

Curve Fit Data Cleaning

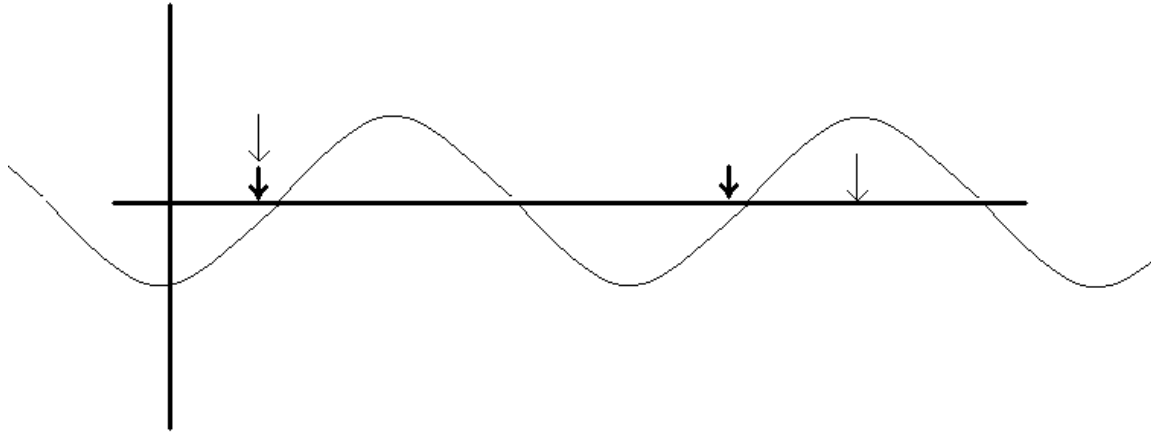
Bob Coldwell

Bernard Whiting

University of Florida

G030129-00-Z

The end points can cause the DFT to differ from the value desired. Obviously in the figure below, a sum between the heavy arrows gives a quite different result than the sum between the light arrows.



In this case a fit to the data, rather than simply a sum gives an end point free estimation of the underlying function.

Tchebyshev Polynomials

$$T_n(t) = \cos(n \cos^{-1}(t))$$

$$T_0(t) = \cos(0) = 1$$

$$T_1(t) = \cos(\cos^{-1}(t)) = t$$

$$\begin{aligned} T_2(t) &= \cos(2 \cos^{-1}(t)) \\ &= 2 \cos^2(\cos^{-1}(t)) - 1 \\ &= 2t^2 - 1 \end{aligned}$$

in the expansion

$$d(t) = \sum_{j=0}^{\infty} c_j T_j(t) = \sum_{j=0}^{N-1} c_j T_j(t) + \sum_{j=N}^{\infty} c_j T_j(t)$$

the error in approximating the function is approximately equal the last term dropped. The orthogonality condition

$$\int_{-1}^1 w(t) T_j(t) T_k(t) dt = \alpha_j \delta_{j,k}$$

Requires the weight function

$$w(t) = 1 / \sqrt{1 - t^2}$$

The c 's which minimize

$$\chi_f^2 = \sum_{i=1}^N w(t_i) \left[d(t_i) - \sum_{j=0}^M c_j T_j(t_i) \right]^2$$

Are the Tchebyshev c 's to the extent that the sum approximates an integral

Extend the definition of chi-square

$$\chi^2 \equiv \sum_{i=1}^N w(t_i) \left(\begin{array}{l} d(t_i) \\ -\cos(2\pi ft_i) \sum_j c_j T_j(t_i) \\ -\sin(2\pi ft_i) \sum_j c_j T_j(t_i) \end{array} \right)^2$$

Note that the sine and cosine are orthogonal and that for large f the sine and cosine merely multiply the ortho-normality by 1/2

Define a few terms

$$A_N = \sum_{i=1}^N w_i \quad \langle d^2 \rangle \equiv \sum_{i=1}^N w_i d^2(t_i)$$

$$\langle dP_k \rangle \equiv \sum_{i=1}^N w_i d(t_i) P_k(t_i)$$

$$\langle P_m P_n \rangle \equiv 2 \sum_{i=1}^N w_i P_m(t_i) P_n(t_i)$$

Note that in principle only the second and third need recalculation for each data set

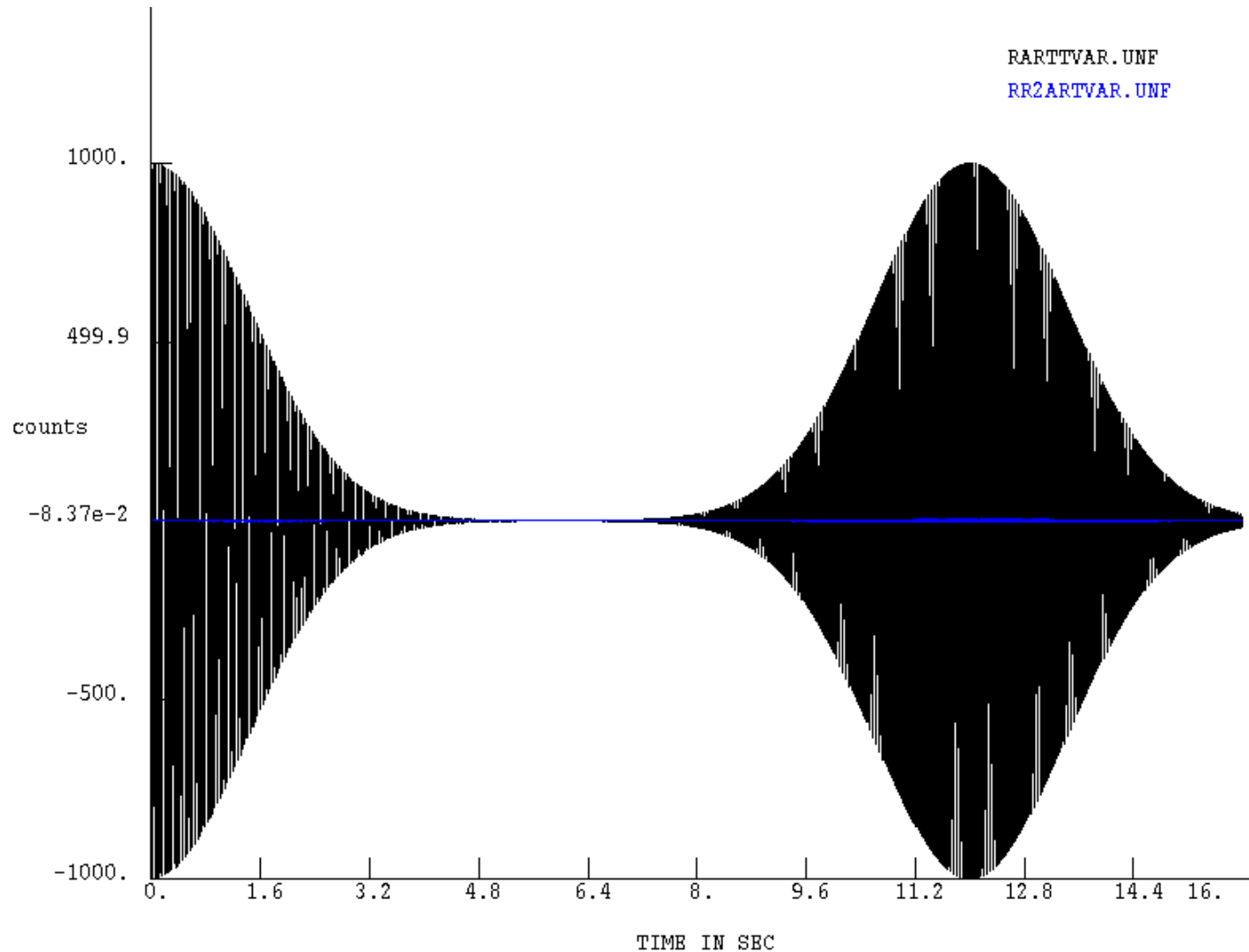
$$\chi^2 = -2 \sum_K c_K \langle dP_K \rangle + \frac{1}{2} \sum_K c_K \sum_M c_M \langle P_K P_M \rangle$$

