

Stochastic gradient descent, semi-analytical optimization

- Analytically calculate filter function for a given $f(t)$ and derivatives wrt pulse spacing
- Adjust pulse spacing in a basis that satisfies certain boundary conditions. Here, adjustments are given by $-\delta t \times \partial_{t_n} \sin^2\left(\frac{\pi m t_n}{T}\right)$ where m is an integer.
 - The total pulse sequence duration is T – therefore, pulses that begin in the domain $0 < t < T$ stay in the domain according to this update rule, for sufficiently small δt .
- Pick the m that maximizes the change in χ .
- Update filter function and repeat.
- Best results start on slide 9, after the algorithm has been updated to numerically calculate χ instead of calculating $d\chi$ for a given df .
 - $\chi = \int |F(\tau, \omega)|^2 S(\omega) d\omega$ is the quantity to be minimized.
 - $f_\tau(t) = \mathcal{F}^{-1}\{F(\tau, \omega)\}$ is the time-domain filter function, where $f(t) = \{-1, 1\}$, depending on the pulse sequence.
 - For good results, see slides 8-12, 14, 15, 18, 21, 22, 24, and 32. Slides 2-7 are at the stage of figuring it out.
- The algorithm does not always work for arbitrary initial conditions. Sometimes the discrete time grid causes points to fall outside of the domain $0 < t < T$.
- The code currently begins with a symmetric sequence; half of the sequence is generated randomly, and the 2nd half is mirrored. This is not necessary. I'm not sure why I did this, but all of the results shown have this starting condition.

Semi-analytical optimization

$$\tilde{f} = \sum_{n=0}^{N-1} (-1)^n e^{-2\pi i b_n v} \frac{\sin(\pi a_n v)}{\pi v} \quad b_n = \frac{t_n + t_{n+1}}{2} \quad a_n = t_{n+1} - t_n$$
$$f(t) = \sum_{n=0}^{N-1} (-1)^n (H(t-t_n) - H(t-t_{n+1}))$$
$$\tilde{f} - \tilde{f}(v, \{a\}, \{b\}) \approx \tilde{f}(v, \{a\}^*, \{b\}^*) + \sum_{n=0}^{N-1} \frac{\partial \tilde{f}}{\partial a_n} \delta a_n + \frac{\partial \tilde{f}}{\partial b_n} \delta b_n$$

update

iteration

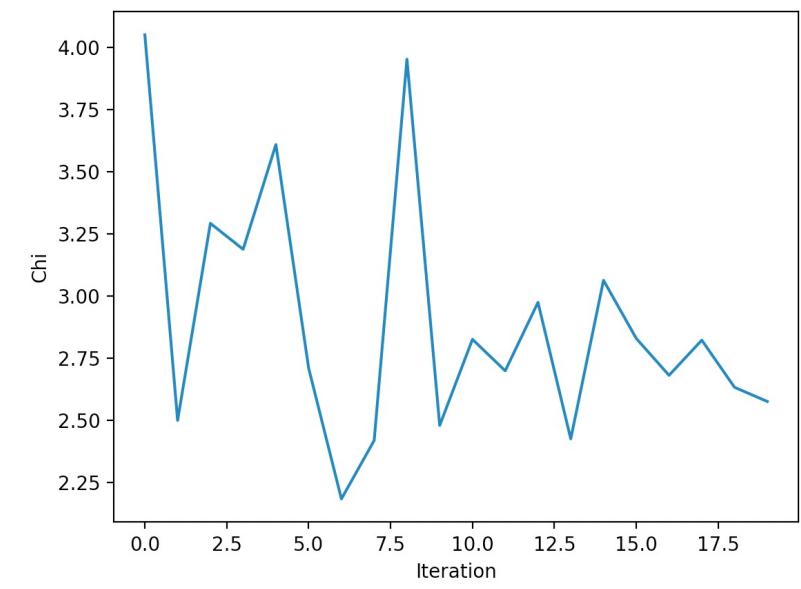
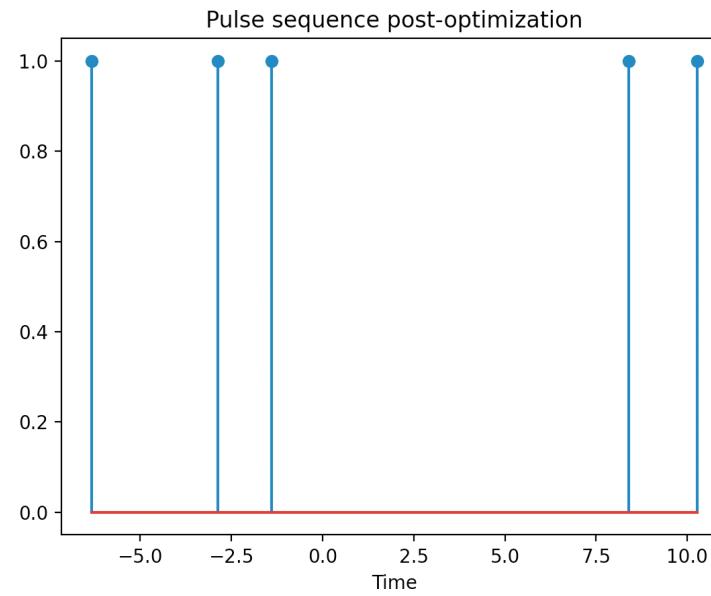
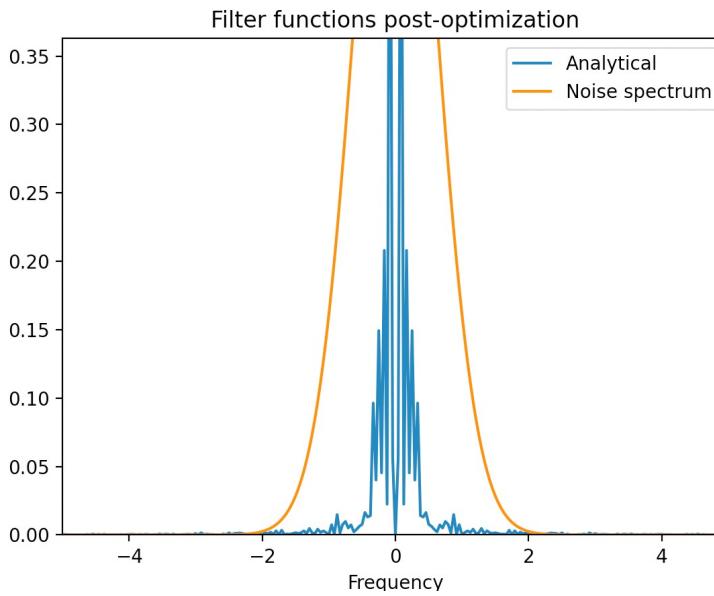
Randomized (symmetric) Starting pulse sequence:

$$\frac{\partial \tilde{f}}{\partial a_n} = (-1)^n e^{-2\pi i b_n v} \cos(\pi a_n v)$$
$$\frac{\partial \tilde{f}}{\partial b_n} = (-1)^n (-2\pi i v) e^{-2\pi i b_n v} \frac{\sin(\pi a_n v)}{\pi v} + (-1)^n (-2i) e^{-2\pi i b_n v} \sin(\pi a_n v)$$
$$t_n \rightarrow t_n - \frac{\partial \tilde{f}}{\partial t} \sin^2\left(\frac{\pi a_n v}{T}\right) \delta t$$
$$= t_n - \frac{2\pi v \delta t}{T} \sin\left(\frac{\pi a_n v}{T}\right) \cos\left(\frac{\pi a_n v}{T}\right)$$
$$a_n \rightarrow a_n - \frac{2\pi v \delta t}{T} \left(\sin\left(\frac{2\pi v t_n}{T}\right) - \sin\left(\frac{2\pi v t_{n+1}}{T}\right) \right) = a_n^* + \delta a_n$$
$$b_n \rightarrow b_n - \frac{2\pi v \delta t}{T} \left(\sin\left(\frac{2\pi v t_n}{T}\right) + \sin\left(\frac{2\pi v t_{n+1}}{T}\right) \right) = b_n^* + \delta b_n$$

Some trials

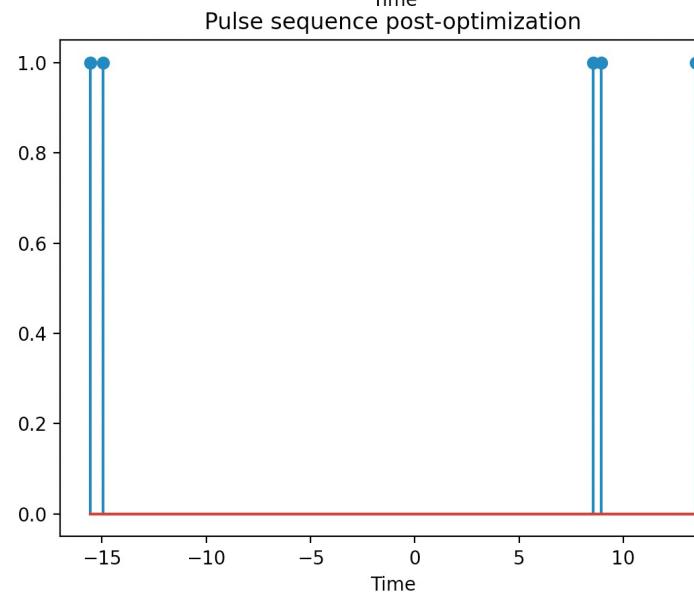
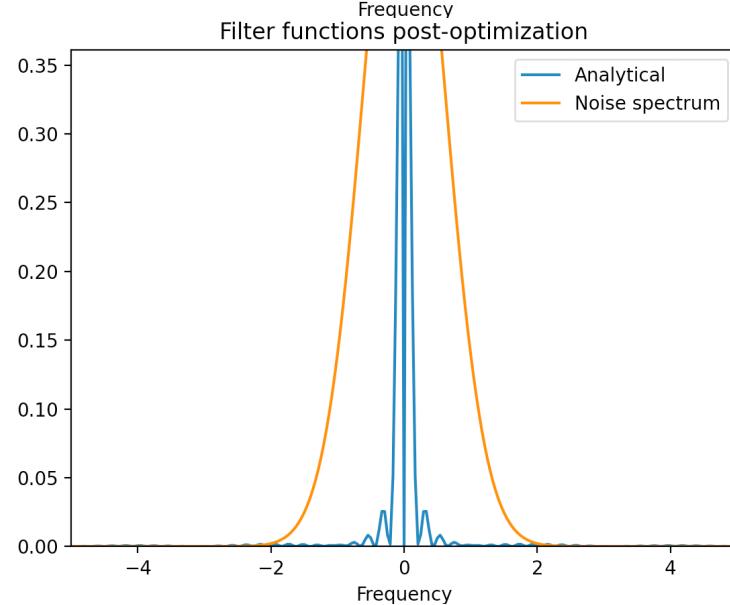
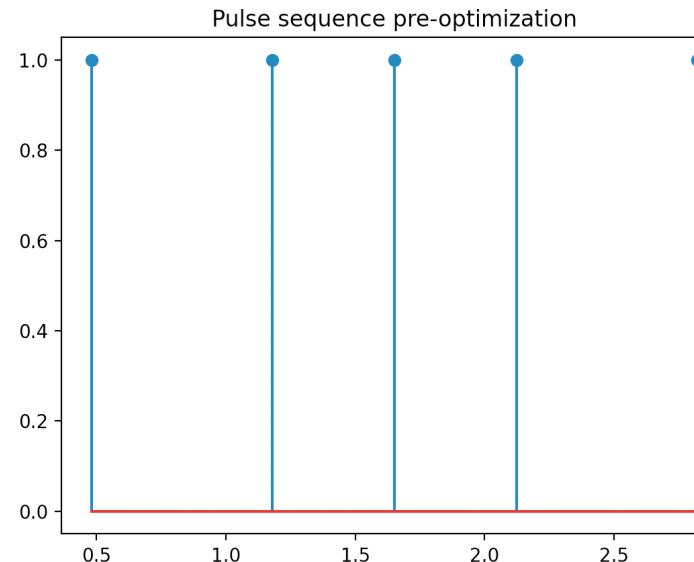
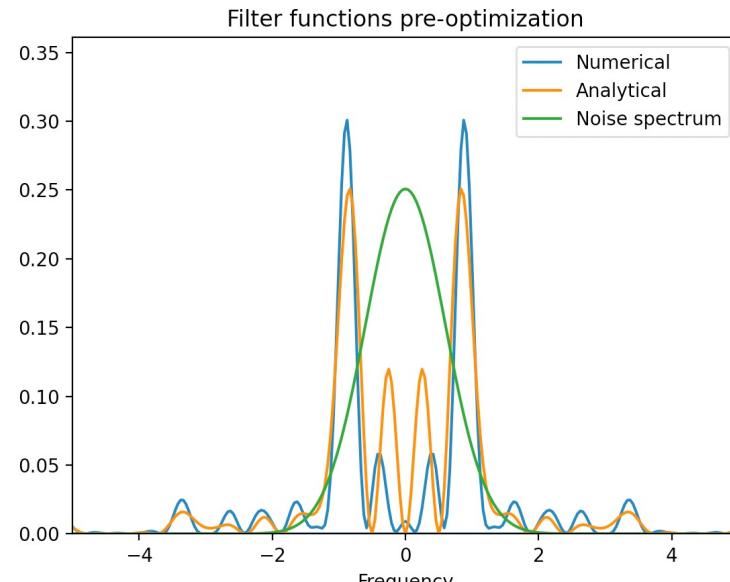
$m=20$ possible harmonics, $\Delta t = 50 dt$, 20 iterations, $N=5$ pulses

I only have the results saved; not the starting pulse sequence

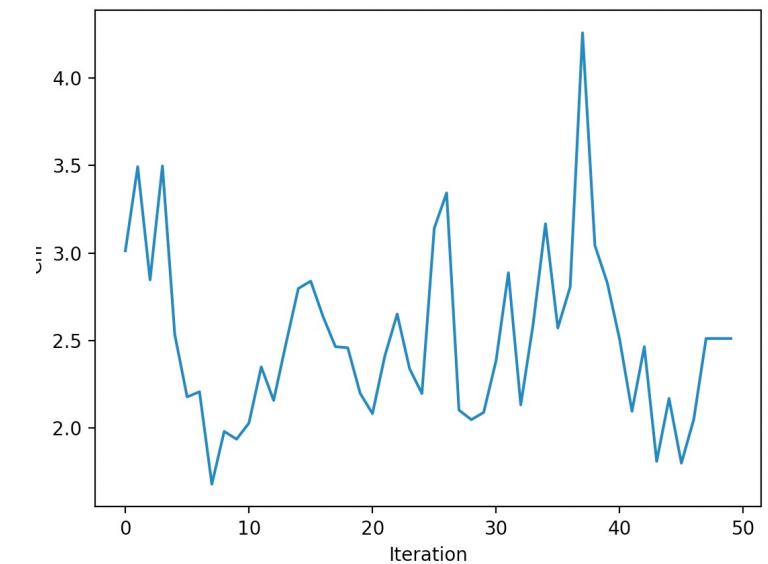


Some trials

$m=20$ possible harmonics, $\Delta t = 50 dt$, 20 iterations, $N=5$ pulses

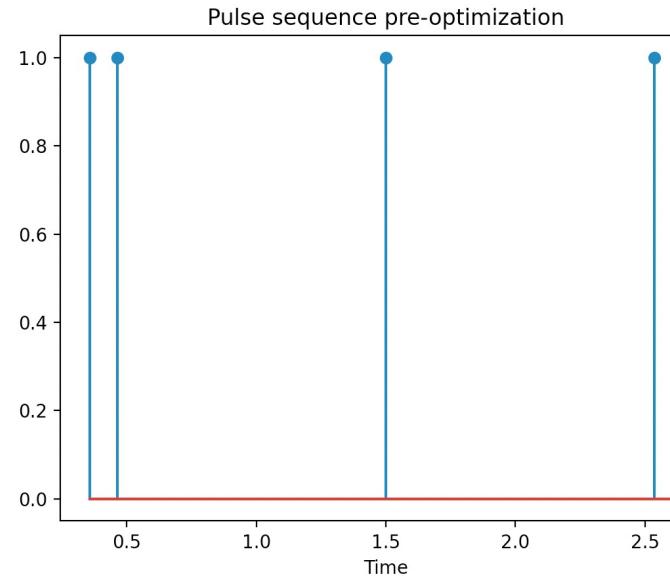
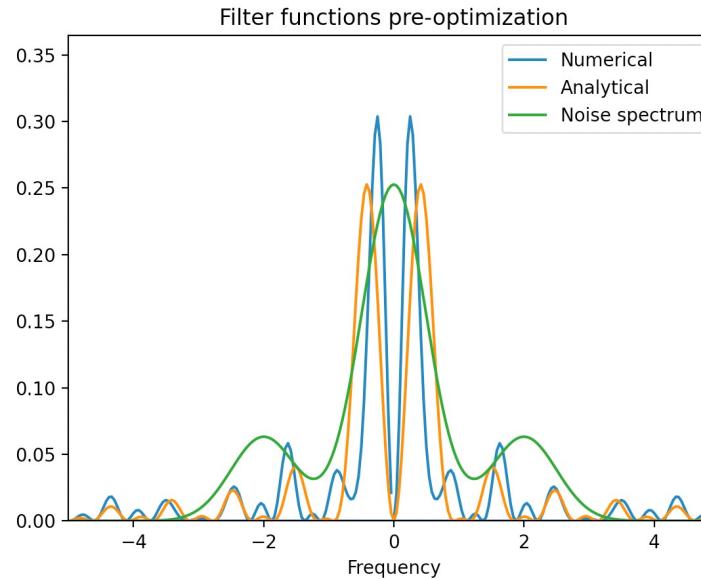


Why such large error?
Pulse timing doesn't satisfy the constraints?
Digital error?

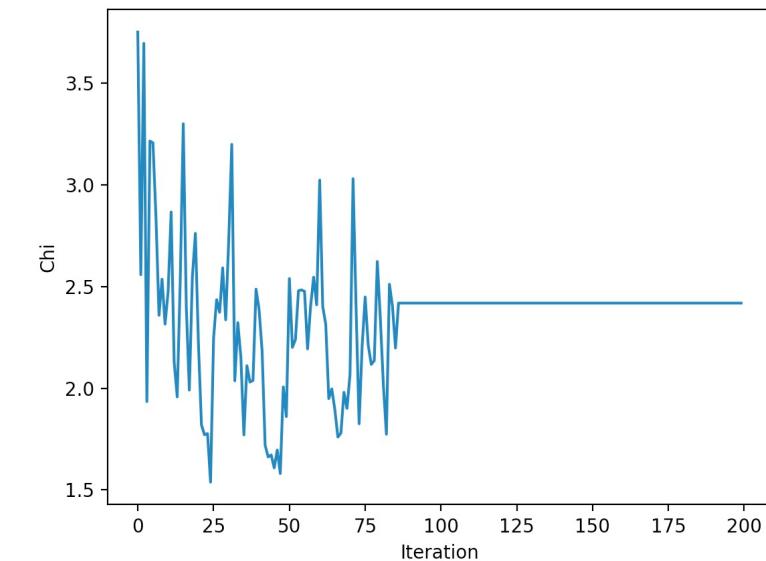
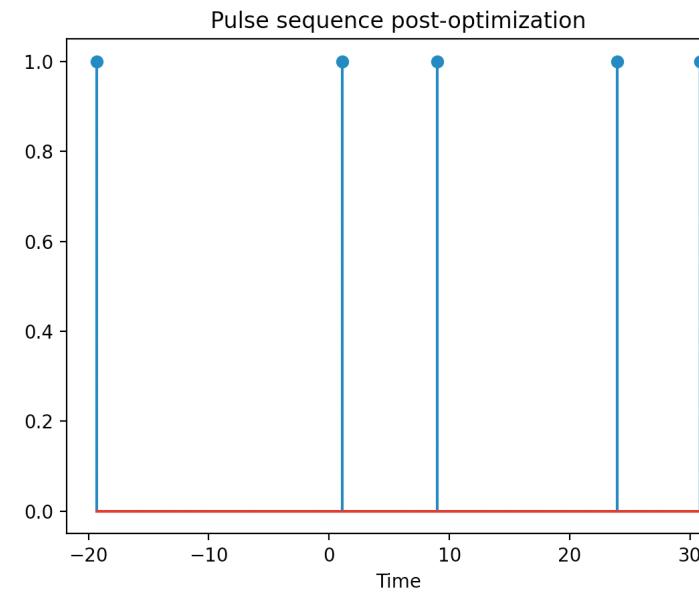
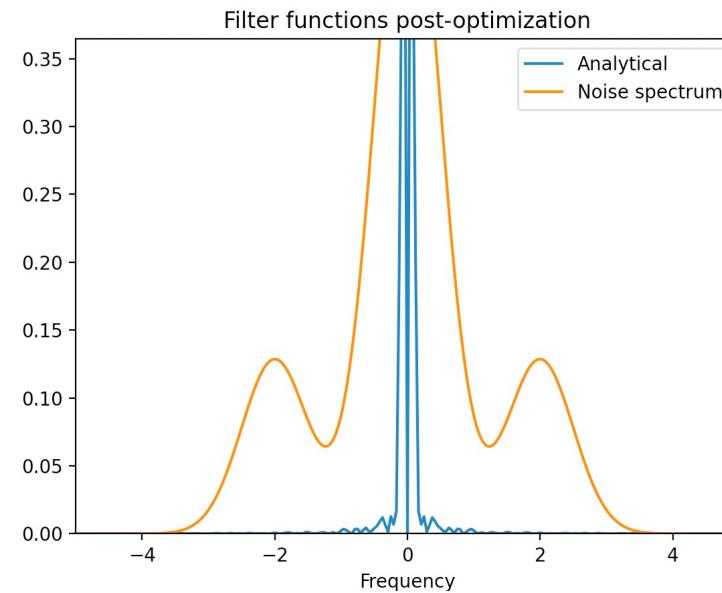


Some trials

$m=20$ possible harmonics, $\delta t = 50 dt$, 200 iterations, $N=5$ pulses; adding structure to the noise spectrum

