Reading:

Lectures 1-24 are on the web page

Lecture 25
  Localization (slides in handout for Lec 24)
Comments on Matched Filtering

• Main ingredient = cross correlation of a template with data under an assumed model
  – template known: easy
  – template not known:
    • perhaps a class of shapes is known
    • can then optimize results against a template bank (i.e. a family of templates)
    • examples: searches for transient signals (bursts, eclipses, transits)
    • orbital searches of periodic emitters:
      – dedisperse + remove candidate orbital modulation + spectral analysis for periodic signal
  – Approximate template ➔ nonoptimal S/N in detection
Matched Filtering

Template too wide

Template too narrow

Template matched
Localization Using Matched Filtering

This handout describes localization of an object in a parameter space. For simplicity we consider localization of a pulse in time. The same formalism applies to localization of a spectral feature in frequency or to an image feature in a 2D image. The results can be extrapolated to a space of arbitrary dimensionality.

I. First consider finding the amplitude of a pulse when the shape and location are known.

Let the data be

\[ I(t) = aA(t) + n(t), \]

where \( a \) = the unknown amplitude and \( n(t) \) is zero mean noise. The known pulse shape is \( A(t) \).

Let the model be \( \hat{I}(t) = \hat{a}A(t) \).

Define the cost function to be the integrated squared error,

\[ Q = \int dt \left[ I(t) - \hat{I}(t) \right]^2. \]
Taking a derivative, we can solve for the estimate of the amplitude, \( \hat{a} \):

\[
\partial_a Q = -2 \int dt \left[ I(t) - \hat{I}(t) \right] \partial_a \hat{I}(t) = 0 
\]

\[\Rightarrow \int dt \left[ I(t) - \hat{I}(t) \right] A(t) = 0\]

\[\Rightarrow \int dt \hat{I}(t) A(t) = \int dt I(t) A(t)\]

\[\Rightarrow \hat{a} \int dt A^2(t) = \int dt I(t) A(t)\]

\[\Rightarrow \hat{a} = \frac{\int dt I(t) A(t)}{\int dt A^2(t)}.\]

Note that:

a. The model is linear in the sole parameter, \( \hat{a} \)

b. The numerator is the zero lag of the crosscorrelation function (CCF) of \( I(t) \) and \( A(t) \).

c. The denominator is the zero lag of the autocorrelation function (ACF) of \( A(t) \).
II. Now consider the case where we don’t know the location of the pulse in time (the time of arrival, TOA) and that it is the TOA we wish to estimate. We still know the pulse shape, \textit{a priori}.

Let the data, model and cost function be

\[ I(t) = aA(t - t_0) + n(t) \]

\[ \hat{I}(t) = \hat{a}A(t - \hat{t}_0). \]

\[ Q = \int dt \left[ I(t) - \hat{I}(t) \right]^2. \]

Note that the model is \textbf{linear} in \( \hat{a} \) but is \textbf{nonlinear} in \( \hat{t}_0 \).

Minimizing \( Q \) with respect to \( \hat{a} \), we have

\[ \partial_{\hat{a}} Q = -2 \int dt \left[ I(t) - \hat{I}(t) \right] \partial_{\hat{a}} \hat{I}(t) = 0 \]

\[ \Rightarrow \int dt \left[ I(t) - \hat{I}(t) \right] A(t - \hat{t}_0) = 0 \]

\[ \Rightarrow \int dt \hat{I}(t) A(t - \hat{t}_0) = \int dt I(t) A(t - \hat{t}_0) \]

\[ \Rightarrow \hat{a} \int dt A^2(t - \hat{t}_0) = \int dt I(t) A(t - \hat{t}_0) \] (3)

\[ \Rightarrow \hat{a} = \frac{\int dt I(t) A(t - \hat{t}_0)}{\int dt A^2(t - \hat{t}_0)}. \]

