsymbol{I} + \Delta t \boldsymbol{A}]\boldsymbol{a} = \boldsymbol{a} + \Delta t \lambda \boldsymbol{a} = [1 + \Delta t \lambda]\boldsymbol{a}. \quad \blacksquare$$

The requirement for the continuous system eigenvalue stability is that λ, eigenvalue of $\boldsymbol{A}$, needs to have the real part strictly negative. If this condition is satisfied, $Re\lambda < 0$, we then move to the next part

The requirement on the eigenvalue stability for discrete time systems translates into

$|1 + \Delta t \lambda| < 1$

$\sqrt{[1 + \Delta t Re\lambda]^2 + [\Delta t Im\lambda]^2} < 1$, always positive I can remove the squareroot

$1 + Re^2\lambda\Delta t^2 + 2Re\lambda\Delta t + Im^2\lambda\Delta t^2 < 1$ we are going to get a parabola in Δt

$\Delta t\{\Delta t[Re(\lambda)^2 + Im(\lambda)^2] + 2[Re(\lambda)]\} < 0$ we can then exclude the negative zone because $\Delta t > 0$

For stability, $\Delta t < -\frac{2[Re(\lambda)]}{[Re(\lambda)^2 + Im(\lambda)^2]}$ $\blacksquare$

Stability of the stochastic discrete-time rate equation system

I have no idea to be honest, there is more literature to read

Condition on the steady states

They need to be strictly positive

Condition on the covariance matrix of the noise

Has to be positive definite

EKF implementation (revise)

Why? Well because it is faster and more stable than the UKF counterpart (if we want to propagate complex units)

What do we need? The Jacobian of the prediction and measure function

$$\mathbf{F_x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \Delta t \begin{bmatrix} \frac{\partial \boldsymbol{f}_N}{\partial N} & \frac{\partial \boldsymbol{f}_N}{\partial N_p} & 0 & 0 \\ \frac{\partial \boldsymbol{f}_{N_p}}{\partial N} & \frac{\partial \boldsymbol{f}_{N_p}}{\partial N_p} & 0 & 0 \\ \frac{\partial \boldsymbol{f}_{\phi}}{\partial N} & \frac{\partial \boldsymbol{f}_{\phi}}{\partial N_p} & 0 & 0 \\ 0 & \frac{\partial \boldsymbol{f}_{P_0}}{\partial N_p} & 0 & 0 \end{bmatrix}$$

$\frac{\partial \boldsymbol{F}_N}{\partial N} = 1 + \Delta t\left[-\frac{1}{\tau} - v_g N_p \frac{\partial g}{\partial N}\right]$ $\frac{\partial \boldsymbol{F}_N}{\partial N_p} = \Delta t\left[-v_g\left(g + N_p \frac{\partial g}{\partial N_p}\right)\right]$ $\frac{\partial \boldsymbol{F}_{N_p}}{\partial N} = \Delta t\Gamma\left[N_p + \frac{\Gamma n_{sp}}{V}\right] v_g \frac{\partial g}{\partial N}$	$\frac{\partial g}{\partial N} = a = \frac{g_0}{(N + N_s)(1 + \epsilon N_p)}$ $\frac{\partial g}{\partial N_p} = -\frac{\epsilon g}{(1 + \epsilon N_p)}$ $g = \frac{g_0}{1 + \epsilon N_p} \ln\left(\frac{N + N_s}{N_{tr} + N_s}\right)$

$\frac{\partial \boldsymbol{F}_{N_p}}{\partial N_p} = 1 + \Delta t\left[\Gamma v_g\left(g + N_p\frac{\partial g}{\partial N_p}\right) - \frac{1}{\tau_p} + \frac{\Gamma^2 v_g n_{sp}}{V}\frac{\partial g}{\partial N_p}\right]$ $\frac{\partial \boldsymbol{F}_{\phi}}{\partial N} = \Delta t\left[\frac{\alpha}{2}\Gamma v_g \Delta t\left(\frac{\partial a}{\partial N}\left\{\frac{\eta_i I}{qV} - \frac{N}{\tau} - v_g N_p g\right\} + a\left\{-\frac{1}{\tau} - v_g N_p\frac{\partial g}{\partial N}\right\}\right)\right]$ $\frac{\partial \boldsymbol{F}_{\phi}}{\partial N_p} = \Delta t\left[\frac{\alpha}{2}\Gamma v_g \Delta t\left(\frac{\partial a}{\partial N_p}\left\{\frac{\eta_i I}{qV} - \frac{N}{\tau} - v_g N_p g\right\} + a\left\{-v_g\left(g + N_p\frac{\partial g}{\partial N_p}\right)\right\}\right)\right]$ $\frac{\partial \boldsymbol{F}_{\phi}}{\partial \phi} = 1$ $\frac{\partial F_{P_0}}{\partial N_p} = \frac{\partial}{\partial N_p}\left[\eta_0 h\nu V\Gamma^{-1}\tau_p^{-1}N_p(t)\right] = \eta_0 h\nu V\Gamma^{-1}\tau_p^{-1}$	$\frac{\partial a}{\partial N} = -\frac{g_0}{(N+N_s)^2(1+\epsilon N_p)}$ $\frac{\partial a}{\partial N_p} = -\frac{\epsilon g_0}{(N+N_s)(1+\epsilon N_p)^2}$

For the measurement equation we have

$$\mathbf{h}_{\mathbf{x}} = \left[0\ 0\ \frac{\partial h}{\partial \phi}\ \frac{\partial h}{\partial P_0}\right]$$

$$\frac{\partial h}{\partial \phi} = -R'\sqrt{P_0(t)}\sin\left(\Delta\omega t T_s + \phi(t)\right)$$

$$\frac{\partial h}{\partial P_0} = \frac{R'}{2\sqrt{P_0(t)}}\, cos\left(\Delta\omega t T_s + \phi(t)\right)$$

Small signal analysis

Necessary to compute RIN and FN from the given static parameters

1. We assume that the states have fluctuations around the steady state $\left(\bar{N}, \bar{N}_p, 0, \bar{P}_0\right)$

$$N(t) = \bar{N} + \delta N(t) = \bar{N} + \int_{-\infty}^{+\infty} N_\omega e^{j\omega t} d\omega$$

$$N_p(t) = \bar{N}_p + \delta N_p(t) = \bar{N}_p + \int_{-\infty}^{+\infty} {N_p}_\omega e^{j\omega t} d\omega$$

$$\Delta\nu(t) = \int_{-\infty}^{+\infty} \Delta\nu_\omega e^{j\omega t} d\omega$$

2. We say that the steady states are much larger than the fluctuations

   $N_\omega \ll \bar{N}$

   ${N_p}_\omega \ll \bar{N}_p$

   The steady state for the optical phase is zero

The first step is to linearize the equations for $\bar{N}, \bar{N}_p, 0, \bar{P}_0$

$$\frac{dN(t)}{dt} = \left.\frac{df_N}{dN}\right|_{\bar{N},\bar{N}_p,0}(N - \bar{N}) + \left.\frac{df_N}{dN_p}\right|_{\bar{N},\bar{N}_p,0}\left(N_p - \bar{N}_p\right) + F_N$$

$$\frac{dN_p(t)}{dt} = \left.\frac{df_{N_p}}{dN}\right|_{\bar{N},\bar{N}_p,0} (N - \bar{N}) + \left.\frac{df_{N_p}}{dN_p}\right|_{\bar{N},\bar{N}_p,0} \left(N_p - \bar{N}_p\right) + F_{N_p}$$

$$\frac{d\phi(t)}{dt} = \left.\frac{df_\phi}{dN}\right|_{\bar{N},\bar{N}_p,0} (N - \bar{N}) + \left.\frac{df_\phi}{dN_p}\right|_{\bar{N},\bar{N}_p,0} \left(N_p - \bar{N}_p\right) + F_\phi$$