$$\frac{\partial \chi^2}{\partial c_L} = -2 \langle dP_L \rangle + \sum_K c_K \langle P_K P_L \rangle$$

$$\frac{\partial^2 \chi^2}{\partial c_L \partial c_M} = \sum_K c_K \langle P_K P_L \rangle \delta_{KM} = \langle P_M P_L \rangle$$

$$Pen = \sum_k 10^4 \left(c_k - 2 \sqrt{\langle d^2 \rangle / A_N} \right)_+^3$$

Artificial data The 16 seconds of data shown were cleaned using 20 sine and cosine constants.



```

NUMBER OF POINTS TO FIT      262144
FR CHI,CHB,CHL .00 .59032E+11 .100000E+67 .500000E+67
FR CHI,CHB,CHL .01 .28750E+11 590323895. .59032 E+11
FR CHI,CHB,CHL .11 .31337E+10 .313377E+10 .287502E+11
FR CHI,CHB,CHL .11 341581387. 341581387. .313377E+10
FR CHI,CHB,CHL .11 37232371.7 37232371.7 341581387.
  AITKIN'S EXTRAPOLATION
FR CHI,CHB,CHL .11 44160.6 4058328.58 37232371.7
FR CHI,CHB,CHL .01 4090.63 4090.63 44160.5993
AT END FR, CHI .1000000000E-01 4090.62882
FREQ 27.1700000000000010 ==> 27.182821790006640
  actual value 27.18282182846
FREQ 41.2400000000000010 ==> 41.231056174760330
actual value 41.2310562562
normal end

```

ARTTVAR.UNF

0 262144 2 40

CONS FOR 27.17000000000001 .0000000000000000

1 141.1877785360938 -153.1200653773168

2 -259.6633918566866 278.1145656944491

3 199.6142932321022 -204.7202555695115

4 -122.5116229808118 113.1739487377875

...

22 -.1239271669839974E-02 -.4565448792598915E-02

23 -.1919799700152570E-02 .4466747019762724E-02

24 .3201049198783860E-02 -.2413425345131505E-02

25 -.2908501269892250E-03 -.5838902543033794E-03

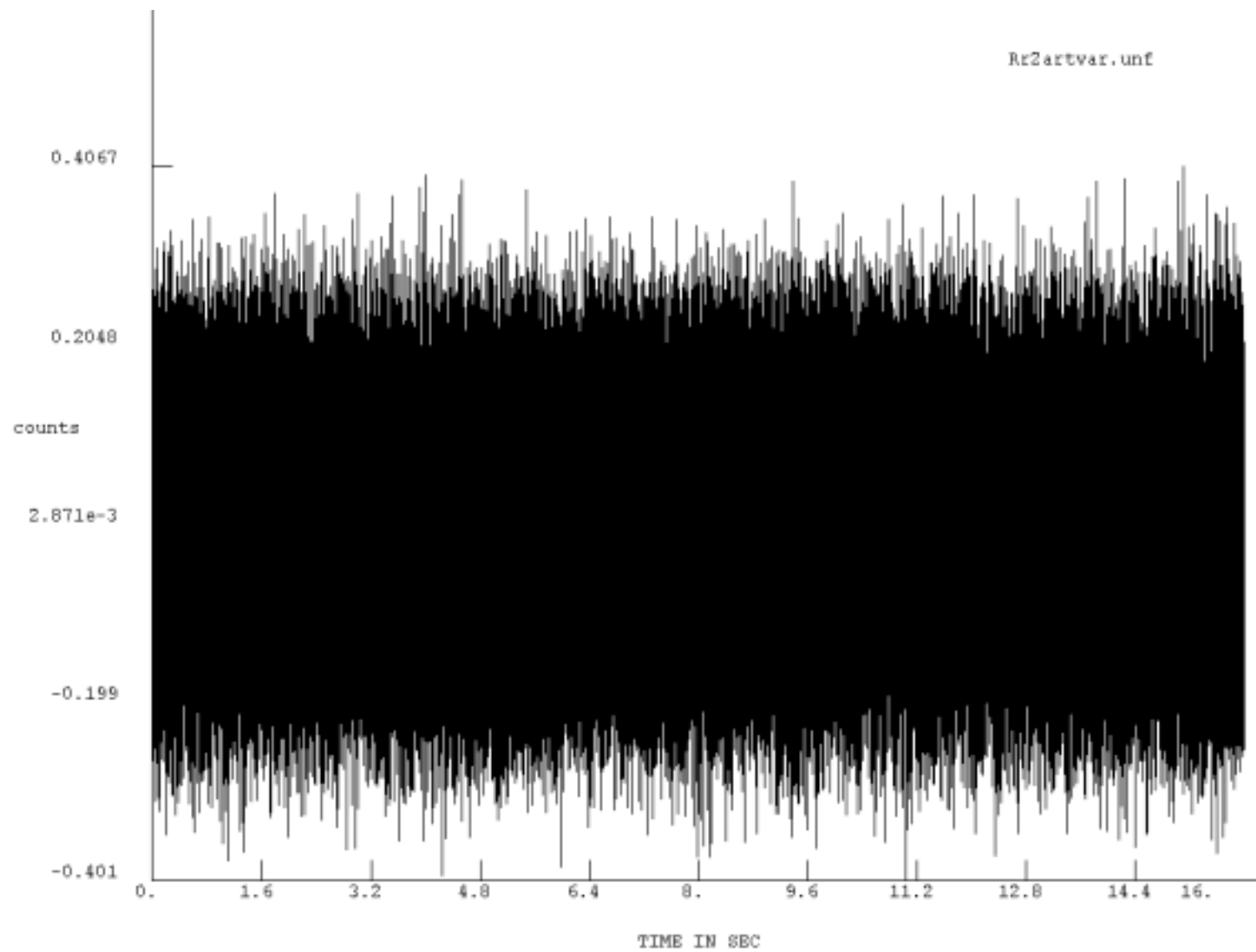
26 .7966868311713653E-03 .3628152278228707E-03

