From SEDFIT to SEDPHAT

SEDPHAT for AUC *SV* Data Analysis for Equilibrium Systems

http://www.analyticalultracentrifugation.com

Ph.D. Chen Hui-Ju

IBC, AS

chenhj@gate.sinica.edu.tw

02-27855696 ext 5010

SEDFIT

- **Size distribution analysis**
- **protein self-association**
- **size-and-shape distributions**
- **macromolecules**
- **sedimentation velocity ultracentrifugation**
- **Lamm equation modeling**

and

SEDPHAT

- multi-site binary and ternary protein interactions
- heterogeneous protein-protein interactions
- multi-protein complexes by multi-signal
- **specific models**

Papers with applications of SEDFIT

- **Uncontrolled Zinc- and Copper-Induced Oligomerisation of the Human Complement Regulator Factor H and Its Possible Implications for Function and Disease.** *Journal of Molecular Biology*, *Volume 384, Issue 5*, *31 December 2008*, *Pages 1341-1352.* Ruodan Nan, Jayesh Gor, Imre Lengyel, Stephen J. Perkins
- **Specificity and Reactivity in Menaquinone Biosynthesis: The Structure of *Escherichia coli* MenD (2-Succinyl-5-Enolpyruvyl-6-Hydroxy-3-Cyclohexadiene-1-Carboxylate Synthase).** *Journal of Molecular Biology*, *Volume 384, Issue 5*, *31 December 2008*, *Pages 1353-1368.* Alice Dawson, Paul K. Fyfe, William N. Hunter
- **Evidence for the Oligomeric State of 'Elastic' Titin in Muscle Sarcomeres.** *Journal of Molecular Biology*, *Volume 384, Issue 2*, *12 December 2008*, *Pages 299-312.* Ahmed Houmeida, Andy Baron, Jeff Keen, G Nasir Khan, Peter J. Knight, Walter F. Stafford III, Kavitha Thirumurugan, Beatrix Thompson, Larissa Tskhovrebova, John Trinick
- **The High-Resolution NMR Structure of the Early Folding Intermediate of the *Thermus thermophilus* Ribonuclease H.** *Journal of Molecular Biology*, *Volume 384, Issue 2*, *12 December 2008*, *Pages 531-539* Zheng Zhou, Hanqiao Feng, Rodolfo Ghirlando, Yawen Bai
- **NUTS and BOLTS: Applications of fluorescence-detected sedimentation.** *Analytical Biochemistry*, ***In Press, Uncorrected Proof***, *Available online 6 December 2008.* Rachel R. Kroe, Thomas M. Laue
- ***Helicobacter pylori* neutrophil-activating protein promotes myeloperoxidase release from human neutrophils.** *Biochemical and Biophysical Research Communications*, *Volume 377, Issue 1*, *5 December 2008*, *Pages 52-56.* Chung-An Wang, Yen-Chun Liu, Shin-Yi Du, Chio-Wen Lin, Hua-Wen Fu
- **The family 52 β-xylosidase from *Geobacillus stearothermophilus* is a dimer: Structural and biophysical characterization of a glycoside hydrolase.** *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics*, *Volume 1784, Issue 12*, *December 2008*, *Pages 1924-1934*

Papers with applications of SEDPHAT

- **Solution Structure of the Complex Formed between Human Complement C3d and Full-length Complement Receptor Type 2.** *Journal of Molecular Biology*, *Volume 384, Issue 1*, *5 December 2008*, *Pages 137-150.* Keying Li, Azubuike I. Okemefuna, Jayesh Gor, Jonathan P. Hannan, Rengasamy Asokan, V. Michael Holers, Stephen J. Perkins
- **Role of tryptophan-208 residue in cytochrome *c* oxidation by ascorbate peroxidase from *Leishmania major*-kinetic studies on Trp208Phe mutant and wild type enzyme**. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics*, *Volume 1784, Issue 5*, *May 2008*, *Pages 863-871.* Rajesh K. Yadav, Subhankar Dolai, Swati Pal, Subrata Adak
- **The dimeric assembly of *Photobacterium leiognathi* and *Salmonella typhimurium* SodC1 Cu,Zn superoxide dismutases is affected differently by active site demetallation and pH: An analytical ultracentrifuge study.** *Archives of Biochemistry and Biophysics*, *Volume 471, Issue 1*, *1 March 2008*, *Pages 77-84.* B. Catacchio, M. D'Orazio, A. Battistoni, E. Chiancone
- **tRNA-dependent asparagine formation in prokaryotes: Characterization, isolation and structural and functional analysis of a ribonucleoprotein particle generating Asn-tRNA^{Asn}.** *Methods*, *Volume 44, Issue 2*, *February 2008*, *Pages 146-163.* Marc Bailly, Mickaël Blaise, Hervé Roy, Marzanna Deniziak, Bernard Lorber, Catherine Birck, Hubert D. Becker, Daniel Kern
- **Characterization and Further Stabilization of Designed Ankyrin Repeat Proteins by Combining Molecular Dynamics Simulations and Experiments.** *Journal of Molecular Biology*, *Volume 375, Issue 3*, *18 January 2008*, *Pages 837-854.* Gianluca Interlandi, Svava K. Wetzel, Giovanni Settanni, Andreas Plückthun, Amedeo Caflisch
- **On the quaternary structure of a C-type lectin from *Bothrops jararacussu* venom – BJ-32 (BjcuL).** *Toxicon*, *Volume 52, Issue 8*, *15 December 2008*, *Pages 944-953.* F.P. Silva Jr., G.M.C. Alexandre, C.H.I. Ramos, S.G. De-Simone

