Phys 512: PS4

1)a) First we consider Newton's method to find optimal parameters for a given fit

↳ Whereas gradient descent simply follows the gradient ∇f along the χ^2 manifold, Newton's method takes into account the curvature as follows $H^{-1}\nabla f$, where H^{-1} is the Hessian (inverse). Thus our step looks like:

$$x \longrightarrow x + \delta x \; : \; \delta x = -\gamma H^{-1}\nabla f \;, \quad \gamma \in \mathbb{R}$$

The following are the derivatives of the Lorentzian:

- $f(t) = \dfrac{a}{1+(t-t_0)^2/w^2} = a\left[1+w^{-2}(t-t_0)^2\right]^{-1}$ · $\partial_a f(t) = \left[1+w^{-2}(t-t_0)^2\right]^{-1}$
- $\partial_{t_0} f(t) = (2a(t-t_0)w^2)/((t-t_0)^2+w^2)^2$
- $\partial_w f(t) = (2a(t-t_0)^2 w)/((t-t_0)^2+w^2)^2$

This will be useful for the calculations:

- $f(t;m) \equiv A(m)$, $r \equiv d - A(m)$ · $\chi^2 = r^T N^{-1} r$ · $\nabla\chi^2 = -2A'^T N^{-1} r$
- $\nabla^2\chi^2 \simeq 2A'^T N^{-1} A'$ · $\delta m = (A'^T N^{-1} A')^{-1}(A'^T N^{-1} r)$

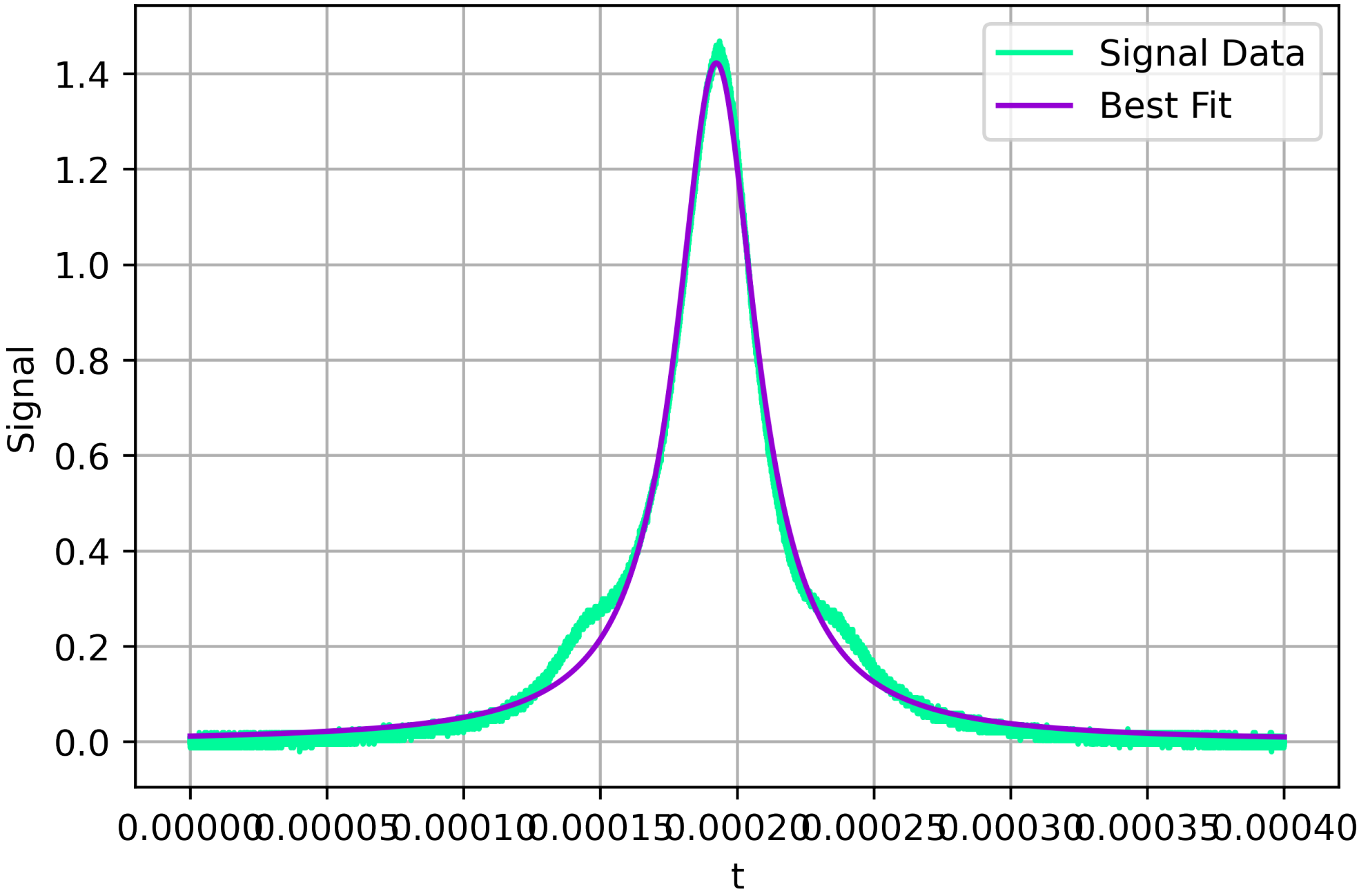
Note: We initially (for 1a)) assume $N = \mathbb{1}$ & thus get:

- $\delta m = \left[\nabla f^T \nabla f\right]^{-1}\left[\nabla f^T r\right] = \left[\nabla f^T \nabla f\right]^{-1}\left[\nabla f^T (d-f)\right]$

For an initial guess of $a, t_0, w = [1.5, 2\times10^{-4}, 1\times10^{-4}]$, we get the following best fit parameters; & the plot looks like:

25 steps

$$a = 1.4228 \;, \quad t_0 = 1.9236\times10^{-4} \;, \quad w = 1.7924\times10^{-5}$$


Signal Data
Best Fit
Signal
t
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
0.00000
0.00005
0.00010
0.00015
0.00020
0.00025
0.00030
0.00035
0.00040

b) Our estimate for computing the error was derived from discussions w/ Guillaume

↳ The procedure for estimating N will be done via taking the standard deviation (at a point x_0) of a small, compact neighborhood of data point surrounding x_0

↳ Moreover we assume the standard error on each data point is $\simeq 5\times10^{-3} \equiv \sigma$ ($\simeq$ the std of 25 points)

With this we can estimate the covariance matrix:

$$N^{-1} = \mathbb{1}/\sigma^2 \Rightarrow N = A^T N^{-1} A' = \frac{1}{\sigma^2}\nabla f^T \nabla f$$

Using this, we get the parameter uncertainties:

$$\sigma_a = 8.4309\times10^{-5}, \quad \sigma_{t_0} = 1.0618\times10^{-9}, \quad \sigma_w = 1.5036\times10^{-4}$$

c) We redo part a) w/ numerical derivatives & get (s the following best-fit parameters:

$$a = 1.4228, \quad t_0 = 1.9236\times10^{-4}, \quad w = 1.7923\times10^{-5}$$

It is somewhat suspicious that they are the exact same values (on python they agree to a lot more digits), but it suffices to say they are not statistically different from the answers in a)

↳ The difference is less than σ we computed in b)