This last equation has the same form as in I. except that the estimate for the arrival time \( \hat{t}_0 \) is involved.
Now, minimizing $Q$ with respect to $\hat{t}_0$, we have
\[
\partial_{\hat{t}_0} Q = -2 \int dt \left[ I(t) - \hat{I}(t) \right] \partial_{\hat{t}_0} \hat{I}(t) = 0
\]
\[
\Rightarrow -\hat{a} \int dt \left( \frac{\hat{I}(t)}{A(t-\hat{t}_0)} \right) A'(t-\hat{t}_0) = -\hat{a} \int dt I(t) A'(t-\hat{t}_0)
\]
\[
\Rightarrow \hat{a} \int dt A(t-\hat{t}_0) A'(t-\hat{t}_0) = \int dt I(t) A'(t-\hat{t}_0).
\] (4)

**Grid Search:** One approach to finding the arrival time is to search over a 2D grid of $\hat{a}, \hat{t}_0$ to find the values that satisfy equations 3 and 4. This approach is inefficient. Instead, one can search over a 1D space for the single nonlinear parameter, $\hat{t}_0$, and then solve for $\hat{a}$ using either equation 3 or 4.

**Linearization + Iteration:** Another method is to find solutions for $\hat{a}$ and $\hat{t}_0$, we can linearize the equations in $\hat{t}_0 - t_0$ by using Taylor-series expansions for $A(t - \hat{t}_0)$ and $A'(t - \hat{t}_0)$.

Let $\hat{t}_0 = t_0 + \delta \hat{t}_0$. Then, to first order in $\delta \hat{t}_0$:
\[
A(t - \hat{t}_0) \approx A(t - t_0) - A'(t - t_0) \delta \hat{t}_0
\]
\[
A'(t - \hat{t}_0) \approx A'(t - t_0) - A''(t - t_0) \delta \hat{t}_0
\]
\[
A^2(t - \hat{t}_0) \approx A^2(t - t_0) - 2A'(t - t_0) A(t - t_0) \delta \hat{t}_0.
\]
Now equations (3) and (4) become

\[ \dot{a} \int dt \left[ A^2(t - t_0) - 2\delta t_0 A(t - t_0)A'(t - t_0) \right] = \int dt I(t)[A(t - t_0) - \delta t_0 A'(t - t_0)] \]

\[ \dot{a} \int dt \left[ A(t - t_0)A'(t - t_0) - \delta t_0 A(t - t_0)A''(t - t_0) - \delta t_0 A'^2(t - t_0) \right] = \int dt I(t)[A'(t - t_0) - \delta t_0 A''(t - t_0)]. \]

Consider the integral

\[ \int dt \, A(t - t_0)A'(t - t_0). \]

The integrand may be written as

\[ A(t - t_0)A'(t - t_0) = \frac{1}{2} \frac{d}{dt} A^2(t - t_0) \]

and so the integral equals

\[ \frac{1}{2} A^2(t - t_0)|_{t_1}^{t_2} \to 0 \]

in the limit of (e.g.) \( t_{1,2} = \mp T/2 \) with \( T \gg \) pulse width.

We then have

\[ \dot{a} \int dt \, A^2(t - t_0) = \int dt I(t)[A(t - t_0) - \delta t_0 A'(t - t_0)] \]

\[ -\delta t_0 \dot{a} \int dt \left[ A(t - t_0)A''(t - t_0) + A'^2(t - t_0) \right] = \int dt I(t)[A'(t - t_0) - \delta t_0 A''(t - t_0)]. \]
Solving for \( \hat{a} \) in both cases we have

\[
\hat{a} = \frac{\int dt \left[ I(t)A(t - t_0) - \delta \hat{t}_0 I(t)A'(t - t_0) \right]}{\int dt \ A^2(t - t_0)}
\]

(5)

\[
\hat{a} = \frac{\int dt \ I(t) \left[ -A'(t - t_0) + \delta \hat{t}_0 A''(t - t_0) \right]}{\delta \hat{t}_0 \int dt \ \left[ A(t - t_0)A''(t - t_0) + A'^2(t - t_0) \right]}
\]

