

Analysis of filtered poisson process of double exponentials

This writeup presents the most important results. For technical reference, see `readme.md`.

Data generation

The time series, $\Phi(\tau)$ was generated with the following conditions:

- $\tau_d = 1\mu s$
- Uniformly distributed arrival times
- Exponentially distributed amplitudes with scale 1
- Predetermined number of events; $n = \gamma T$
- Four different values of γ ; 0.1, 1, 10 and ∞ .
- For each γ , six different values of λ ; 0, 0.1, 0.2, 0.3, 0.4 and 0.5
- A total duration of $T = 100000\mu s$
- $dt = 0.01\mu s$

The realizations with real numbered γ were generated using the [superposed pulses](#) package. The one with $\gamma = \infty$ was generated by convolving the pulse shape $\phi(\tau)$ with a series of normally distributed amplitudes.

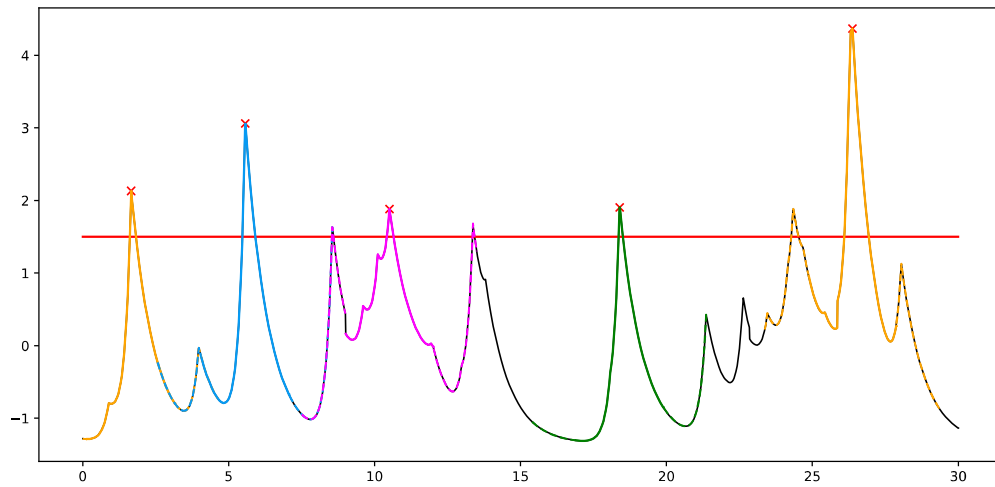
For each value pair $\{\gamma, \lambda\}$, 100 realizations of the series were made.

Conditional averaging

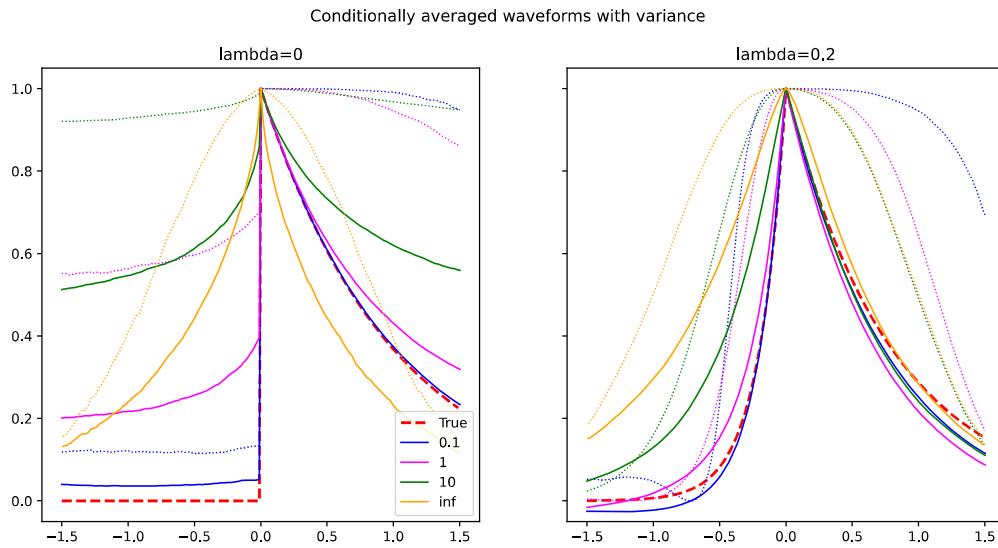
The conditional averaging algorithm used is a modified version of from the one supplied in the [fpp-analysis-tools](#) package. It was modified to support normalizing the amplitudes of the conditional events as well as plotting an illustration of the analysis process. The base conditions used in the analysis were:

- Size of conditional events, $\Delta = 3\mu s$
- Threshold of $2.5\Phi_{rms}$
- Do not allow overlapping conditional events (`window=True`)
- Normalize the amplitudes of each conditional event

Below is an illustration of the process. Note that the threshold in the figure for illustrational purposes is set to $1.5\Phi_{rms}$. Each valid peak is marked with a red cross. The colored solid colored curves show the size of each event. The dotted lines show the domain within which other peaks are not counted as conditional events.

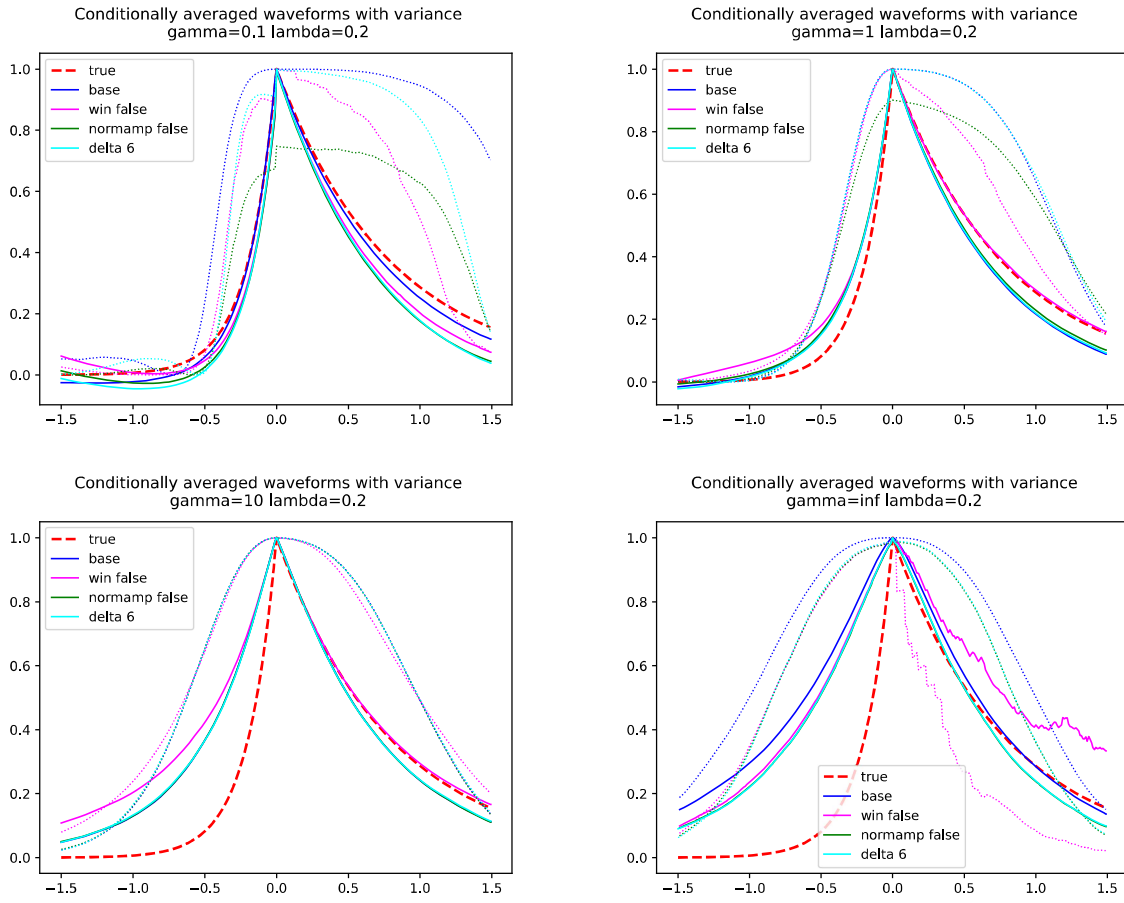


The resulting waveforms, with $1 - \text{conditional_variance}$ represented by the dotted lines, are shown here for two values of λ with each color representing a different γ :



Subsequently, three variations of the base conditions were analyzed. One allowing overlap, one in which the events were not normalized and one with $\Delta = 6$.