Step by step

Content

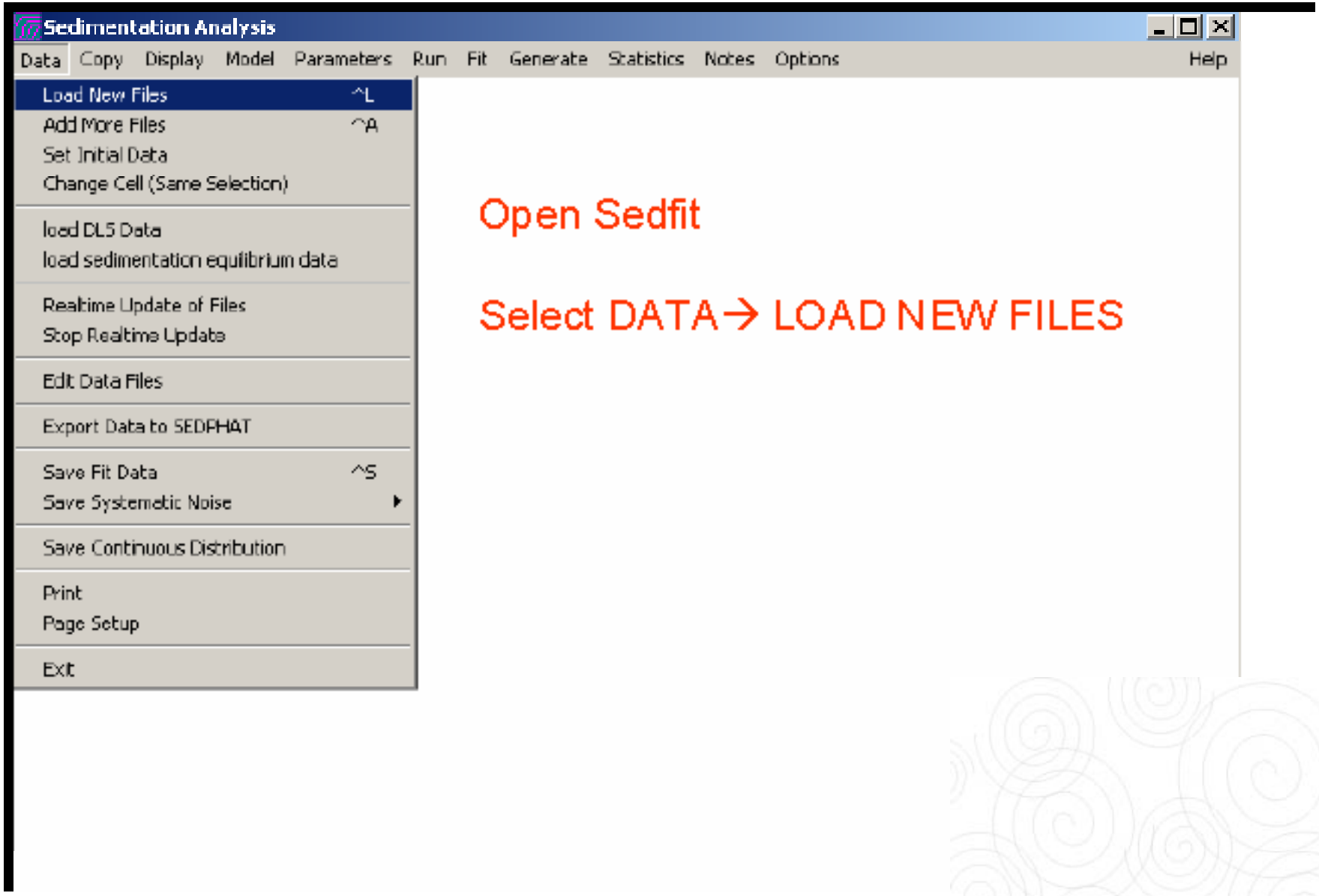
- SedFit C(S) analysis
- SedFit C(M) analysis
- SedFit C(f/fo) analysis
- Transform to SEDPHAT
- SEDPHAT models :