(6)

Using the notation

\[
i_0 \equiv \int dt \ I(t)A(t - t_0)
\]

\[
i_1 \equiv \int dt \ I(t)A'(t - t_0)
\]

\[
i_2 \equiv \int dt \ I(t)A^2(t - t_0)
\]

\[
i_3 \equiv \int dt \ I(t)A''(t - t_0)
\]

\[
i_4 \equiv \int dt \ I(t) \left[ A(t - t_0)A''(t - t_0) + A'^2(t - t_0) \right]
\]

we have

\[
\hat{a} = \frac{i_0 - \delta \hat{t}_0 i_1}{i_2}
\]

\[
\hat{a} = \frac{-i_1 + \delta \hat{t}_0 i_3}{\delta \hat{t}_0 \ i_4}.
\]

Solving for \( \delta \hat{t}_0 \) (to first order) we have

\[
\delta \hat{t}_0 = \frac{-i_1 i_2}{i_0 i_4 + i_2 i_3}.
\]
Iterative Solution for $\hat{t}_0$

This equation can be solved iteratively for $\delta\hat{t}_0$:

0. choose a starting value for $\hat{t}_0$.
1. calculate $\delta\hat{t}_0$ using the linearized equations.
2. is $\delta\hat{t}_0 = 0$?
3a. if yes, stop.
3b. if no, update $\hat{t}_0 \rightarrow \hat{t}_0 + \delta\hat{t}_0$ and go back to step 1.

For the best fit value for $\hat{t}_0$, the change is zero, $\delta\hat{t}_0 = 0$ (top of the hill) and $\hat{a}$ can be calculated using one of the equations 5 or 6.

Correlation Function Approach

The iterative solution for $\hat{t}_0$ is similar to the following procedure that uses a crosscorrelation approach more directly:

1. cross correlate the template $A(t)$ with $I(t)$ to get a CCF.
2. find the lag of peak correlation as an estimate for the arrival time, $\hat{t}_0 = \tau_{\text{max}}$.
3. calculate $\hat{a}$ if needed.

Subtleties of the Cross Correlation Method

The CCF is calculated using sampled data and therefore is itself a discrete quantity. Often one wants greater precision on the arrival time than is given by the sample interval. I.e. we want a floating point
number for $\hat{t}_0$, not an integer index. Therefore we want to calculate
the peak of the CCF by interpolating near its peak. The interpolation
should be done properly by using the appropriate interpolation
formula for sampled data (using the sinc function). Using parabolic
interpolation yields excessive errors for the arrival time.

In practice, the proper interpolation is effectively done in the frequency
domain by calculating the phase shift of the Fourier transform of the
CCF, which is the product of the Fourier transform of the template
and the Fourier transform of the data.
Timing Error from Radiometer Noise

rms TOA error from template fitting with additive noise:

$$\Delta t_{S/N} = \frac{\left[ \int \int dt \, dt' \, \rho(t-t')U(t)U'(t') \right]^{1/2}}{\text{SNR} \int dt \, [U(t)]^2} = \frac{W_{\text{eff}} \left( \frac{\Delta}{W_{\text{eff}}} \right)^{1/2}}{\text{SNR}}$$

Gaussian shaped pulse:

$$\Delta t_{S/N} = \frac{W}{(2\pi \ln 2)^{1/4} \text{SNR} \sqrt{N}} \left( \frac{\Delta}{W} \right)^{1/2}$$

$$\Delta t_{S/N} = 0.69 \mu s W_{\text{ms}} N_6^{-1/2} \text{SNR}_{1}^{-1} (\Delta/W)^{1/2}$$

Interstellar pulse broadening, when large, increases $\Delta t_{S/N}$ in two ways:

- SNR decreases by a factor $W / [W^2 + \tau_d^2]^{1/2}$
- $W$ increases to $[W^2 + \tau_d^2]^{1/2}$