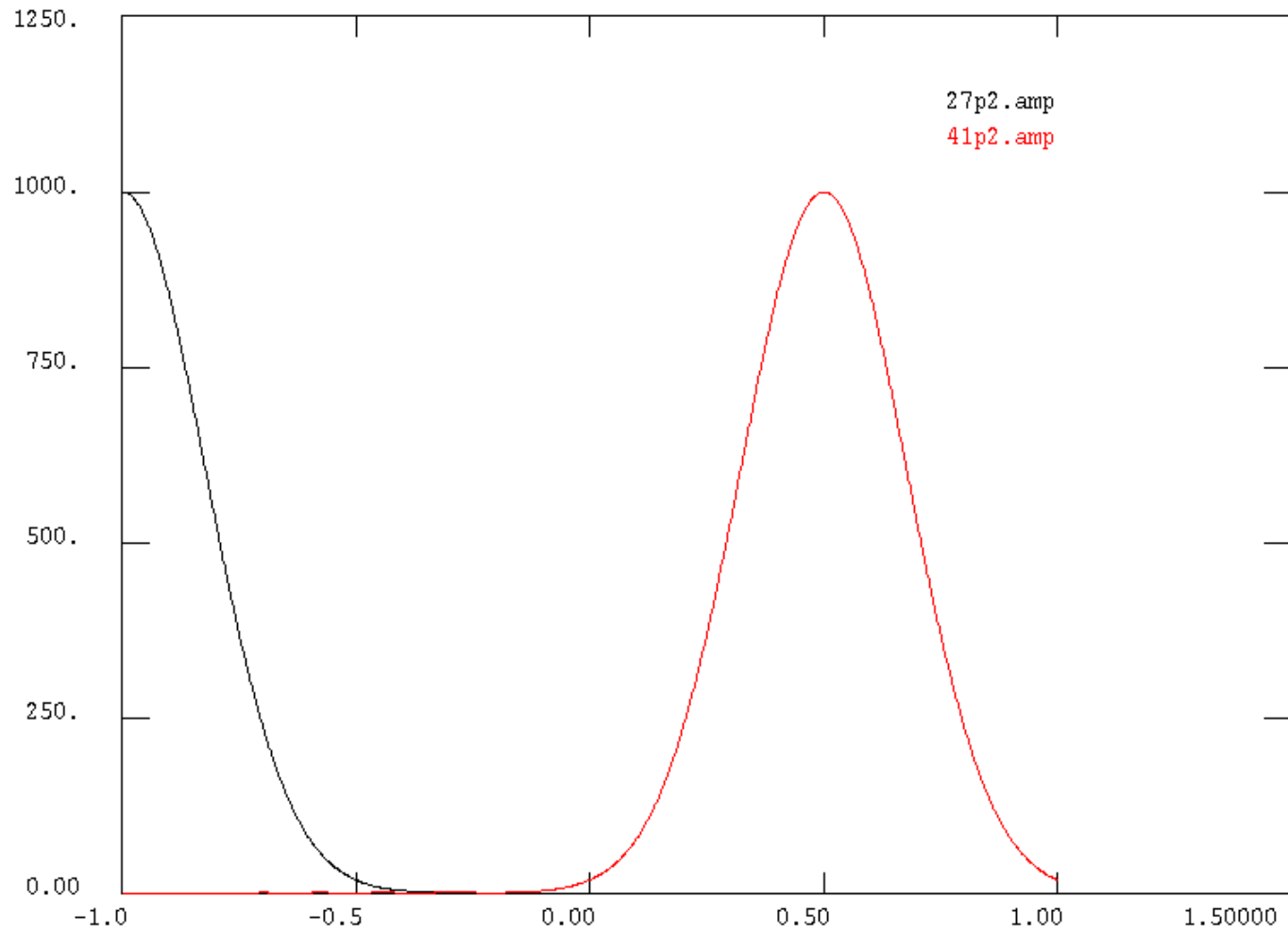
27 .5178381587015188E-03 -.2964034810670433E-03

...