Following are comparisons of the base case and the three variations for different values of γ and $\lambda = 0.2$:

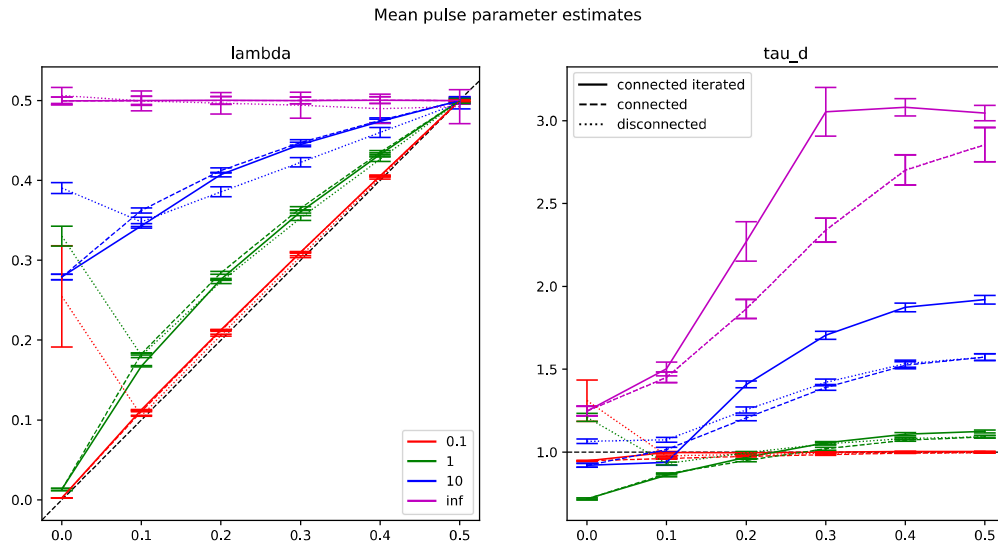


One qualitative conclusion to be drawn from these plots is that the base conditions seem to give overall lower conditional variance. Another observation is that for $\gamma \in \{1, 10\}$ allowing overlap seems to improve the estimate of the falling curve at the cost of the estimate of the rising curve.

Fit analysis

A least squares fit was performed to extract shape parameters from the conditionally averaged waveforms. Three different variations of this method were tested. First of all, two different functions were used; One double exponential, where the constant term and scaling factor were shared between the rising and falling flank, and one with both flanks having an independent constant term and scaling factor. Both of these were tested using the entire conditional waveform, $\tau \in [-\Delta/2, \Delta/2]$. Furthermore, the function with shared parameters was tested with a domain of $\tau \in [-\tau_r, \tau_f]$, where τ_r and τ_f were estimated iteratively: First, the function was fitted for $\tau \in [-\Delta/2, \Delta/2]$, from which $\tau_{r,i}$ and $\tau_{f,i}$ were estimated. Then, $\tau_{r,i+1}$ and $\tau_{f,i+1}$ were estimated on a domain of $[\tau_{r,i}, \tau_{f,i}]$ until $\tau_{d,i} = \tau_{r,i} + \tau_{f,i}$ converged. The vast majority of estimates converged within 100 iterations.

All three methods were tested on the base conditional waveform for each 100 realizations for each $\{\gamma, \lambda\}$ pair. The mean estimates of λ and τ_d with related standard variations are shown in the below plot, along with the true value in black:



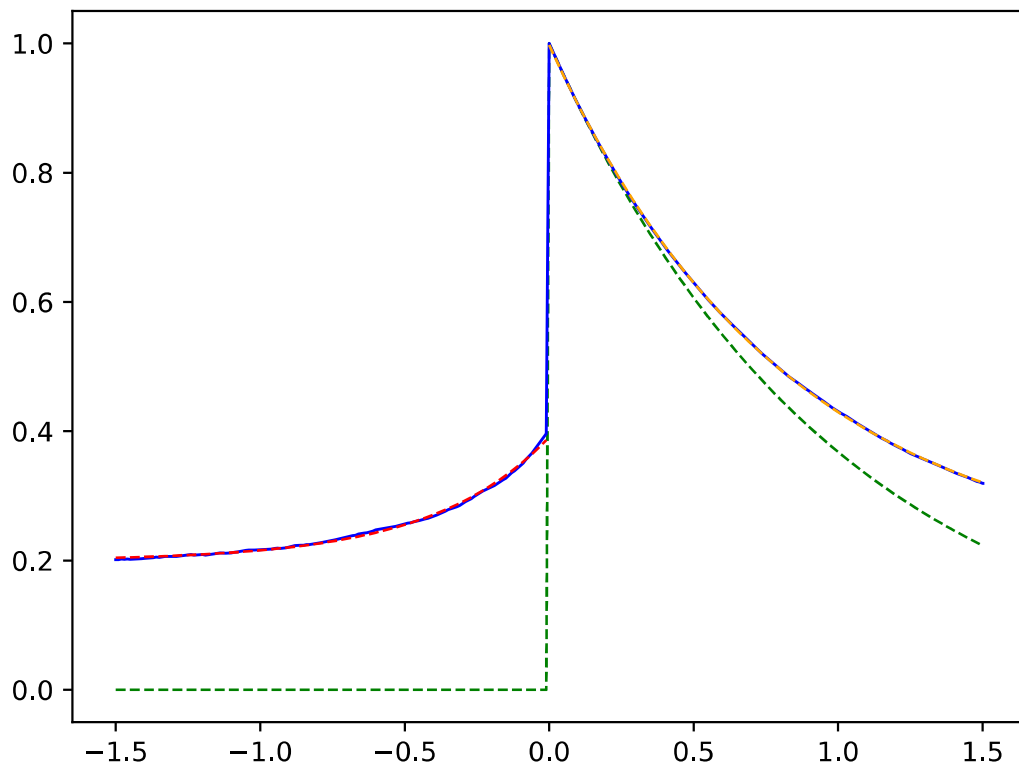
The first takeaway is that all methods give results within the same order of magnitude, and they all suggest the same trend (except for $\lambda = 0$); The slope of the estimate curve of λ decreases with higher γ , and the slope of the τ_d estimate increases, both giving less accurate estimates for higher γ .

We can also see that the method with shared parameters and constant domain ("connected") gives overall best performance for estimates of both parameters.

The method with shared parameters and iteratively estimated domain ("connected iterated") significantly underperforms in estimating τ_d for higher γ .

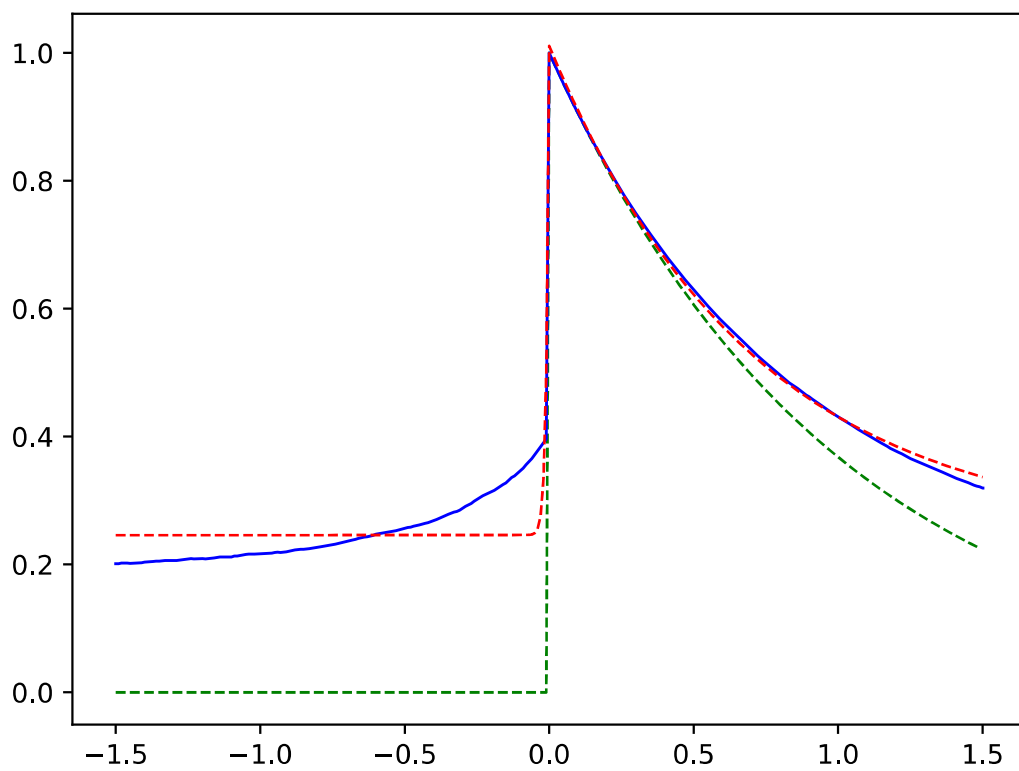
The method with independent parameters and constant domain ("disconnected") underperforms in estimating both parameters for $\lambda = 0$. This is due to the rising flank being too short, only consisting of 2-3 data points. Therefore the floor curve leading up to the flank, consisting of a lot more data points, is estimated instead of the actual flank. This is illustrated below, where the green dashed line is the true waveform, the blue line is the conditionally averaged waveform, and the red and yellow dashed lines represent the least squares fit:

least squares fit for $\gamma=1$, $\lambda=0$



Conversely, with shared parameters, the fit is more representative of the waveform:

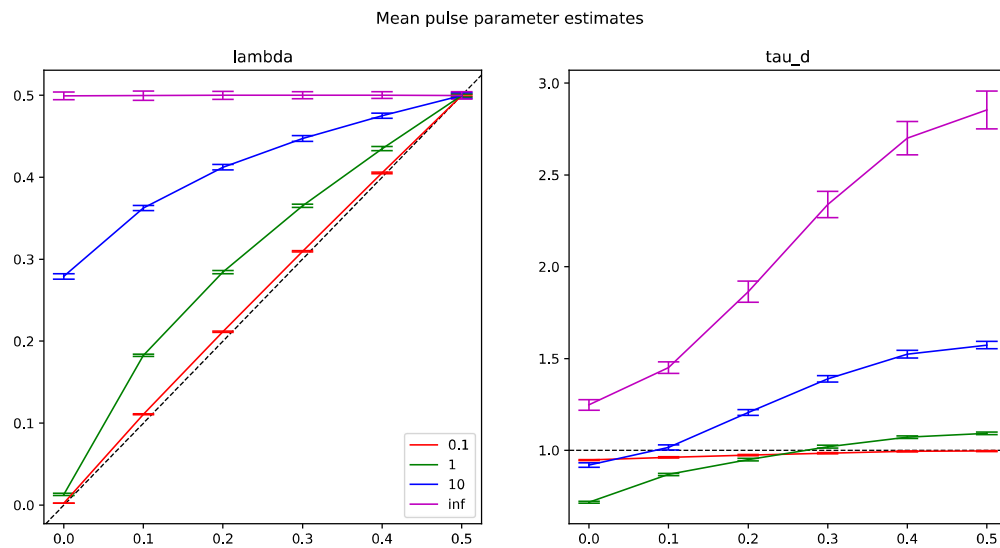
least squares fit for $\gamma=1$, $\lambda=0$



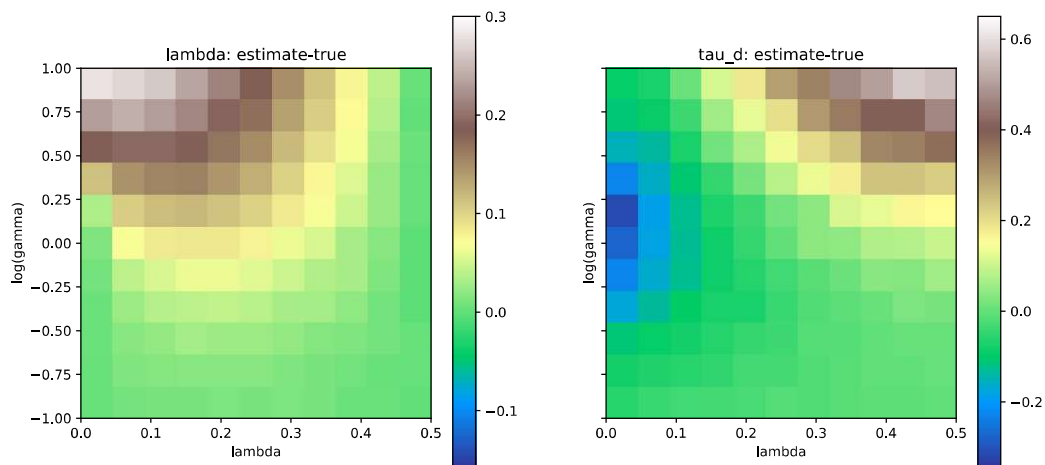
As a result of these conclusions, only the method with shared parameters and a constant domain is used in further analysis.