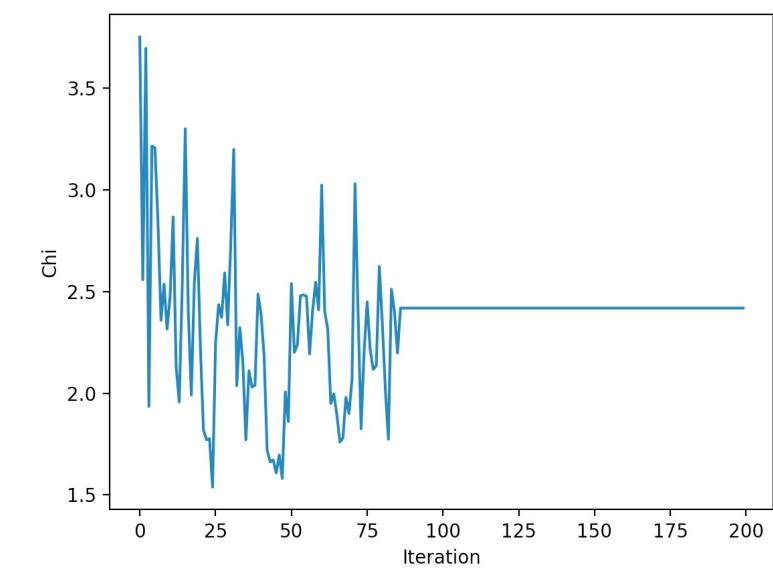
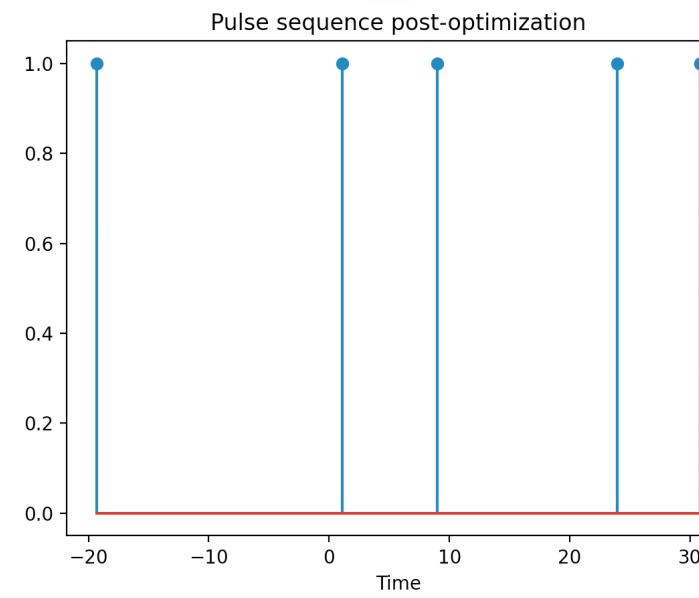
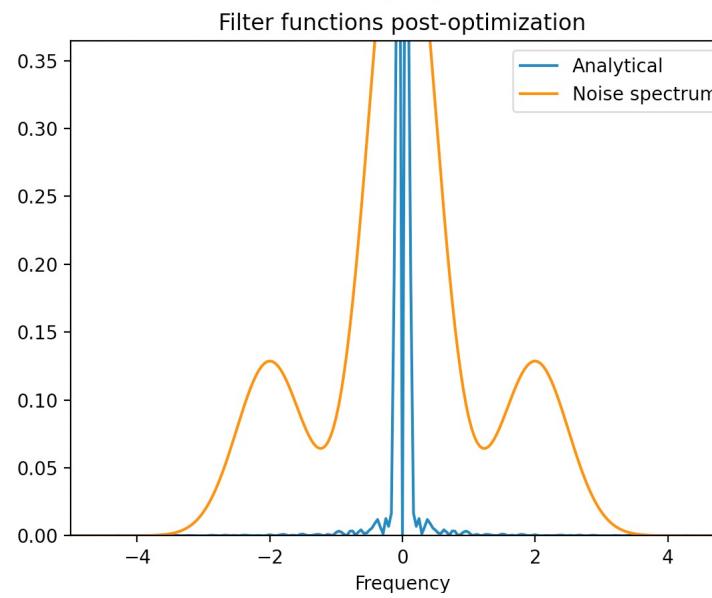
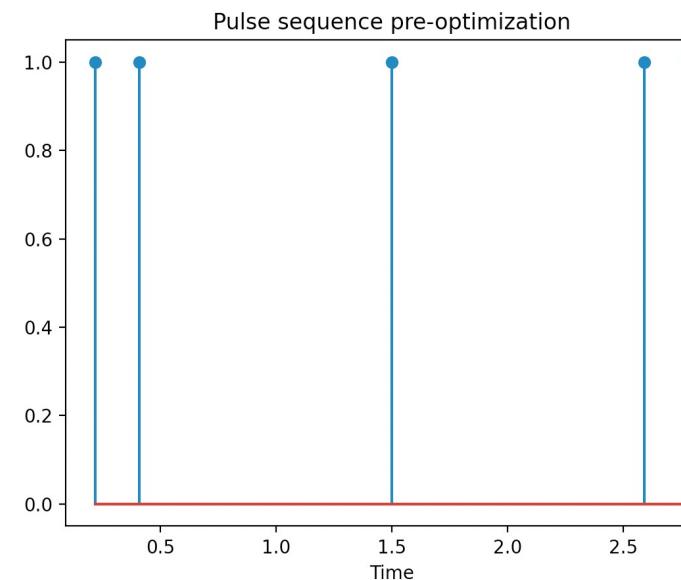
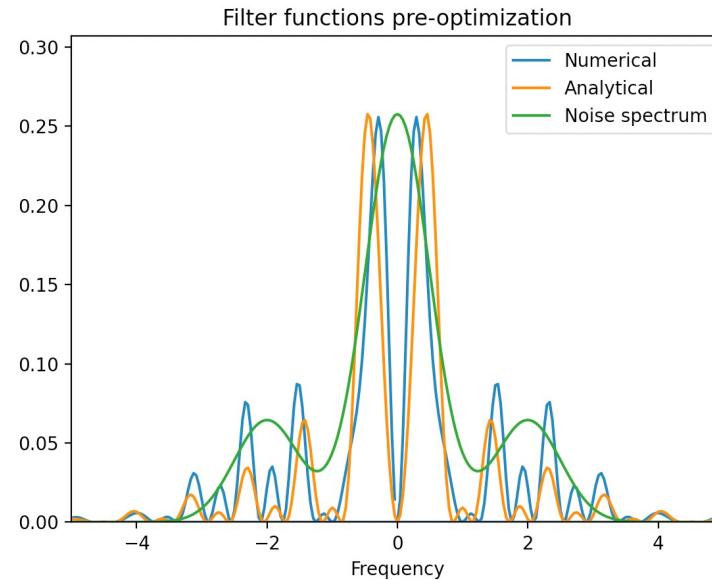


Pulses keep going less than zero. Why?
It looks like it's spreading the pulses out as far as possible, since the largest shift per pulse per iteration is of order 1, so that the pulses can go below zero. Then it approaches a delta function at the origin for chi.
Try smaller pulse shift per step δt .



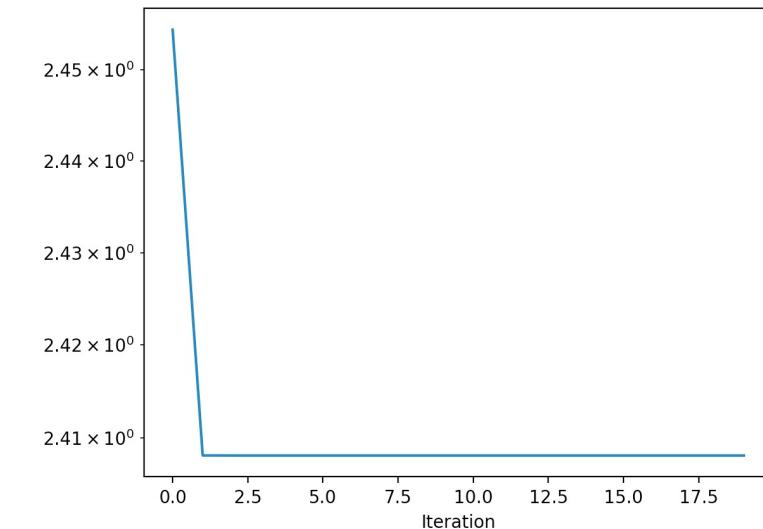
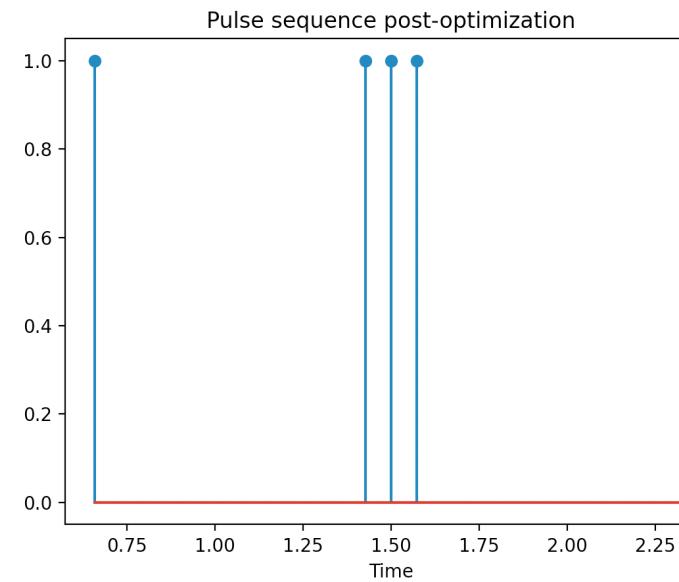
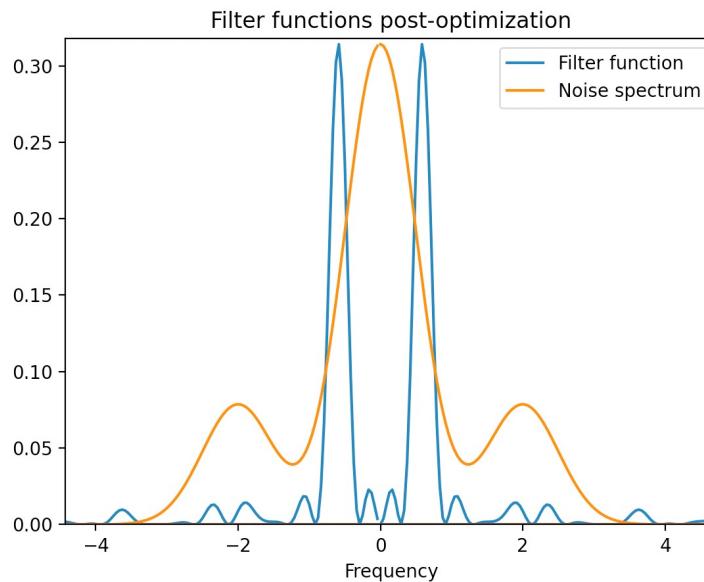
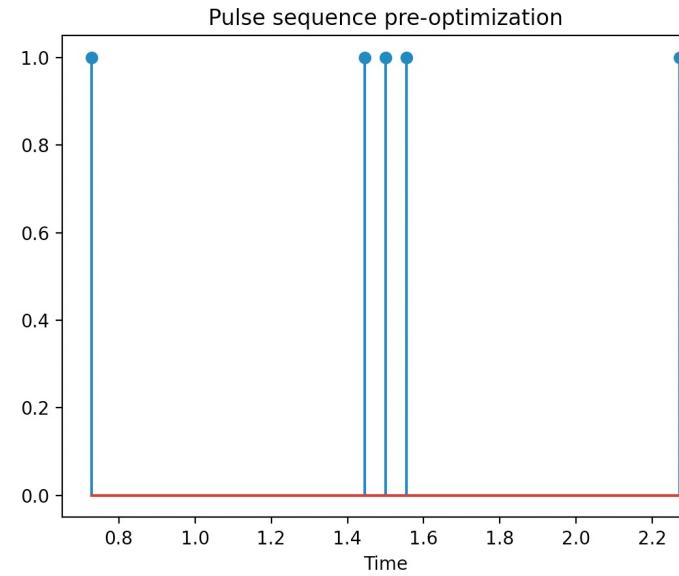
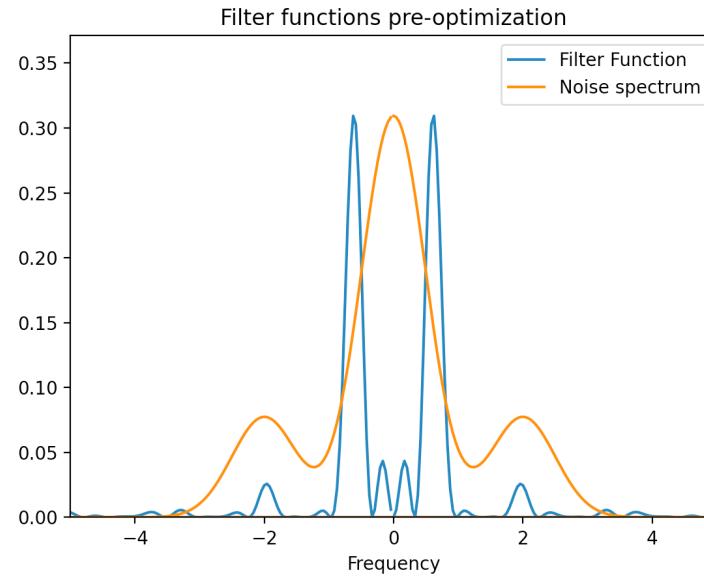
Some trials

$m=20$ possible harmonics, $\Delta t = 5 dt$, 200 iterations, $N=5$ pulses



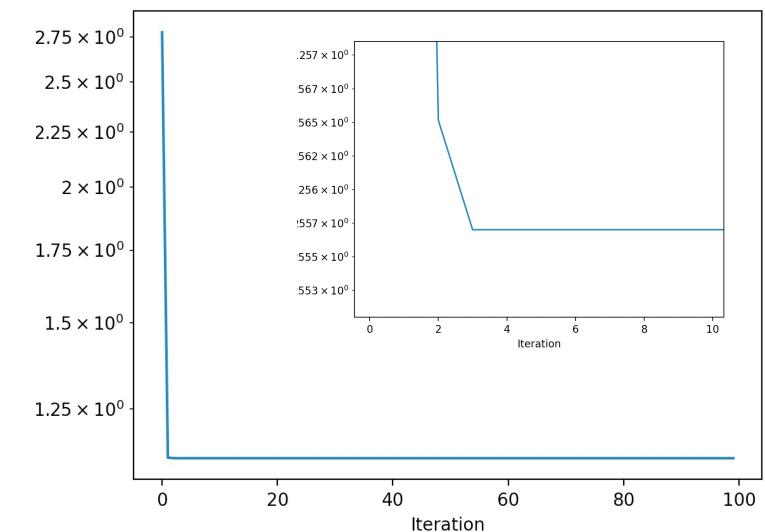
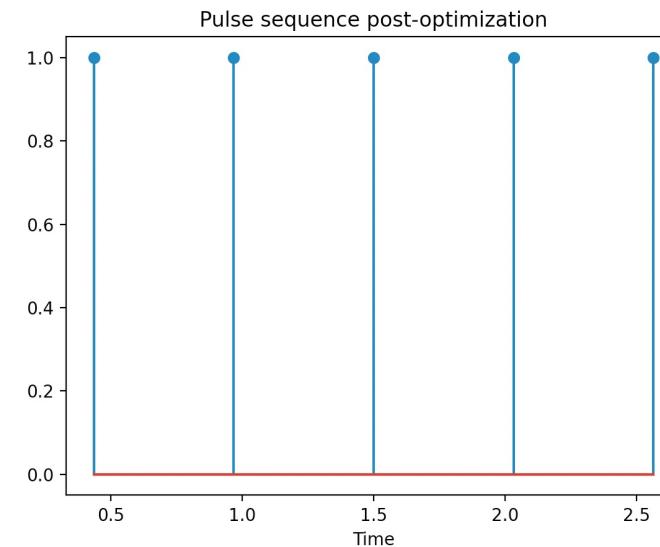
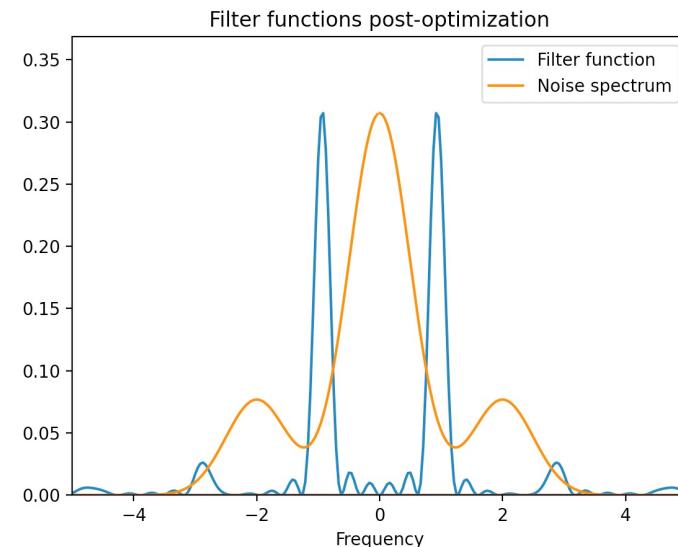
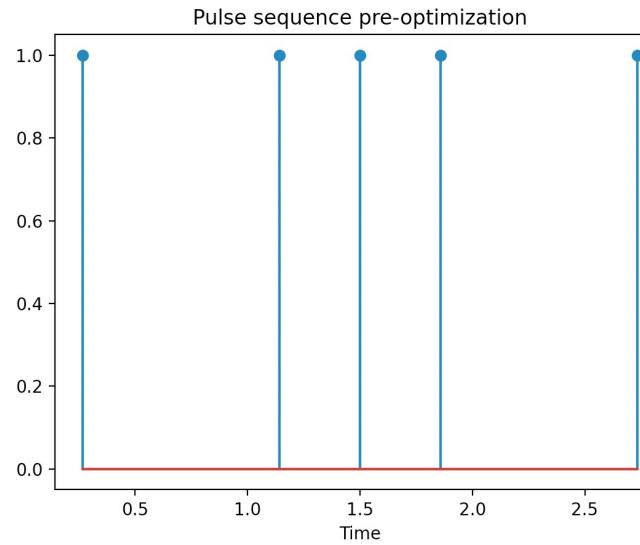
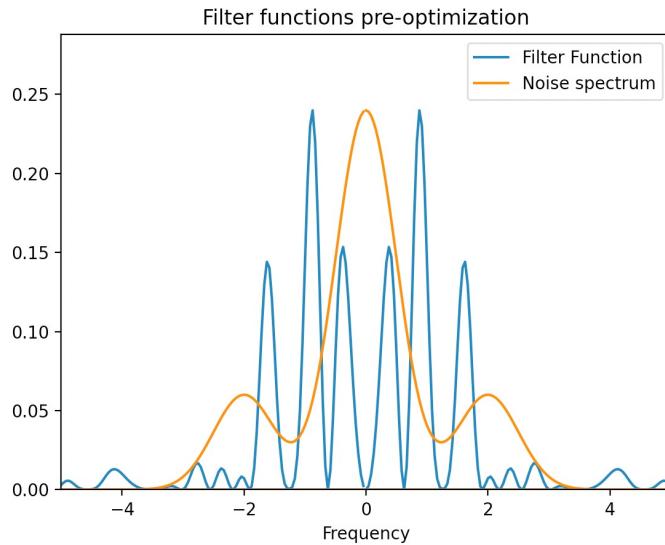
Some trials – cutting out analytical part

$m=20$ possible harmonics – now including the possibility of negative m - , $\Delta t = dt$, 200 iterations, $N=5$ pulses