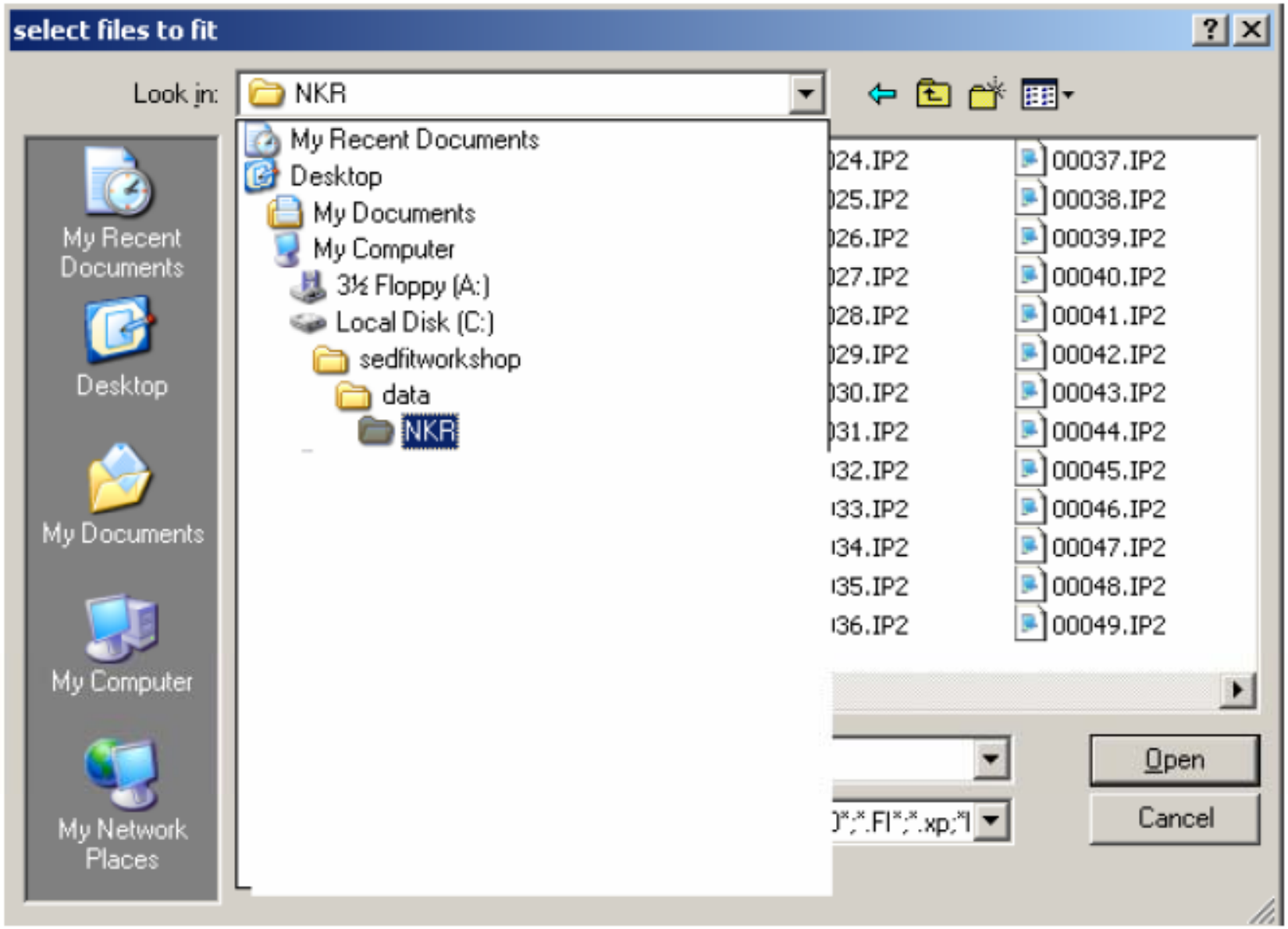
(1) Single species & heterogenesis

(2) Monomer to dimer

Parameter setup

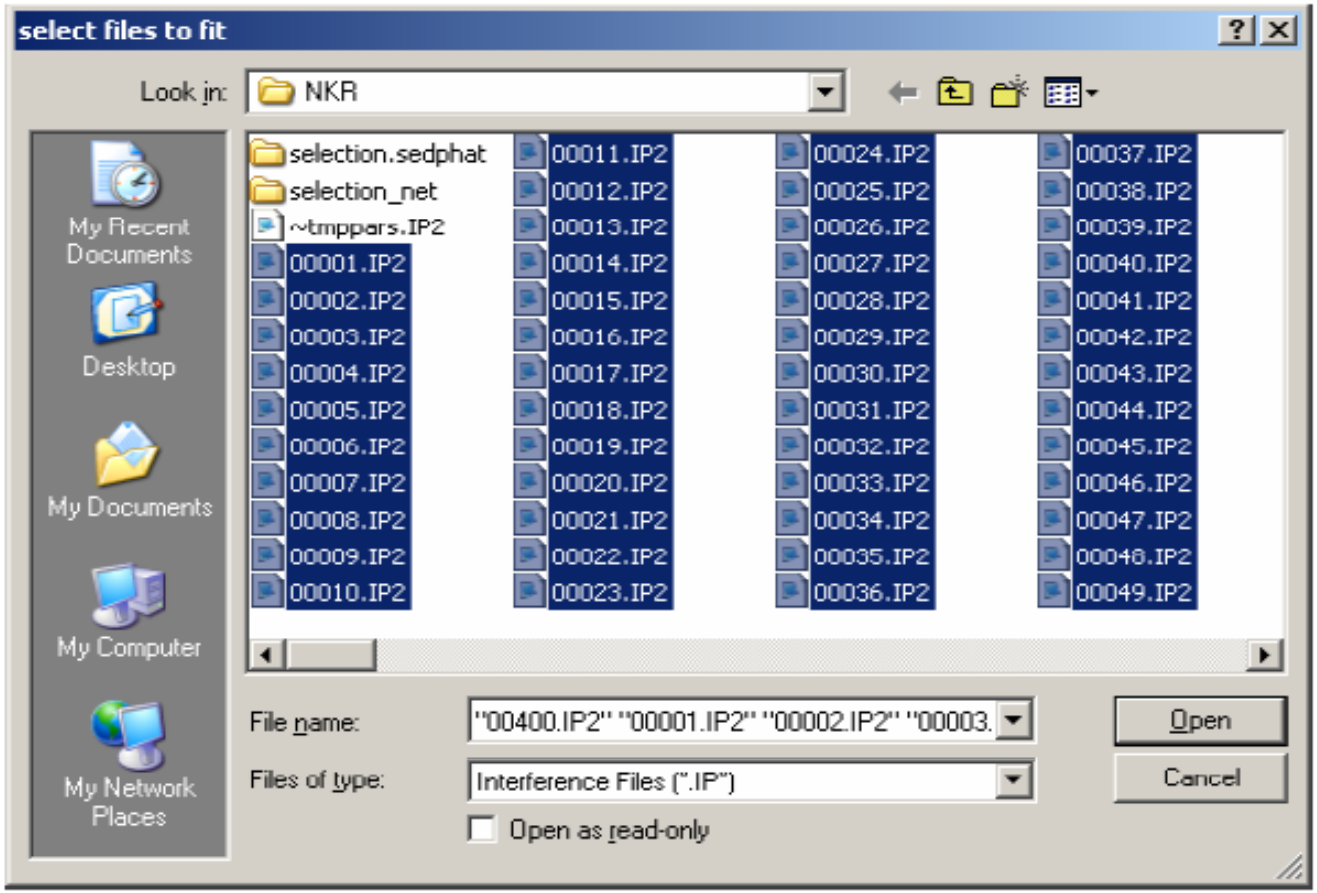
- Buffer density and viscosity
- Protein sample v-bar
- Rmsd
- Z value
- Boundary observation
- Statistics analysis