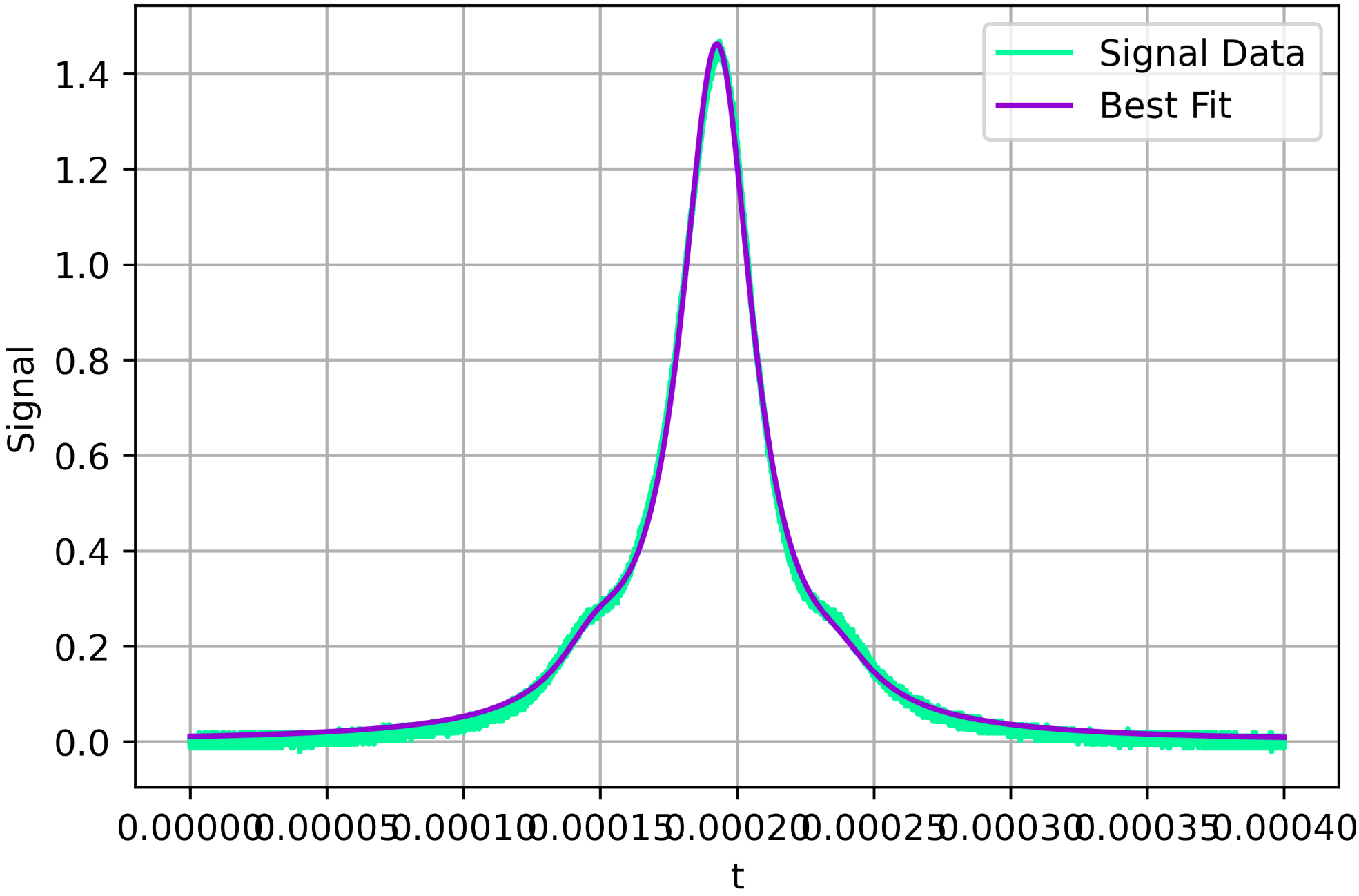
d) We repeat what is done in c) (numerical derivatives), except now with:

$$f(t) = a\left[1 + (t-t_0)^2/w^2\right]^{-1} + b\left[1 + (t-t_0+dt)^2/w^2\right]^{-1} + c\left[1 + (t-t_0-dt)^2/w^2\right]^{-1}$$