→ Large errors for high DM pulsars and low-frequency observations

Low-DM pulsars: DISS (and RISS) will modulate SNR

$N_6 = N / 10^6$
Timing Error from Pulse-Phase Jitter

\[ U(\phi) \propto \int d\phi' f_\phi(\phi') a(\phi - \phi') \]

\[ \Delta t_j = N_i^{-1/2} (1 + m_j^2)^{1/2} P(\phi^2)^{1/2} \]

\[ = N_i^{-1/2} (1 + m_j^2)^{1/2} P \left[ \int d\phi \phi^2 f_\phi(\phi) \right]^{1/2} \]

- \( f_\phi \) = PDF of phase variation
- \( a(\phi) \) = individual pulse shape
- \( N_i \) = number of independent pulses summed
- \( m_i \) = intensity modulation index \( \approx 1 \)
- \( f_j \) = fraction jitter parameter = \( \phi_{rms} / W \) \( \approx 1 \)

Gaussian shaped pulse:

\[ \Delta t_J = \frac{f_J W_i (1 + m_i^2)^{1/2}}{2(2N_i \ln 2)^{1/2}} \]

\[ N_6 = N_i / 10^6 \]

\[ \Delta t_J = 0.28 \mu s W_{i,ms} N_6^{-1/2} \left( \frac{f_J}{1/3} \right) \left( \frac{1 + m_i^2}{2} \right)^{1/2} \]
Arrival Time Errors

Here we wish to localize the occurrence of a function $A(t)$. We will consider this to be a pulse whose arrival time $t_0$ we want to estimate, along with its expected error.

Let the signal be

$$I(t) = a_0 A(t - t_0) + n(t),$$

where $a_0$ is the amplitude and $n(t)$ is zero mean noise.

We will find the time of arrival (TOA) by cross correlating the presumed known pulse shape $A(t)$ with the signal:

$$C_{AI}(\tau) = \int dt \ I(t) A(t - \tau).$$

First, assume that the signal has been coarsely shifted so that the template is already aligned with the signal and that template is centered on $t = 0$. This way we can assume that the arrival time estimate is a small correction to the coarse estimate.
We can expand the signal as

\[ I(t) \approx a_0 A(t) - a_0 t_0 A'(t) + n(t). \]

We also expand the template, but to second order in \( \tau \) because we will be taking a derivative to find the lag of maximum correlation:

\[
\begin{align*}
A(t - \tau) &\approx A(t) - \tau A'(t) + \frac{\tau^2}{2} A''(t) \\
A'(t - \tau) &\approx A'(t) - \tau A''(t).
\end{align*}
\]

Then we can write

\[
C''_{IA}(\tau) = \frac{d}{d\tau} C'_{AI}(\tau) = 0 \\
= \int dt \, I(t) A'(t - \tau) \\
\approx C'_{IA'}(0) - \tau C''_{IA'}(0)
\]

An estimator for \( \tau \) is then

\[
\tau = \frac{C_{IA'}(0)}{C''_{IA'}(0)} = \frac{a_0 C_{AA'}(0) - a_0 t_0 C_{A'A'}(0) + C_{nA'}(0)}{a_0 C_{AA'} - a_0 t_0 C_{A'A'}(0) + C_{nA'}(0)}.
\]

Using previous approximations we encounter the terms \( C'_{AA'} \) and \( C_{A'A'}(0) \) that vanish for pulses that are zero at \( \pm \infty \). Also the \( C'_{nA'} \) term in the denominator yields a second-order term that can be ignored. Then the TOA estimator becomes

\[
\tau = \frac{-a_0 t_0 C_{A'A'}(0) + C_{nA'}(0)}{a_0 C_{AA'}(0)}.
\]
Noiseless case:

When there is no noise we have

\[ \tau = -t_0 \frac{C_{A'A'}(0)}{C_{AA''}(0)}. \]

Using a trial function, such as a Gaussian shape, it can be shown that \( \tau = t_0 \), as we would expect!