Sedimentation Analysis
Data
Copy
Display
Model
Parameters
Run
Fit
Generate
Statistics
Notes
Options
Help
Load New Files ^L
Add More Files ^A
Set Initial Data
Change Cell (Same Selection)
load DLS Data
load sedimentation equilibrium data
Realtime Update of Files
Stop Realtime Update
Edit Data Files
Export Data to SEDPHAT
Save Fit Data ^S
Save Systematic Noise
Save Continuous Distribution
Print
Page Setup
Exit
Open Sedfit
Select DATA→ LOAD NEW FILES

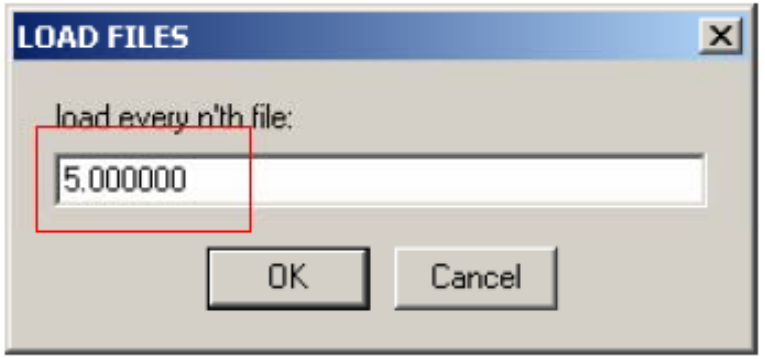


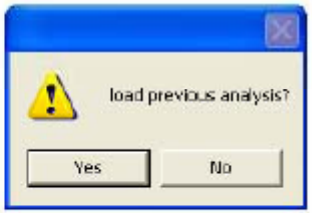
Choose scans within the NKR folder

Select all the interference scans for Cell 2 (highlight scans 1-400)



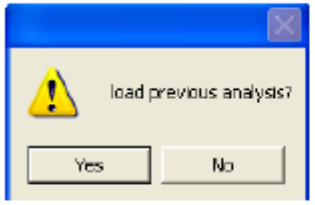
and load every 5th scan





If the data has been loaded before, Sedfit recognizes this and asks if the user wants to use the same parameters as the previous analysis (Sedfit had automatically saved this information in a temporary parameters file "tmppars.ip2" located with the raw data.

Select "No"