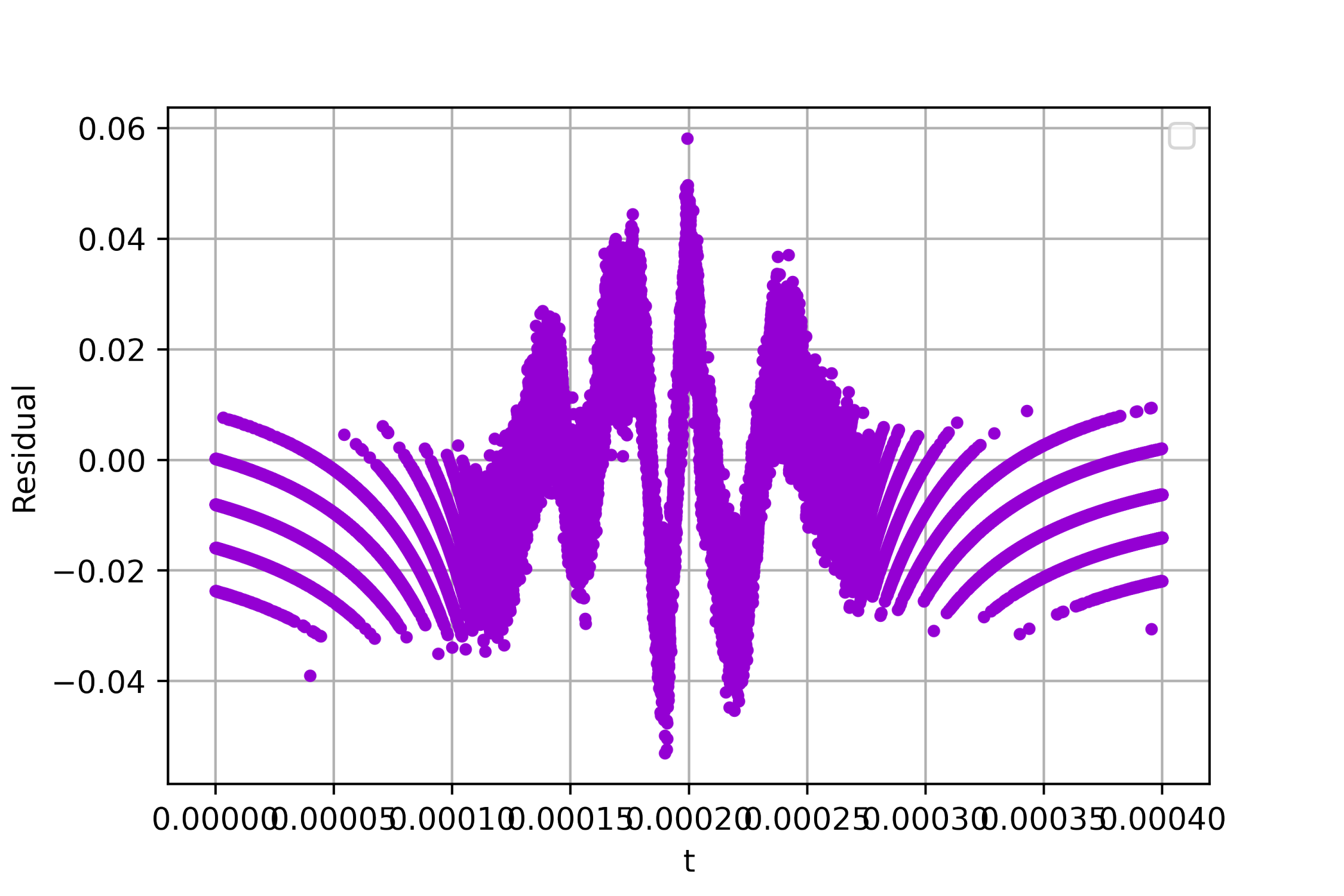
Following the procedure we've done in a-c), we have the following best fit parameters & their uncertainties:

$a = 1.4430 \pm 9.1390 \times 10^{-5}$, $b = 1.0391 \times 10^{-1} \pm 8.7167 \times 10^{-5}$
$c = 6.4733 \times 10^{-2} \pm 8.5391 \times 10^{-5}$, $t_0 = 1.9258 \times 10^{-4} \pm 1.0820 \times 10^{-9}$
$w = 1.6065 \times 10^{-5} \pm 1.9378 \times 10^{-9}$, $dt = 4.457 \times 10^{-5} \pm 1.3044 \times 10^{-8}$

Moreover, the prediction fits the data better, as seen in:

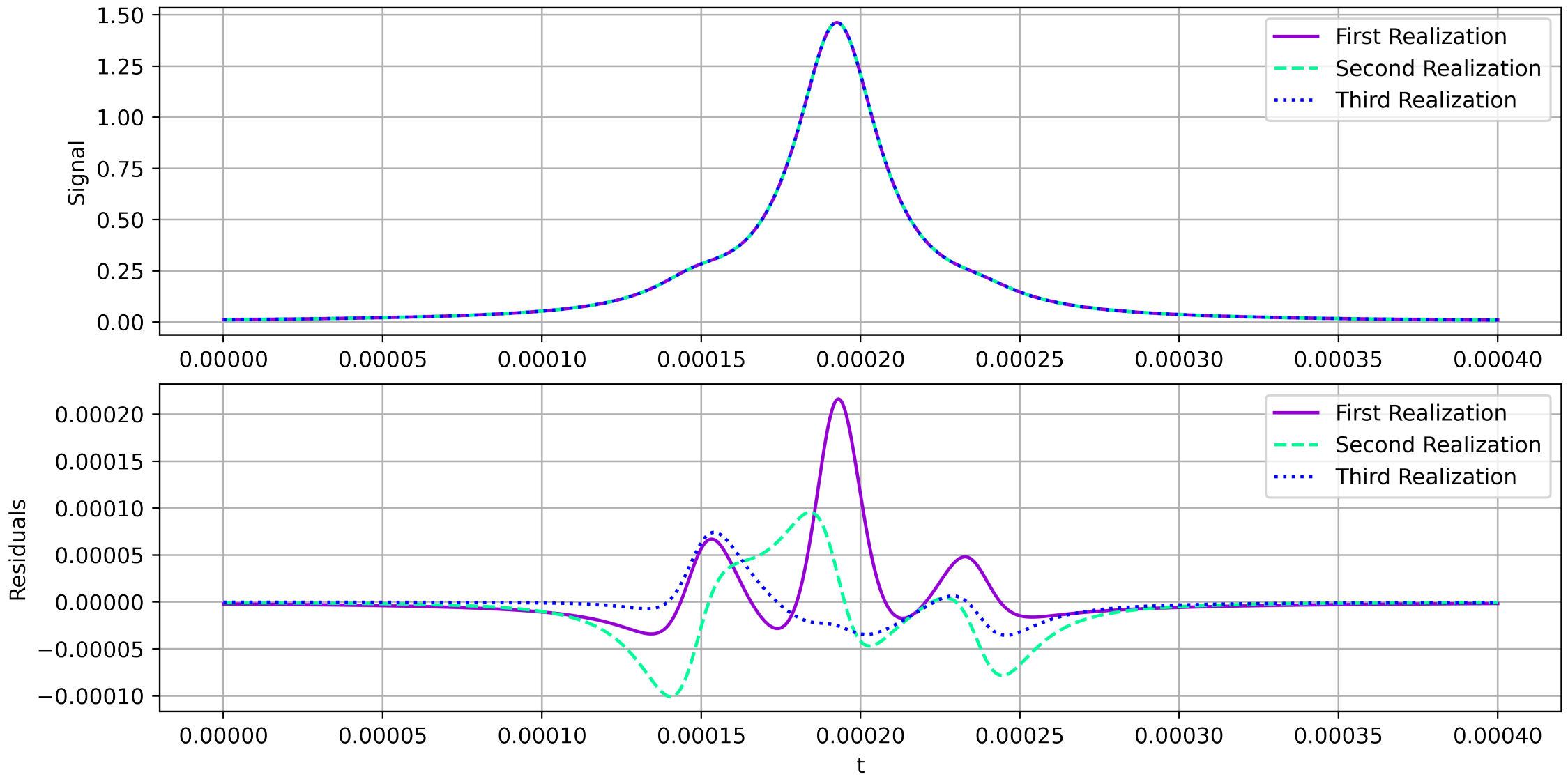

Signal Data
Best Fit
Signal
t
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
0.00000
0.00005
0.00010
0.00015
0.00020
0.00025
0.00030
0.00035
0.00040

e) The residuals (data - fit) have the following plot: (2)


0.06
0.04
0.02
0.00
−0.02
−0.04
Residual
0.00000
0.00005
0.00010
0.00015
0.00020
0.00025
0.00030
0.00035
0.00040
t