40 .4511806768964878E-04 -.4811085379703161E-04

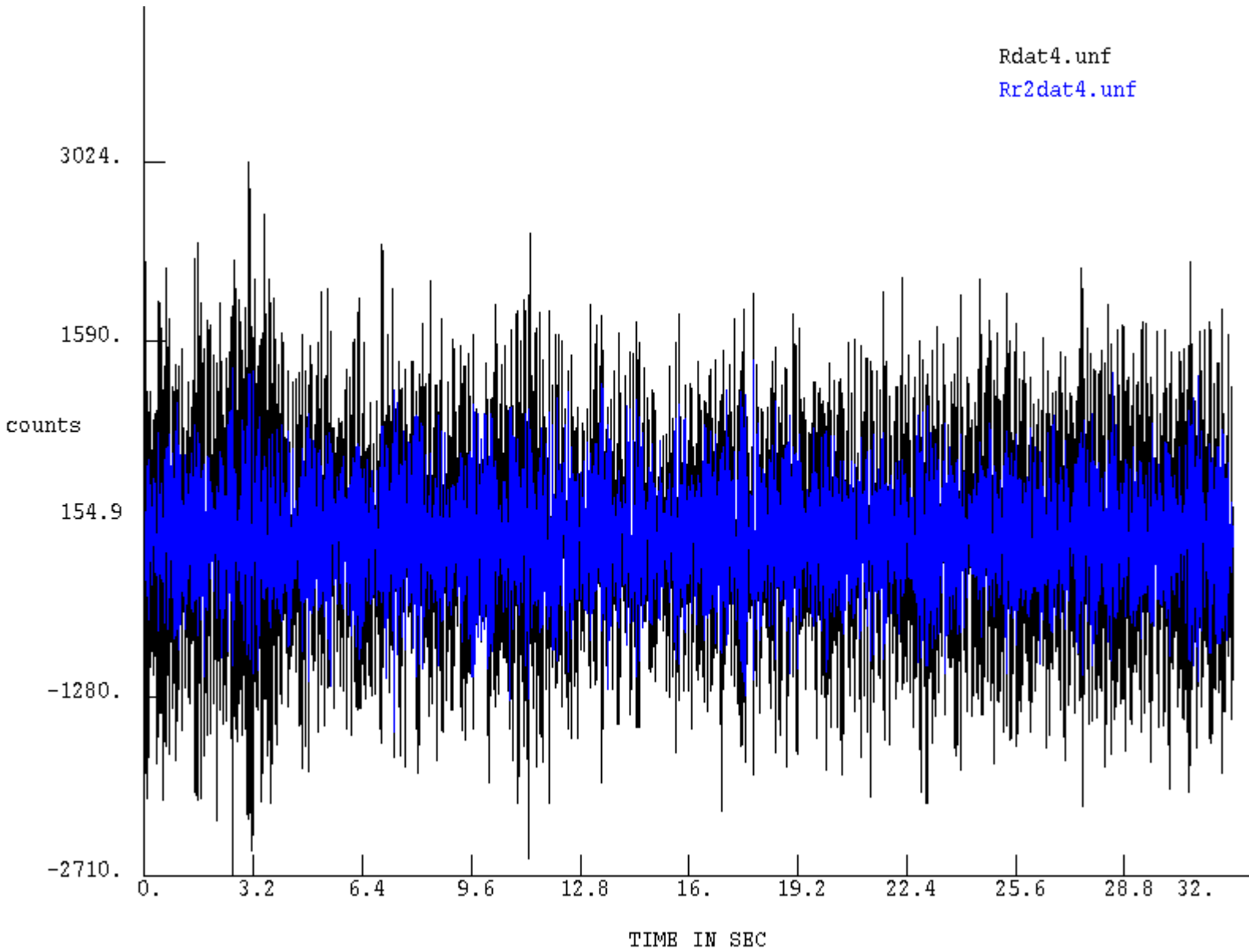
Rr2artvar.unf

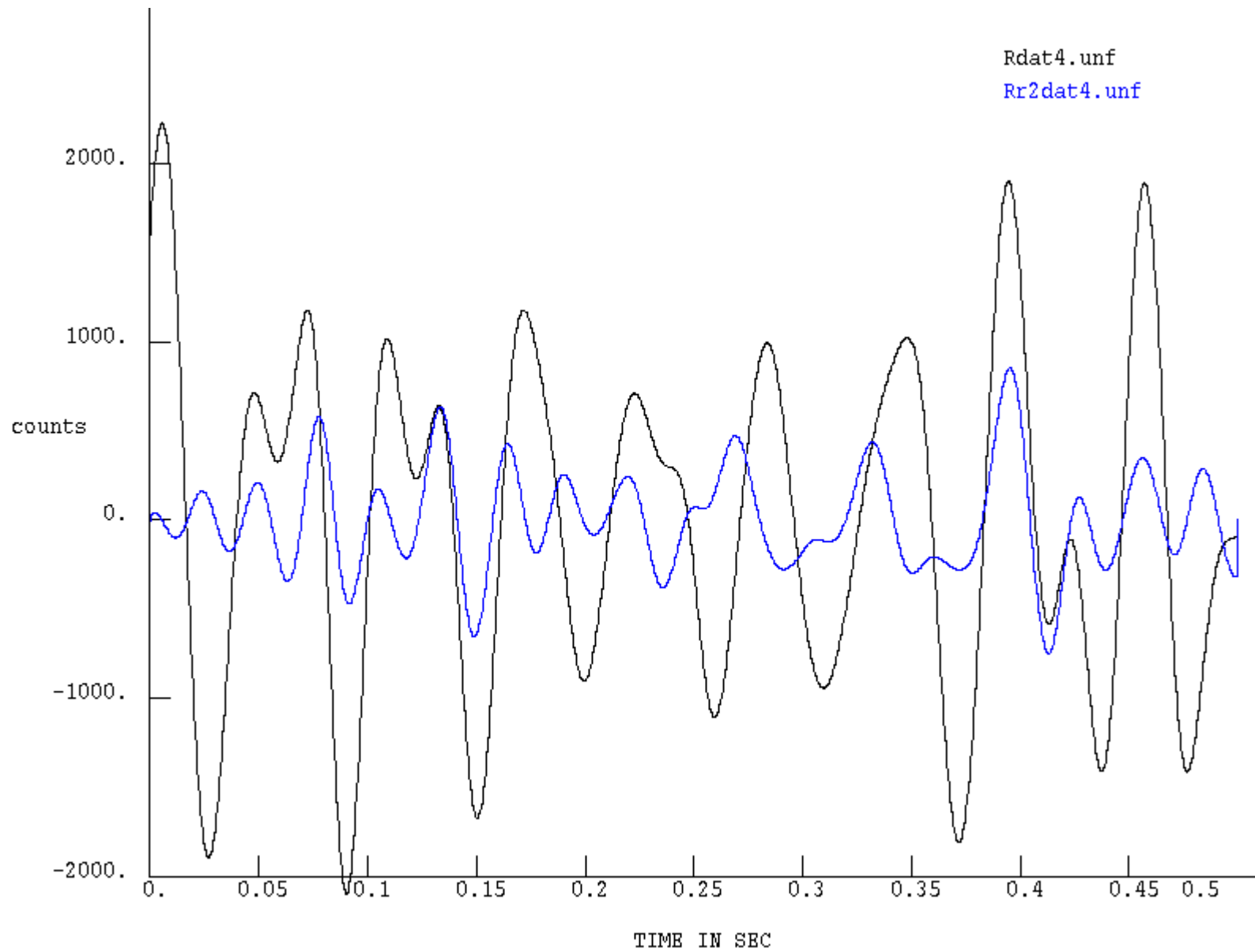


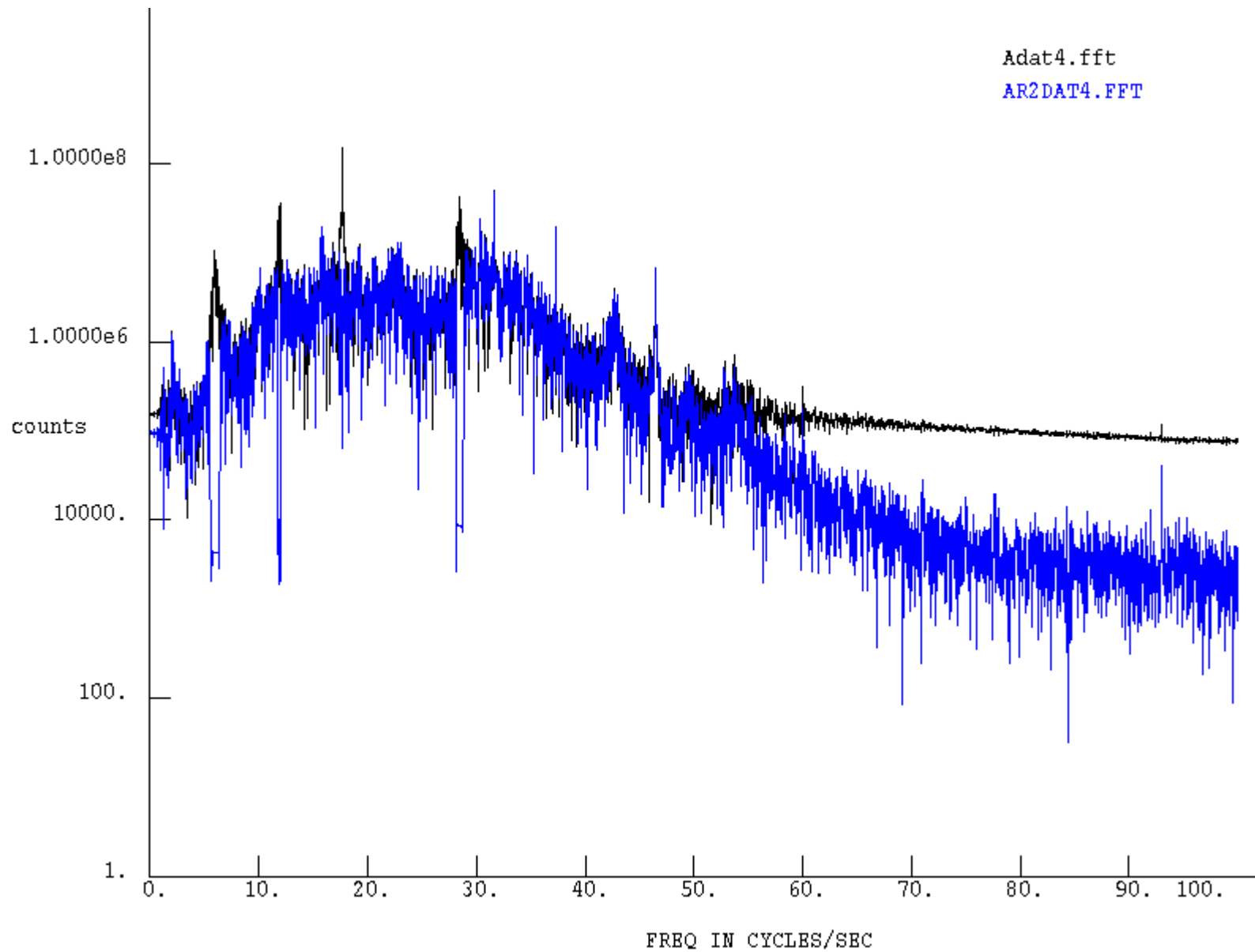