By doing some algebra

$$\frac{dN(t)}{dt} = \left.\left[-\frac{1}{\tau} - v_g N_p \frac{\partial g}{\partial N}\right]\right|_{\bar{N},\bar{N}_p,0} (N - \bar{N}) + \left.\left[-v_g \left(N_p \frac{\partial g}{\partial N_p} + g\right)\right]\right|_{\bar{N},\bar{N}_p,0} \left(N_p - \bar{N}_p\right) + F_N$$

$$\frac{dN(t)}{dt} = \left[-\frac{1}{\tau} - \bar{N}_p \frac{v_g g_0}{(\bar{N} + N_s)\left(1 + \epsilon \bar{N}_p\right)}\right] (N - \bar{N}) + \left[\frac{v_g g_0}{1 + \epsilon \bar{N}_p} ln\left(\frac{\bar{N} + N_s}{N_{tr} + N_s}\right)\left(\frac{\bar{N}_p \epsilon}{\left(1 + \epsilon \bar{N}_p\right)} - 1\right)\right] \left(N_p - \bar{N}_p\right) + F_N$$

$$\frac{dN_p(t)}{dt} = \left.\left[\Gamma v_g \left(N_p + \frac{\Gamma n_{sp}}{V}\right) \frac{\partial g}{\partial N}\right]\right|_{\bar{N},\bar{N}_p,0} (N - \bar{N}) + \left.\left[\Gamma v_g g \left(1 - \frac{\epsilon}{1 + \epsilon N_p}\left[N_p + \frac{\Gamma n_{sp}}{V}\right]\right) - \frac{1}{\tau_p}\right]\right|_{\bar{N},\bar{N}_p,0} \left(N_p - \bar{N}_p\right) + F_{N_p}$$

$$\frac{dN_p(t)}{dt} = \left[\frac{\Gamma v_g g_0 \left(\bar{N}_p + \frac{\Gamma n_{sp}}{V}\right)}{(\bar{N} + N_s)\left(1 + \epsilon \bar{N}_p\right)}\right] (N - \bar{N}) + \left[\frac{\Gamma v_g g_0}{1 + \epsilon \bar{N}_p} ln\left(\frac{\bar{N} + N_s}{N_{tr} + N_s}\right)\left(1 - \frac{\epsilon}{1 + \epsilon \bar{N}_p}\left[\bar{N}_p + \frac{\Gamma n_{sp}}{V}\right]\right) - \frac{1}{\tau_p}\right] \left(N_p - \bar{N}_p\right) + F_{N_p}$$

$$\frac{d\phi(t)}{dt} = \left.\left[\frac{\alpha}{2}\ \Gamma v_g \left(\frac{\partial^2 g}{\partial N^2}(N - \bar{N}) + \frac{\partial g}{\partial N}\right)\right]\right|_{\bar{N},\bar{N}_p,0} (N - \bar{N}) + \left.\left[\frac{\alpha}{2}\ \Gamma v_g (N - \bar{N}) \frac{\partial^2 g}{\partial N \partial N_p}\right]\right|_{\bar{N},\bar{N}_p,0} \left(N_p - \bar{N}_p\right) + F_\phi$$

$$\frac{d\phi(t)}{dt} = \left[\frac{\alpha}{2}\ \Gamma v_g g_0 \frac{1}{(\bar{N} + N_s)\left(1 + \epsilon \bar{N}_p\right)}\right] (N - \bar{N}) + F_\phi$$

$$\left(P_0 - \bar{P}_0\right) = \eta_0 h\nu V_p \tau_p^{-1} \left(N_p - \bar{N}_p\right) + F_0(t)$$

After we linearized the equations, we can start to solve for the linear spectrum. Meaning that we do a Fourier transform of the linearized equations

$$j\omega N_\omega = \left[-\frac{1}{\tau} - \frac{v_g \bar{N}_p g_0}{(\bar{N} + N_s)\left(1 + \epsilon \bar{N}_p\right)}\right] N_\omega + \left[\frac{v_g g_0}{1 + \epsilon \bar{N}_p} ln\left(\frac{\bar{N} + N_s}{N_{tr} + N_s}\right)\left(\frac{\bar{N}_p \epsilon}{\left(1 + \epsilon \bar{N}_p\right)} - 1\right)\right] N_{p,\omega} + F_{N,\omega}$$

$$j\omega N_{p,\omega} = \left[\frac{\Gamma v_g g_0 \left(\bar{N}_p + \frac{\Gamma n_{sp}}{V}\right)}{(\bar{N} + N_s)\left(1 + \epsilon \bar{N}_p\right)}\right] N_\omega + \left[\frac{\Gamma v_g g_0}{1 + \epsilon \bar{N}_p} ln\left(\frac{\bar{N} + N_s}{N_{tr} + N_s}\right)\left(1 - \frac{\epsilon}{1 + \epsilon \bar{N}_p}\left[\bar{N}_p + \frac{\Gamma n_{sp}}{V}\right]\right) - \frac{1}{\tau_p}\right] N_{p,\omega} + F_{N_{p,\omega}}$$

$$\Delta\nu_\omega = \left[\frac{1}{2\pi} \frac{\alpha}{2}\ \Gamma v_g g_0 \frac{1}{(\bar{N} + N_s)\left(1 + \epsilon \bar{N}_p\right)}\right] N_\omega + \left[\frac{1}{2\pi}\right] F_{\phi,\omega}$$

$$P_{0,\omega} = \eta_0 h\nu V \Gamma^{-1} \tau_p^{-1} N_{p,\omega} + F_{0,\omega}$$

It's a linear system so it can be solved easily. Just express $N_\omega, N_{p,\omega}$ as a function of $F_{N,\omega}, F_{N_{p,\omega}}$

$$\left[j\omega + \frac{1}{\tau} + \frac{v_g \bar{N}_p g_0}{(\bar{N} + N_s)(1 + \epsilon \bar{N}_p)}\right] N_\omega + \left[\frac{v_g g_0}{1 + \epsilon \bar{N}_p} ln\left(\frac{\bar{N} + N_s}{N_{tr} + N_s}\right)\left(1 - \frac{\bar{N}_p \epsilon}{(1 + \epsilon \bar{N}_p)}\right)\right] N_{p,\omega} = F_{N,\omega}$$

$$\left[-\frac{\Gamma v_g g_0 \left(\bar{N}_p + \frac{\Gamma n_{sp}}{V}\right)}{(\bar{N} + N_s)(1 + \epsilon \bar{N}_p)}\right] N_\omega + \left[j\omega - \frac{\Gamma v_g g_0}{1 + \epsilon \bar{N}_p} ln\left(\frac{\bar{N} + N_s}{N_{tr} + N_s}\right)\left(1 - \frac{\epsilon}{1 + \epsilon \bar{N}_p}\left[\bar{N}_p + \frac{\Gamma n_{sp}}{V}\right]\right) + \frac{1}{\tau_p}\right] N_{p,\omega} = F_{N_p,\omega}$$

$$\boldsymbol{A}\begin{bmatrix} N_\omega \\ N_{p,\omega} \end{bmatrix} = \begin{bmatrix} F_{N,\omega} \\ F_{N_p,\omega} \end{bmatrix}$$

$$\boldsymbol{A} = \begin{bmatrix} \left[j\omega + \frac{1}{\tau} + \frac{v_g \bar{N}_p g_0}{(\bar{N} + N_s)(1 + \epsilon \bar{N}_p)}\right] & \left[\frac{v_g g_0}{1 + \epsilon \bar{N}_p} ln\left(\frac{\bar{N} + N_s}{N_{tr} + N_s}\right)\left(1 - \frac{\bar{N}_p \epsilon}{(1 + \epsilon \bar{N}_p)}\right)\right] \\ \left[-\frac{\Gamma v_g g_0 \left(\bar{N}_p + \frac{\Gamma n_{sp}}{V}\right)}{(\bar{N} + N_s)(1 + \epsilon \bar{N}_p)}\right] & \left[j\omega - \frac{\Gamma v_g g_0}{1 + \epsilon \bar{N}_p} ln\left(\frac{\bar{N} + N_s}{N_{tr} + N_s}\right)\left(1 - \frac{\epsilon}{1 + \epsilon \bar{N}_p}\left[\bar{N}_p + \frac{\Gamma n_{sp}}{V}\right]\right) + \frac{1}{\tau_p}\right] \end{bmatrix}$$