Even better, the denominator can be integrated by parts to show that \( C_{AA'}(0) = -C_{A'A'}(0) \) for pulses that vanish at \( \pm \infty \), so the equality is general.

With noise:

We can now write

\[ \tau = t_0 + \frac{C_{nA'}(0)}{a_0C_{AA''}(0)}. \]

Then the mean-square TOA error is, using \( \delta \tau = \tau - t_0 \),

\[ \langle (\delta \tau)^2 \rangle = \frac{C_{nA'}^2(0)}{a_0^2C_{AA''}^2(0)} \]

\[ = \frac{C_{nA'}^2(0)}{a_0^2C_{A'A'}^2(0)} \]

\[ = \int \int dtdt' \langle n(t)n(t') \rangle A(t)A(t'). \]
Now assume \textit{white noise} (for specificity) so that
\[
\langle n(t)n(t') \rangle = \sigma_n^2 w_n \delta(t - t')
\]
where \( w_n \) is a short characteristic time scale (such as an inverse bandwidth) to keep the units correct. Then
\[
\left\langle (\delta \tau)^2 \right\rangle = \left( \frac{\sigma_n}{a_0} \right)^2 \frac{w_n \int dt A'^2(t)}{C_{A'A'}(0)}
\]
\[
= \left( \frac{\sigma_n}{a_0} \right)^2 \frac{w_n}{C_{A'A'}(0)}
\]
We can then write this out as a TOA error
\[
\sigma_\tau = \left( \frac{\sigma_n}{a_0} \right) \left[ \frac{w_n}{\int dt \ [A'(t)]^2} \right]^{1/2}
\]
\[
= \frac{1}{\text{SNR}} \left[ \frac{w_n}{\int dt \ [A'(t)]^2} \right]^{1/2}.
\]
We see that the error scales as the inverse of the \textit{signal to noise ratio} (SNR). The denominator also involves the integral of the squared derivative of the pulse shape, suggesting that \textit{sharper} pulses with larger derivatives will produce smaller arrival time errors.
AN OPTIMIZED ALGORITHM
FOR DETERMINING PULSAR ARRIVAL TIMES

J. H. Taylor
Joseph Henry Laboratories and Physics Department
Princeton University

Berkeley Workshop on Pulsar Timing, June 7-9, 1990

Binary and millisecond pulsars are thought to be relatively old objects, recycled
into active status by the accretion of mass and angular momentum from an evolving
companion star. They have relatively short periods and small period derivatives, and
they make particularly attractive targets for precision timing experiments — partly
because they are even better time-keepers than ordinary pulsars, and partly because
of the special characteristics associated with their complicated evolutionary histo-
ries. My collaborators and I have recently been using accurate measurements of pulse
times of arrival (TOA's) for a wide range of applications including milli-arcsecond
astrometry, tests of fundamental physical laws, cosmology, and timekeeping metro-
logy. Accumulated experience has only strengthened our motivation for achieving the
highest possible accuracies in pulsar timing measurements. In this paper, I describe
an algorithm used at Princeton for the optimum extraction of information from pulsar
timing data.

To a reasonably good approximation, the integrated profiles for a given pulsar at a
particular observing frequency are all the same except for a DC bias, a multiplicative
scale factor, a shift of time origin, and additive random noise. In other words, an
observed profile \( p(t) \) for a given pulsar is related to the corresponding standard profile,
\( s(t) \), by an equation of the form

\[
p(t) = a + b s(t - \tau) + g(t),
\]

where \( a \), \( b \), and \( \tau \) are constants and \( g(t) \) is a random variable representing radiometer
and background noise. To measure TOA's we want to determine \( \tau \) as accurately as
possible in the presence of a given amount of noise. This might be done by using a
"least squares" or "maximum likelihood" method to fit a scaled, shifted version of
the standard profile to the observed profile, obtaining optimized estimates of the bias,
scale factor, and time offset. The parameter uncertainties would also be calculated;
their magnitudes would depend on the pulse shape and the signal-to-noise ratio.