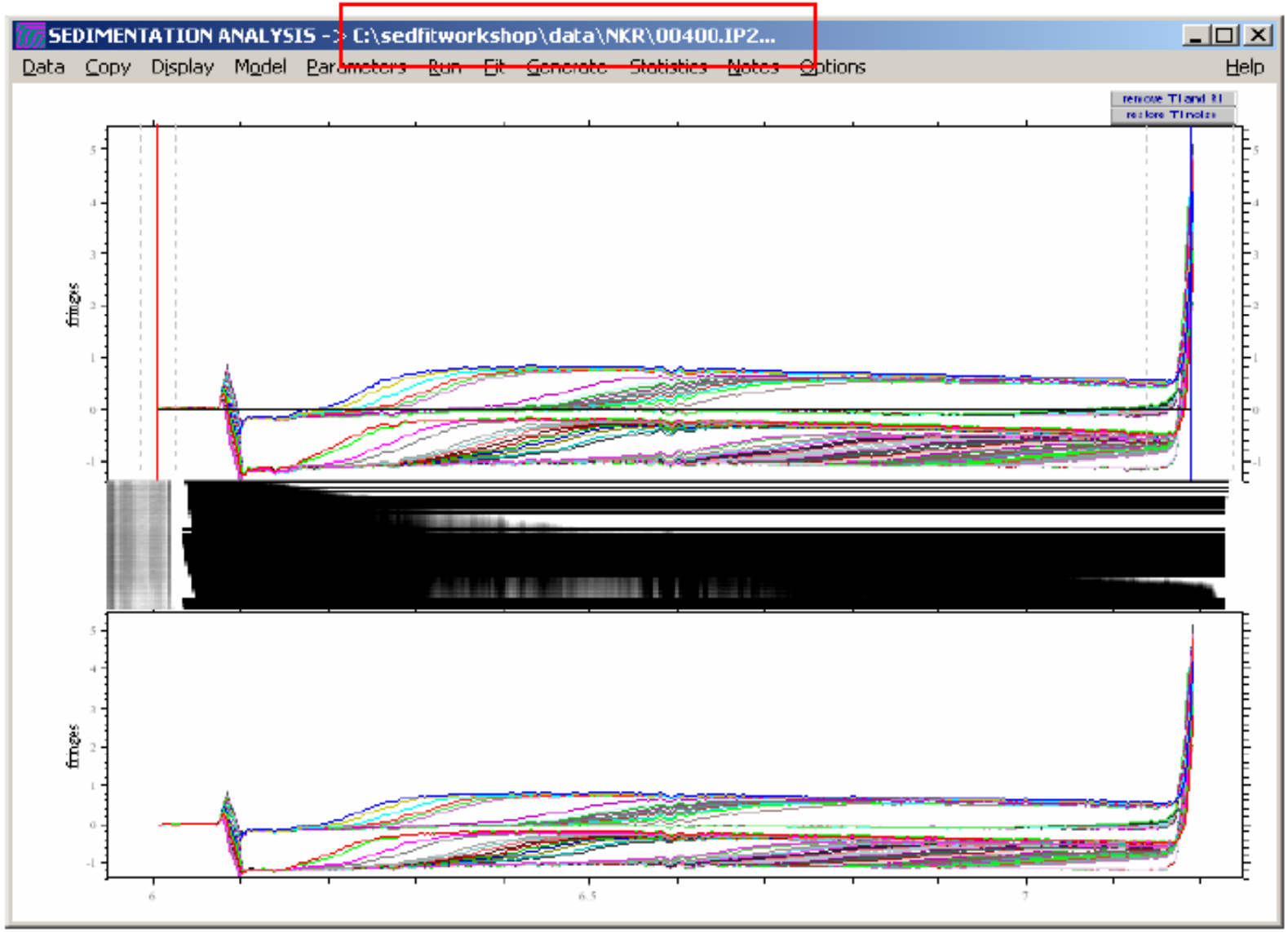


Selecting YES will load all the parameters that we used before (ie. Meniscus, bottom, fitting limits,etc.)