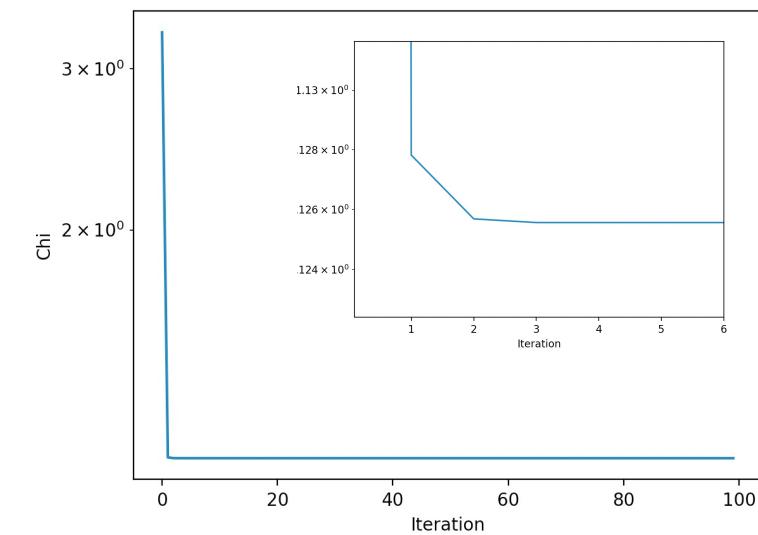
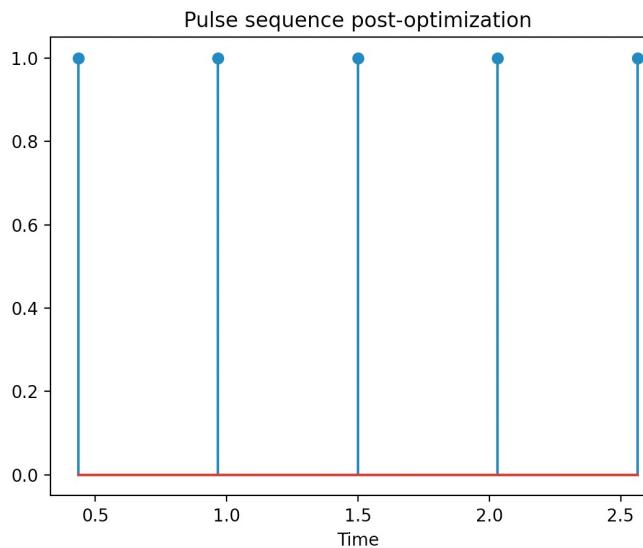
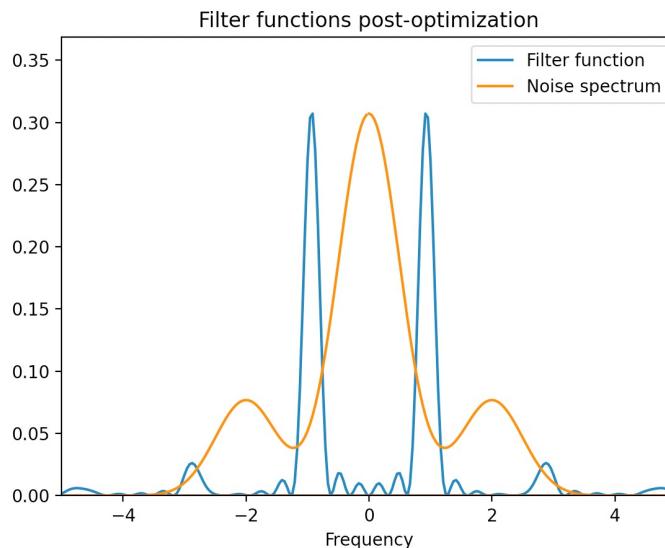
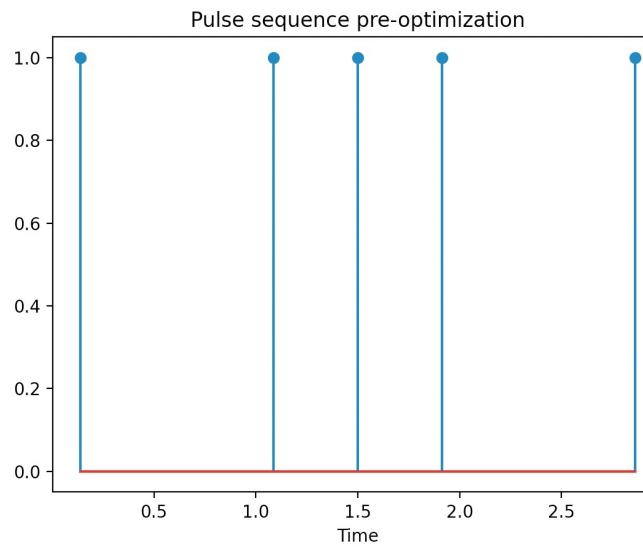
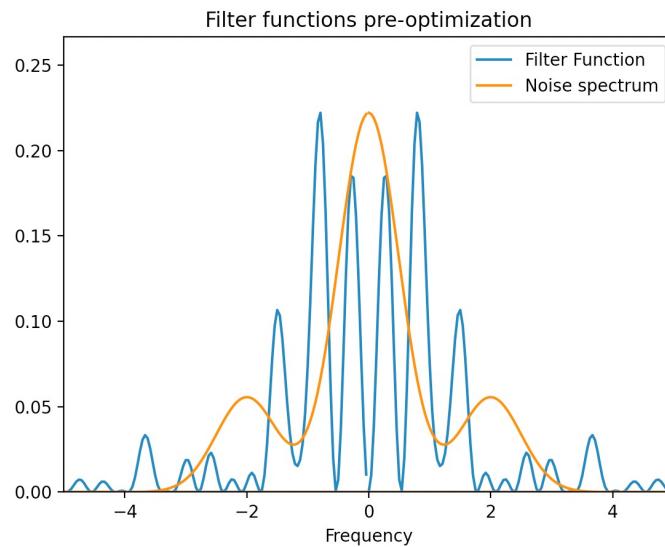
Some trials – cutting out analytical part

$m=100$ possible harmonics – now including the possibility of negative m - , delta t = dt, 100 iterations, N=5 pulses; smaller grid 2×10^3 data points (previously 10^4 data points for the last 2 slides)



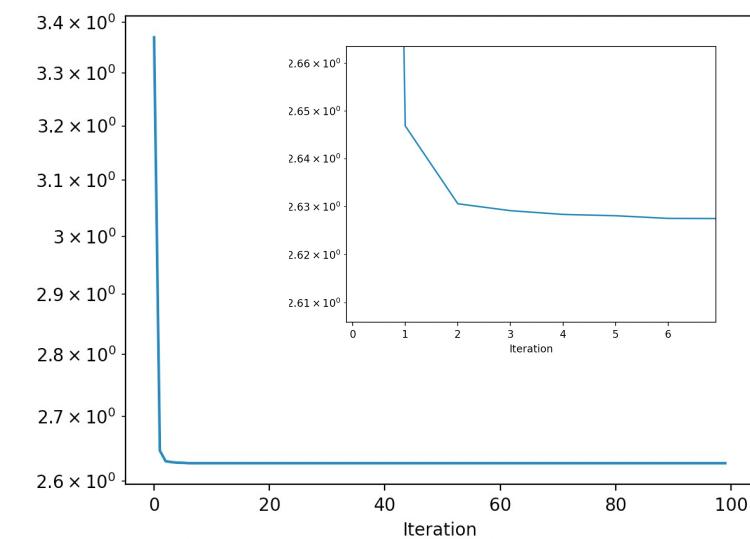
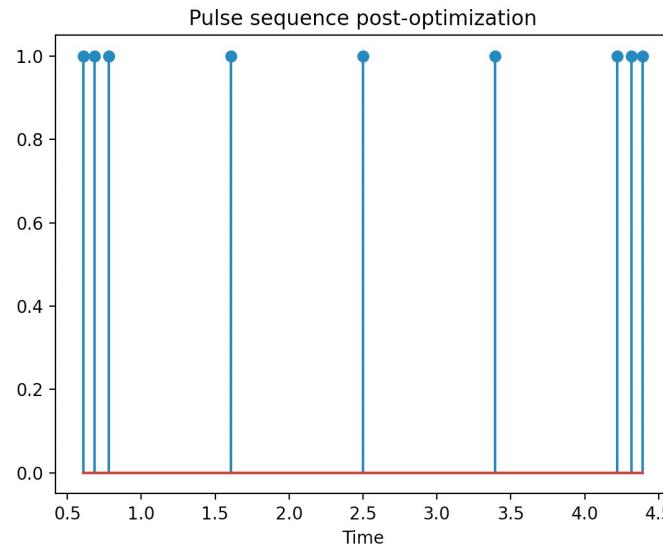
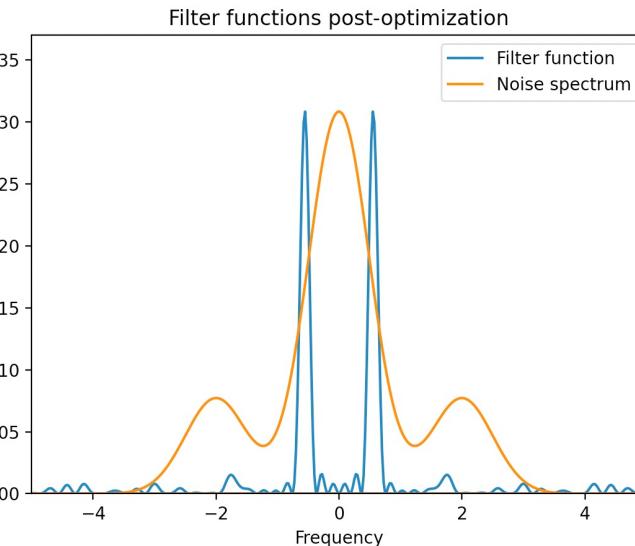
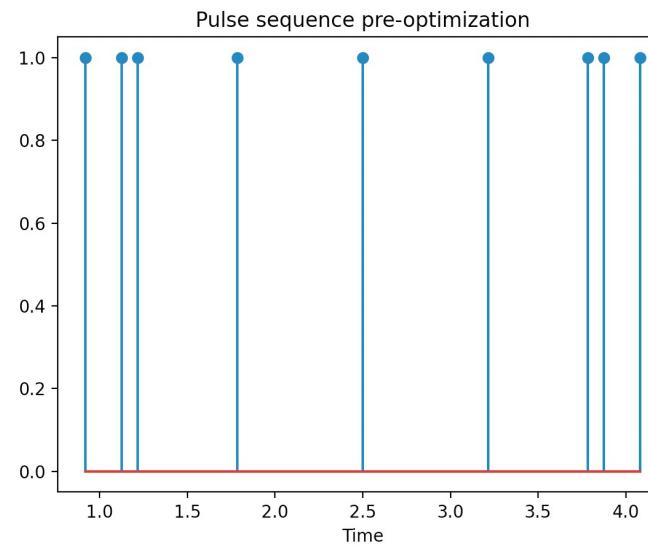
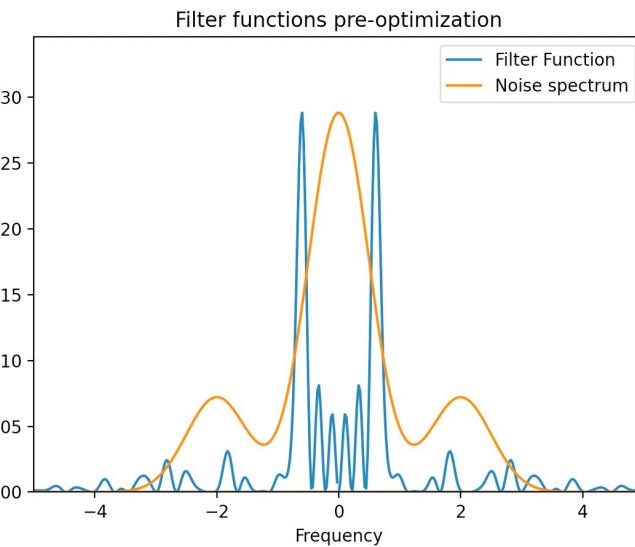
Some trials – cutting out analytical part

$m=100$ possible harmonics – now including the possibility of negative m - , delta $t = dt$, 100 iterations, N=5 pulses; smaller grid 2×10^3 data points (previously 10^4 data points for the last 2 slides)



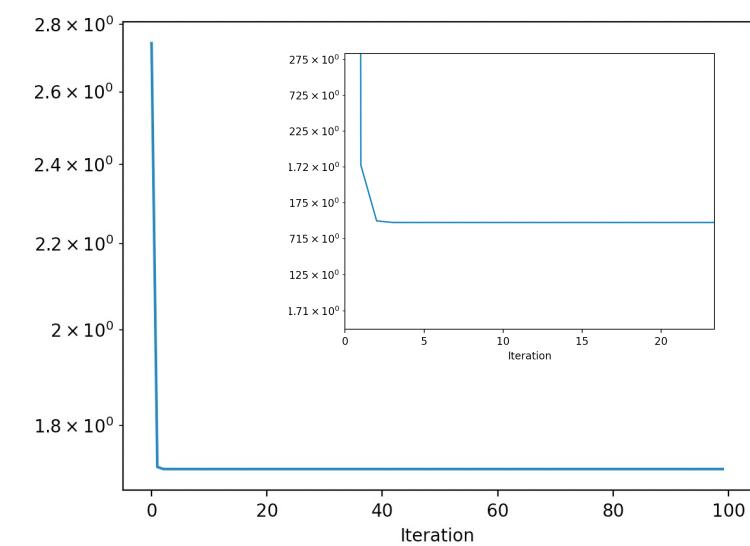
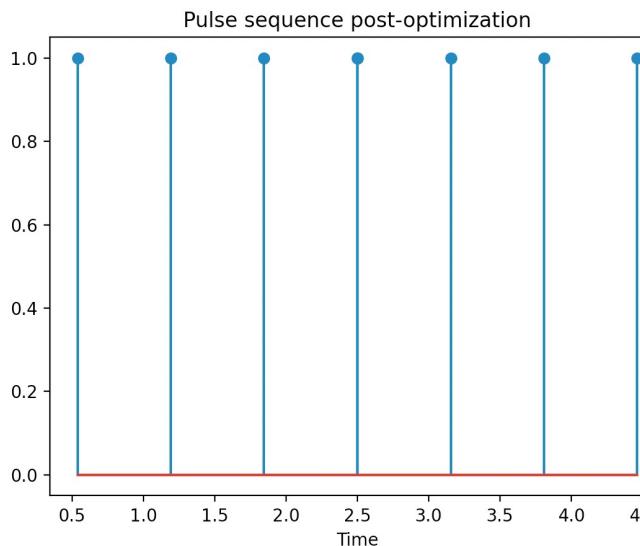
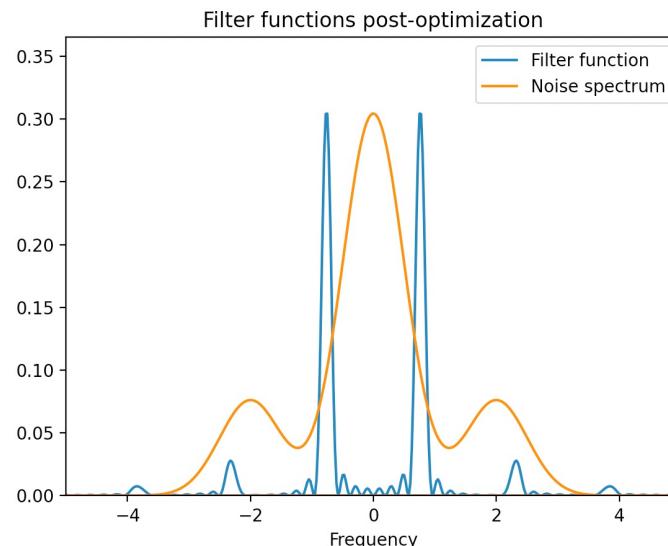
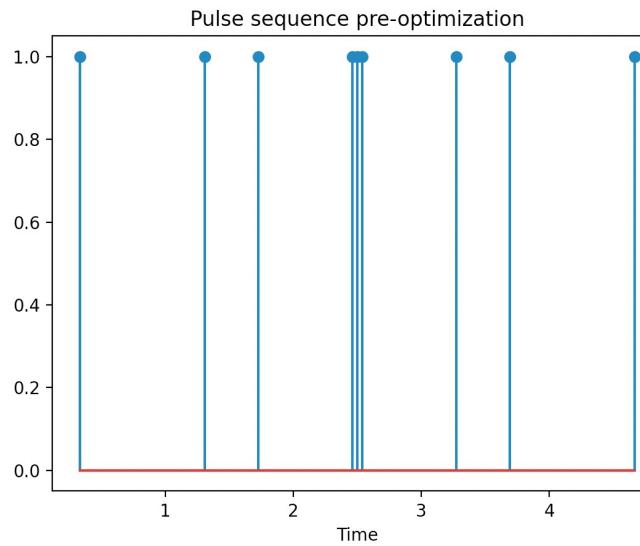
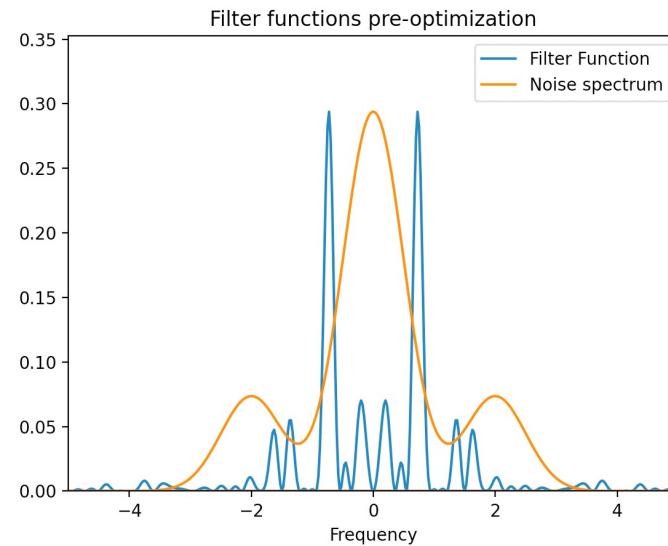
Some trials – cutting out analytical part

$m=100$ possible harmonics – now including the possibility of negative m - , delta $t = dt$, 100 iterations, $N=9$ pulses; smaller grid 2×10^3 data points (previously 10^4 data points for the last 2 slides)



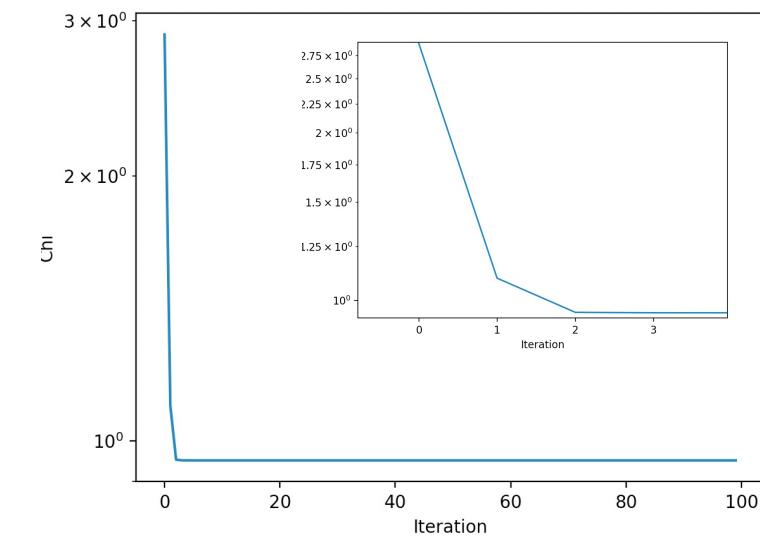
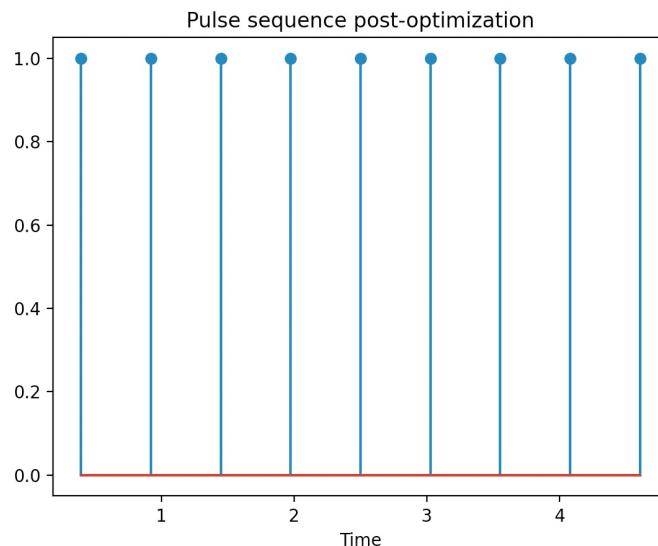
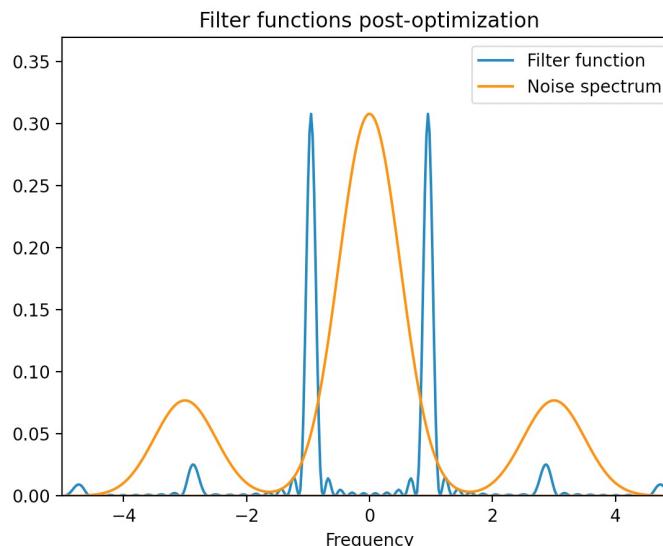
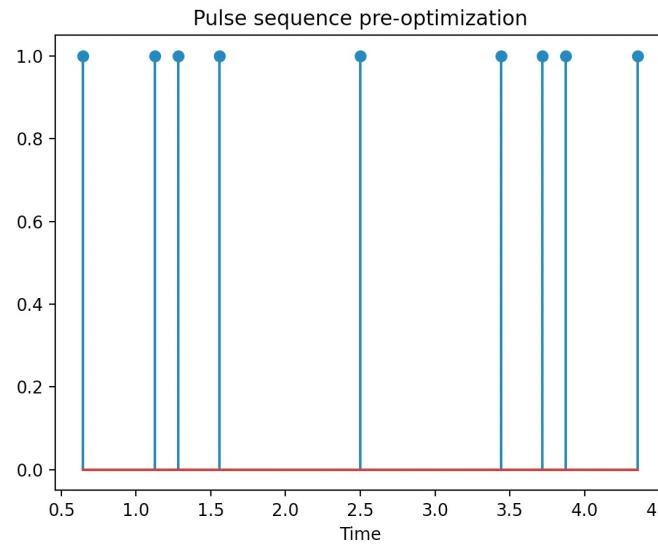
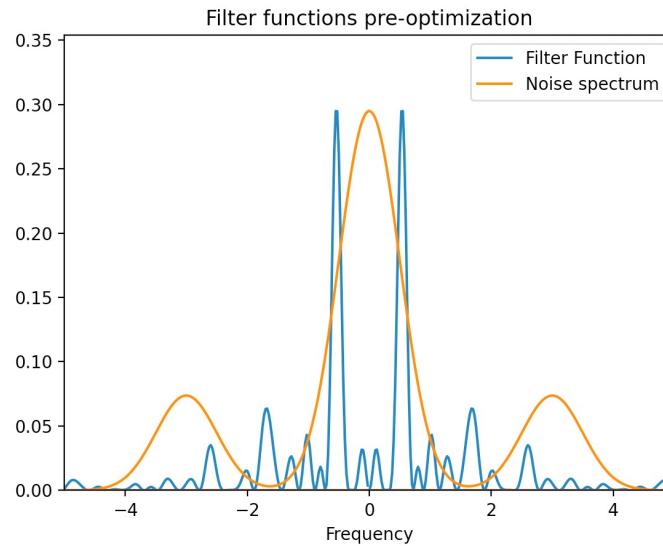
Some trials – cutting out analytical part

$m=100$ possible harmonics – now including the possibility of negative m - , delta t = dt, 100 iterations, N=9 pulses; smaller grid 2×10^3 data points (previously 10^4 data points for the last 2 slides)