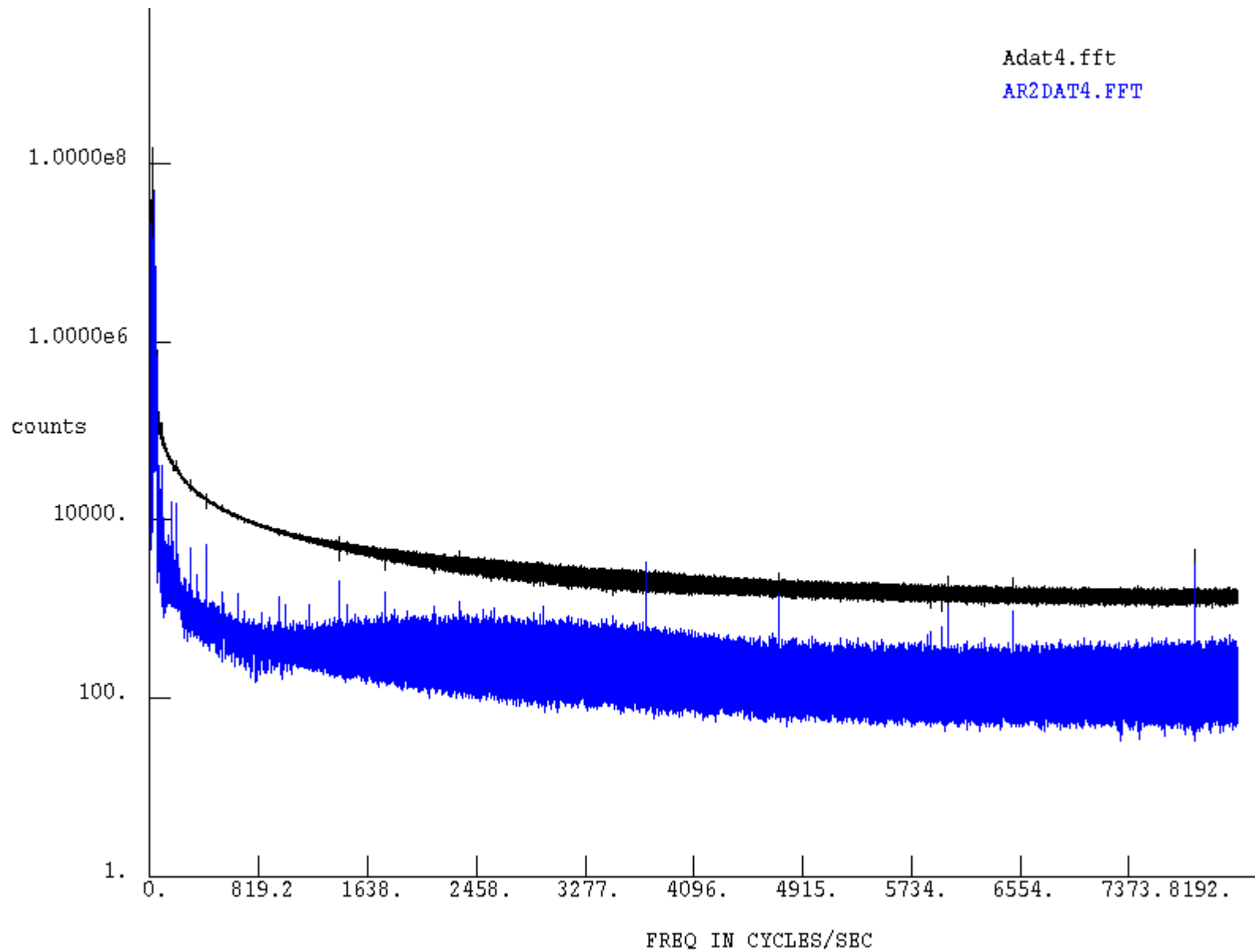
Amplitudes as fitted together.

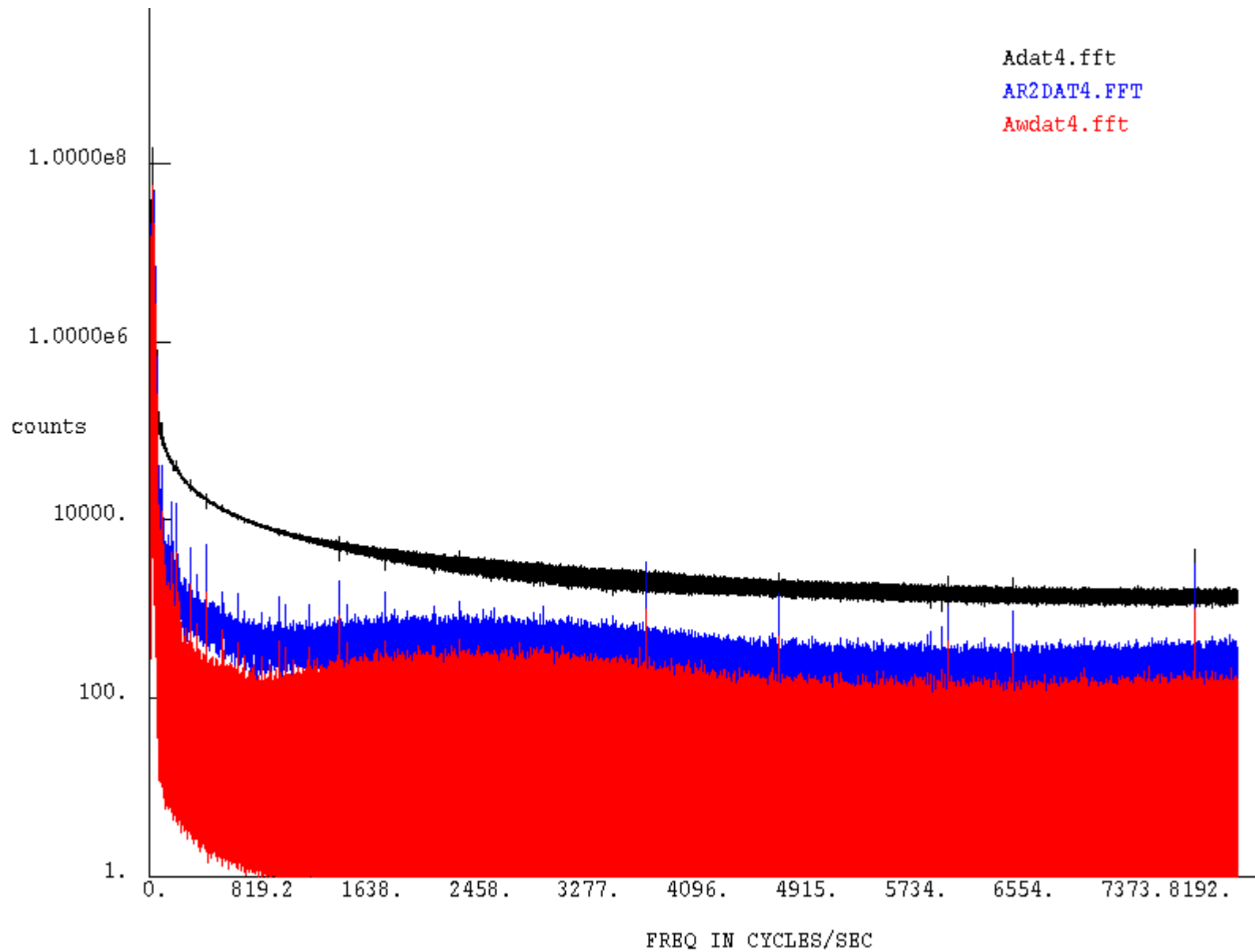
This data is from a locked, but noisy section of data on 02/09/2002.