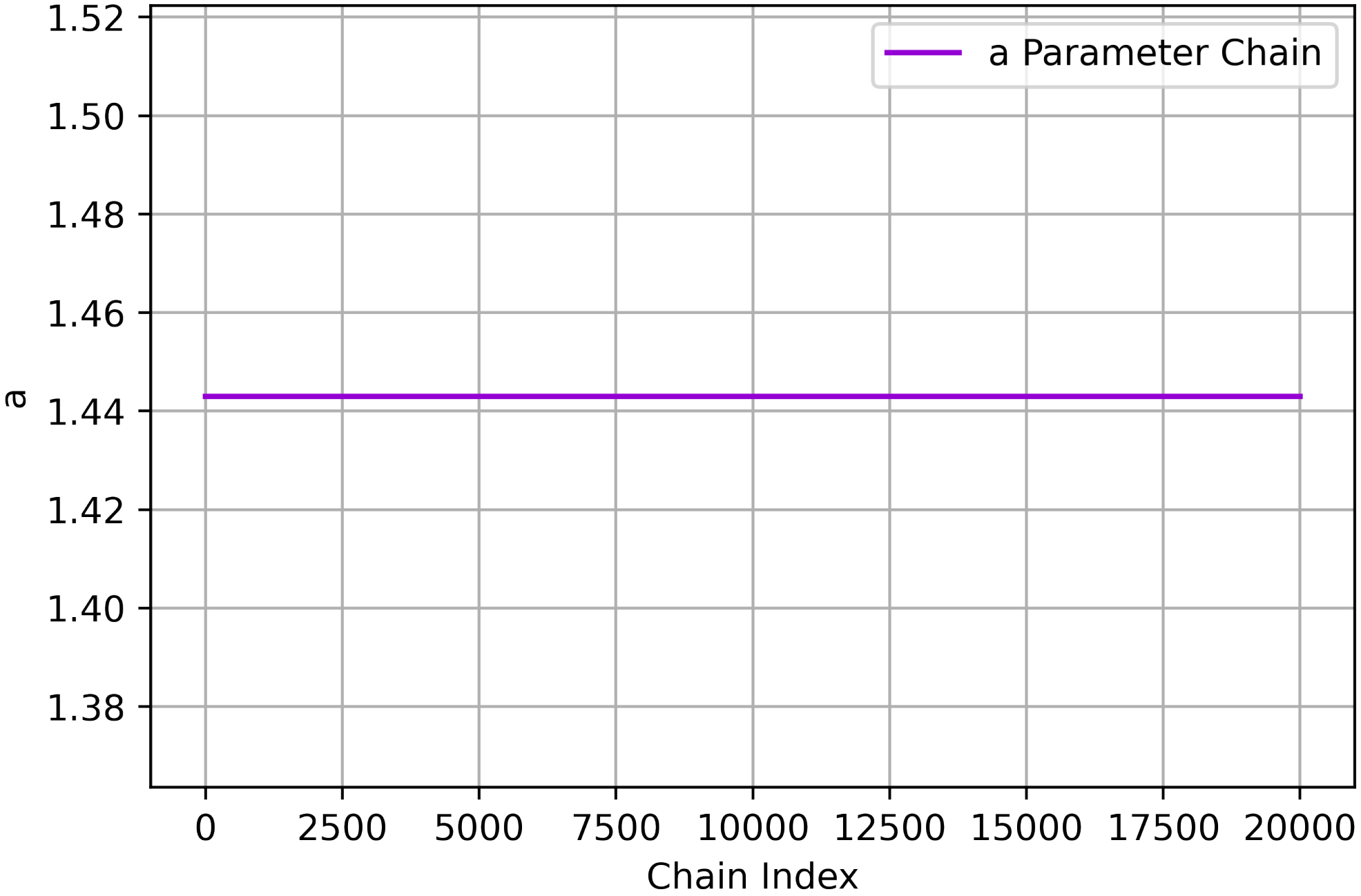
The residuals clearly contain a structure & are thus correlated. This tells us the error bars we got by assuming independence is not a full description of the data!

f) We consider three different realizations for randomly (1 perturbed parameters & it has the following plot (& their corresponding residuals):


Signal
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0.00000
0.00005
0.00010
0.00015
0.00020
0.00025
0.00030
0.00035
0.00040
First Realization
Second Realization
Third Realization
Residuals
0.00020
0.00015
0.00010
0.00005
0.00000
−0.00005
−0.00010
t

Computing the typical difference in χ^2 gives us 8.5200×10^{-5} which is reasonable as it is much less than 1 & we only perturbed the parameters by a small, random amount.

g) After computing the MCMC algorithm for 20,000 steps, we have the chain plot for the 'a' parameter:


a Parameter Chain
1.52
1.50
1.48
1.46
1.44
1.42
1.40
1.38
a
0
2500
5000
7500
10000
12500
15000
17500
20000
Chain Index

As we can see from above, the chain converges as there is no structure on the scale of the full chain.*

↳ Moreover, the best-fit parameters & their uncertainty is:

- $a = 1.4430 \pm 6.6613 \times 10^{-16}$
- $b = 0.1039 \pm 4.1633 \times 10^{-17}$
- $c = 0.0647 \pm 1.3878 \times 10^{-17}$
- $t_0 = 0.0002 \pm 5.4210 \times 10^{-20}$
- $\omega = 1.6065 \times 10^{-5} \pm 1.8061 \times 10^{-16}$
- $dt = 4.4567 \times 10^{-5} \pm 1.2351 \times 10^{-17}$

As we can see, the error bars are much, much smaller. This is a very good result but it makes me think there is some error w/ my code as my peers have much higher errors (I have not yet found the bug).

*The fact that it is constant raises some red flags but I cannot find the bug in my code that would produce such a result.

h) Using our best fit parameters, we can compute the (?
width μ of the resonance cavity:

$$\mu = dt/w \cdot 9\,GHz = (24.9674 \pm 0.0280)\,GHz$$