Estimation results

The results from estimating λ and τ_d for the base conditional averaging parameters are seen below:



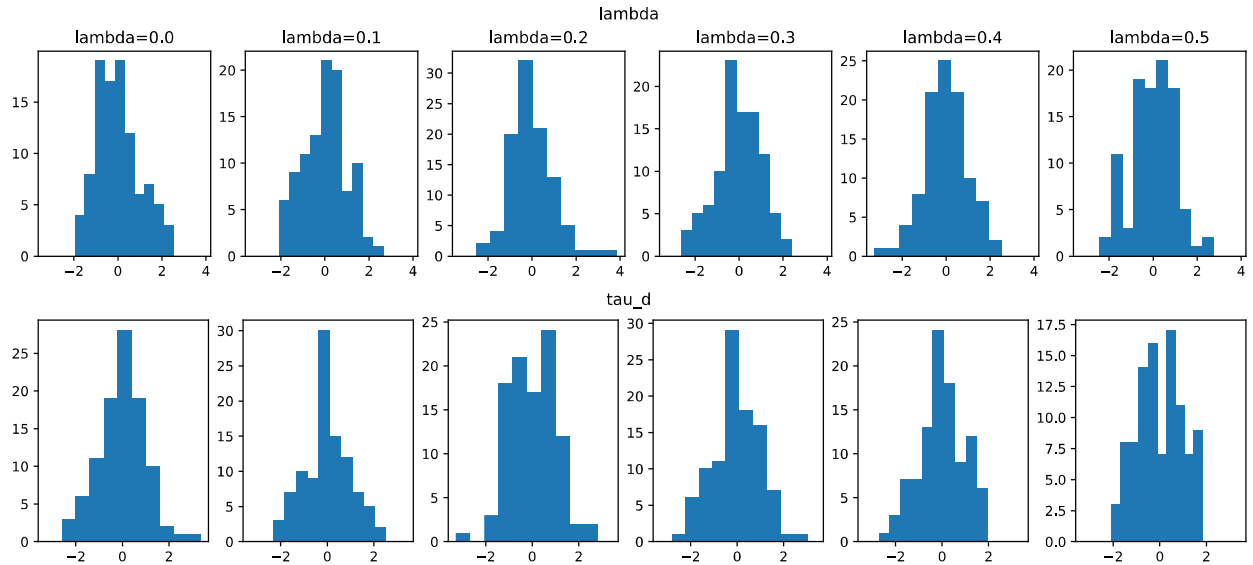
The following plot gives higher resolution, but each datapoint represents the estimate from one series only - not a mean like the plot above.



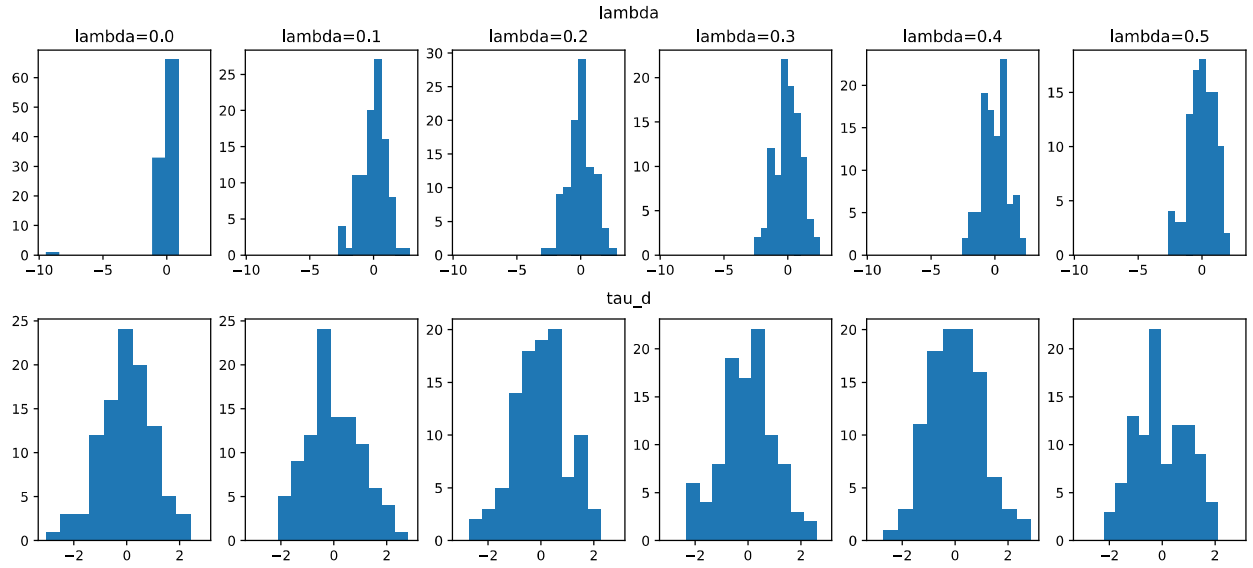
From both plots, some main conclusions can be drawn: The estimates are consistently correct within an order of magnitude, and for lower γ they provide a qualitatively good estimate.