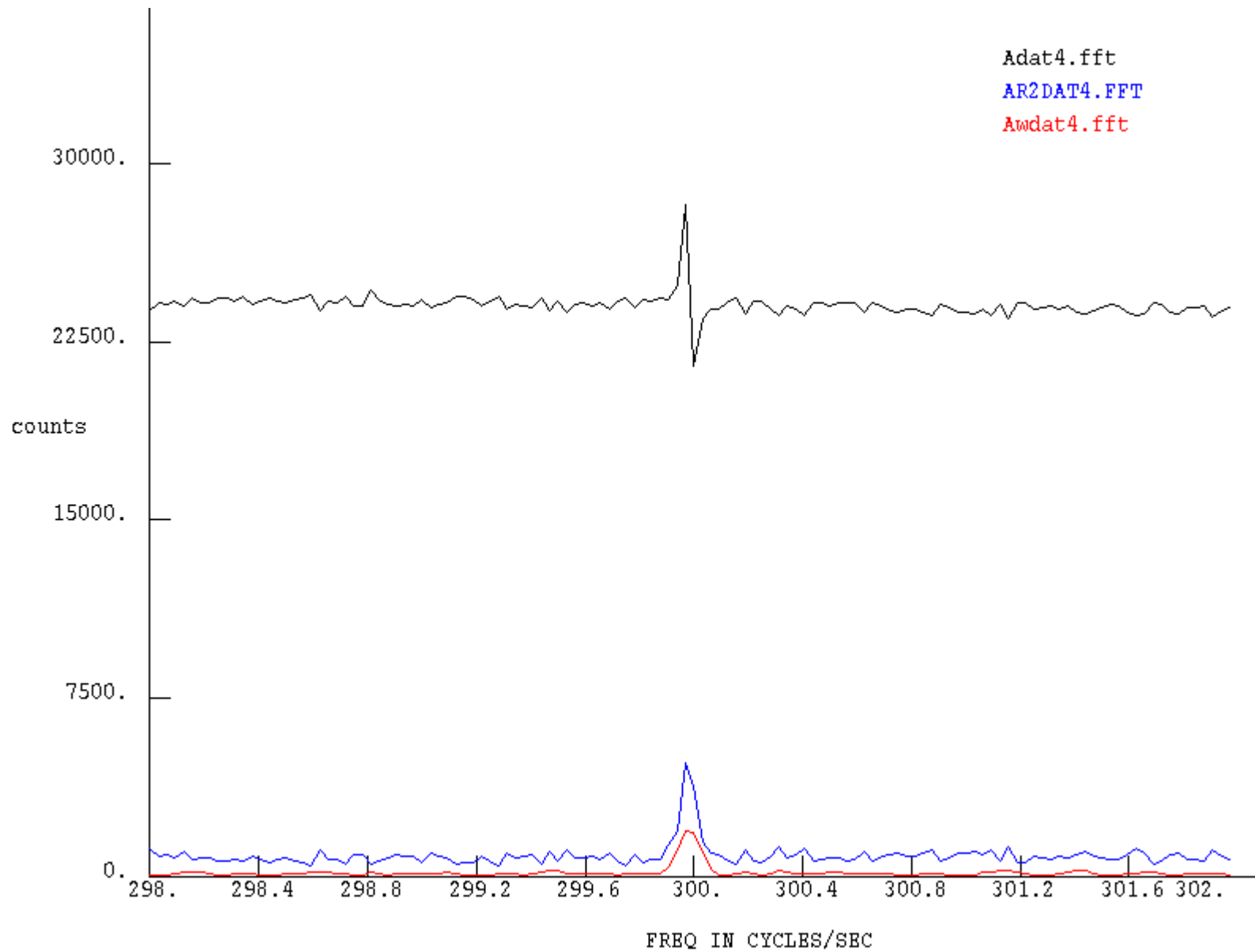


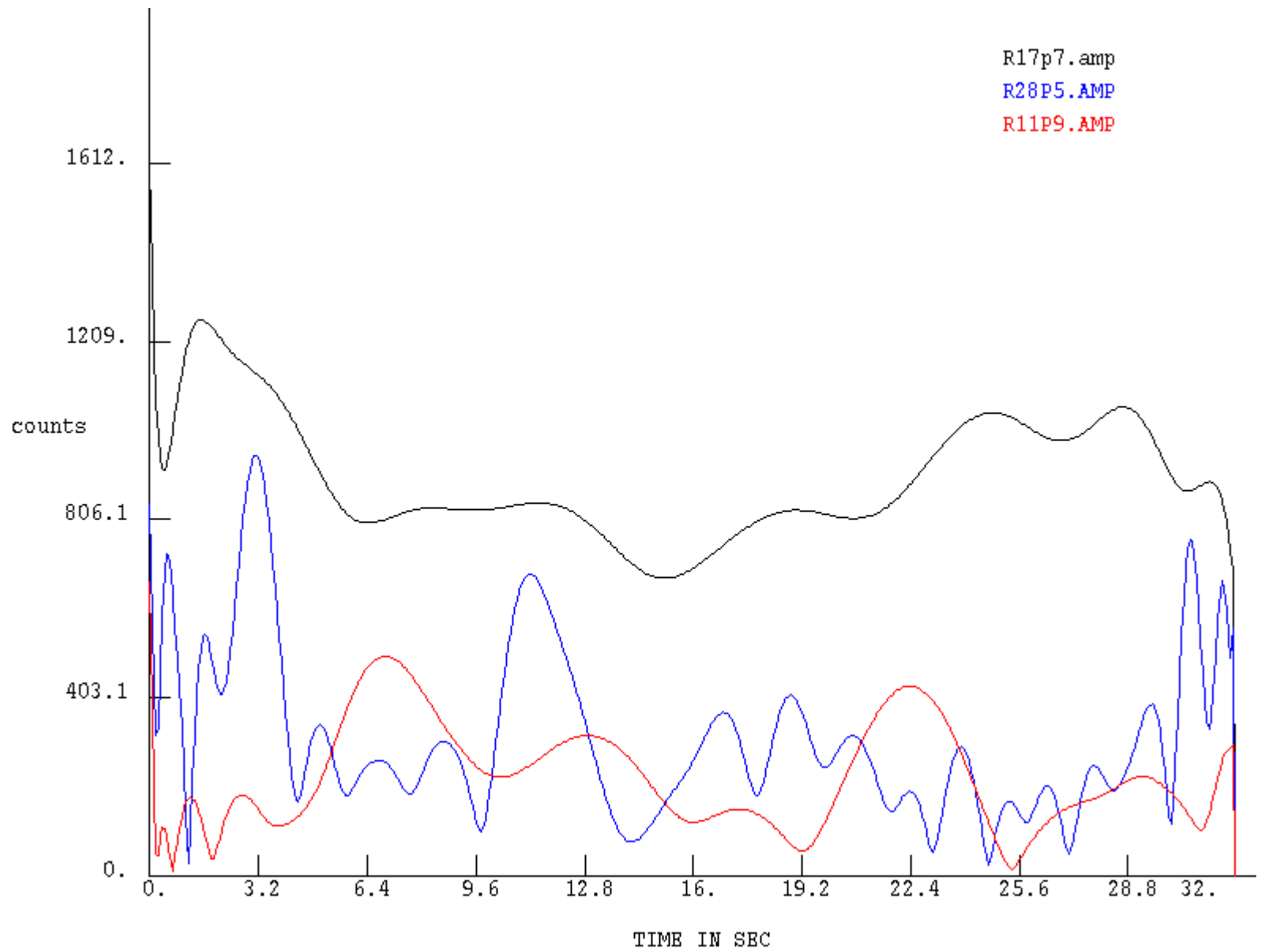




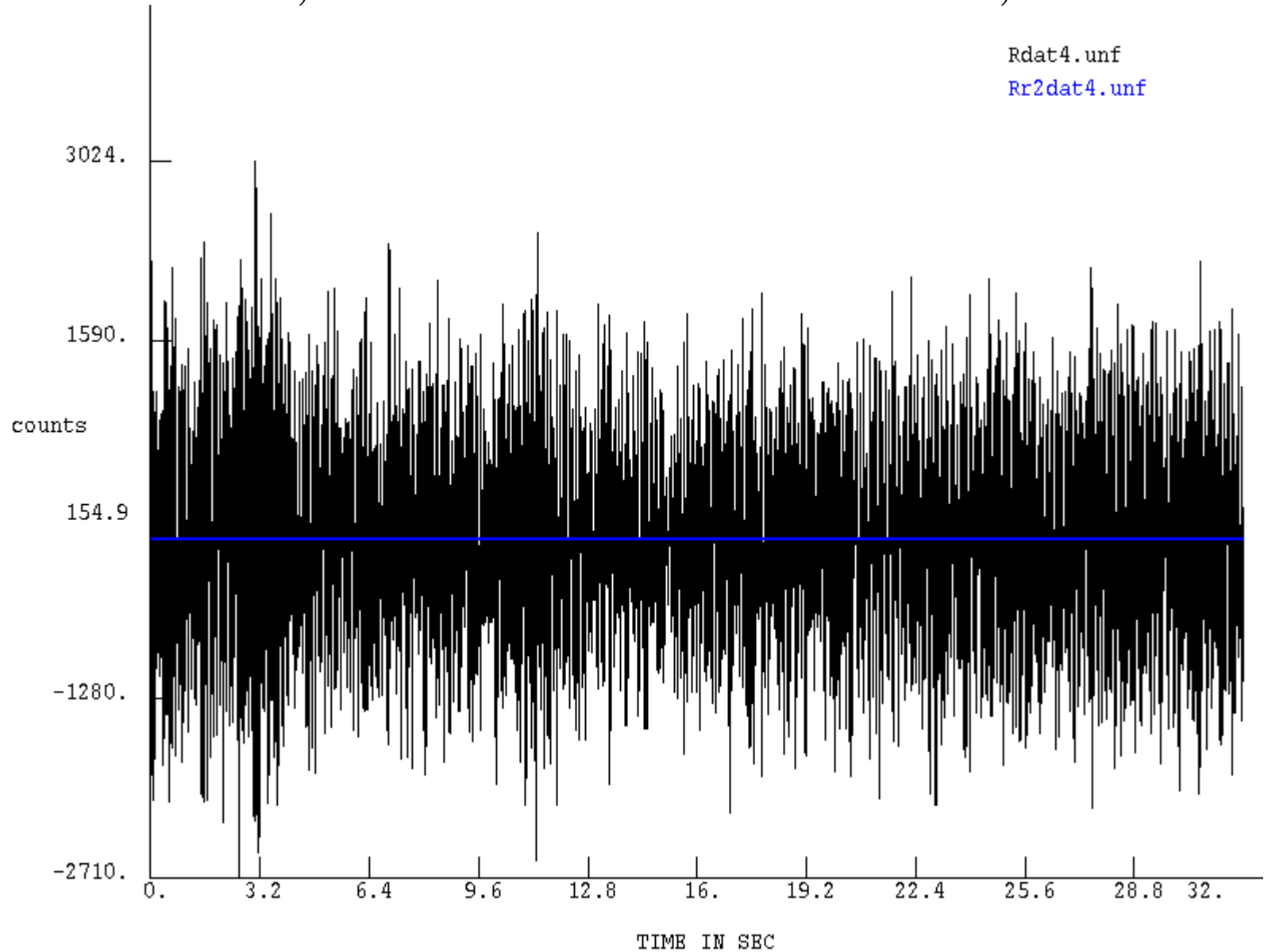