The solutions is easily written as

$\begin{bmatrix} N_\omega \\ N_{p,\omega} \end{bmatrix} = \boldsymbol{A^{-1}} \begin{bmatrix} F_{N,\omega} \\ F_{N_p,\omega} \end{bmatrix}$. If we define $\boldsymbol{A^{-1}} = \boldsymbol{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$, then

$$N_\omega = b_{11} F_{N,\omega} + b_{12} F_{N_p,\omega}$$

$$N_{p,\omega} = b_{21} F_{N,\omega} + b_{22} F_{N_p,\omega}$$

We can express the terms of B by using the inverse formula for a 2x2 matrix

$$\boldsymbol{A^{-1}} = \boldsymbol{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \frac{1}{|\boldsymbol{A}|}\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

Therefore

$$\begin{aligned} \Delta\nu_\omega &= \left[\frac{1}{2\pi}\frac{\alpha}{2}\, \Gamma v_g g_0 \frac{1}{(\bar{N} + N_s)(1 + \epsilon \bar{N}_p)}\right] N_\omega + \left[\frac{1}{2\pi}\right] F_{\phi,\omega} \\ &= \left[\frac{1}{2\pi}\frac{\alpha}{2}\, \Gamma v_g g_0 \frac{1}{(\bar{N} + N_s)(1 + \epsilon \bar{N}_p)}\right] \left(b_{11} F_{N,\omega} + b_{12} F_{N_p,\omega}\right) + \left[\frac{1}{2\pi}\right] F_{\phi,\omega} \end{aligned}$$

The expressions for the RIN and the FN noise are as follow

$$FN(\omega) = \langle \Delta\nu_\omega^2 \rangle = \Delta\nu_\omega \cdot \Delta\nu_\omega^*$$

$$PN(\omega) = \frac{FN(\omega)}{f^2}$$

$RIN(\omega) = \frac{\langle N_{p,\omega}^2 \rangle}{\bar{N}_p{}^2} = \frac{N_{p,\omega} \cdot N_{p,\omega}^*}{\bar{N}_p{}^2}$ or even better just use the optical power $RIN(\omega) = \frac{\langle P_{0,\omega}^2 \rangle}{\bar{P}_0{}^2} = \frac{P_{0,\omega} \cdot P_{0,\omega}^*}{\bar{P}_0{}^2}$

Theoretical FN

Let's define, for sake of compact notation $\frac{1}{2\pi}\frac{\alpha}{2}\ \Gamma v_g g_0 \frac{1}{(\bar{N}+N_s)(1+\epsilon\bar{N}_p)} = c_1$

$$FN(\omega) = \langle \Delta v_\omega^2 \rangle = \Delta v_\omega \cdot \Delta v_\omega^* = \left(c_1 b_{11} F_{N,\omega} + c_1 b_{12} F_{N_p,\omega} + \left[\frac{1}{2\pi}\right] F_{\phi,\omega} \right)\left(c_1 b_{11}^* F_{N,\omega}^* + c_1 b_{12}^* F_{N_p,\omega}^* + \left[\frac{1}{2\pi}\right] F_{\phi,\omega}^* \right)$$

$$\begin{aligned} &= (c_1 b_{11} F_{N,\omega})(c_1 b_{11}^* F_{N,\omega}^*) + (c_1 b_{11} F_{N,\omega})\left(c_1 b_{12}^* F_{N_p,\omega}^*\right) + (c_1 b_{11} F_{N,\omega})\left(\left[\frac{1}{2\pi}\right] F_{\phi,\omega}^*\right) \\ &\qquad + \left(c_1 b_{12} F_{N_p,\omega}\right)(c_1 b_{11}^* F_{N,\omega}^*) + \left(c_1 b_{12} F_{N_p,\omega}\right)\left(c_1 b_{12}^* F_{N_p,\omega}^*\right) + \left(c_1 b_{12} F_{N_p,\omega}\right)\left(\left[\frac{1}{2\pi}\right] F_{\phi,\omega}^*\right) \\ &\qquad + \left(\left[\frac{1}{2\pi}\right] F_{\phi,\omega}\right)(c_1 b_{11}^* F_{N,\omega}^*) + \left(\left[\frac{1}{2\pi}\right] F_{\phi,\omega}\right)\left(c_1 b_{12}^* F_{N_p,\omega}^*\right) + \left(\left[\frac{1}{2\pi}\right] F_{\phi,\omega}\right)\left(\left[\frac{1}{2\pi}\right] F_{\phi,\omega}^*\right) \end{aligned}$$

Some product are zeros because there is no correlation and we are left with

$$\boxed{FN(\omega) = [c_1]^2 b_{11} b_{11}^* \langle F_N F_N \rangle + [c_1]^2 (b_{11} b_{12}^* + b_{11}^* b_{12}) \langle F_N F_{N_p} \rangle + [c_1]^2 b_{12} b_{12}^* \langle F_{N_p} F_{N_p} \rangle + \left[\frac{1}{2\pi}\right]^2 \langle F_\phi F_\phi \rangle}$$

Theoretical RIN

Let's define for the sake of a compact notation $\eta_0 h\nu V \Gamma^{-1} \tau_p^{-1} \ = c_2$

$$\begin{aligned} RIN(\omega) &= \frac{\langle P_{0,\omega}^2 \rangle}{\bar{P}_0^{\ 2}} = \frac{1}{\bar{P}_0^{\ 2}} \left(P_{0,\omega} \cdot P_{0,\omega}^* \right) \\ &= \frac{1}{\bar{P}_0^{\ 2}} \left(c_2 b_{21} F_{N,\omega} + c_2 b_{22} F_{N_p,\omega} + F_{0,\omega} \right)\left(c_2 b_{21}^* F_{N,\omega}^* + c_2 b_{22}^* F_{N_p,\omega}^* + F_{0,\omega}^* \right) \\ &= \frac{1}{\bar{P}_0^{\ 2}} \Big[(c_2 b_{21} F_{N,\omega})(c_2 b_{21}^* F_{N,\omega}^*) + (c_2 b_{21} F_{N,\omega})\left(c_2 b_{22}^* F_{N_p,\omega}^*\right) + (c_2 b_{21} F_{N,\omega})(F_{0,\omega}^*) \\ &\quad + \left(c_2 b_{22} F_{N_p,\omega}\right)(c_2 b_{21}^* F_{N,\omega}^*) + \left(c_2 b_{22} F_{N_p,\omega}\right)\left(c_2 b_{22}^* F_{N_p,\omega}^*\right) + \left(c_2 b_{22} F_{N_p,\omega}\right)(F_{0,\omega}^*) \\ &\quad + (F_{0,\omega})(c_2 b_{21}^* F_{N,\omega}^*) + (F_{0,\omega})\left(c_2 b_{22}^* F_{N_p,\omega}^*\right) + (F_{0,\omega})(F_{0,\omega}^*) \Big] \end{aligned}$$

Some product are zeros because there is no correlation and we are left with

$$\boxed{RIN(\omega) = \frac{1}{\bar{P}_0^{\ 2}} [(c_2)^2 b_{21} b_{21}^* \langle F_N F_N \rangle + (c_2)^2 (b_{21} b_{22}^* + b_{22} b_{21}^*) \langle F_P F_N \rangle + \ (c_2)^2 b_{22} b_{22}^* \langle F_P F_P \rangle + c_2 (b_{22} + b_{22}^*) \langle F_P F_0 \rangle + \langle F_0 F_0 \rangle]}$$