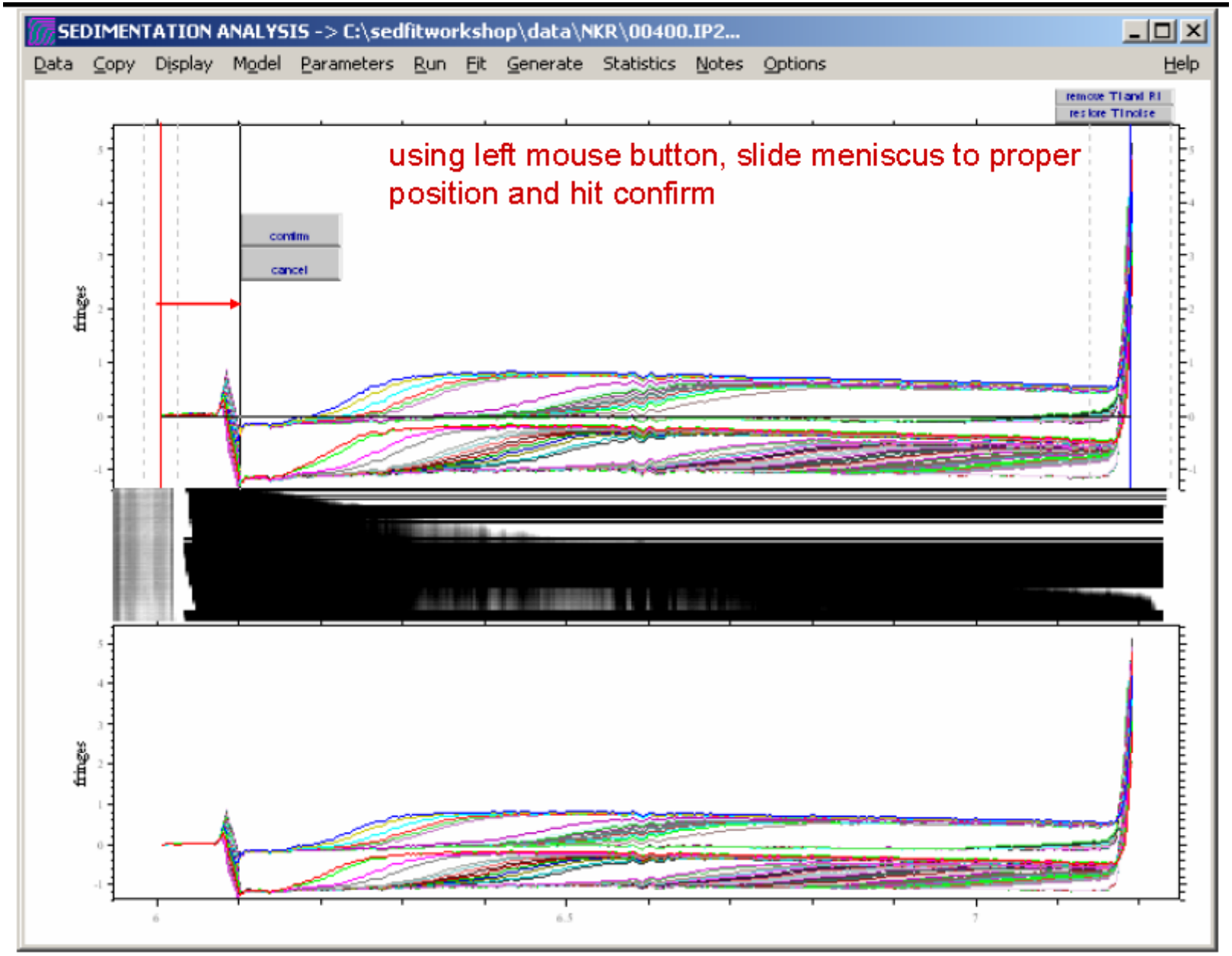
Selecting NO will result in the user needing to set these parameters again.