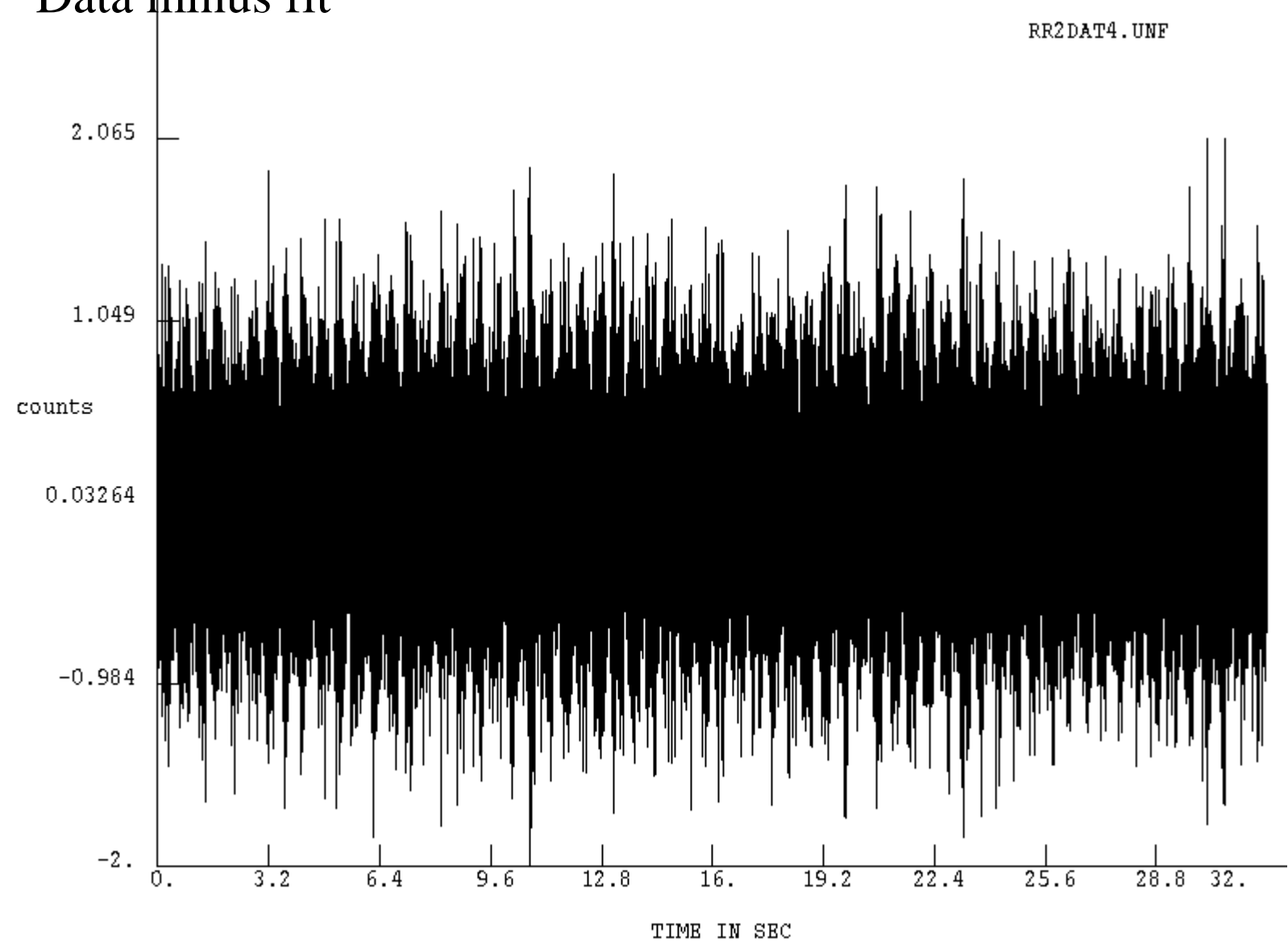


120 constants, half second intervals. Data black, data-fit blue

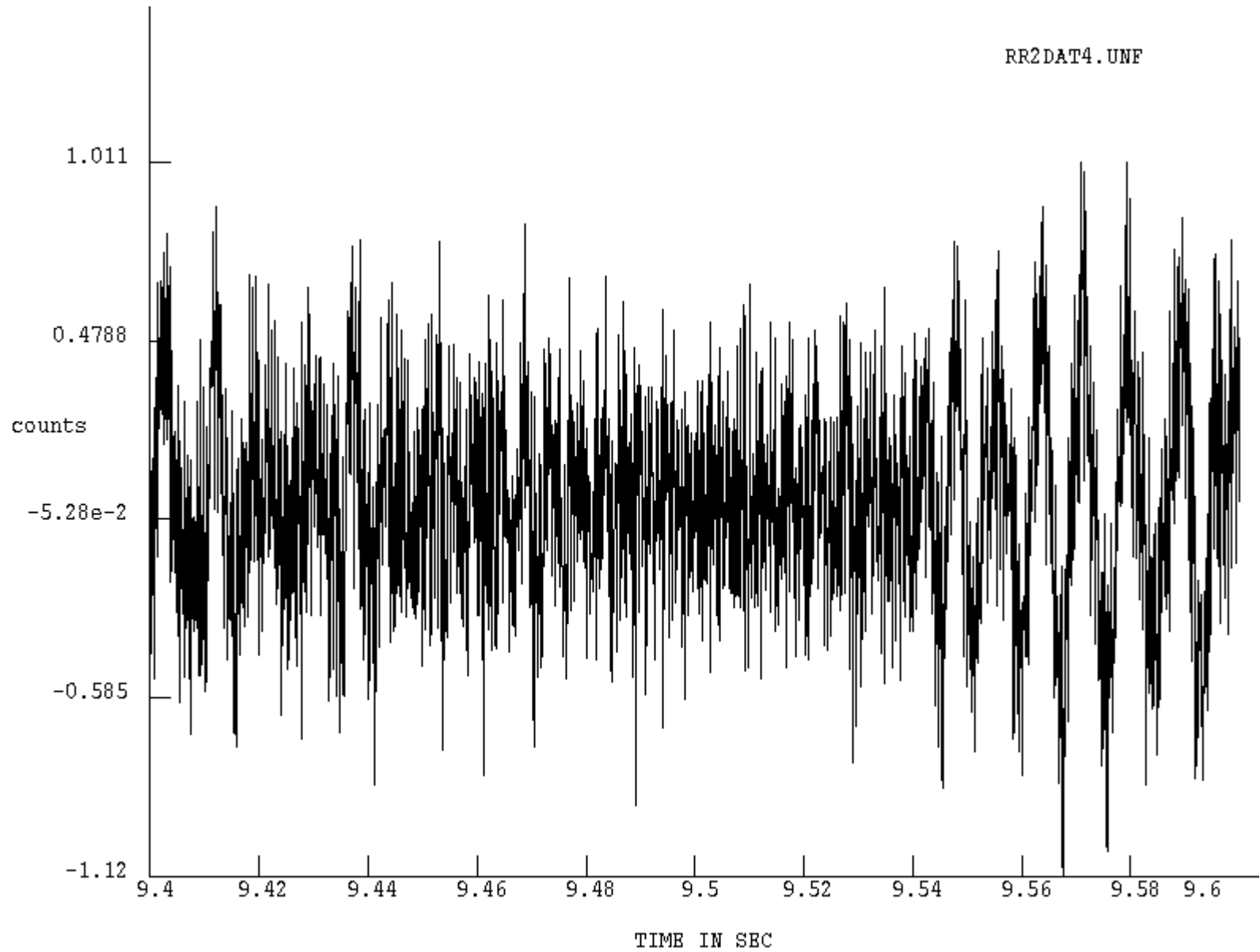


Data minus fit