Some trials – cutting out analytical part

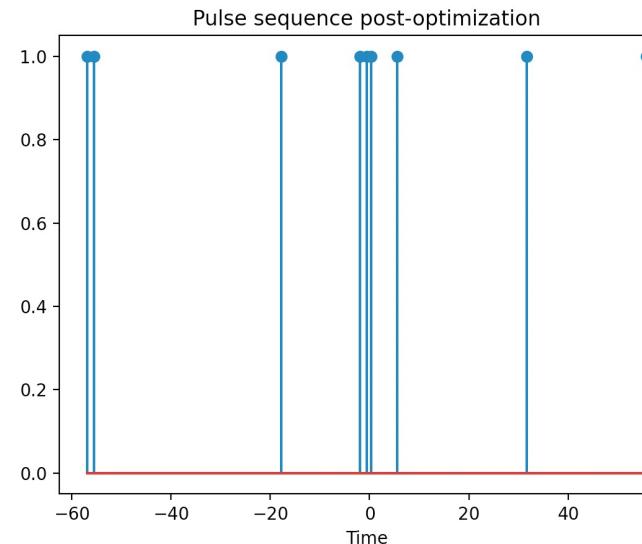
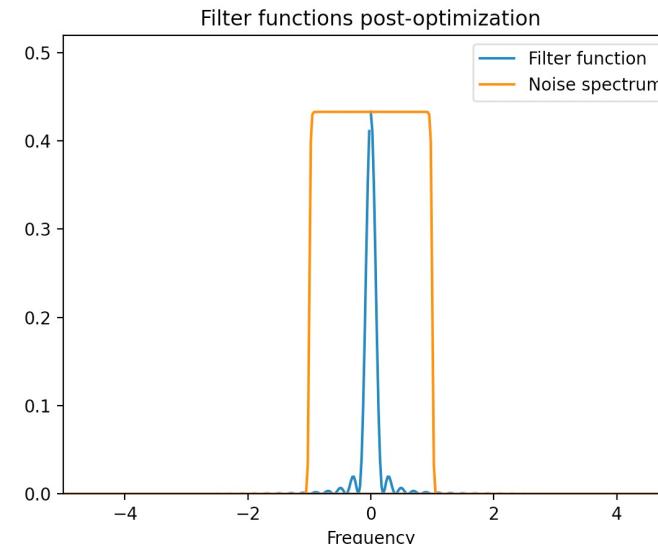
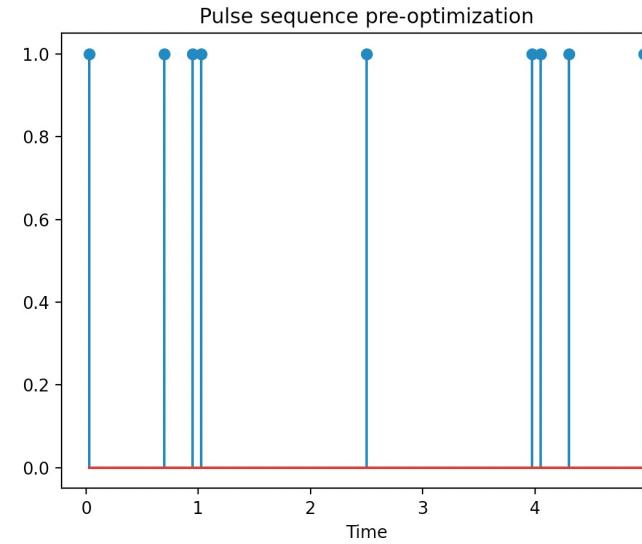
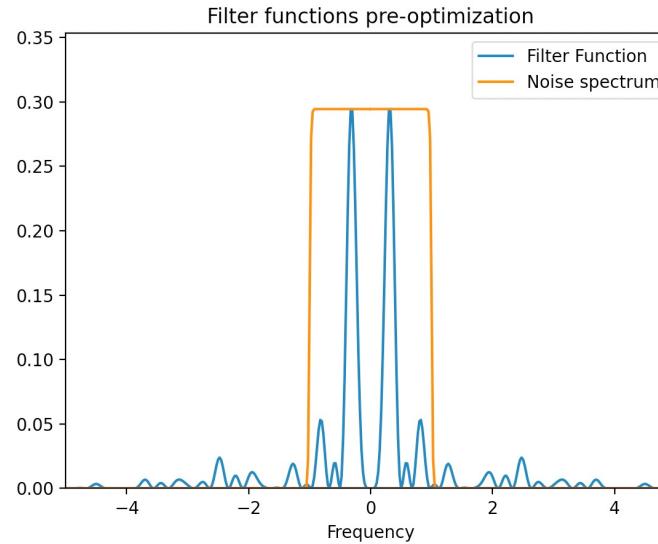
$m=100$ possible harmonics – now including the possibility of negative m - , delta $t = dt$, 100 iterations, $N=9$ pulses; smaller grid 2×10^3 data points (previously 10^4 data points for the last 2 slides)



Some trials – cutting out analytical part

Fermi-Dirac noise; mu = 1, T = 0.01

m=100 possible harmonics – now including the possibility of negative m - , delta t = dt, 100 iterations, N=9 pulses; smaller grid 2×10^3 data points (previously 10^4 data points for the last 2 slides)

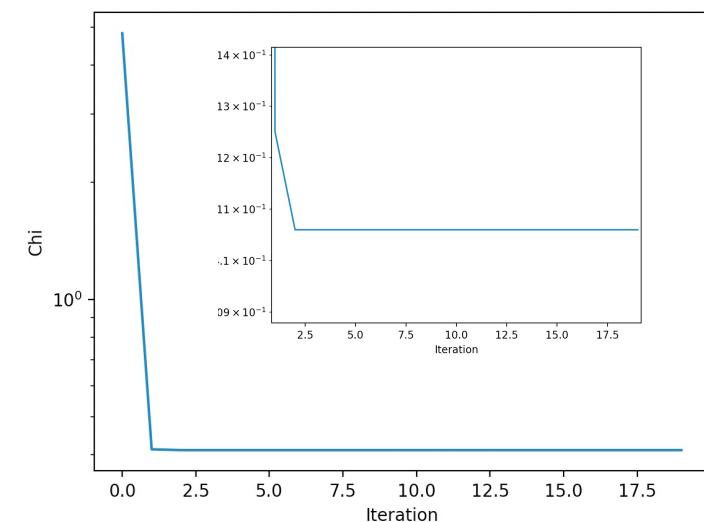
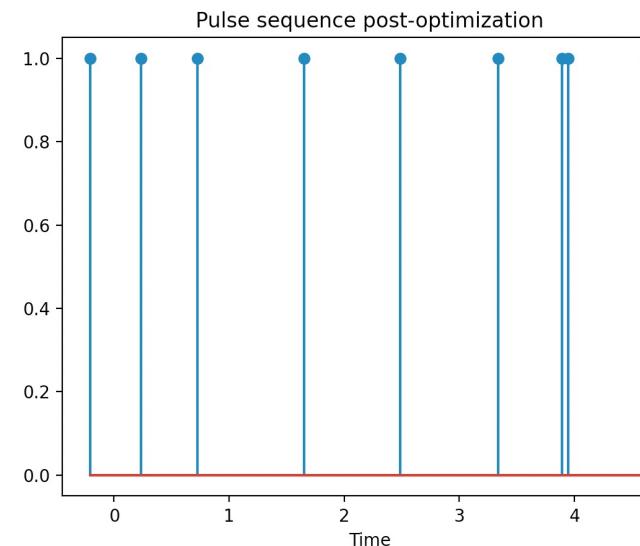
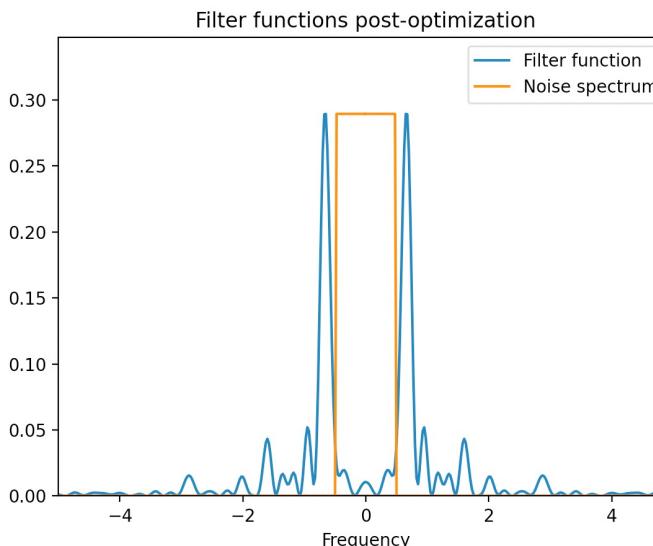
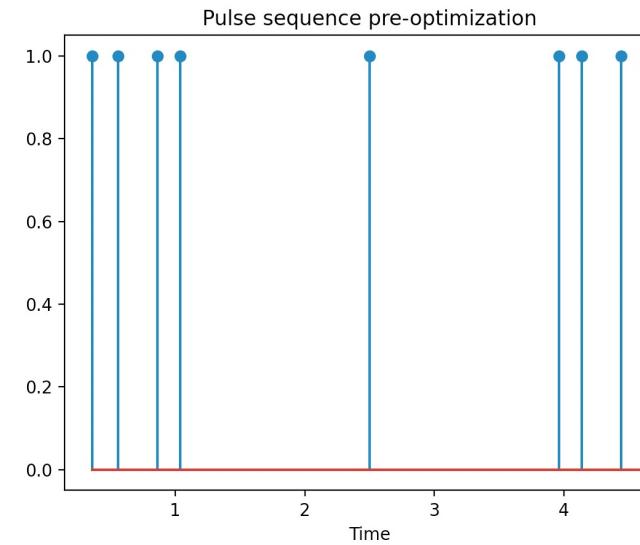
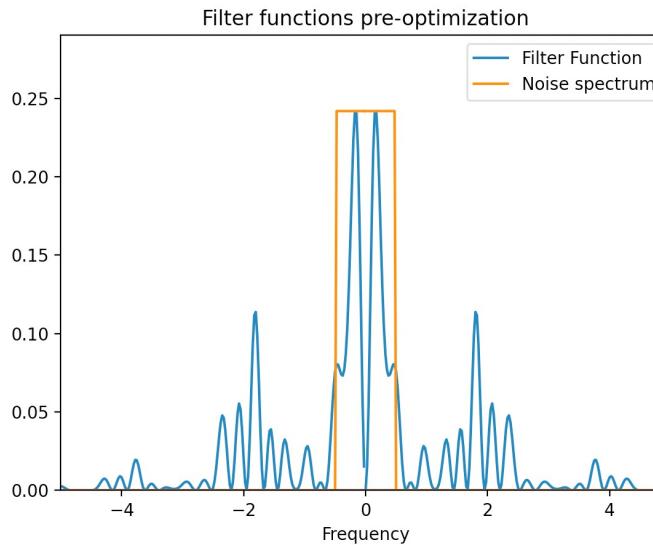


Didn't work

Some trials – cutting out analytical part

Fermi-Dirac noise; mu = 1, T = 0.01

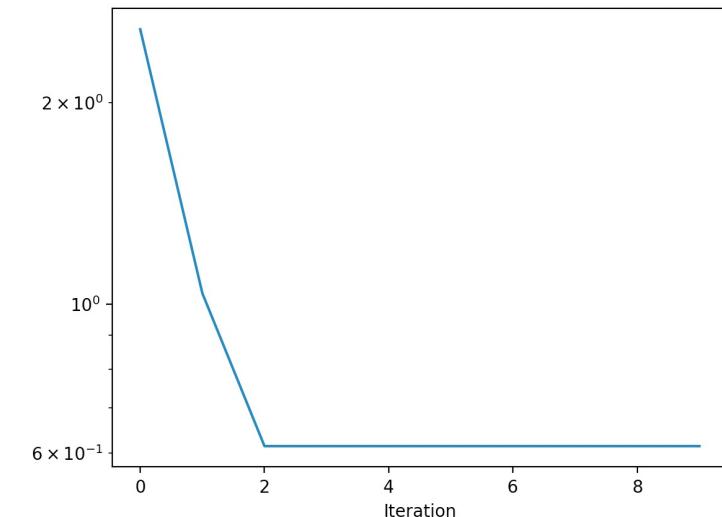
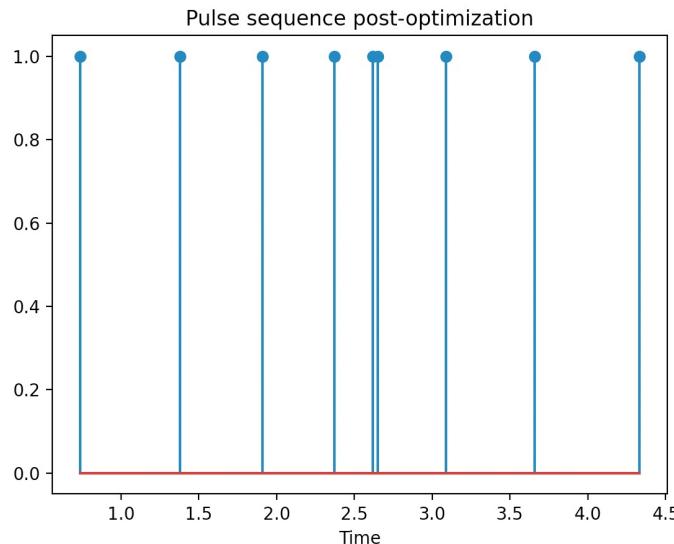
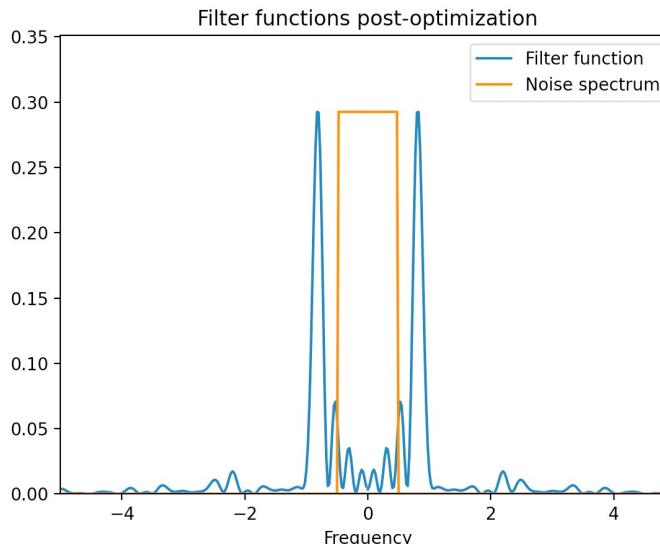
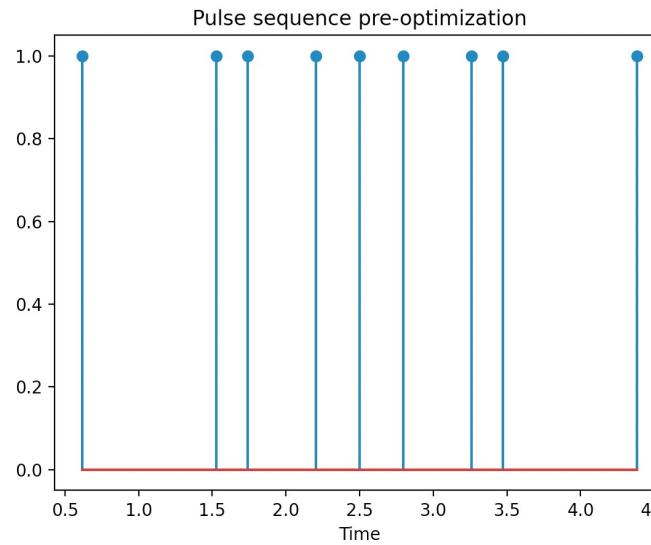
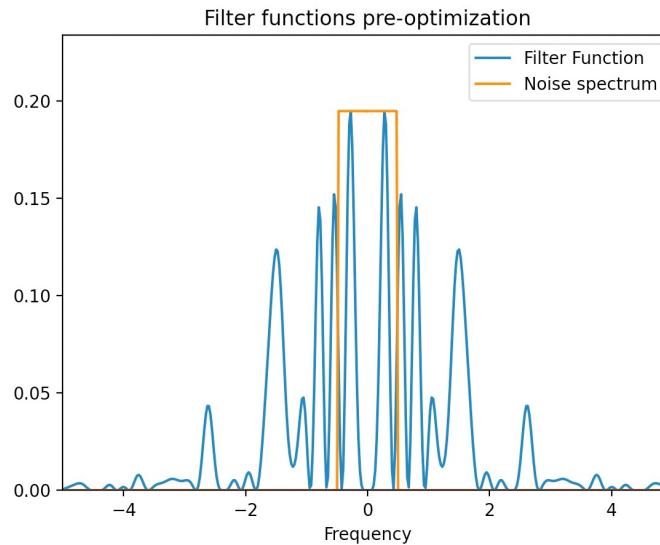
m=100 possible harmonics – now including the possibility of negative m - , delta t = dt, 20 iterations, N=9 pulses; smaller grid 2×10^4 data points



Some trials – cutting out analytical part

Fermi-Dirac noise; mu = 1, T = 0.01

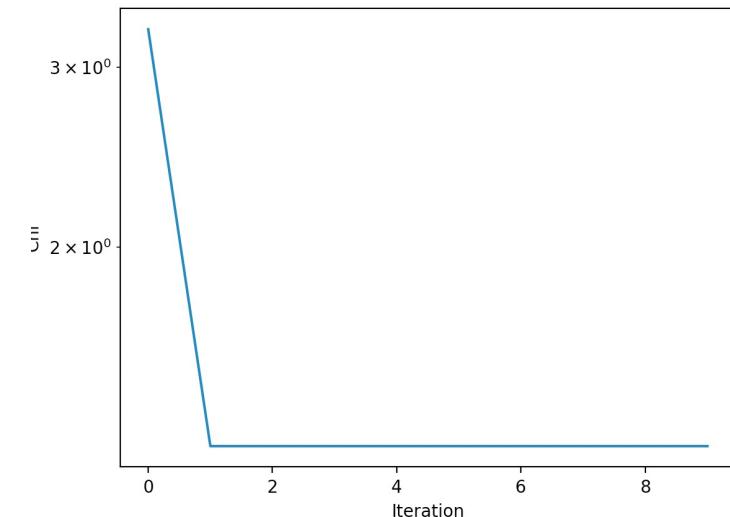
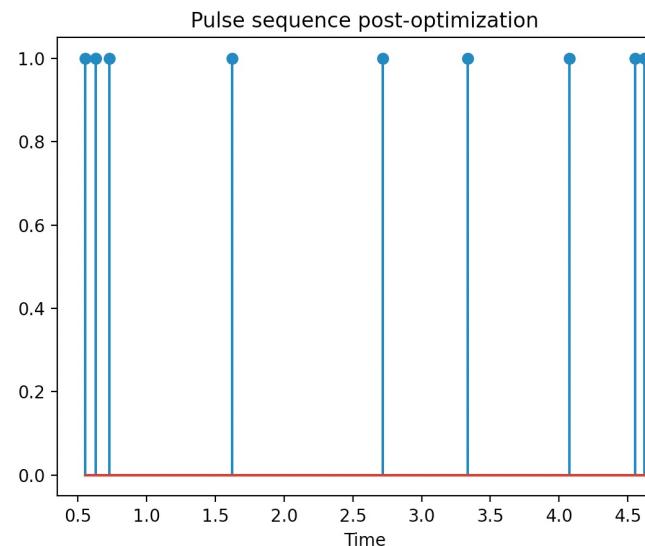
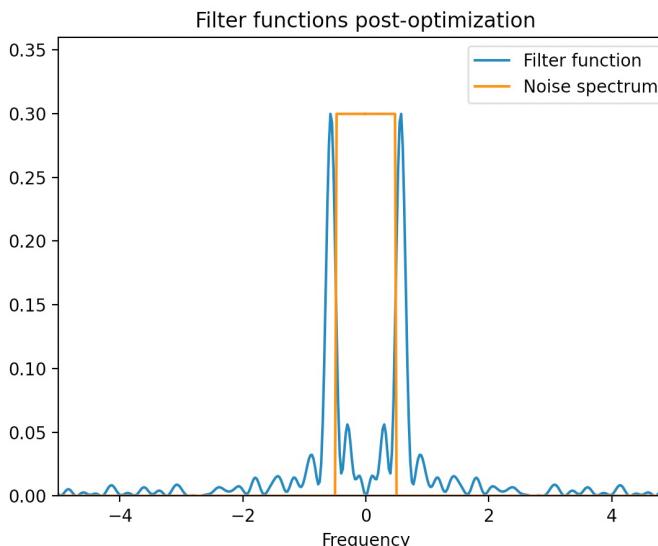
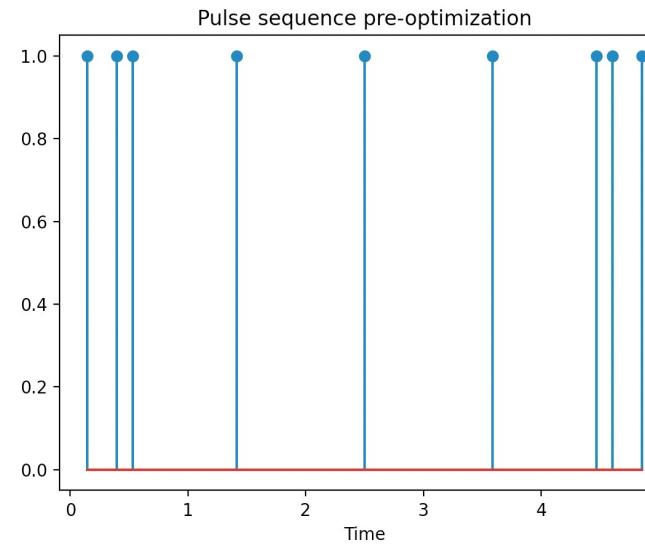
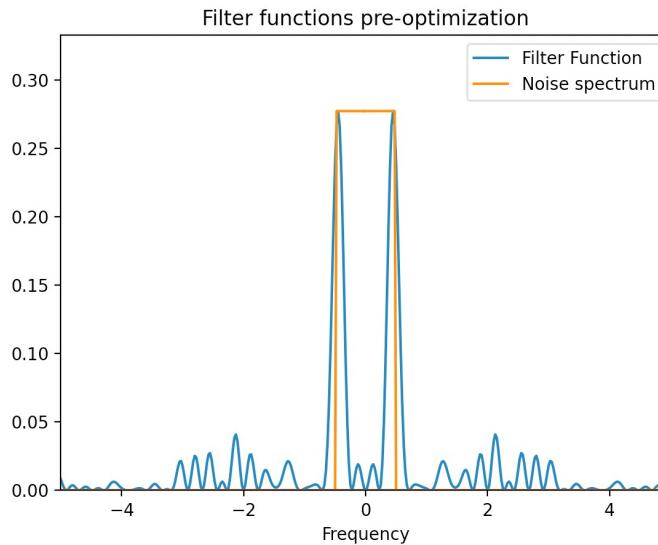
m=100 possible harmonics – now including the possibility of negative m - , delta t = dt, 20 iterations, N=9 pulses; smaller grid 2×10^4 data points