In practice, observed profiles are digitized representations of the detected signals,
with a well defined sampling interval and known amount of instrumental smoothing.
Let's assume that \( p(t) \) and \( s(t) \) have been recorded at \( N \) equally spaced intervals
covering the full period, say \( t_j = j \Delta t, j = 0, 1, \ldots, N - 1 \). Good experiment design
requires that the detected signals be low-pass filtered at a cutoff frequency \( f_c \leq \)
Notice that the sample interval, $\Delta t = P/N$, appears nowhere in Eqs. (3-12). Our formalism places no limit on timing accuracy expressed as a fraction of the sampling interval. In contrast, experience has shown that the time-domain methods widely in use do not readily produce arrival-time accuracies much smaller than about $0.1\Delta t$ (see, for example, Rawley 1986).

Figure 1 illustrates the results of fitting a set of 1000 artificially generated pulse profiles, with known time offsets, using both time-domain and frequency-domain algorithms. The frequency-domain uncertainties, represented by filled circles in Figure 1, scale faithfully with the signal-to-noise ratio, reaching values around $0.005\Delta t$ at signal-to-noise-ratios of several hundred. On the other hand, the time-domain solutions, in which $\tau$ was estimated by parabolic interpolation of the $\chi^2$ minimum, achieve accuracies no better than about $0.05\Delta t$ even at very high signal-to-noise ratios.

![Graph showing uncertainty in TOA vs signal to noise ratio](image_url)

**Figure 1:** Average uncertainties obtained for arrival times from simulated observations, plotted as a function of signal-to-noise ratio. Open circles represent TOA's determined by parabolic interpolation in the time domain; filled circles are based on data using the algorithm described in this paper.
AN OPTIMIZED ALGORITHM FOR DETERMINING TOA's

J. H. Taylor

Berkeley Workshop on Pulsar Timing
June 7–9, 1990

Standard profile, $s(t)$

Observed profiles, $p(t) = a + b s(t - \tau) + g(t)$

The problem: for each observed profile, find $\tau$ as accurately as possible.
1. Least squares in the time domain:

- Sampled profiles, \( s_j = s(t_j) = s(j \Delta t) \), 
  \( j = 0, \ldots, N - 1 \)  \((N \Delta t = \text{period})\).

- Find minimum with respect to \( \tau \) of

\[
\chi^2(a, b, \tau) = \sum_{j=0}^{N-1} \left( \frac{p_j - a - bs_j - \tau}{\sigma} \right)^2
\]

- Equivalent: find maximum of CCF,

\[
c(\tau) = p_j * s_j
\]

\( \Rightarrow \) Problem: profiles are quantized in time; cannot get analytic derivative \( \partial \chi^2 / \partial \tau \).
• Partial solution: use parabolic interpolation to find the extrema.

• Better solution: use an interpolating function appropriate to band-limited functions,

\[
sinc(x) \equiv \frac{\sin(\pi x)}{\pi x}
\]

• This suggests transforming to the frequency domain.
A tell-tale signature of difficulties with the time domain method is contained in the error histograms reproduced in Figure 2. For the low signal-to-noise cases near the top of the diagram, histograms for both time-domain and frequency-domain results have the same width and the expected gaussian shape. For stronger signal levels, however, the time-domain results show much more scatter and their distribution becomes bimodal owing to the tendency of TOA's to be "pulled" toward bin alignment with the standard profile. In extreme cases the advantage of the frequency-domain algorithm can be as much as an order of magnitude.