Clearly the noise sources are referred to the steady state value

If we do the same but for the number of photons… what would be the difference?

$$RIN(\omega) = \frac{\langle N_{p,\omega}^2 \rangle}{\bar{N}_p^{\ 2}} = \frac{N_{p,\omega} \cdot N_{p,\omega}^*}{\bar{N}_p^{\ 2}}$$

$$= \frac{1}{\bar{N}_p^{\,2}}\left(b_{21}F_{N,\omega} + b_{22}F_{N_p,\omega}\right)\left(b_{21}^{*}F_{N,\omega} + b_{22}^{*}F_{N_p,\omega}\right)$$

$$= \frac{1}{\bar{N}_p^{\,2}}\left[(b_{21}b_{21}^{*})\langle F_N F_N\rangle + (b_{21}b_{22}^{*} + b_{22}b_{21}^{*})\langle F_P F_N\rangle + (b_{22}b_{22}^{*})\langle F_P F_P\rangle\right]$$

Carriers and photons power spectra

$$NPSD(\omega) = N_\omega \cdot N_\omega^{*} = \left(b_{11}F_{N,\omega} + b_{12}F_{N_p,\omega}\right)\left(b_{11}^{*}F_{N,\omega} + b_{12}^{*}F_{N_p,\omega}\right)$$
$$= b_{11}b_{11}^{*}\langle F_N F_N\rangle + \; b_{12}b_{12}^{*}\langle F_P F_P\rangle + (b_{11}b_{12}^{*} + b_{12}b_{11}^{*})\langle F_P F_N\rangle$$

$$N_pPSD(\omega) = N_{p,\omega} \cdot N_{p,\omega}^{*} = \left(b_{21}F_{N,\omega} + b_{22}F_{N_p,\omega}\right)\left(b_{21}^{*}F_{N,\omega} + b_{22}^{*}F_{N_p,\omega}\right)$$
$$= b_{21}b_{21}^{*}\langle F_N F_N\rangle + \; b_{22}b_{22}^{*}\langle F_P F_P\rangle + (b_{21}b_{22}^{*} + b_{22}b_{21}^{*})\langle F_P F_N\rangle$$

Relaxation frequency

It is found by setting to zero the frequency derivatives of the power spectra (in order to find the peak)

$RIN'(\omega) = \frac{\partial RIN}{\partial \omega}$ then $RIN'(\omega_r) = 0$ we work with the RIN that has less terms

Some easy calculation first

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} j\omega + a & b \\ c & j\omega + d \end{bmatrix}$$

Where $a = \frac{1}{\tau} + \frac{v_g \bar{N}_p g_0}{(\bar{N}+N_s)(1+\epsilon\bar{N}_p)}$, $b = \left[\frac{v_g g_0}{1+\epsilon\bar{N}_p} ln\left(\frac{\bar{N}+N_s}{N_{tr}+N_s}\right)\left(1 - \frac{\bar{N}_p\epsilon}{(1+\epsilon\bar{N}_p)}\right)\right]$,$c = -\frac{\Gamma v_g g_0\left(\bar{N}_p+\frac{\Gamma n_{sp}}{V}\right)}{(\bar{N}+N_s)(1+\epsilon\bar{N}_p)}$,
$d = -\frac{\Gamma v_g g_0}{1+\epsilon\bar{N}_p} ln\left(\frac{\bar{N}+N_s}{N_{tr}+N_s}\right)\left(1 - \frac{\epsilon}{1+\epsilon\bar{N}_p}\left[\bar{N}_p + \frac{\Gamma n_{sp}}{V}\right]\right) + \frac{1}{\tau_p}$

$$\boldsymbol{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \frac{1}{|\boldsymbol{A}|}\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} = \frac{1}{ad - bc - \omega^2 + j\omega(a+d)}\begin{bmatrix} j\omega + d & -b \\ -c & j\omega + a \end{bmatrix}$$

The un-normalized RIN will have the same peak so

$$\widetilde{RIN}(\omega) = N_{p,\omega} \cdot N_{p,\omega}^{*} = b_{21}b_{21}^{*}\langle F_N F_N\rangle + \; b_{22}b_{22}^{*}\langle F_P F_P\rangle + (b_{21}b_{22}^{*} + b_{22}b_{21}^{*})\langle F_P F_N\rangle =$$

$$= \frac{1}{|\boldsymbol{A}||\boldsymbol{A}|^{*}}\{a_{21}a_{21}^{*}\langle F_N F_N\rangle + \; a_{11}a_{11}^{*}\langle F_P F_P\rangle - (a_{21}a_{11}^{*} + a_{11}a_{21}^{*})\langle F_P F_N\rangle\} =$$

$$= \frac{c^2\langle F_N F_N\rangle + (\omega^2 + a^2)\langle F_P F_P\rangle - (c(a - j\omega) + (a + j\omega)c)\langle F_P F_N\rangle}{(ad - bc - \omega^2)^2 + \omega^2(a+d)^2} =$$

$$= \frac{c^2\langle F_N F_N\rangle + (\omega^2 + a^2)\langle F_P F_P\rangle - 2ac\langle F_P F_N\rangle}{(ad - bc - \omega^2)^2 + \omega^2(a+d)^2} =$$

$$= \frac{(\omega^2)\langle F_P F_P\rangle + c^2\langle F_N F_N\rangle + a^2\langle F_P F_P\rangle - 2ac\langle F_P F_N\rangle}{\omega^4 + \omega^2(a^2 + d^2 + 2bc) + (ad - bc)^2} = \frac{N(\omega)}{D(\omega)}$$

$$\frac{\partial \widetilde{RIN}(\omega)}{\partial \omega} = \frac{N'(\omega)D(\omega) - D'(\omega)N(\omega)}{D(\omega)^2}$$

$$
\begin{aligned}
D(\omega)^2 = (\omega^4 + \omega^2(a^2+d^2+2bc) + (ad-bc)^2)^2 \\
&= \omega^8 + \omega^4(a^2+d^2+2bc)^2 + (ad-bc)^4 + 2\omega^6(a^2+d^2+2bc) \\
&+ 2\omega^2(a^2+d^2+2bc)(ad-bc)^2 + 2\omega^4(ad-bc)^2 \\
&= \omega^8 + 2\omega^6(a^2+d^2+2bc) + \omega^4[(a^2+d^2+2bc)^2 + 2(ad-bc)^2] \\
&+ 2\omega^2(a^2+d^2+2bc)(ad-bc)^2 + (ad-bc)^4
\end{aligned}
$$