Some trials – cutting out analytical part

Fermi-Dirac noise; mu = 1, T = 0.01

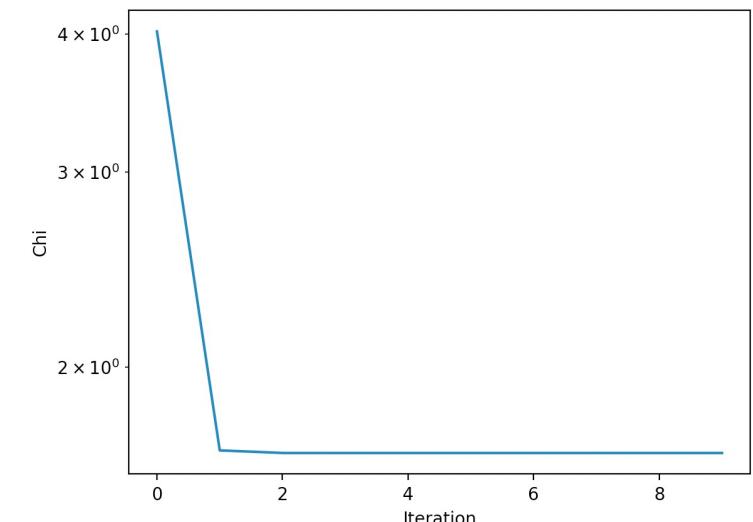
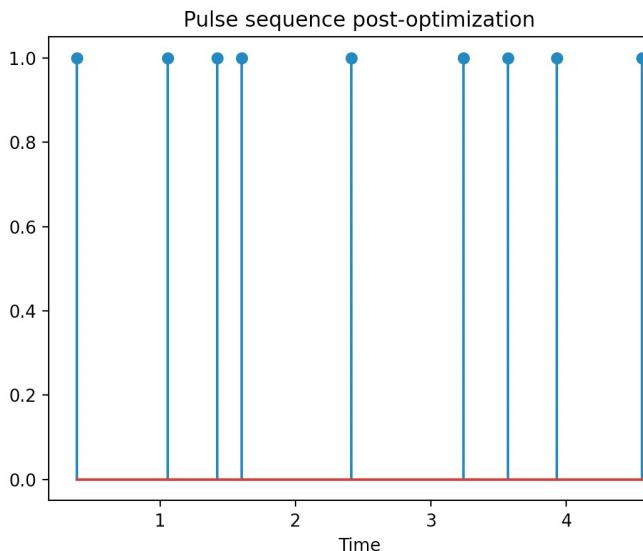
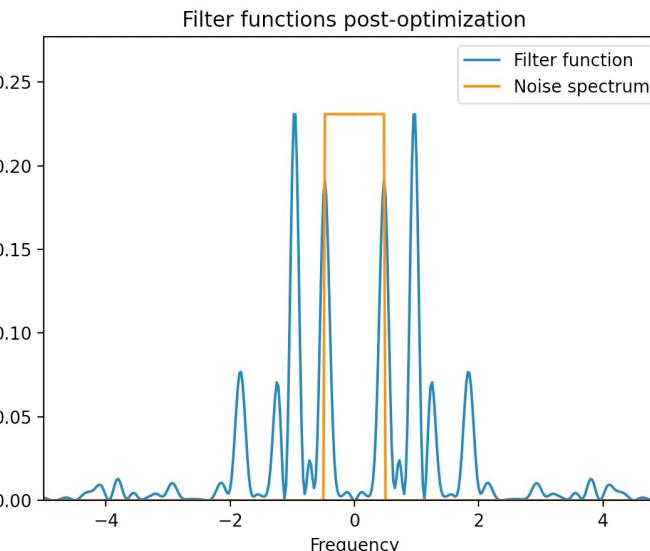
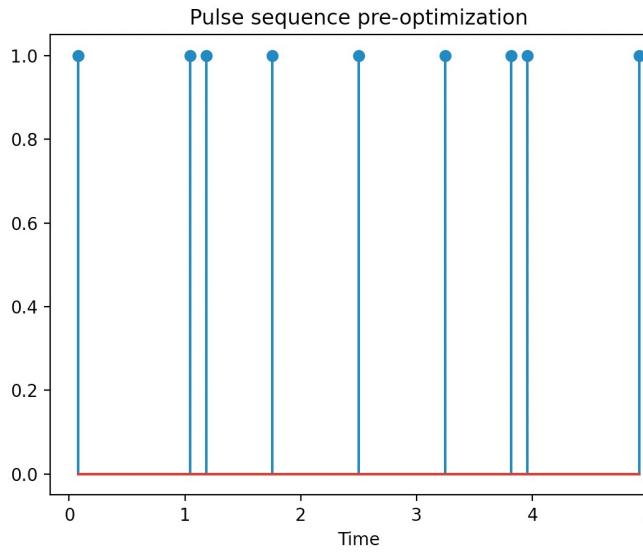
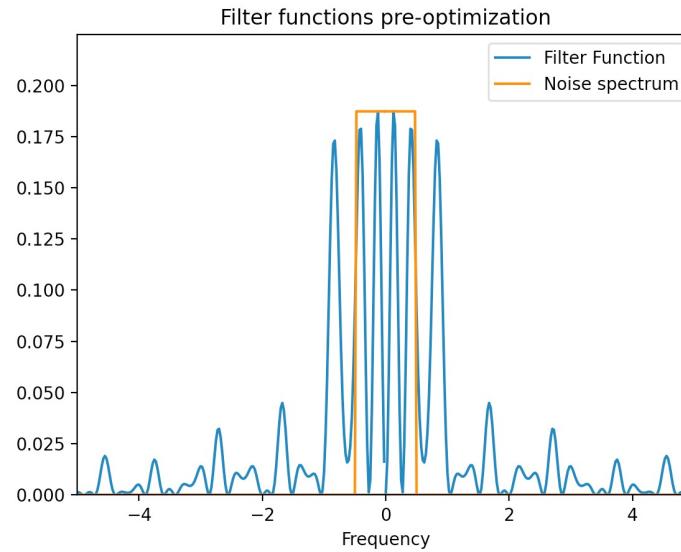
m=100 possible harmonics – now including the possibility of negative m - , delta t = dt, 20 iterations, N=9 pulses; smaller grid 2×10^4 data points



Some trials – cutting out analytical part

Step function noise $T \rightarrow 0$, $\mu = 0.5$

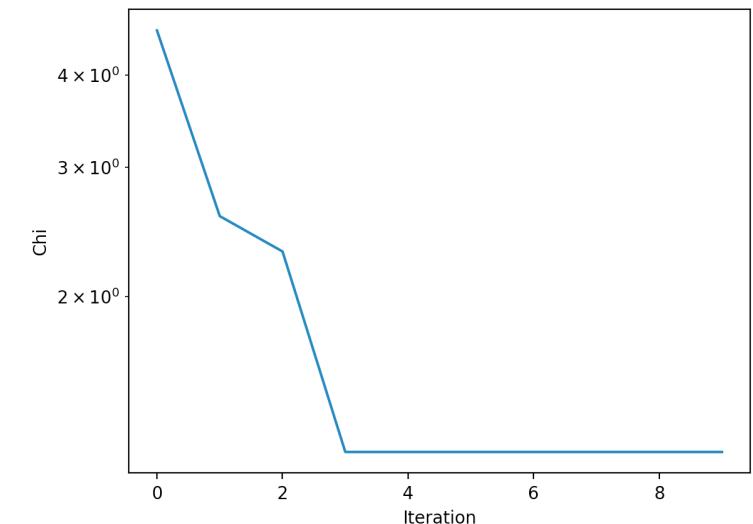
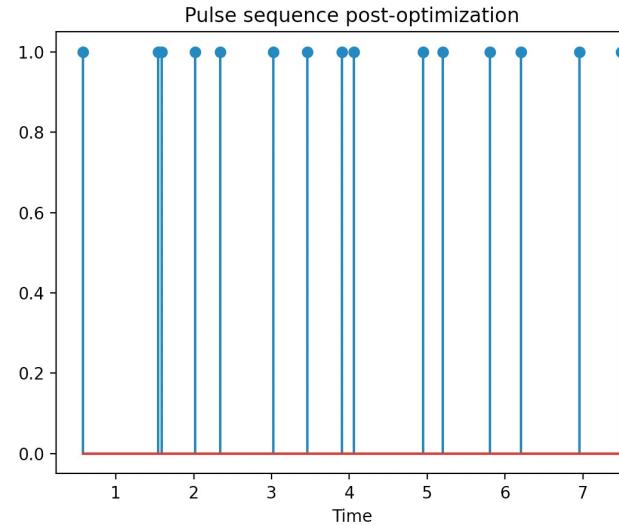
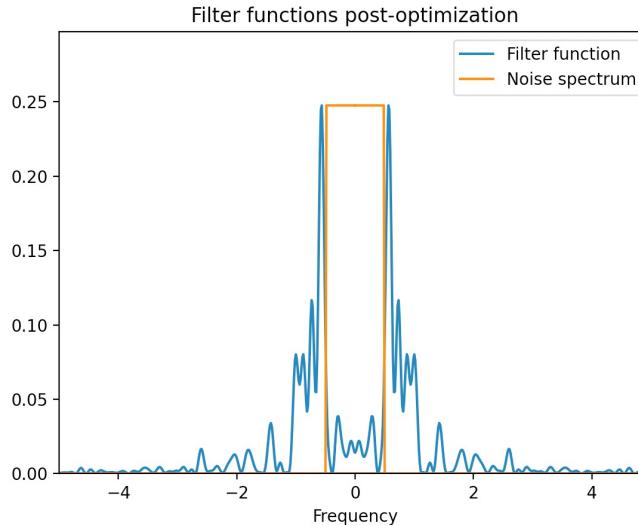
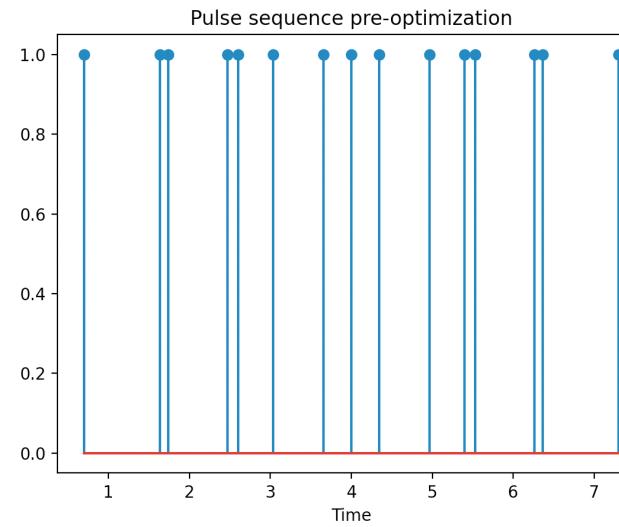
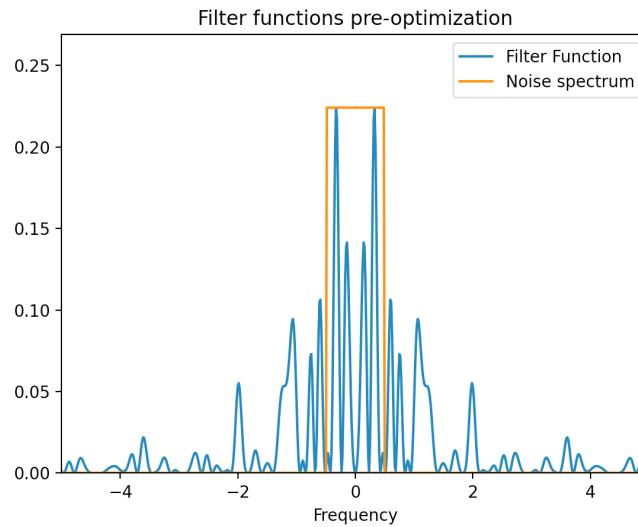
$m=100$ possible harmonics – now including the possibility of negative m - , $\Delta t = dt$, 20 iterations, $N=9$ pulses; smaller grid 2×10^4 data points



Some trials – cutting out analytical part

Step function noise $T \rightarrow 0$, $\mu = 0.5$

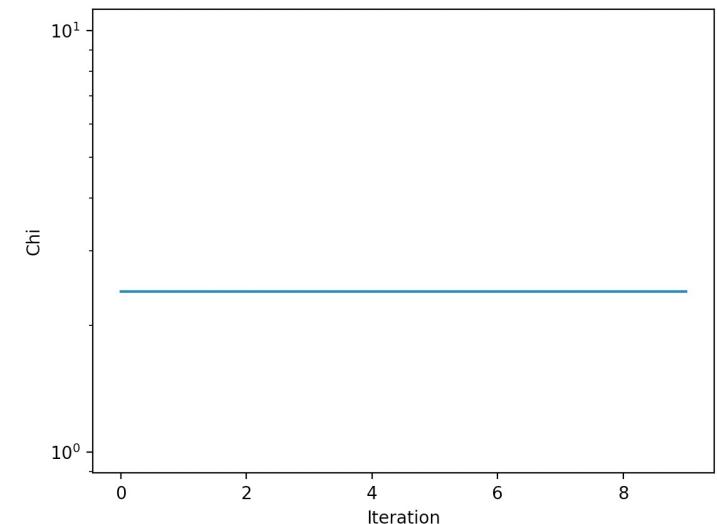
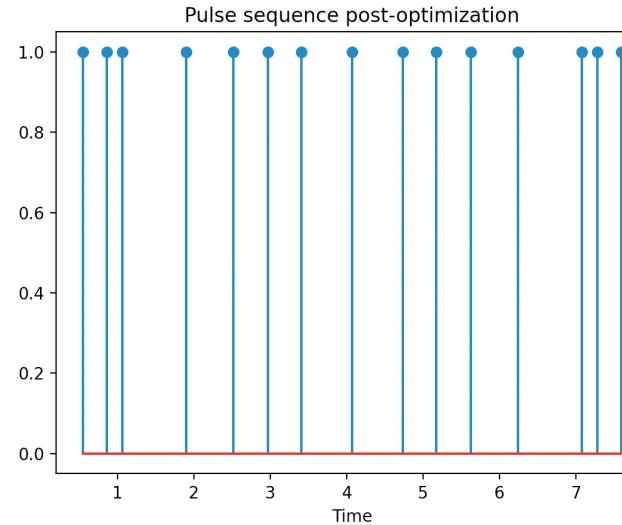
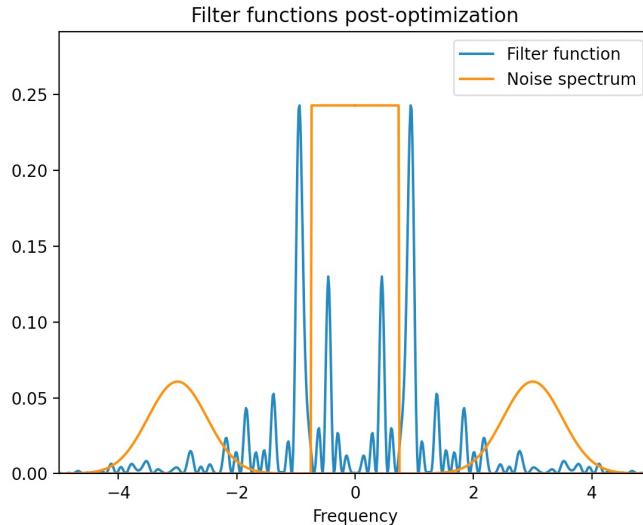
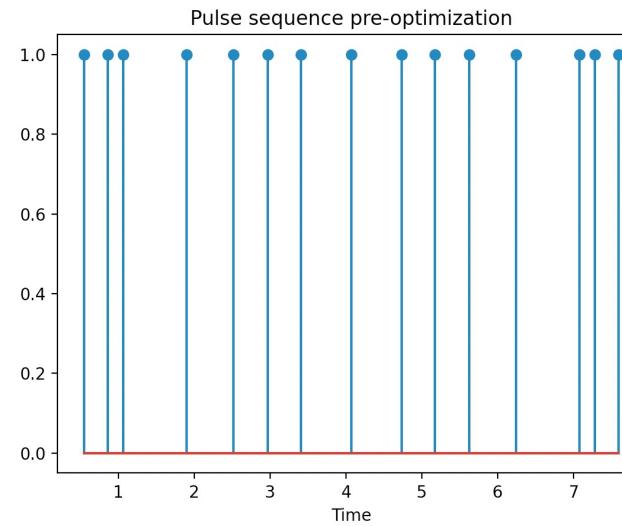
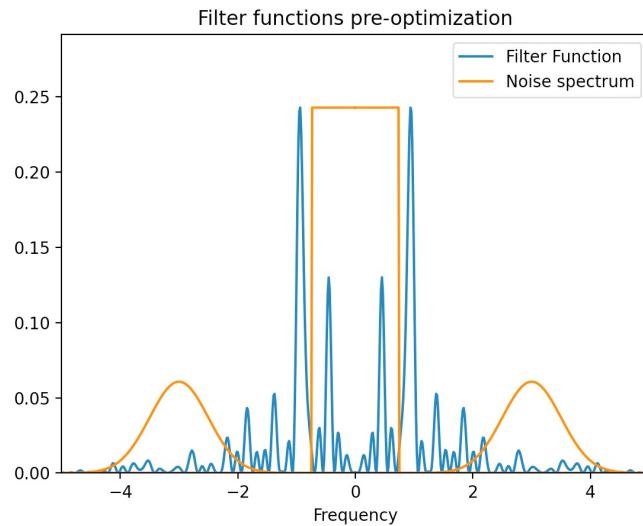
$m=100$ possible harmonics – now including the possibility of negative m - , $\Delta t = dt$, 20 iterations, $N=15$ pulses; smaller grid 2×10^4 data points



Some trials – cutting out analytical part

Step function noise $T \rightarrow 0$, $\mu = 0.5$, plus some “side bands”

$m=100$ possible harmonics – now including the possibility of negative m - , $\Delta t = dt$, 20 iterations, $N=15$ pulses; smaller grid 2×10^5 data points

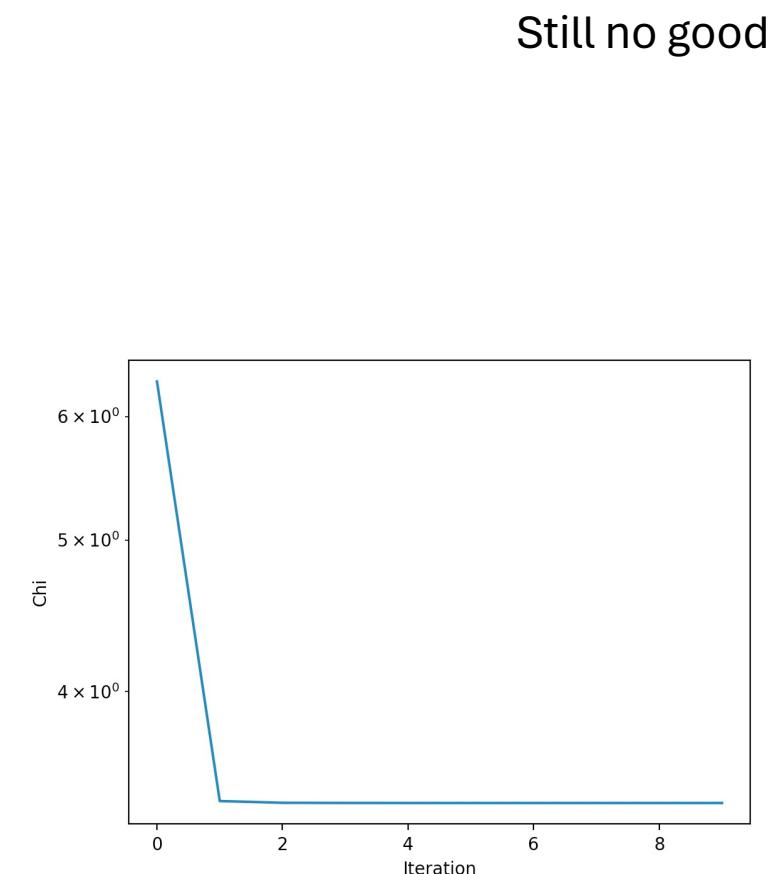
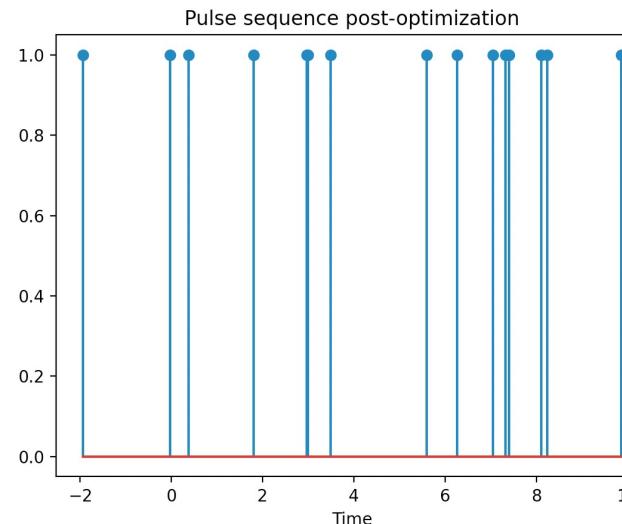
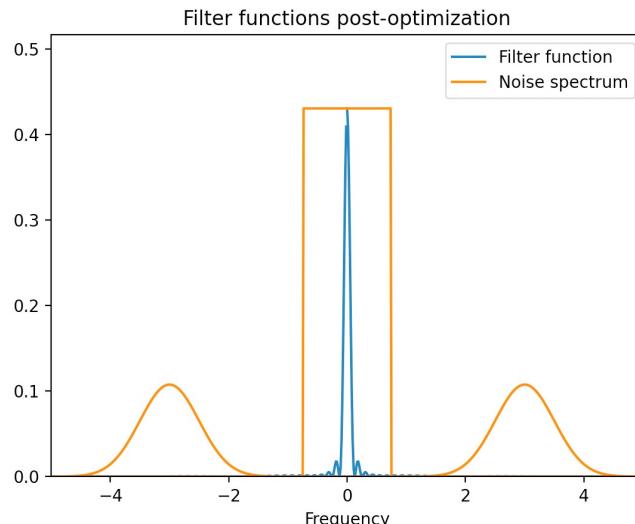
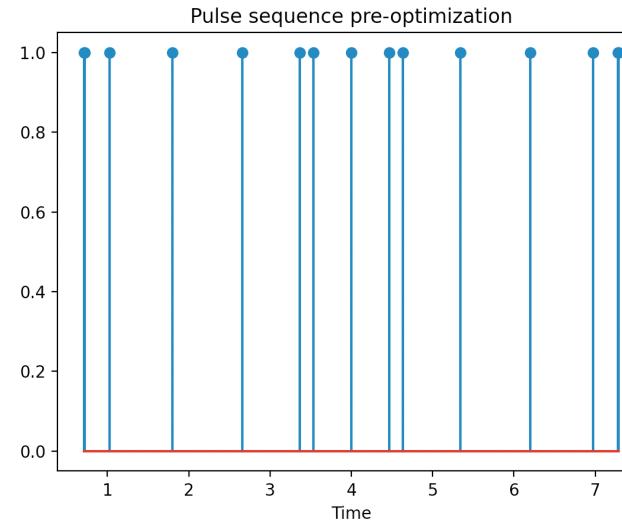
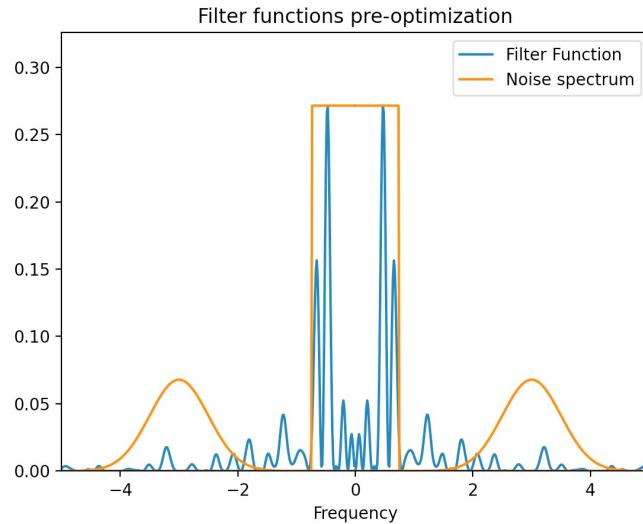