Figure 2: Histograms of individual errors making up the averages represented in Figure 1. The plot widths have been scaled with the signal-to-noise ratio, so the frequency-domain histograms show approximately constant width. The degraded performance of the time-domain algorithm at high signal-to-noise ratios is clearly evident.
Of course it must be true that anything that can be done in one domain can also be done in the other. In this case, the problem with the commonly used time-domain procedure is not in the binned $\chi^2$ or cross-correlation function, but rather in the parabolic interpolation scheme used to find the extremum. The correct procedure would be to use the interpolation function appropriate for band-limited functions, namely $\text{sinc}(z) \equiv \sin(\pi z)/\pi z$. The interpolation operation then becomes a convolution, and for computational efficiency it (as well as the cross correlation of the observed and standard profiles) would best be carried out by transforming to the frequency domain and multiplying transforms. With our preferred algorithm, there is no good reason ever to transform back into the time domain.

REFERENCES


PSR 1913+16
Mark I

Relative amplitude

log [Frequency bin number]
Localization: time vs frequency domain

For a Nyquist sampled, band limited process the sampling theorem is:

\[ x(t) = \sum_{n=-\infty}^{\infty} x_n \text{sinc}(t-n\Delta t)/\Delta t \]  \( \star \)

Consider a model where:

\[ x(t) = a \cdot T(t-t_0) + n(t) \]

where:

- \( a \) = amplitude
- \( T \) = known template
- \( n \) = noise

Suppose we have sampled data and a sampled template \( X_n, T_n \).

It would be tempting to find the discrete CCF

\[ C_{XT}(\tau) = \sum_{n} X_n \cdot T_{n-\tau} \]

and find the maximum distance to \( a \cdot T(t-t_0) \).

However, we really want the CCF of the continuous function \( x(t) \) and \( T(t) \):

\[ C_{XT}(\tau) = \int_{-\infty}^{\infty} x(t) T(t-\tau) \]

We cannot get the true lag \( T_0 = \tau \) by interpolating the sampled data unless we interpolate very finely, e.g., \( \star \).

If we interpolate sufficiently we will get non-optimal results.
or we could cross correlate the properly interpolated functions

or we could calculate \( \hat{x} \) in the frequency domain.

**Frequency Domain**

\[
\hat{x}(f) = a \hat{T}(f) e^{-j2\pi ft_0} + \hat{v}(f)
\]

1. Consider a case with no noise and define:

\[
\hat{x}(f) = \hat{x}_r(f) + j\hat{x}_i(f)
\]

\[
\phi(f) = \tan^{-1} \left( \frac{\hat{x}_i(f)}{\hat{x}_r(f)} \right) = \tan^{-1} \left( \frac{\sin(2\pi ft_0)}{\cos(2\pi ft_0)} \right) = -2\pi ft_0
\]

2. With noise

\[
\phi(f) \quad \text{phase wrap. for } \tan^{-1} x \in [-\pi, \pi]
\]

\[
\phi(f) \quad \text{phase wrap. induced by noise}
\]
fitting a line to $\phi(t)$ is a bad idea.

Better fit $X_r$ and $\tilde{X}_i$:

Let $\hat{X}(t) = \tilde{X}(t) e^{-\text{unit} \cdot t}$.

Let $\tilde{J}(t) = \tilde{X}(t) \cdot \hat{X}^\ast(t)$

$$= a \hat{\tilde{\eta}}(t) e^{-\text{unit} \cdot t} + \tilde{\eta}(t) \hat{\tilde{\eta}}^\ast(t)$$

Interpret the partial:

$$\tilde{J}(t_0) = \int \tilde{J}(t) \, dt = \int \left[ a \hat{\tilde{\eta}}(t) e^{-\text{unit} \cdot t} + \tilde{\eta}(t) \hat{\tilde{\eta}}^\ast(t) \right] \, dt$$

Note $\tilde{\eta}(t)$ has real phase.

Maximize w.r.t. $t_0$: $\hat{J}(t_0) = \tilde{J}(t_0)$ ⇒ maximum.

N.B. that we are comparing a continuous partial here.

With sampled data we have (for interpolating samples)

$$\tilde{X}(t) = \Pi \left( \frac{t}{\Delta} \right) \sum X_n e^{-\text{unit} \cdot n \Delta t}$$

Can and DFT $\tilde{X}_n = (2\pi)^{-1} \Delta t$.