Setting the derivative equal zero, we just care about the numerator (assuming that the zero of the numerator does not correspond to a pole in the denumerator (we can check later once found the solution)

$$N'(\omega)D(\omega) - D'(\omega)N(\omega) = 0$$

$$2\omega\langle F_PF_P\rangle[\omega^4 + \omega^2(a^2+d^2+2bc) + (ad-bc)^2] - [4\omega^3 + 2\omega(a^2+d^2+2bc)][(\omega^2+a^2)\langle F_PF_P\rangle + c^2\langle F_NF_N\rangle - 2ac\langle F_PF_N\rangle] = 0$$

$$2\langle F_PF_P\rangle[\omega^5 + \omega^3(a^2+d^2+2bc) + \omega(ad-bc)^2] - [4\omega^3 + 2\omega(a^2+d^2+2bc)][(\omega^2+a^2)\langle F_PF_P\rangle + c^2\langle F_NF_N\rangle - 2ac\langle F_PF_N\rangle] = 0$$

$$[\omega^5 2\langle F_PF_P\rangle + \omega^3(a^2+d^2+2bc)2\langle F_PF_P\rangle + \omega(ad-bc)^2 2\langle F_PF_P\rangle] - [4\omega^3 + 2\omega(a^2+d^2+2bc)][\omega^2\langle F_PF_P\rangle + a^2\langle F_PF_P\rangle + c^2\langle F_NF_N\rangle - 2ac\langle F_PF_N\rangle] = 0$$

$$\omega^5 2\langle F_PF_P\rangle + \omega^3(a^2+d^2+2bc)2\langle F_PF_P\rangle + \omega(ad-bc)^2 2\langle F_PF_P\rangle - \big[[4\omega^3 + 2\omega(a^2+d^2+2bc)]\omega^2\langle F_PF_P\rangle + [4\omega^3 + 2\omega(a^2+d^2+2bc)](a^2\langle F_PF_P\rangle + c^2\langle F_NF_N\rangle - 2ac\langle F_PF_N\rangle)\big] = 0$$

$$\omega^5 2\langle F_PF_P\rangle + \omega^3(a^2+d^2+2bc)2\langle F_PF_P\rangle + \omega(ad-bc)^2 2\langle F_PF_P\rangle - 4\omega^5\langle F_PF_P\rangle - 2\omega^3\langle F_PF_P\rangle(a^2+d^2+2bc) - 4\omega^3(a^2\langle F_PF_P\rangle + c^2\langle F_NF_N\rangle - 2ac\langle F_PF_N\rangle) - 2\omega(a^2+d^2+2bc)(a^2\langle F_PF_P\rangle + c^2\langle F_NF_N\rangle - 2ac\langle F_PF_N\rangle) = 0$$

$$\omega^5\{-2\langle F_PF_P\rangle\} + \omega^3\{-4(a^2\langle F_PF_P\rangle + c^2\langle F_NF_N\rangle - 2ac\langle F_PF_N\rangle)\} + \omega\{(ad-bc)^2 2\langle F_PF_P\rangle - 2(a^2+d^2+2bc)(a^2\langle F_PF_P\rangle + c^2\langle F_NF_N\rangle - 2ac\langle F_PF_N\rangle)\} = 0$$

$$\omega^5\{-2\langle F_PF_P\rangle\} + \omega^3\{-4a^2\langle F_PF_P\rangle - 4c^2\langle F_NF_N\rangle + 8ac\langle F_PF_N\rangle\} + \omega\{(2b^2c^2 - 4adbc - 2a^4 - 4bca^2)\langle F_PF_P\rangle - (2a^2c^2 + 2d^2c^2 + 4bc^3)\langle F_NF_N\rangle + (4a^3c + d^2 4ac + 8bc^2a)\langle F_PF_N\rangle\} = 0$$

We can put this in a Matlab solver, to find the root for ω. The resulting root will be the relaxation frequency.

Slope change

To find it we need the second derivative

$$\widetilde{RIN}(\omega) = \frac{(\omega^2)\langle F_PF_P\rangle + c^2\langle F_NF_N\rangle + a^2\langle F_PF_P\rangle - 2ac\langle F_PF_N\rangle}{\omega^4 + \omega^2(a^2+d^2+2bc) + (ad-bc)^2} = \frac{N(\omega)}{D(\omega)}$$

$$\frac{\partial\widetilde{RIN}(\omega)}{\partial\omega} = \frac{N'(\omega)D(\omega) - D'(\omega)N(\omega)}{D(\omega)^2}$$

$$\frac{\partial^2\widetilde{RIN}(\omega)}{\partial\omega^2} = \frac{\big(N'(\omega)D(\omega) - D'(\omega)N(\omega)\big)' D(\omega)^2 - 2\big(N'(\omega)D(\omega) - D'(\omega)N(\omega)\big)D(\omega)D(\omega)'}{D(\omega)^4}$$

$$
\begin{aligned}
&\big(N'(\omega)D(\omega) - D'(\omega)N(\omega)\big)' \\
&\quad = 5\omega^4\{-2\langle F_PF_P\rangle\} + 3\omega^2\{-4a^2\langle F_PF_P\rangle - 4c^2\langle F_NF_N\rangle + 8ac\langle F_PF_N\rangle\} \\
&\quad + \{(2b^2c^2 - 4adbc - 2a^4 - 4bca^2)\langle F_PF_P\rangle - (2a^2c^2 + 2d^2c^2 + 4bc^3)\langle F_NF_N\rangle \\
&\quad + (4a^3c + d^2 4ac + 8bc^2a)\langle F_PF_N\rangle\}
\end{aligned}
$$

$$D(\omega)' = 4\omega^3 + 2\omega(a^2+d^2+2bc)$$

Alternative, the books suggest to have a simplified expression for it

OR I may suspect there is a relation between the relaxation frequency and the imaginary part of the eigenvalues of the linearized system

Cost function E

We can think about a general cost function that includes many terms (multidimensional). Therefore the problem of parameter identification is solved by minimizing the norm of such function (by assuring all terms are positive)

E(1)	E(2)	E(3)	E(4)
Positive log likelihood	Distance between theoretical curves and reference	Relaxation frequency within the bounds	Discrete system eigenvalues showing stability
LL^{+}			

Phase noise, frequency noise and amplitude noise: spectral definitions and conversions

$$s(t) = \sqrt{P(t)}\cos\big(\Delta\omega t + \phi(t)\big) = A(t)\cos\big(\Delta\omega t + \phi(t)\big) = A_0\big(1 + a(t)\big)\cos\big(\Delta\omega t + \phi(t)\big)$$

$$\delta v(t) = \frac{1}{2\pi}\frac{d\phi(t)}{dt}$$

$PN(f) = PSD_{\phi}(f)$

$FN(f) = PSD_{\delta v}(f) = f^2 PSD_{\phi}(f)$

$AN(f) = \ PSD_{a}(f)$ – is this the same as RIN?

$$RIN(t) = \frac{\delta P(t)}{P_{avg}} = \frac{P(t) - P_{avg}}{P_{avg}}$$

$$RIN(f) = \frac{PSD_{\delta P}(f)}{P_{avg}^2}$$

Phase noise

$$PN\left[\frac{\mu rad}{\sqrt{Hz}}\right] = 10^6 PN\left[\frac{rad}{\sqrt{Hz}}\right]$$

$$PN\left[\frac{rad}{\sqrt{Hz}}\right] = \sqrt{PN\left[\frac{rad^2}{Hz}\right]}$$

$$PN\left[dB\left[\frac{rad}{\sqrt{Hz}}\right]\right] = 10\log_{10}\left(PN\left[\frac{rad}{\sqrt{Hz}}\right]\right)$$

$$PN\left[dB\left[\frac{rad^2}{Hz}\right]\right] = 10\log_{10}(PN\left[\frac{rad^2}{Hz}\right]) = \ 20\log_{10}\left(PN\left[\frac{rad}{\sqrt{Hz}}\right]\right)$$

$$PN\left[\frac{dBc}{Hz}\right] = PN\left[dB\left[\frac{rad^2}{Hz}\right]\right] - 3dB$$

Frequency noise

$$FN\left[\frac{Hz}{\sqrt{Hz}}\right] = \sqrt{FN\left[\frac{Hz^2}{Hz}\right]}$$

$$FN\left[dB\left[\frac{Hz}{\sqrt{Hz}}\right]\right] = 10\log_{10}\left(FN\left[\frac{Hz}{\sqrt{Hz}}\right]\right)$$

$$FN\left[dB\left[\frac{Hz^2}{Hz}\right]\right] = 10\log_{10}(FN\left[\frac{Hz^2}{Hz}\right]) = 20\log_{10}\left(FN\left[\frac{Hz}{\sqrt{Hz}}\right]\right)$$

$$FN\left[\frac{dBc}{Hz}\right] = FN\left[dB\left[\frac{Hz^2}{Hz}\right]\right] - 3dB$$

Amplitude noise – and RIN?