Some trials – cutting out analytical part

Step function noise $T \rightarrow 0$, $\mu = 0.5$, plus some “side bands”

$m=100$ possible harmonics – now including the possibility of negative m - , $\Delta t = dt$, 20 iterations, $N=15$ pulses; smaller grid 2×10^5 data points

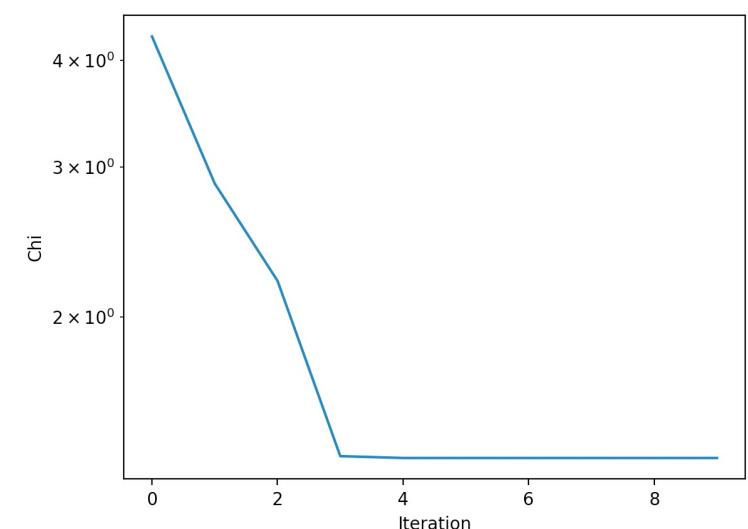
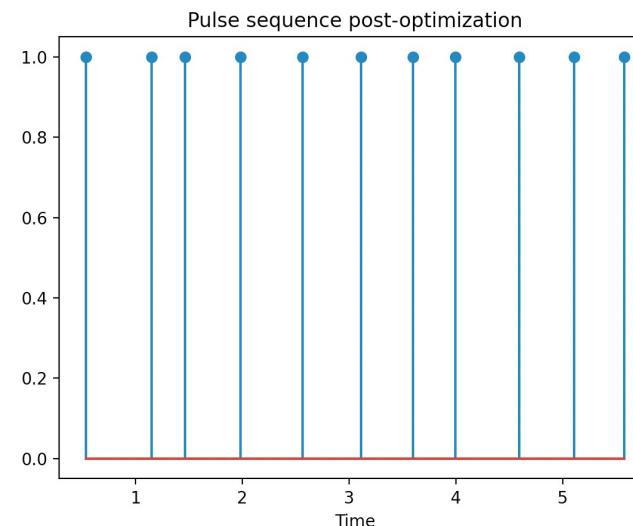
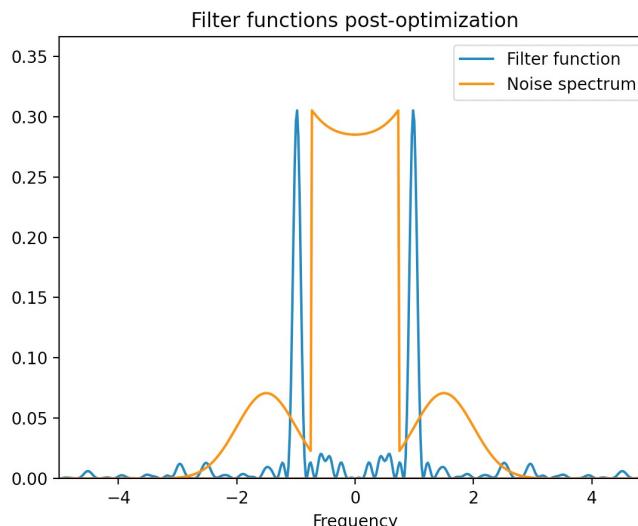
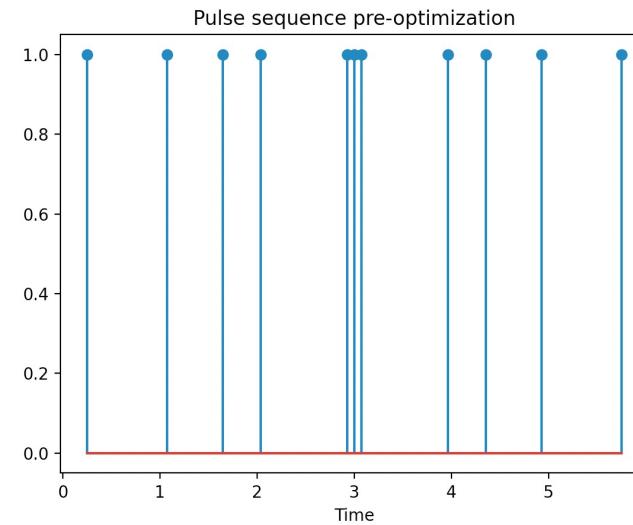
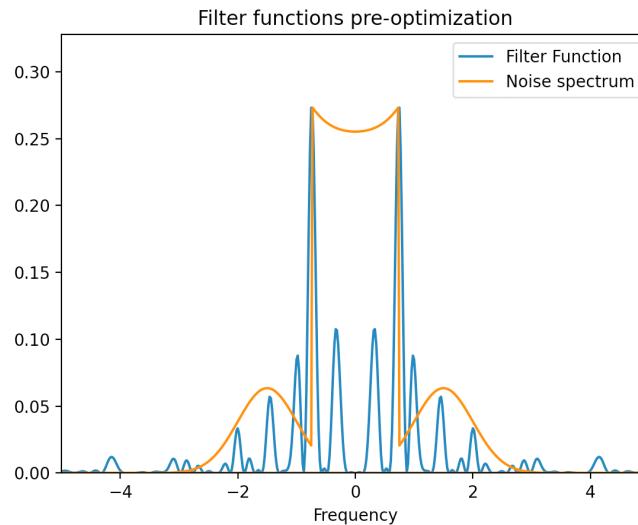
The pulses keep going negative – try adding in more data points to make the time adjustments smaller.



Some trials – cutting out analytical part

Step function noise $T \rightarrow 0$, mu = 0.5, plus some “side bands”

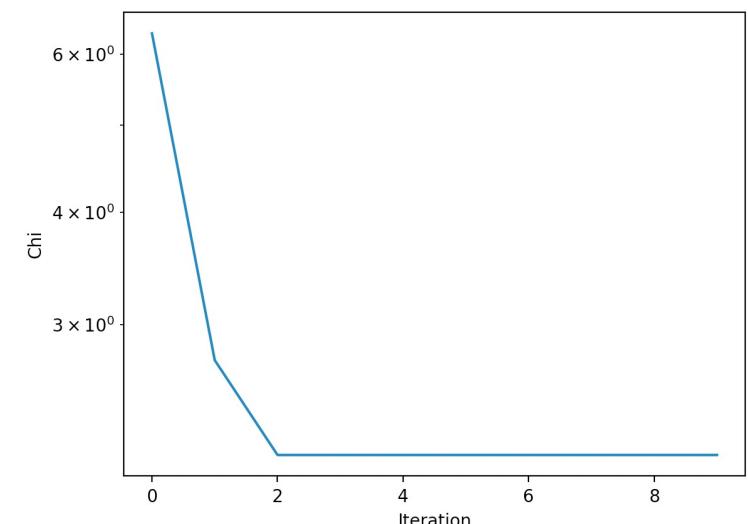
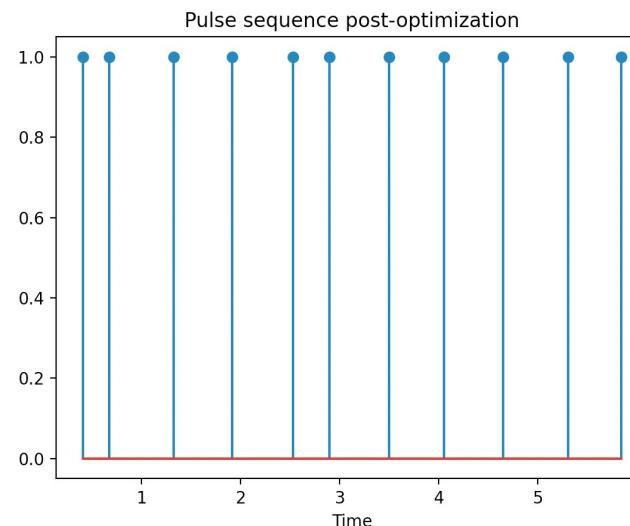
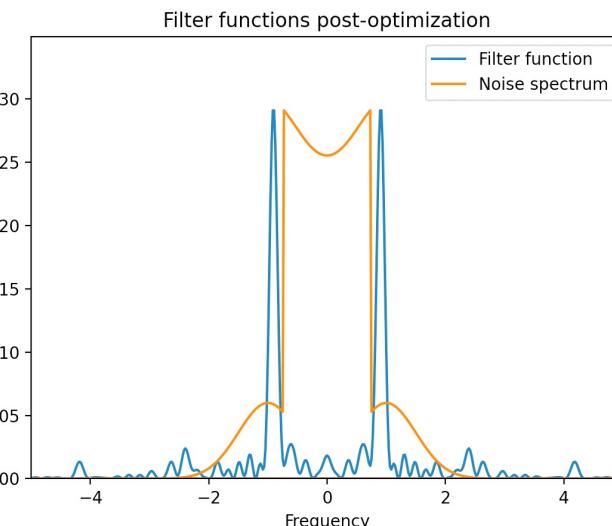
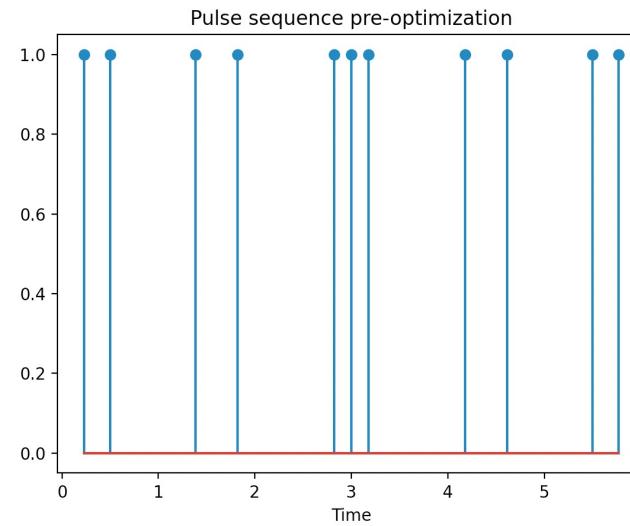
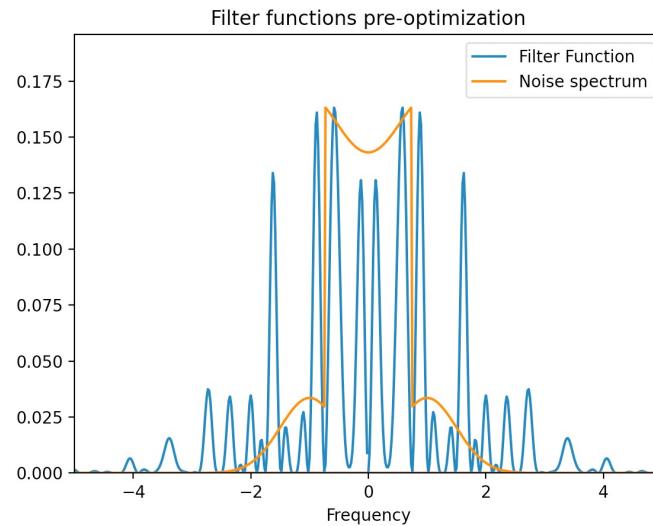
m=100 possible harmonics – now including the possibility of negative m - , delta t = dt, 20 iterations, N=11 pulses; grid 2×10^5 data points



Some trials – cutting out analytical part

Step function noise $T \rightarrow 0$, $\mu = 0.5$, plus some “side bands”

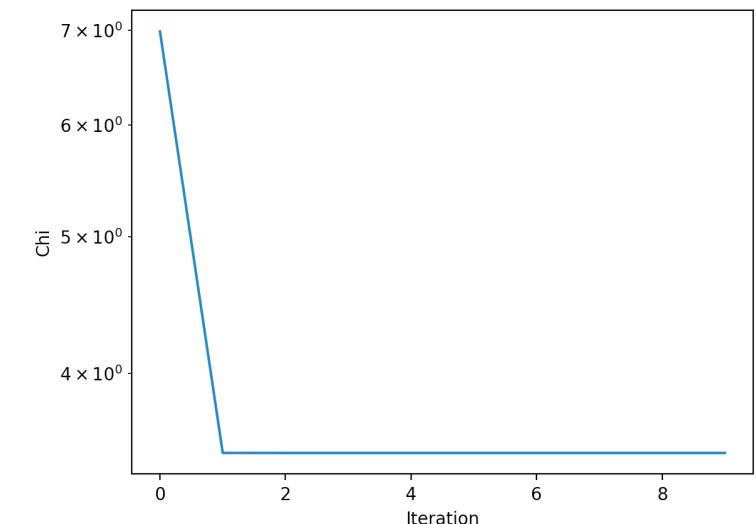
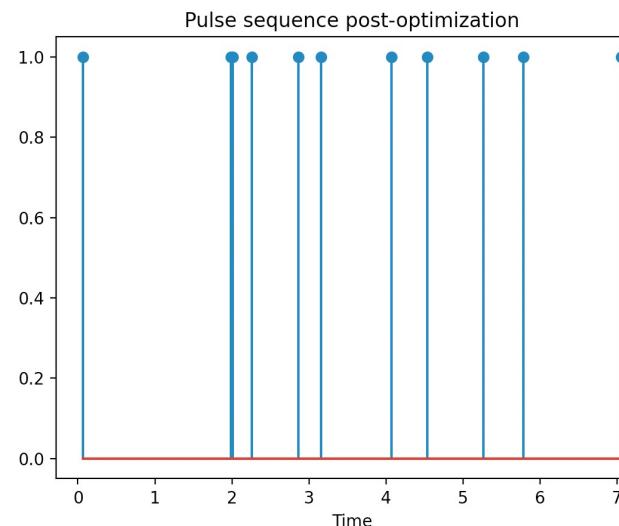
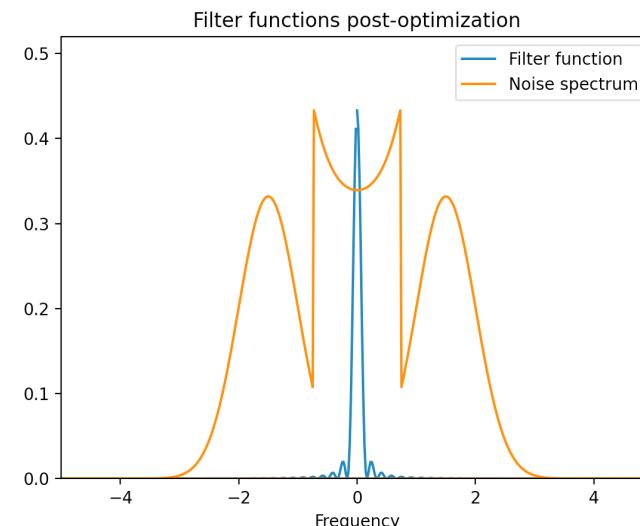
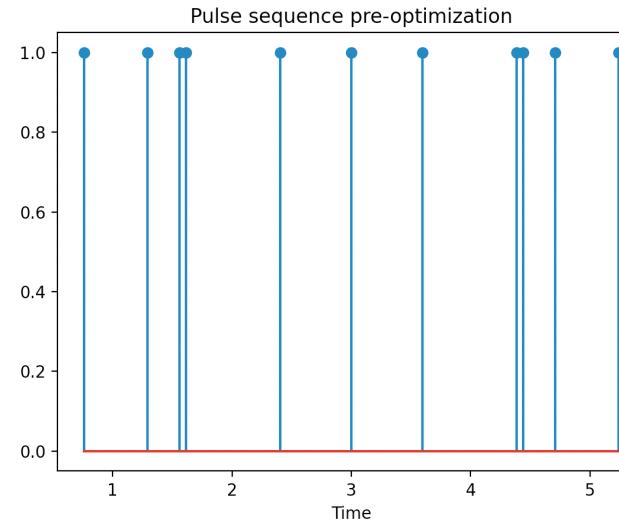
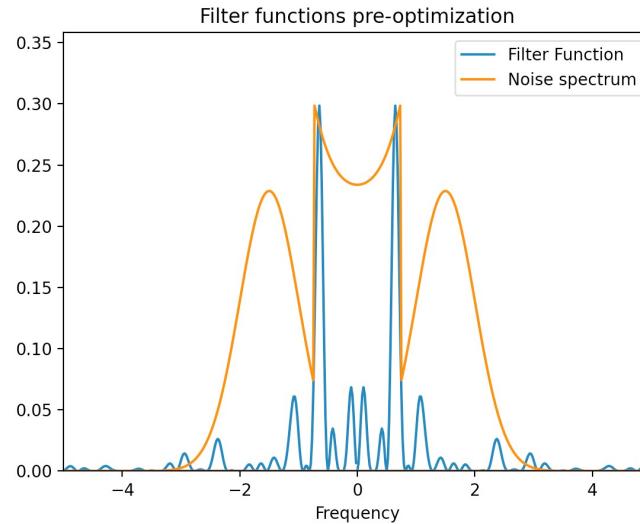
$m=100$ possible harmonics – now including the possibility of negative m –, $\Delta t = dt$, 20 iterations, $N=11$ pulses; grid 2×10^5 data points



Some trials – cutting out analytical part

Step function noise $T \rightarrow 0$, mu = 0.5, plus some “side bands”