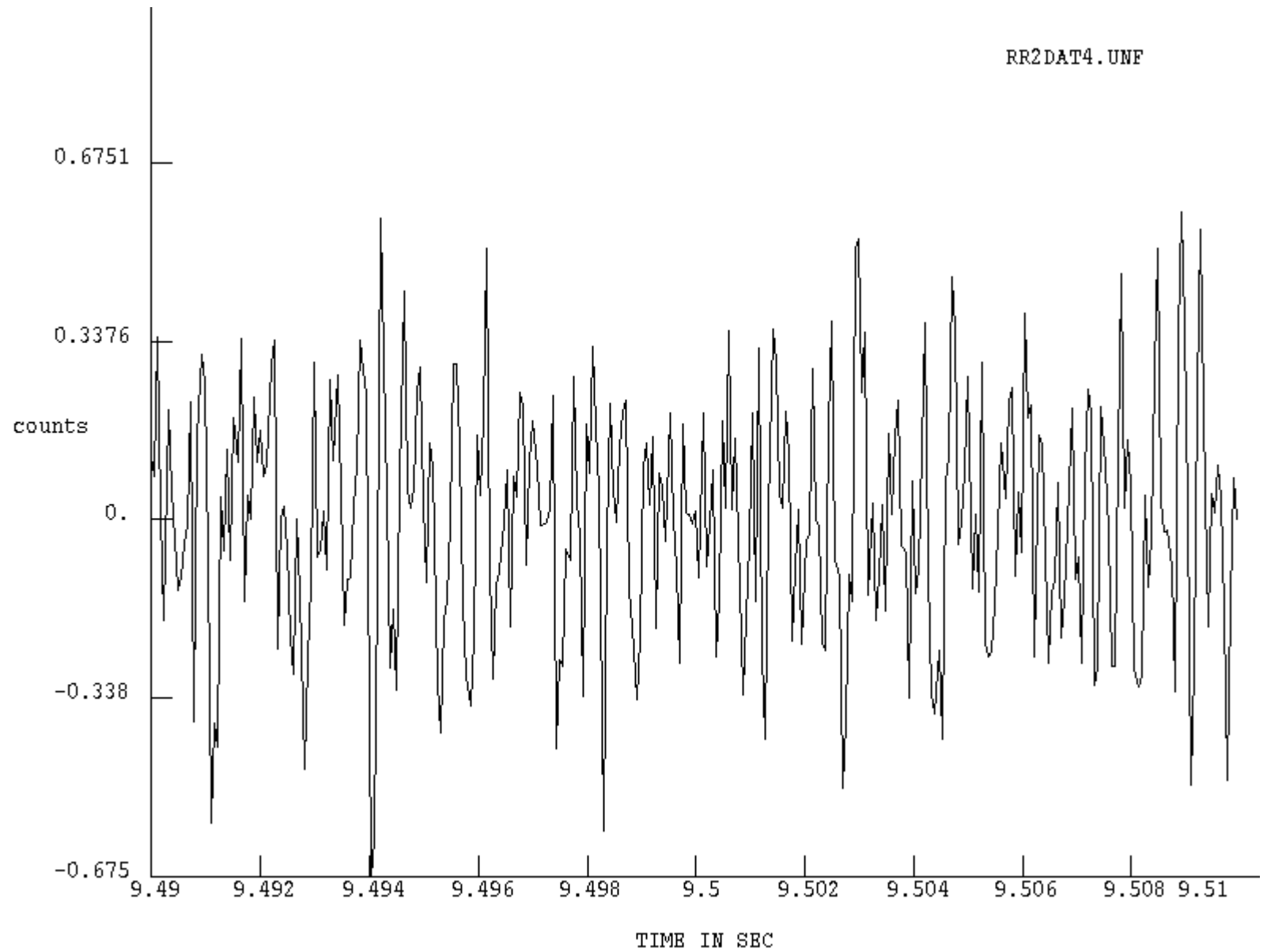
RR2DAT4.UNF



RR2DAT4.UNF



RR2DAT4.UNF



Uniform FFT of the data in black and data minus fit in blue.