$$AN\left[\frac{dBc}{Hz}\right] = 10\log_{10} AN\left[\frac{V^2}{Hz}\right] - 3$$

$$RIN\left[\frac{dBc}{Hz}\right] = 10\log_{10} RIN\left[\frac{V^2}{Hz}\right]$$

I am not super sure about it

Theoretical spectra

White noise level

Theoretical level of white noise, given the process variance σ_r^2

The power spectral density (single side) is constant for all spectra and equal to

$$PSD_r(f) = 2T_s\sigma_r^2$$

Note the factor 2 in front of the spectra

If we see discrepancy is because we are using a Hanning window

Brown phase noise level – Random walk phase noise – Lorentzian phase noise

When we generate the phase noise according to the random walk model

$$\phi_{t+1} = \phi_t + q_{\phi,t} \qquad\qquad q_{\phi,t} \sim \mathcal{N}\left(0, \sigma_\phi^2\right)$$

Then the PSD of the process ϕ correspond to (single side)

$$PSD_\phi(f) = 2\frac{LW}{\pi}\frac{1}{f^2} = \frac{\sigma_\phi^2}{\pi^2 T_s}\frac{1}{f^2}$$

Where $LW = \frac{\sigma_\phi^2}{2\pi T_s}$, $\sigma_\phi^2 = LW2\pi T_s$ is the theoretical linewidth. Note the factor 2 in front of the spectra

For a brown phase noise, the correspondent frequency noise is white. This is the relation between the two (single side psd, remember the factor 2):

$$[rad^2/Hz] \quad PSD_{\phi}(f) = \frac{1}{f^2} PSD_{\delta v}(f) \quad [Hz^2/Hz]$$

$$PSD_{\delta v}(f) = \frac{2LW}{\pi}$$

Brown amplitude noise and RIN

When we generate the amplitude noise according to the random walk model

$$a_{t+1} = a_t + q_{a,t} \qquad\qquad q_{a,t} \sim \mathcal{N}(0, \sigma_a^2)$$

Then the PSD of the amplitude noise has a the same shape

$$PSD_a(f) = \frac{\sigma_a^2}{2\pi^2 T_s} \frac{1}{f^2}$$

The question is : what shape will the RIN have? – Answer: no idea ☹

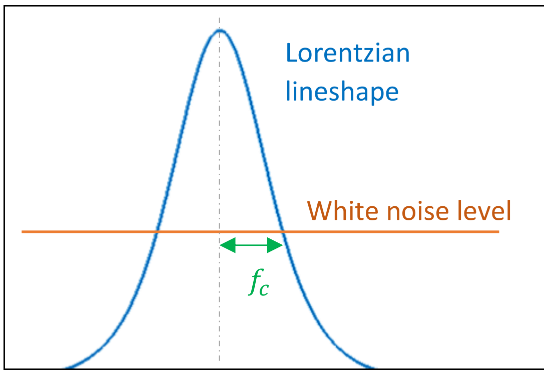
Lorentzian profile (when Lorentzian phase noise is applied)

Aka: the shape of the signal in frequency domain

$$PSD_y(f) = E_0^2 \frac{LW}{\pi} \frac{2}{\left(f - \frac{\omega_r}{2\pi}\right)^2 + \left(\frac{LW}{2}\right)^2} = \frac{E_0^2 \sigma_\phi^2}{\pi^2 T_s} \frac{1}{\left(f - \frac{\omega_r}{2\pi}\right)^2 + \left(\frac{\sigma_\phi^2}{4\pi T_s}\right)^2}$$

Where E_0^2 is the squared mean amplitude.

I hypothesize that the crossing between the Lorentzian profile and the white noise power should give us the confidence interval regarding spectra estimation. We solve for f



$$PSD_y(f) = PSD_r(f)$$

The frequency up which the spectra overlap is

$$f_c = \sqrt{\frac{E_0^2 \sigma_\phi^2}{2\sigma_r^2 \pi^2 T_s^2} - \left(\frac{\sigma_\phi^2}{4\pi T_s}\right)^2}$$

Non-trivial power spectral density (numerical and theoretical) from linear dynamical systems

Suppose that we have the following linear dynamical stochastic system (discrete time) with additive noise

$$\boldsymbol{x}_{t+1} = A\boldsymbol{x}_t + \boldsymbol{q}_t$$

Where:

- A is a $n \times n$ matrix, the system matrix. Requirements of this matrix should be : real, and stable (eigenvalues with absolute value within the unit circle)
- $\boldsymbol{q}_t$ is a multivariate random vector, $\boldsymbol{q}_t \sim \mathcal{N}(\boldsymbol{0}, \Sigma)$, with Σ a symmetric and positive definite covariance matrix

If we want to estimate, using the small signal analysis, the power spectral density of the system state $\boldsymbol{x}$ for a given frequency ω, we just compute

$$PSD_{\boldsymbol{x},\omega} = [j\omega I - A]^{-1}\Sigma\{[j\omega I - A]^{-1}\}^*$$

Where I is the unit matrix $n \times n$, j is the imaginary unit and * denotes the transpose conjugate. $PSD_{\boldsymbol{x},\omega}$ is then a matrix whose diagonal contains the power spectral density of the single states, and the other elements represent the cross power spectral density.

Exploiting the properties of the inverse and the determinant we obtain

$$PSD_{\boldsymbol{x},\omega} = \underbrace{\frac{-1}{\det[j\omega I - A]\det[j\omega I + A]}}_{scalar} \cdot \underbrace{adj[j\omega I - A]\Sigma\, adj[j\omega I + A^\top]}_{matrix}$$

Now, since the aim is to design the system to generate just one colored noise following our specifications, the other states are there only for support. The objective is then to have a closed expression for the i-th $PSD_{\boldsymbol{x}_i,\omega}$ which depends on the frequency ω and the system parameters A and Σ

Writing it down

$$PSD_{\boldsymbol{x}_i,\omega} = \frac{-1}{\det[j\omega I - A]\det[j\omega I + A]} \cdot \underbrace{adj[j\omega I - A]}_{i-th\ row}\Sigma\ \underbrace{adj[j\omega I + A^\top]}_{i-th\ column}$$

What we have left to do is:

- Get $\det[j\omega I - A]\det[j\omega I + A]$ in a closed form
- Get $\underbrace{adj[j\omega I - A]}_{i-th\ row}\Sigma\ \underbrace{adj[j\omega I + A^\top]}_{i-th\ column}$ also in a closed form

$\det[j\omega I - A]$ is the characteristic polynomial of A, and $\det[j\omega I + A]$ is the characteristic polynomial of $-A$.

They should have equal and opposite eigenvalues (A and –A) because

$A\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i \ \ \forall i$ if then we multiply by -1 both sides $(-A)\boldsymbol{v}_i = (-\lambda_i)\boldsymbol{v}_i \ \ \forall i$ we have the eigenvalues equals in magnitude but with opposite sign

Therefore, the characteristic polynomials can be factorized as follow

$\det[j\omega I - A] = (j\omega - \lambda_1)(j\omega - \lambda_2) \dots (j\omega - \lambda_n)$

$\det[j\omega I + A] = (j\omega + \lambda_1)(j\omega + \lambda_2) \dots (j\omega + \lambda_n)$

And we can write the product of both determinant as

$$\begin{aligned}\det[j\omega I - A]\det[j\omega I + A] &= (j\omega - \lambda_1)(j\omega + \lambda_1)(j\omega - \lambda_2)(j\omega + \lambda_2) \dots (j\omega - \lambda_n)(j\omega + \lambda_n) \\ &= ([j\omega]^2 - [\lambda_1]^2)([j\omega]^2 - [\lambda_2]^2) \dots ([j\omega]^2 - [\lambda_n]^2) \\ &= (-)^n(\omega^2 + \lambda_1^2)(\omega^2 + \lambda_2^2) \dots (\omega^2 + \lambda_n^2) \blacksquare\end{aligned}$$

$$PSD_{x_i,\omega} = \frac{(-)^{1-n}}{(\omega^2 + \lambda_1^2)(\omega^2 + \lambda_2^2) \dots (\omega^2 + \lambda_n^2)} \cdot \underbrace{adj[j\omega I - A]}_{i-th\, row} \Sigma \underbrace{adj[j\omega I + A^{\mathsf{T}}]}_{i-th\, column}$$