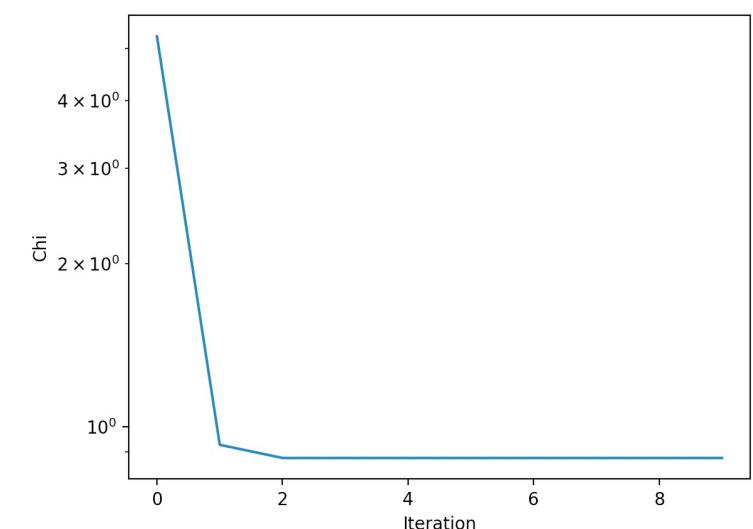
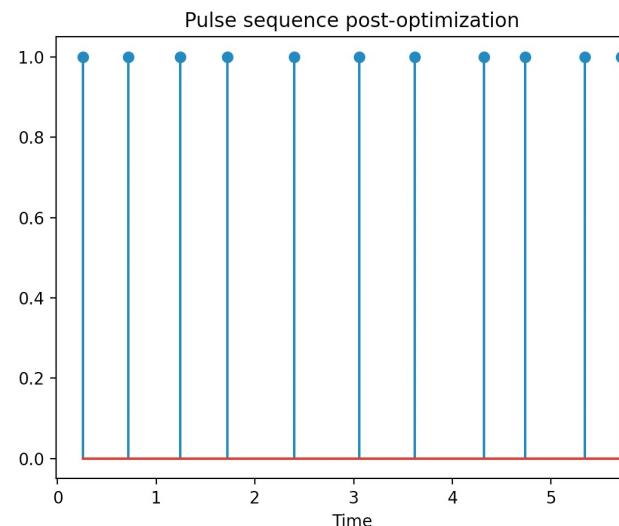
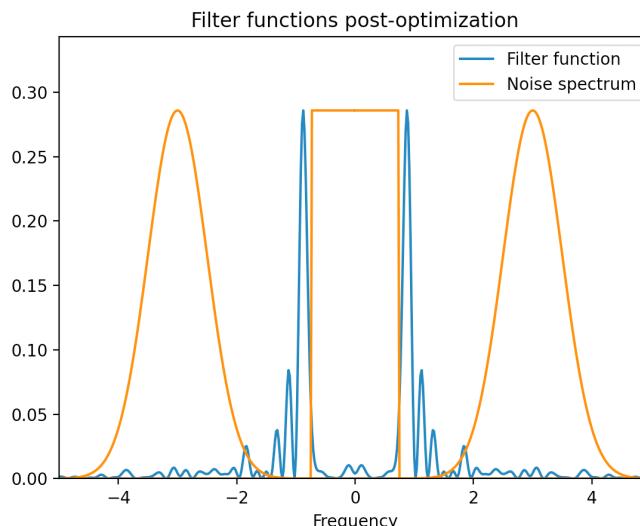
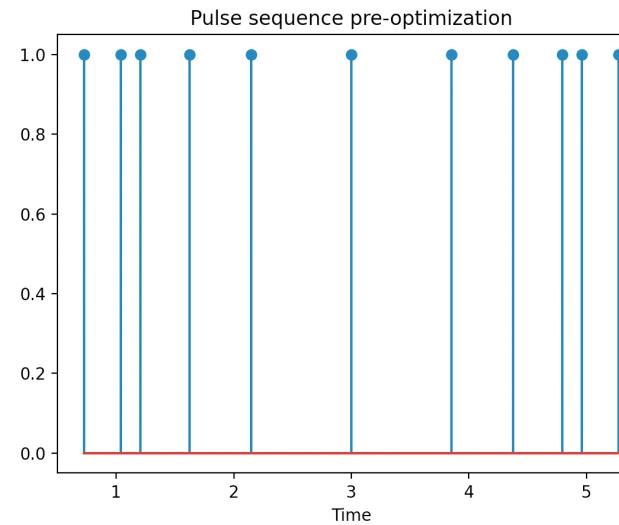
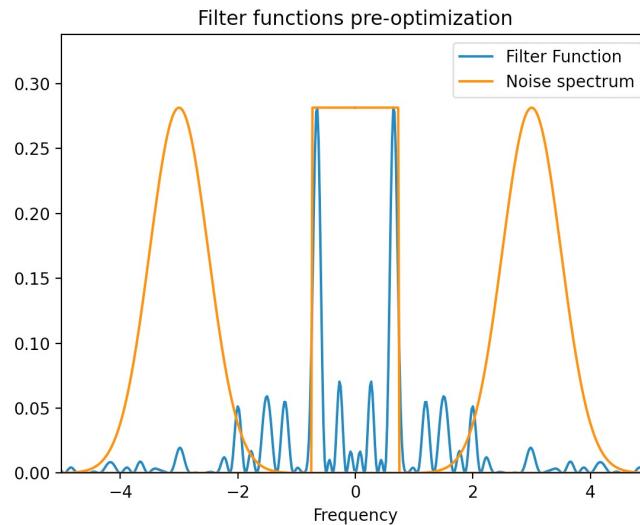
m=100 possible harmonics – now including the possibility of negative m - , delta t = dt, 20 iterations, N=11 pulses; grid 2×10^5 data points



Some trials – cutting out analytical part

Step function noise $T \rightarrow 0$, mu = 0.5, plus some “side bands”

m=100 possible harmonics – now including the possibility of negative m - , delta t = dt, 10 iterations, N=11 pulses; grid 2×10^5 data points

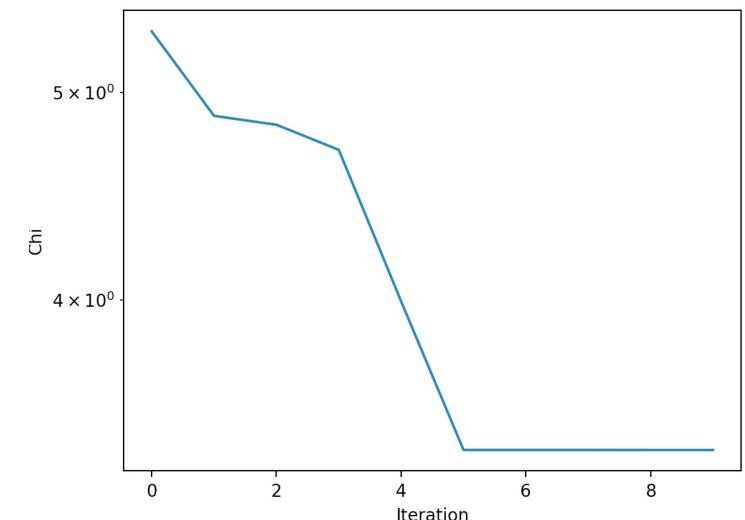
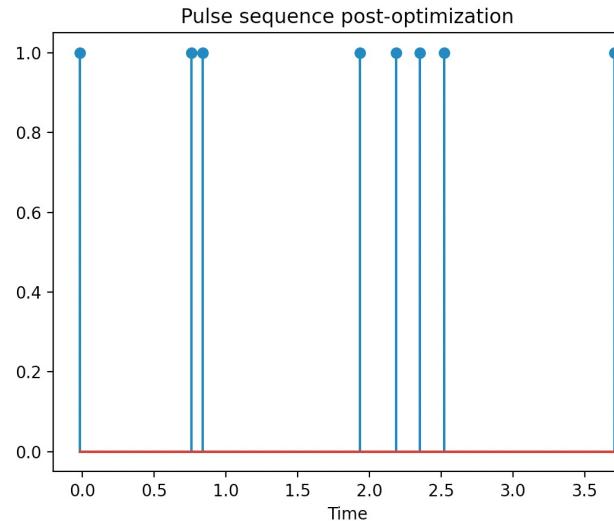
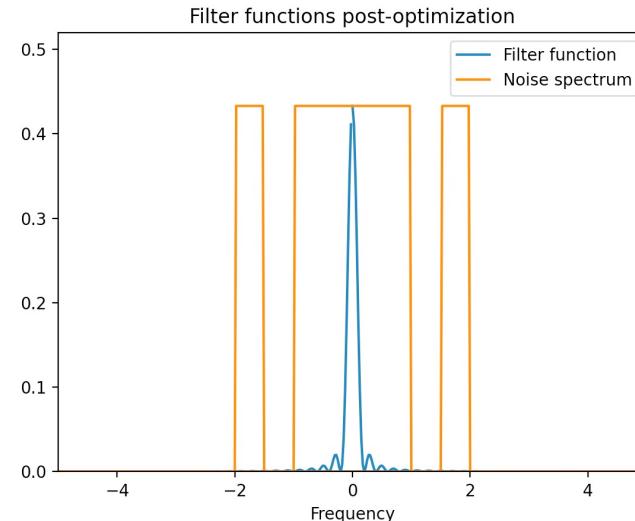
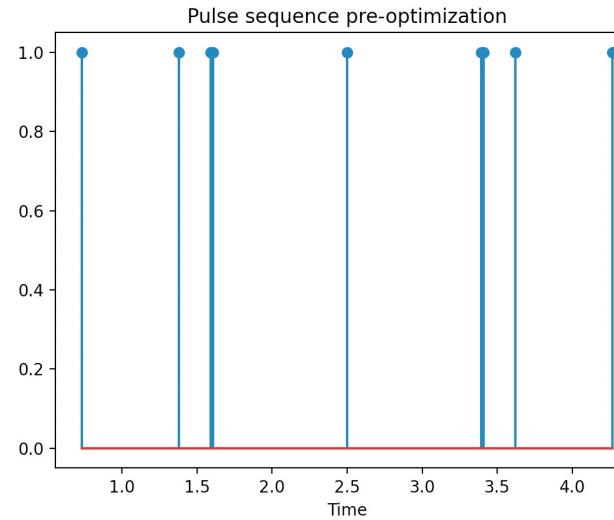
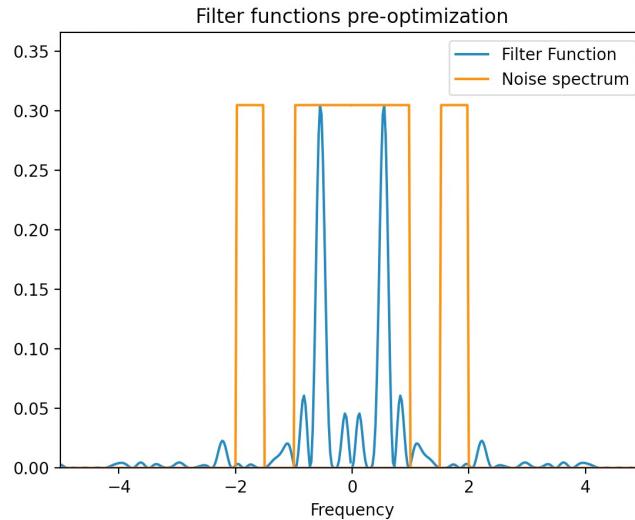


Some trials – cutting out analytical part

Arbitrary new noise spectrum

$m=25$ possible harmonics – now including the possibility of negative m - , $\Delta t = dt$, 10 iterations, $N=9$ pulses; grid 5×10^4 data points.

Add in small perturbation to the pulse timings $= N(0, 1) * dt$ when stuck at a local minimum.

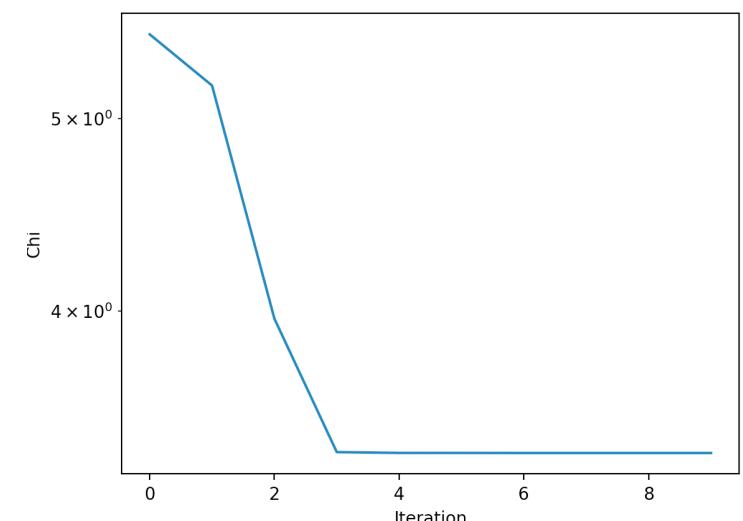
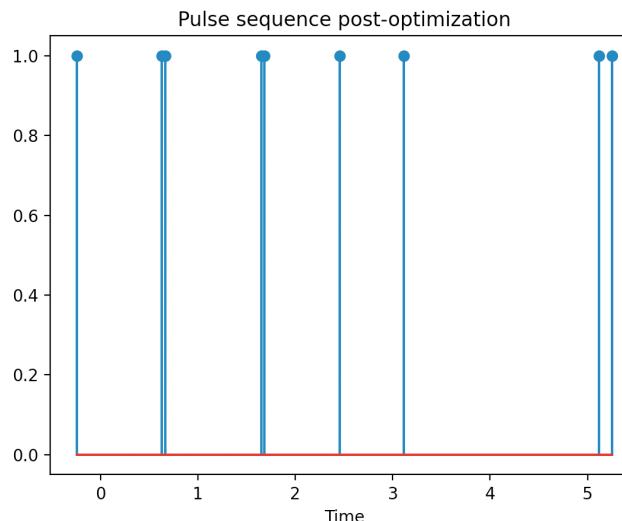
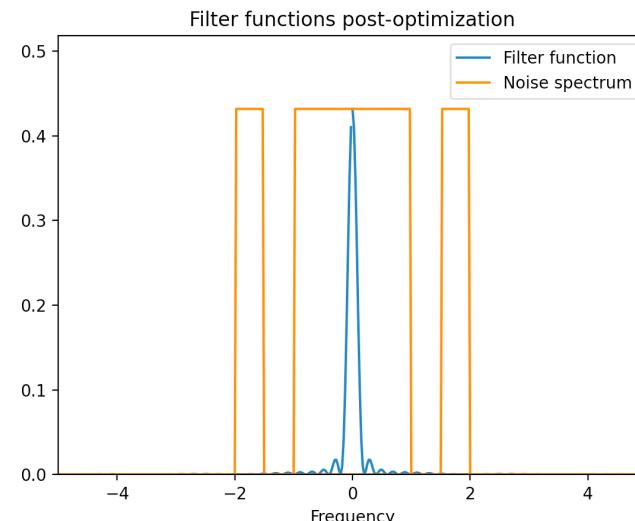
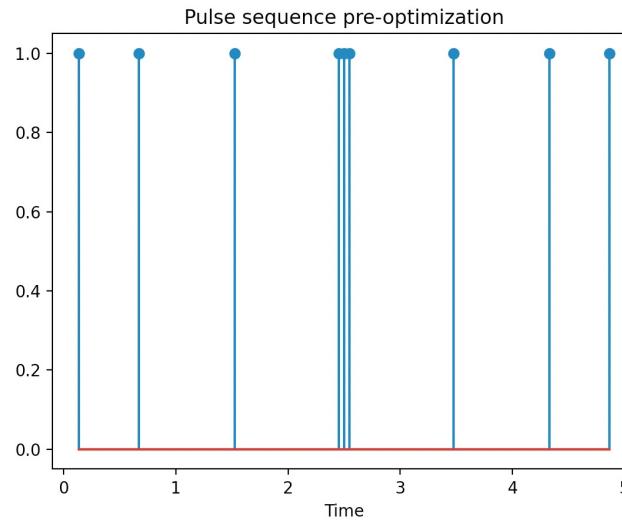
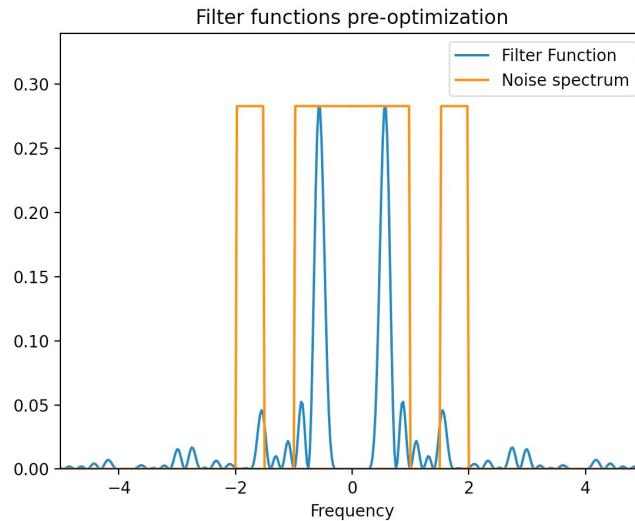


Some trials – cutting out analytical part

Arbitrary new noise spectrum

$m=25$ possible harmonics – now including the possibility of negative m - , $\Delta t = dt$, 10 iterations, $N=9$ pulses; grid 5×10^4 data points.

Add in small perturbation to the pulse timings $= N(0, 1) * dt$ when stuck at a local minimum.

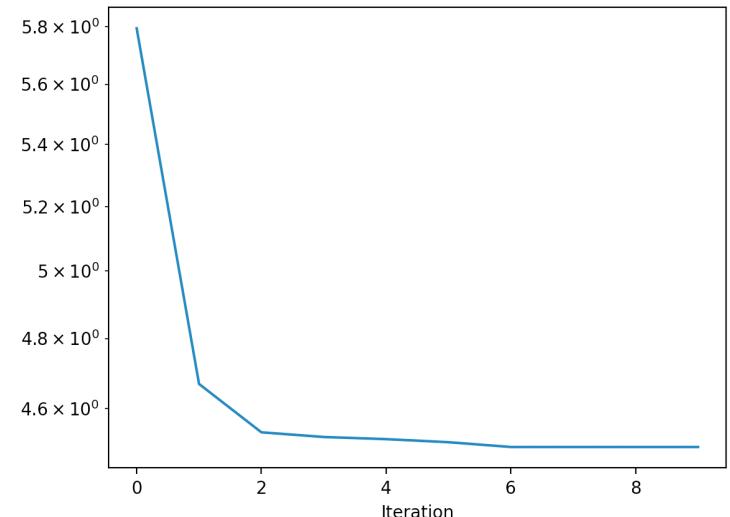
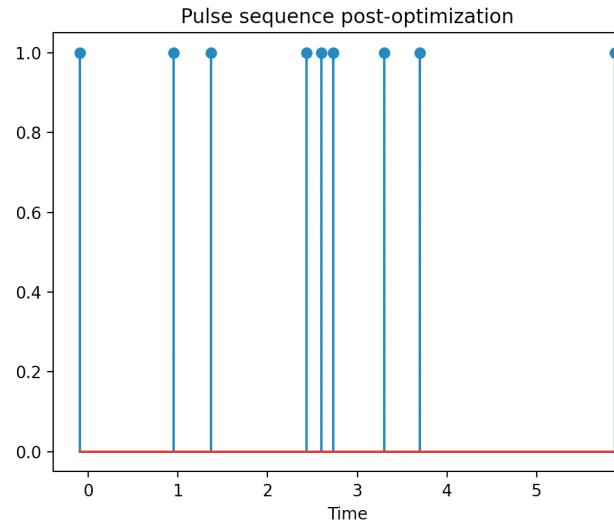
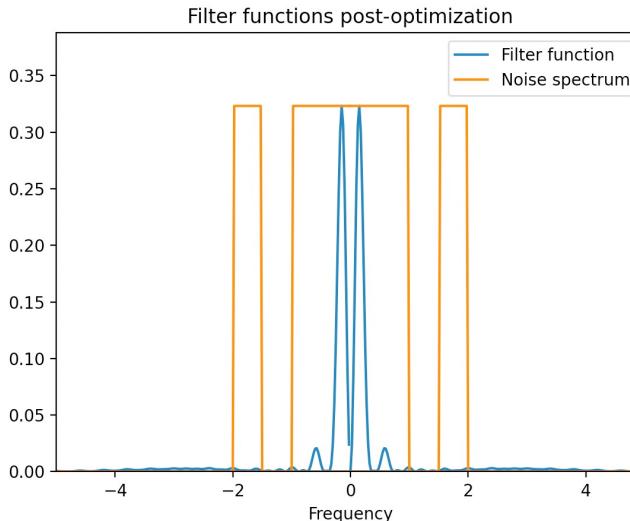
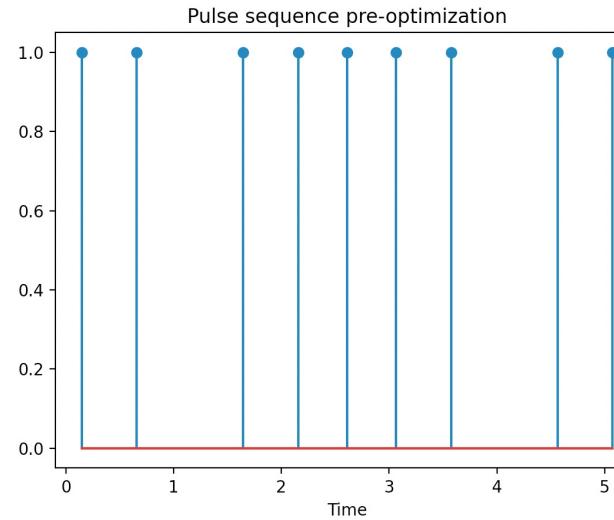
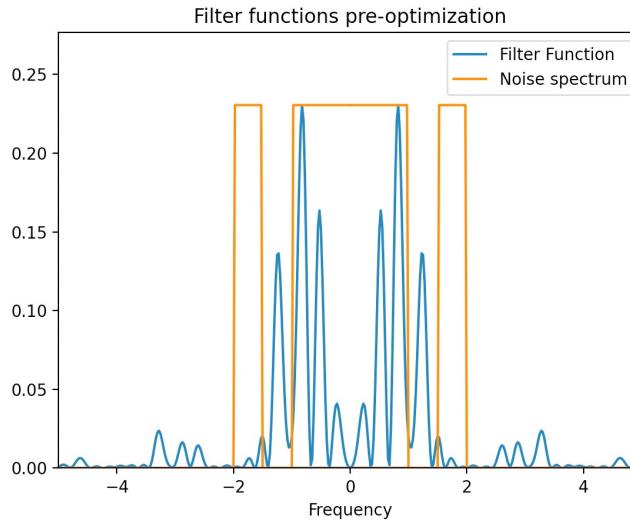


Some trials – cutting out analytical part

Arbitrary new noise spectrum

$m=25$ possible harmonics – now including the possibility of negative m - , $\Delta t = dt$, 10 iterations, $N=9$ pulses; grid 5×10^4 data points.

Add in small perturbation to the pulse timings $= N(0, 1) * 5dt$ when stuck at a local minimum.