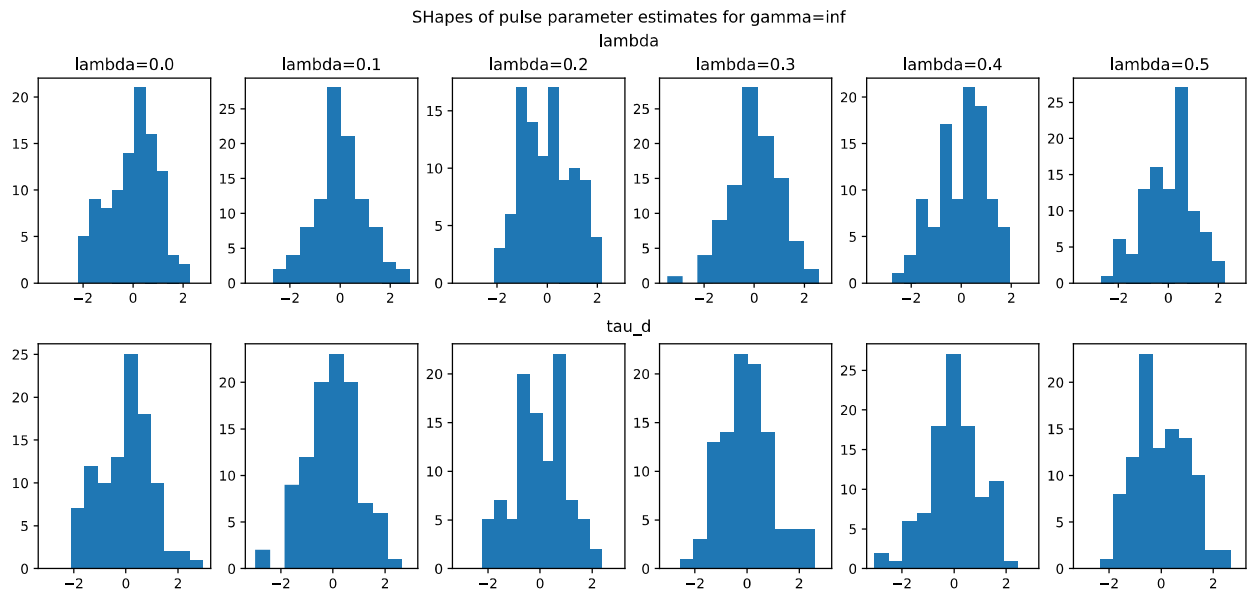
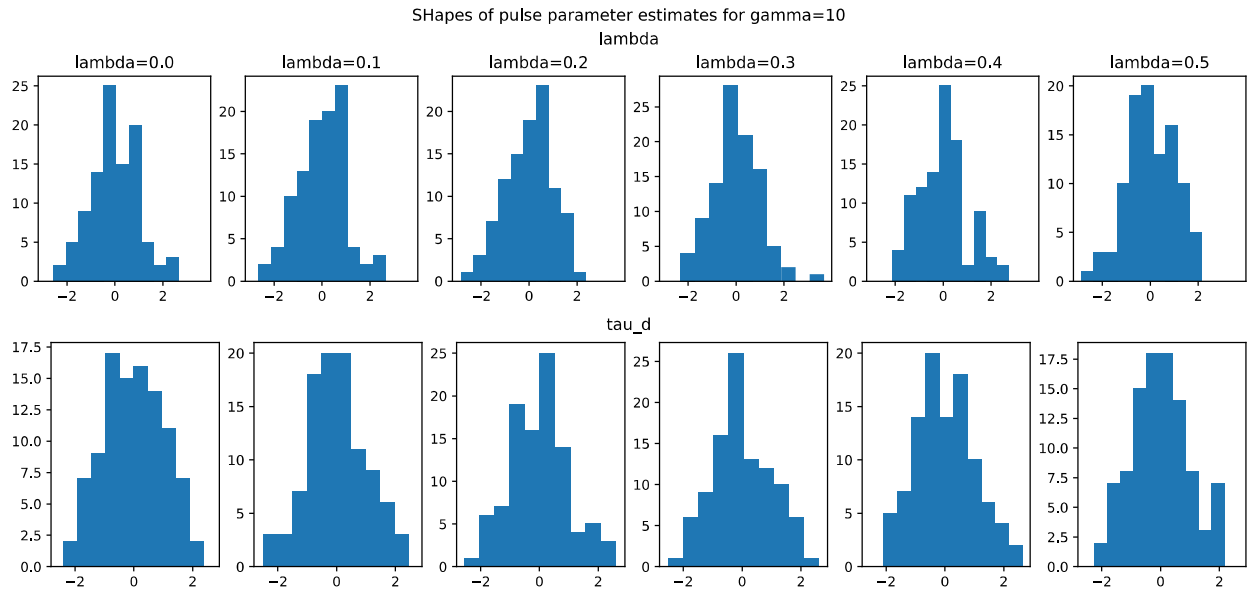
The shapes of the sets of estimates are shown below:

SHapes of pulse parameter estimates for gamma=0.1

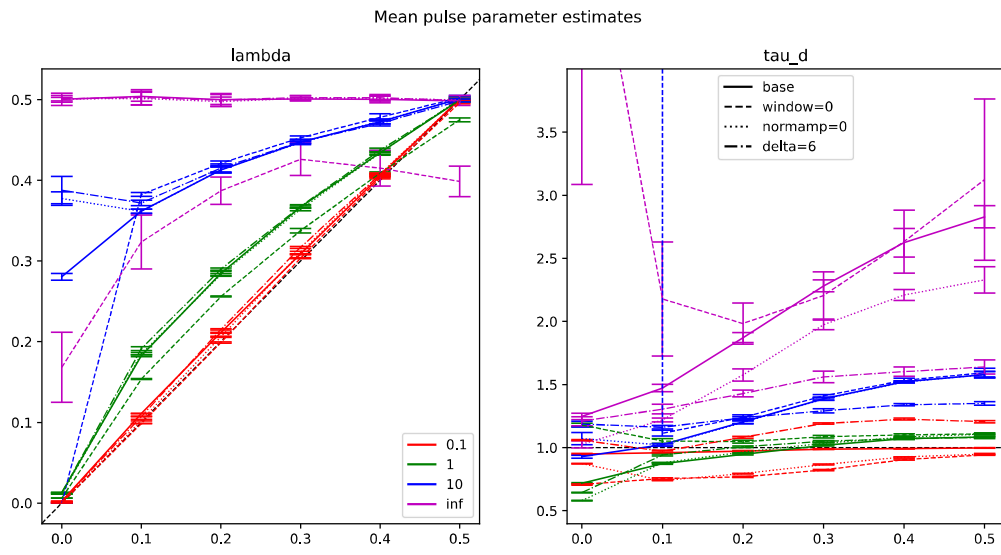


SHapes of pulse parameter estimates for gamma=1





Further investigations of the conditional averaging parameters show that the "base" parameters provide overall best performance. Allowing overlap does show promise for $\gamma = 1$, but is too unstable for higher values, and badly misses for $\gamma = 0$. Increasing Δ seems to improve estimates of τ_d for higher γ , but it also leads to a significant error for $\gamma = 0$. This is all visible in the fairly chaotic plot below:



Notice how the estimate for τ_d when allowing overlap shoots up for $\lambda = 0$, $\gamma = 10$; it ends up around $\tau_d = 10^7$.