There is a way to obtain the same expression by using the Woodbury identity

$$PSD_{x,\omega} = [j\omega I - A]^{-1}\Sigma\{[j\omega I - A]^{-1}\}^*$$

Let $B = -A$. The difference in this matrix are the eigenvalues (opposite sign)

Therefore we can write

$$PSD_{x,\omega} = [j\omega I + B]^{-1}\Sigma\{[j\omega I + B]^{-1}\}^*$$

Now we can use the Woodbury identity

3.2.2 The Woodbury identity

The Woodbury identity comes in many variants. The latter of the two can be found in [12]

$$(\mathbf{A} + \mathbf{C}\mathbf{B}\mathbf{C}^T)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{C}(\mathbf{B}^{-1} + \mathbf{C}^T\mathbf{A}^{-1}\mathbf{C})^{-1}\mathbf{C}^T\mathbf{A}^{-1} \quad (156)$$

$$(\mathbf{A} + \mathbf{U}\mathbf{B}\mathbf{V})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{B}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}\mathbf{A}^{-1} \quad (157)$$

The matrix B can be decomposed in terms of its eigenvalues and eigenvector (which are the opposite of the eigenvalues of A)

$$B = V\Lambda V^{-1}$$

$$\begin{aligned}[j\omega I + B]^{-1} = [j\omega I + V\Lambda V^{-1}]^{-1} &= (j\omega)^{-1}I - (j\omega)^{-1}V[\Lambda^{-1} + V^{-1}(j\omega)^{-1}V]^{-1}(j\omega)^{-1} \\ &= (j\omega)^{-1}I - (j\omega)^{-2}V[\Lambda^{-1} + (j\omega)^{-1}I]^{-1}V^{-1}\end{aligned}$$

We can express $\Lambda^{-1} + (j\omega)^{-1}I$ as

$$\Lambda^{-1} + (j\omega)^{-1}I = diag\{(\lambda_{1:n}^B)^{-1} + (j\omega)^{-1}\}$$

$$\{\Lambda^{-1} + (j\omega)^{-1}I\}^{-1} = diag\left\{\frac{j\omega\lambda_{1:n}^B}{\lambda_{1:n}^B + j\omega}\right\}$$

Then

$$
\begin{aligned}
[j\omega I + B]^{-1} &= (j\omega)^{-1} I - (j\omega)^{-2} V \mathrm{diag}\{j\omega \lambda_{1:n}^{B} (\lambda_{1:n}^{B} + j\omega)^{-1}\} V^{-1} \\
&= (j\omega)^{-1} I - (j\omega)^{-1} V \mathrm{diag}\{j\omega \lambda_{1:n}^{A} (\lambda_{1:n}^{A} - j\omega)^{-1}\} V^{-1} \\
&= (j\omega)^{-1} [I - V \mathrm{diag}\{j\omega \lambda_{1:n}^{A} (\lambda_{1:n}^{A} - j\omega)^{-1}\} V^{-1}]
\end{aligned}
$$

The conjugate transpose is pretty similar

$$
\begin{aligned}
\{[j\omega I + B]^{-1}\}^* &= (-j\omega)^{-1} [I - \{V^{-1}\}^* \mathrm{diag}\{-j\omega \lambda_{1:n}^{A*} (\lambda_{1:n}^{A*} + j\omega)^{-1}\} V^*] \\
&= (-j\omega)^{-1} [I + \{V^{-1}\}^* \mathrm{diag}\{j\omega \lambda_{1:n}^{A*} (\lambda_{1:n}^{A*} + j\omega)^{-1}\} V^*]
\end{aligned}
$$

Let's define the matrix Z as

$$
Z = \mathrm{diag}\{j\omega \lambda_{1:n}^{A} (\lambda_{1:n}^{A} - j\omega)^{-1}\}
$$

Each diagonal element can be expressed as

$$
z_i = j\omega \lambda_i^A \left(\lambda_i^A - j\omega\right)^{-1} = j\omega (a_i + jb_i)(a_i + jb_i - j\omega)^{-1}
$$

$$
z_i^* = -j\omega \lambda_i^{A*} \left(\lambda_i^{A*} + j\omega\right)^{-1} = -j\omega (a_i - jb_i)(a_i - jb_i + j\omega)^{-1}
$$

And finally, we can simplify the PSD expression

$$
\begin{aligned}
PSD_{\boldsymbol{x},\omega} &= [j\omega I - A]^{-1} \Sigma \{[j\omega I - A]^{-1}\}^* = (j\omega)^{-1}[I - V\mathrm{Z}V^{-1}]\Sigma(-j\omega)^{-1}[I - \{V^{-1}\}^*\mathrm{Z}^*V^*] = \omega^{-2}[I - V\mathrm{Z}V^{-1}]\Sigma[I - \{V^{-1}\}^*\mathrm{Z}^*V^*] = \omega^{-2}[\Sigma - V\mathrm{Z}V^{-1}\Sigma - \Sigma\{V^{-1}\}^*\mathrm{Z}^*V^* + V\mathrm{Z}V^{-1}\Sigma\{V^{-1}\}^*\mathrm{Z}^*V^*] \\
&= \omega^{-2}\left[\Sigma - \left(\sum_{i=1}^{n} \boldsymbol{v}_i z_i [\boldsymbol{v}_i^{-1}]^\top\right)\Sigma - \Sigma\left(\sum_{i=1}^{n} [\boldsymbol{v}_i^{-1}]^* z_i^* [\boldsymbol{v}_i^*]^\top\right) + \left(\sum_{i=1}^{n} \boldsymbol{v}_i z_i [\boldsymbol{v}_i^{-1}]^\top\right)\Sigma\left(\sum_{i=1}^{n} [\boldsymbol{v}_i^{-1}]^* z_i^* [\boldsymbol{v}_i^*]^\top\right)\right] \\
&= \omega^{-2}\left[\Sigma - \left(\sum_{i=1}^{n} \boldsymbol{v}_i \frac{j\omega(a_i + jb_i)}{a_i + jb_i - j\omega} [\boldsymbol{v}_i^{-1}]^\top\right)\Sigma - \Sigma\left(\sum_{i=1}^{n} [\boldsymbol{v}_i^{-1}]^* \frac{-j\omega(a_i - jb_i)}{a_i - jb_i + j\omega} [\boldsymbol{v}_i^*]^\top\right) + \left(\sum_{i=1}^{n} \boldsymbol{v}_i \frac{j\omega(a_i + jb_i)}{a_i + jb_i - j\omega} [\boldsymbol{v}_i^{-1}]^\top\right)\Sigma\left(\sum_{i=1}^{n} [\boldsymbol{v}_i^{-1}]^* \frac{-j\omega(a_i - jb_i)}{a_i - jb_i + j\omega} [\boldsymbol{v}_i^*]^\top\right)\right] \\
&= \omega^{-2}\left[\Sigma - j\omega\left(\sum_{i=1}^{n} \boldsymbol{v}_i \frac{(a_i + jb_i)}{a_i + jb_i - j\omega} [\boldsymbol{v}_i^{-1}]^\top\right)\Sigma + j\omega\Sigma\left(\sum_{i=1}^{n} [\boldsymbol{v}_i^{-1}]^* \frac{(a_i - jb_i)}{a_i - jb_i + j\omega} [\boldsymbol{v}_i^*]^\top\right) + \left(\sum_{i=1}^{n} \boldsymbol{v}_i \frac{j\omega(a_i + jb_i)}{a_i + jb_i - j\omega} [\boldsymbol{v}_i^{-1}]^\top\right)\Sigma\left(\sum_{i=1}^{n} [\boldsymbol{v}_i^{-1}]^* \frac{-j\omega(a_i - jb_i)}{a_i - jb_i + j\omega} [\boldsymbol{v}_i^*]^\top\right)\right]
\end{aligned}
$$