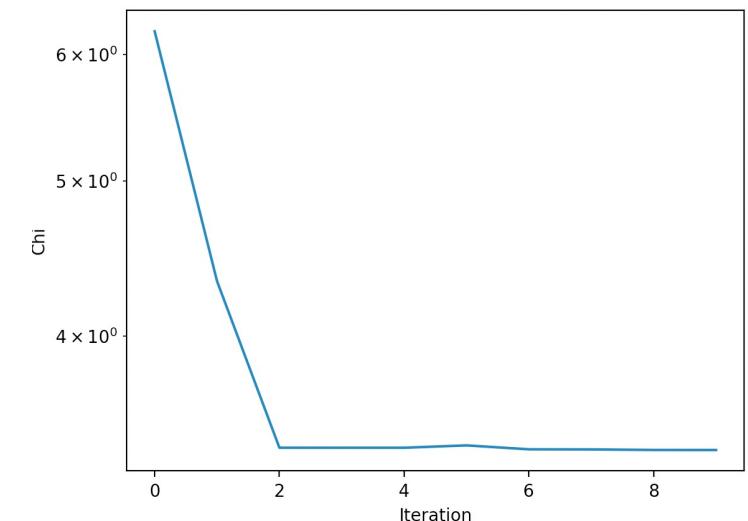
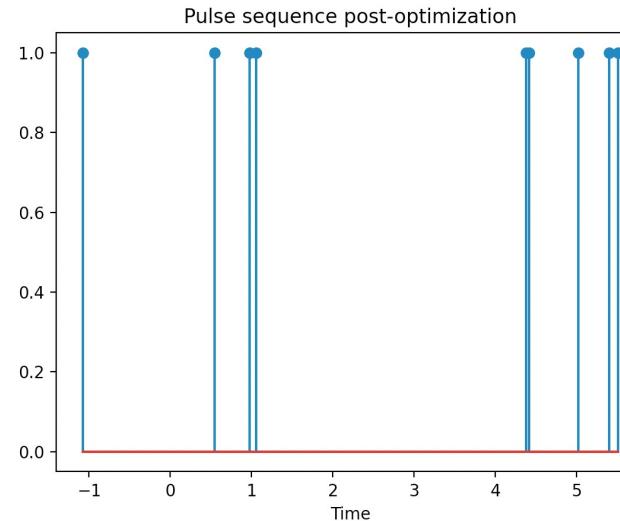
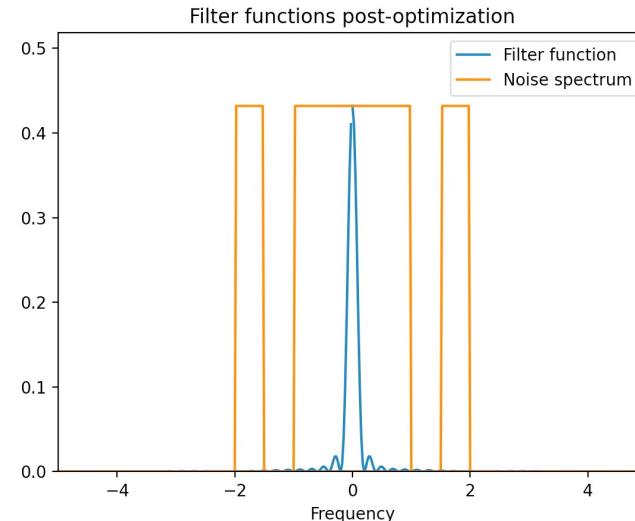
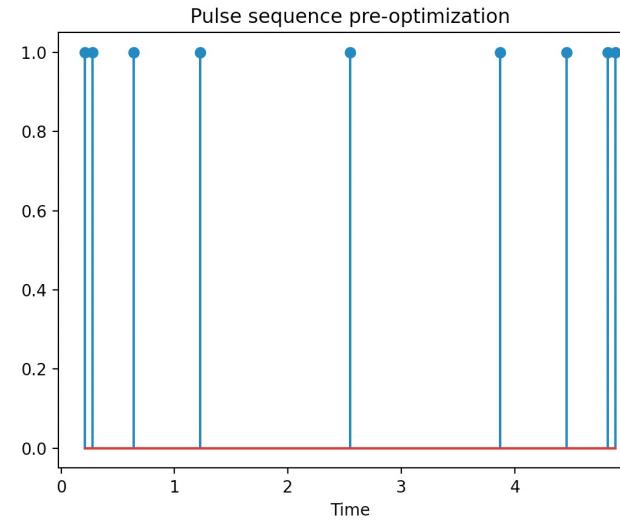
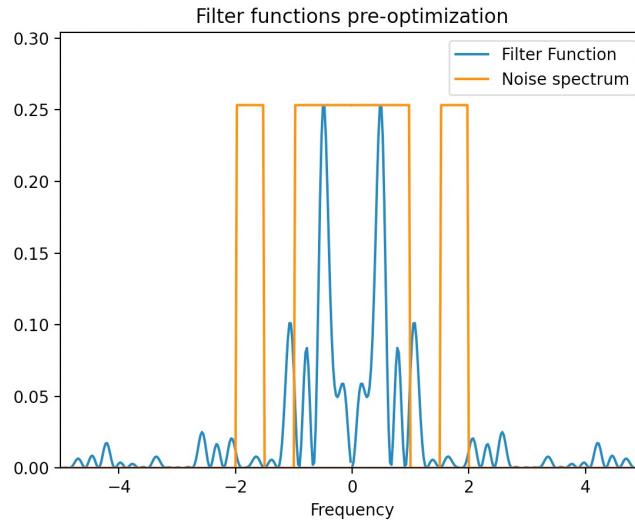


Some trials – cutting out analytical part

Arbitrary new noise spectrum

$m=25$ possible harmonics – now including the possibility of negative m - , $\Delta t = dt$, 10 iterations, $N=9$ pulses; grid 5×10^4 data points.

Add in small perturbation to the pulse timings $= N(0, 1) * 5dt$ when stuck at a local minimum.

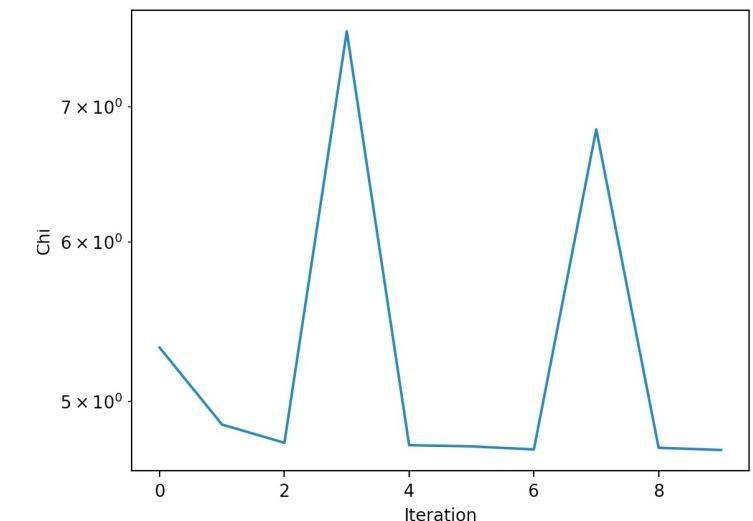
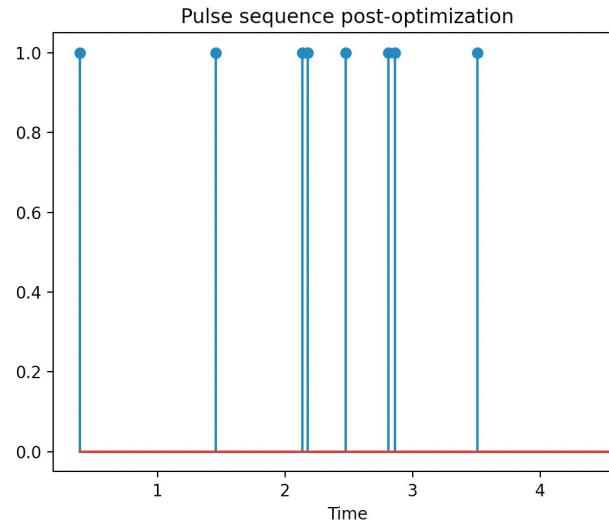
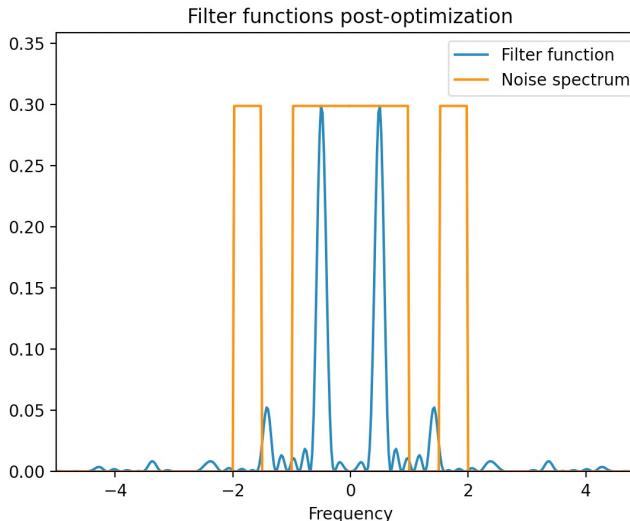
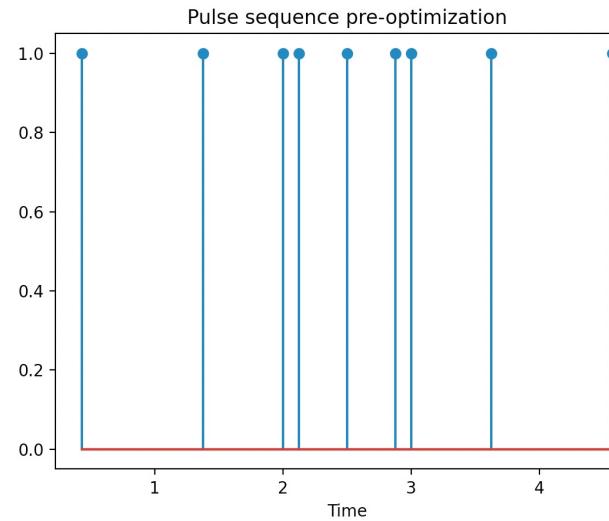
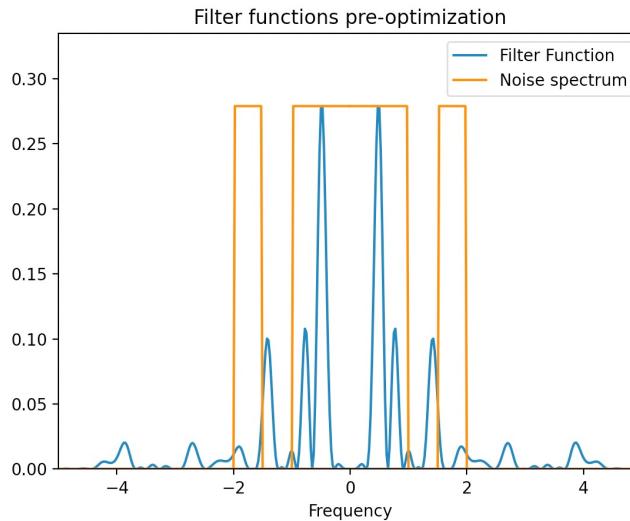


Some trials – cutting out analytical part

Arbitrary new noise spectrum

$m=25$ possible harmonics – now including the possibility of negative m - , $\Delta t = dt$, 10 iterations, $N=9$ pulses; grid 10^5 data points.

Add in small perturbation to the pulse timings $= N(0, 1) * 5dt$ when stuck at a local minimum.

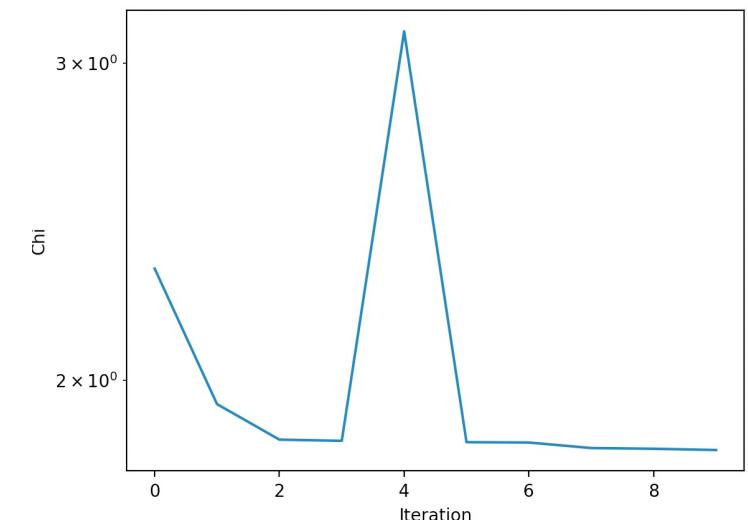
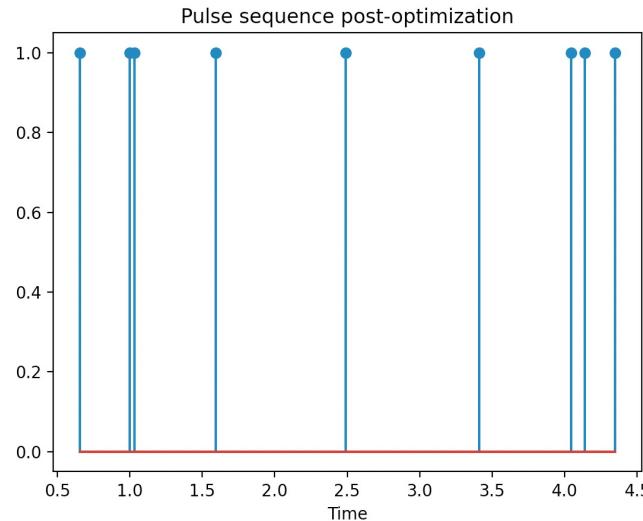
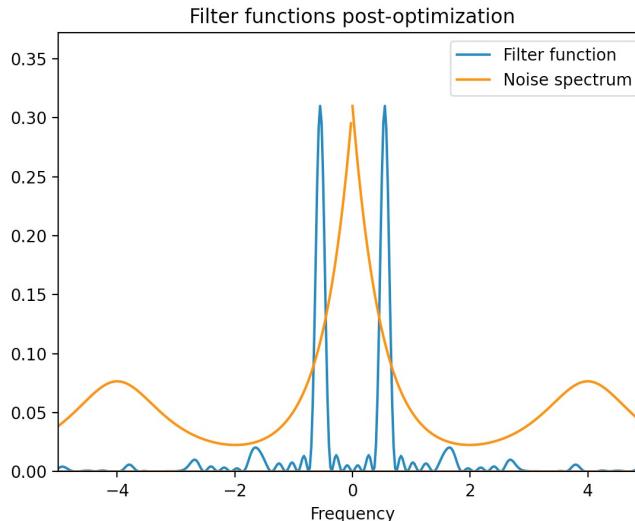
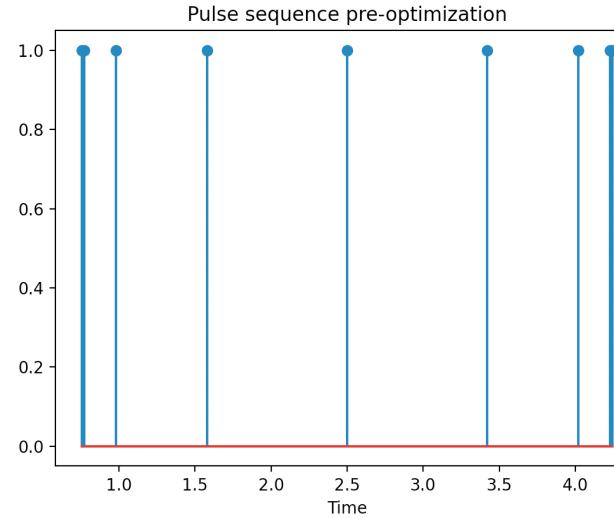
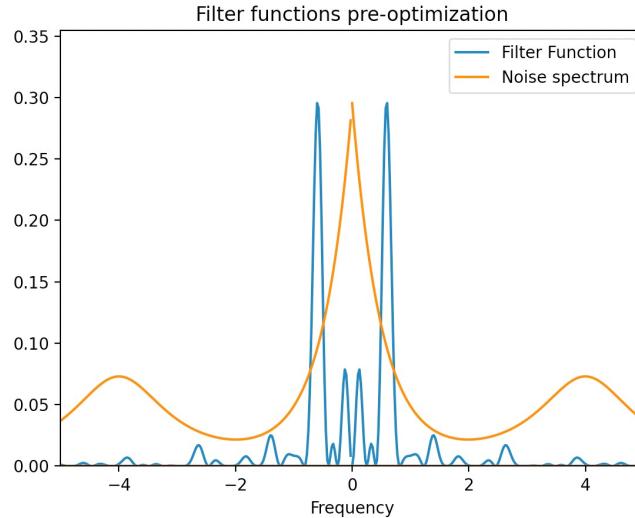


Some trials – cutting out analytical part

Arbitrary new noise spectrum

$m=25$ possible harmonics – now including the possibility of negative m - , $\Delta t = dt$, 10 iterations, $N=9$ pulses; grid 10^5 data points.

Add in small perturbation to the pulse timings $= N(0, 1) * 5dt$ when stuck at a local minimum.

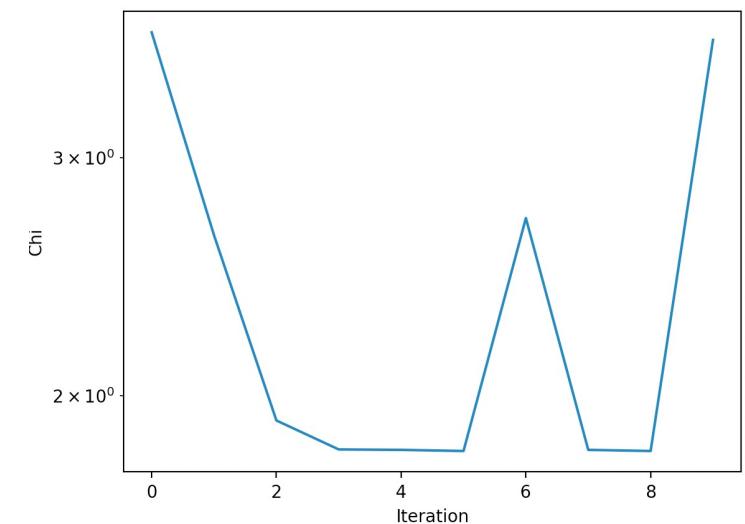
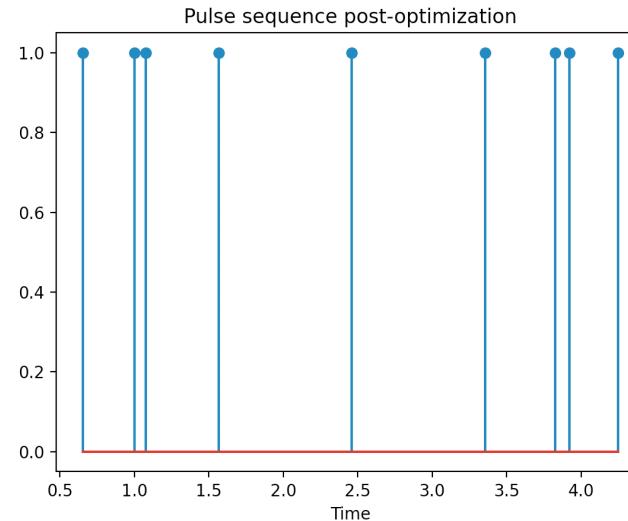
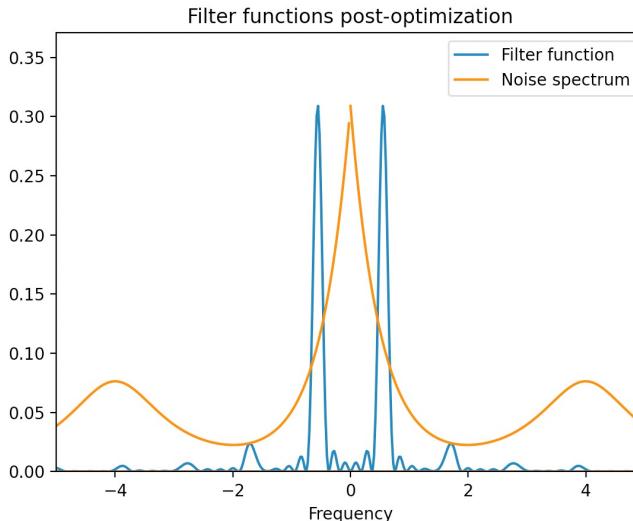
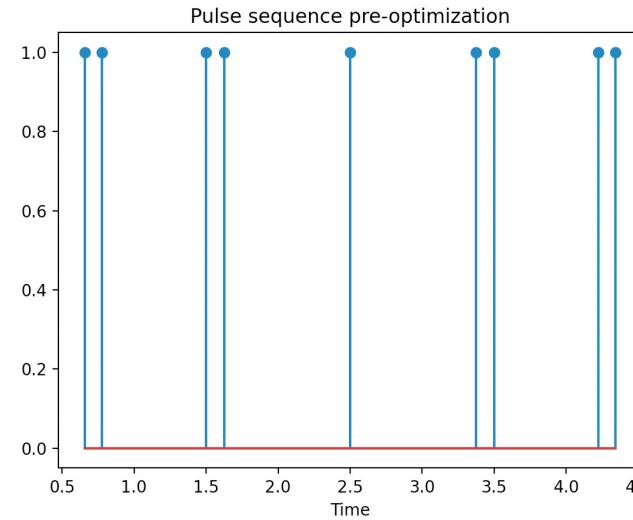
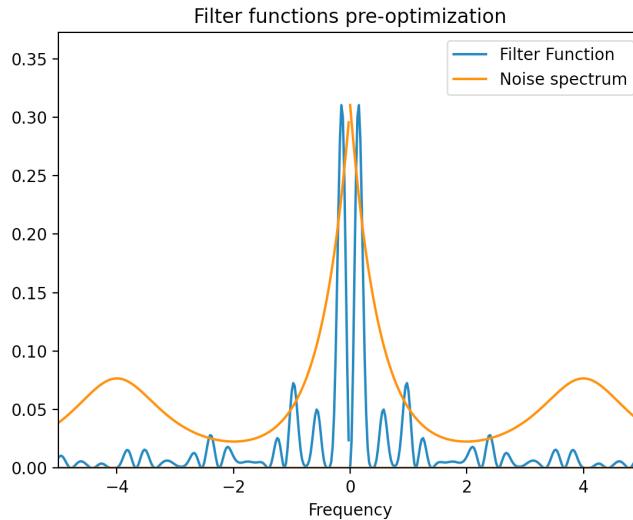


Some trials – cutting out analytical part

Arbitrary new noise spectrum

$m=25$ possible harmonics – now including the possibility of negative m - , $\Delta t = dt$, 10 iterations, $N=9$ pulses; grid 10^5 data points.

Add in small perturbation to the pulse timings $= N(0, 1) * dt$ when stuck at a local minimum.



Some trials – cutting out analytical part

Arbitrary new noise spectrum

$m=25$ possible harmonics – now including the possibility of negative m - , $\Delta t = dt$, 15 iterations, $N=9$ pulses; grid 10^5 data points.

Add in small perturbation to the pulse timings $= N(0, 1) * dt$ when stuck at a local minimum.