In the hypothesis that A is symmetric, then the eigenvalue decomposition becomes

$$
A = V\Lambda V^*
$$

So we just replace V^{-1} with V^*

$$
PSD_{\boldsymbol{x},\omega} = \omega^{-2}[\Sigma - V\mathrm{Z}V^*\Sigma - \Sigma V\mathrm{Z}^*V^* + V\mathrm{Z}V^*\Sigma V\mathrm{Z}^*V^*]
$$

And we can express the eigenvalue product as

$$
V\mathrm{Z}V^* = \sum_{i=1}^{n} \boldsymbol{v}_i z_i \boldsymbol{v}_i^* = j\omega \sum_{i=1}^{n} \frac{\boldsymbol{v}_i \boldsymbol{v}_i^* \lambda_i^A}{\left(\lambda_i^A - j\omega\right)}
$$

In this sum the product of the eigenvectors $\boldsymbol{v}_i \boldsymbol{v}_i^*$ is going to be an external product

$$PSD_{\boldsymbol{x},\omega} = \omega^{-2}\left[\Sigma - \left(\sum_{i=1}^{n} \boldsymbol{v}_i z_i [\boldsymbol{v}_i^*]^\top\right)\Sigma - \Sigma\left(\sum_{i=1}^{n} \boldsymbol{v}_i z_i^* [\boldsymbol{v}_i^*]^\top\right) + \left(\sum_{i=1}^{n} \boldsymbol{v}_i z_i [\boldsymbol{v}_i^*]^\top\right)\Sigma\left(\sum_{i=1}^{n} \boldsymbol{v}_i z_i^* [\boldsymbol{v}_i^*]^\top\right)\right]$$
$$= \omega^{-2}\left[\Sigma - \left(\sum_{i=1}^{n} \boldsymbol{v}_i j\omega\lambda_i^A\left(\lambda_i^A - j\omega\right)^{-1}[\boldsymbol{v}_i^*]^\top\right)\Sigma - \Sigma\left(\sum_{i=1}^{n} \boldsymbol{v}_i\left[-j\omega\lambda_i^{A*}\left(\lambda_i^{A*} + j\omega\right)^{-1}\right][\boldsymbol{v}_i^*]^\top\right)\right.$$
$$\left. + \left(\sum_{i=1}^{n} \boldsymbol{v}_i j\omega\lambda_i^A\left(\lambda_i^A - j\omega\right)^{-1}[\boldsymbol{v}_i^*]^\top\right)\Sigma\left(\sum_{i=1}^{n} \boldsymbol{v}_i\left[-j\omega\lambda_i^{A*}\left(\lambda_i^{A*} + j\omega\right)^{-1}\right][\boldsymbol{v}_i^*]^\top\right)\right]$$
$$= \omega^{-2}\Sigma - j\omega^{-1}\left[\left(\sum_{i=1}^{n} \boldsymbol{v}_i\lambda_i^A\left(\lambda_i^A - j\omega\right)^{-1}[\boldsymbol{v}_i^*]^\top\right)\Sigma + \Sigma\left(\sum_{i=1}^{n} \boldsymbol{v}_i\left[\lambda_i^{A*}\left(\lambda_i^{A*} + j\omega\right)^{-1}\right][\boldsymbol{v}_i^*]^\top\right)\right]$$
$$+ \left(\sum_{i=1}^{n} \boldsymbol{v}_i\lambda_i^A\left(\lambda_i^A - j\omega\right)^{-1}[\boldsymbol{v}_i^*]^\top\right)\Sigma\left(\sum_{i=1}^{n} \boldsymbol{v}_i\left[\lambda_i^{A*}\left(\lambda_i^{A*} + j\omega\right)^{-1}\right][\boldsymbol{v}_i^*]^\top\right)$$

Maybe we can exploit the relationship among the eigenvalues of A

$\lambda_i^A = a_i + jb_i,\ \ \lambda_i^{A*} = a_i - jb_i$

$$PSD_{\boldsymbol{x},\omega} = \omega^{-2}\left[\Sigma - j\omega\left(\sum_{i=1}^{n} \frac{\boldsymbol{v}_i\boldsymbol{v}_i^*(a_i + jb_i)}{(a_i + jb_i - j\omega)}\right)\Sigma + j\omega\Sigma\left(\sum_{i=1}^{n} \frac{\boldsymbol{v}_i\boldsymbol{v}_i^*(a_i - jb_i)}{(a_i - jb_i + j\omega)}\right)\right.$$
$$\left. + \omega^2\left(\sum_{i=1}^{n} \frac{\boldsymbol{v}_i\boldsymbol{v}_i^*(a_i + jb_i)}{(a_i + jb_i - j\omega)}\right)\Sigma\left(\sum_{i=1}^{n} \frac{\boldsymbol{v}_i\boldsymbol{v}_i^*(a_i - jb_i)}{(a_i - jb_i + j\omega)}\right)\right]$$

Another property that we can exploit is the fact that A is real, and its eigenvalues comes in conjugate pairs – and so the eigenvectors, meaning that $\left(\sum_{i=1}^{n} \frac{\boldsymbol{v}_i\boldsymbol{v}_i^*\lambda_i^A}{(\lambda_i^A - j\omega)}\right)$ and $\left(\sum_{i=1}^{n} \frac{\boldsymbol{v}_i\boldsymbol{v}_i^*\lambda_i^{A*}}{(\lambda_i^{A*} + j\omega)}\right)$ should be identical, or better saying

$$VZV^* = (VZV^*)^* = V^*Z^*V$$

That the product VZV^* is Hermitian. We already hypothesized that $A = V\Lambda V^*$ is symmetric, but is it true also for VZV^*?

General Bayesian filtering stuff

Lower bound to the negative log-likelihood

The general expression for the negative log-likelihood is (when the measurements are scalar)

$$LL^- = \frac{1}{2}\sum_{k=1}^{T}[\log|2\pi S| + v^2S^{-1}]$$

The LL is lower bounded in the best case scenario, when $S_k = \sigma_n^2$ and $v_k \sim \mathcal{N}(0, \sigma_n^2)$

$$LL^{-} = \frac{1}{2}\sum_{k=1}^{T}\left[\log|2\pi S_k| + v_k^2 S_k^{-1}\right] \leq \frac{1}{2}\sum_{k=1}^{T}\log|2\pi\sigma_n^2| + \frac{1}{2}\sum_{k=1}^{T}\left(\frac{v}{\sigma_n}\right)^2 = \frac{T}{2}\log|2\pi\sigma_n^2| + \frac{1}{2}\underbrace{\sum_{k=1}^{T}\left(\frac{v_k}{\sigma_n}\right)^2}_{\chi_{T-1}^2}$$

If we calculate the expected values on both sides

$$\mathbb{E}[LL^{-}] \leq \frac{T}{2}\log|2\pi\sigma_n^2| + \frac{1}{2}\mathbb{E}[\chi_{T-1}^2] = \frac{1}{2}\{T(\log 2\pi\sigma_n^2 + 1) - 1\}$$

Having a lower bound allows us to use the log-likelihood as a cost function that is always positive.

This would allow to define a more general cost function, that has other terms, spanning over more than one dimension. Then we do not need to worry about composing an eventual energy function, we just need to minimize every term, hence take the norm

We can define as positive log likelihood

$$LL^{+} = LL^{-} - \frac{1}{2}\{T(log\ 2\pi\sigma_n^2 + 1) - 1\}$$