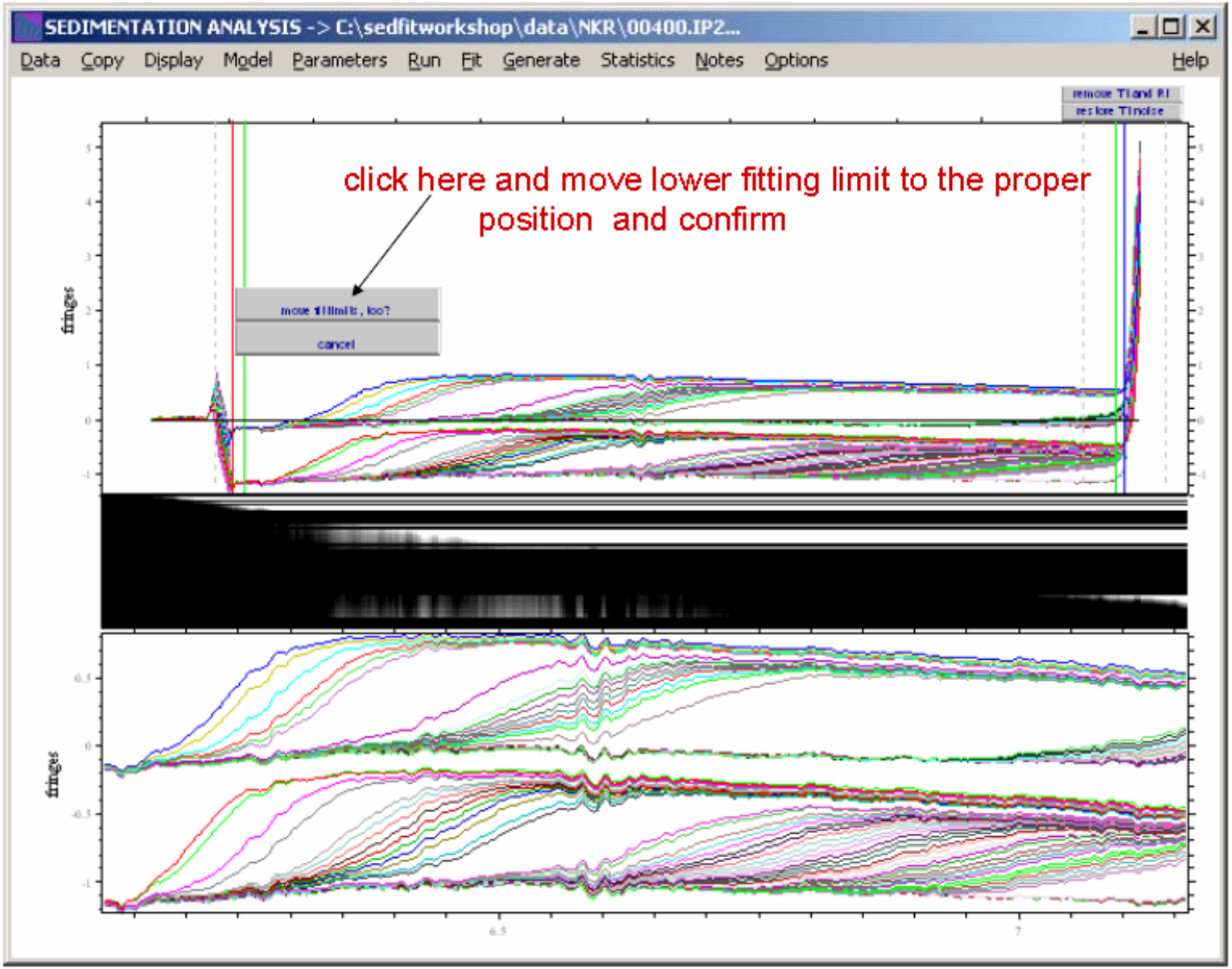
Select YES

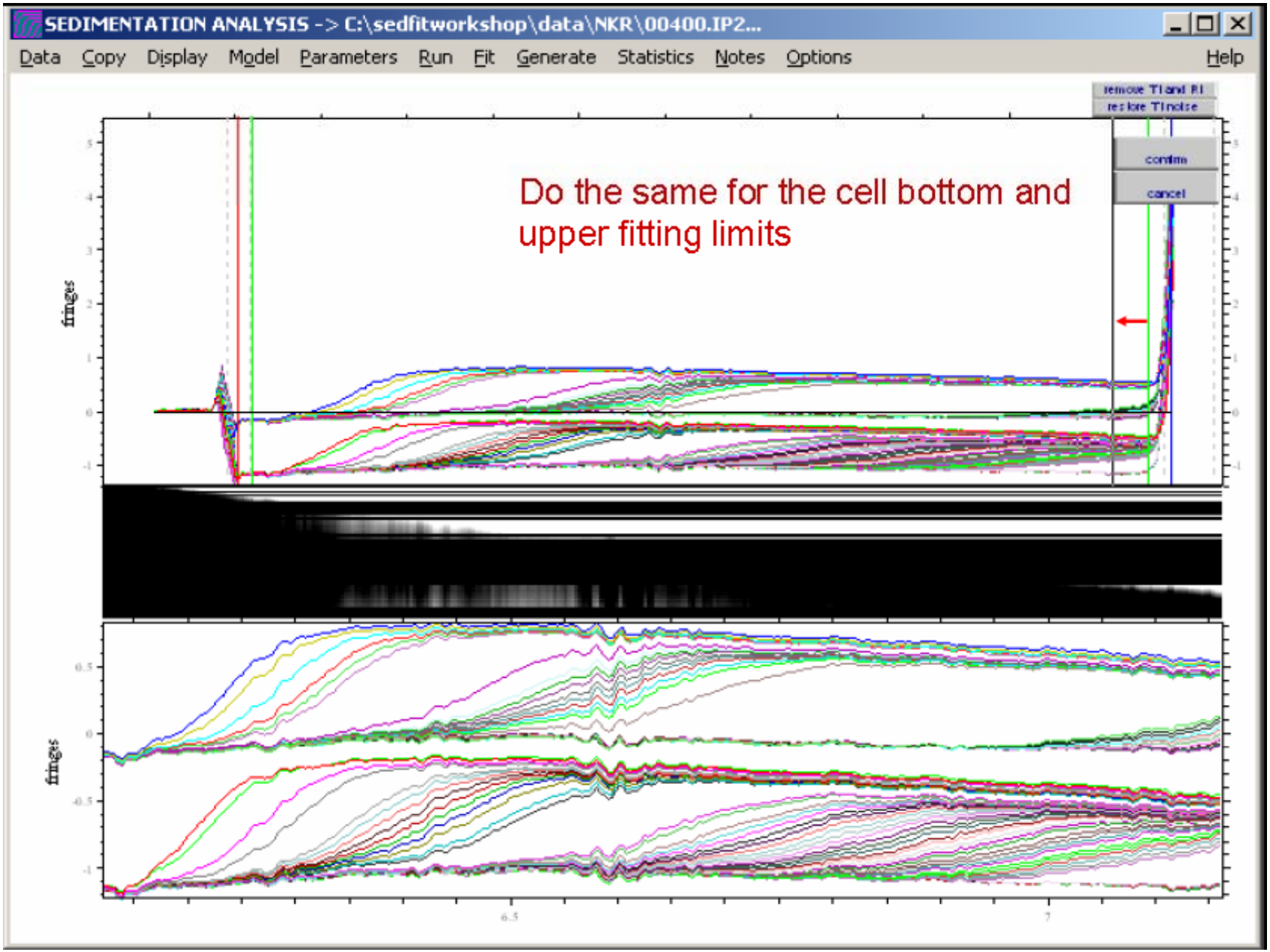
Every 5th scan loaded.

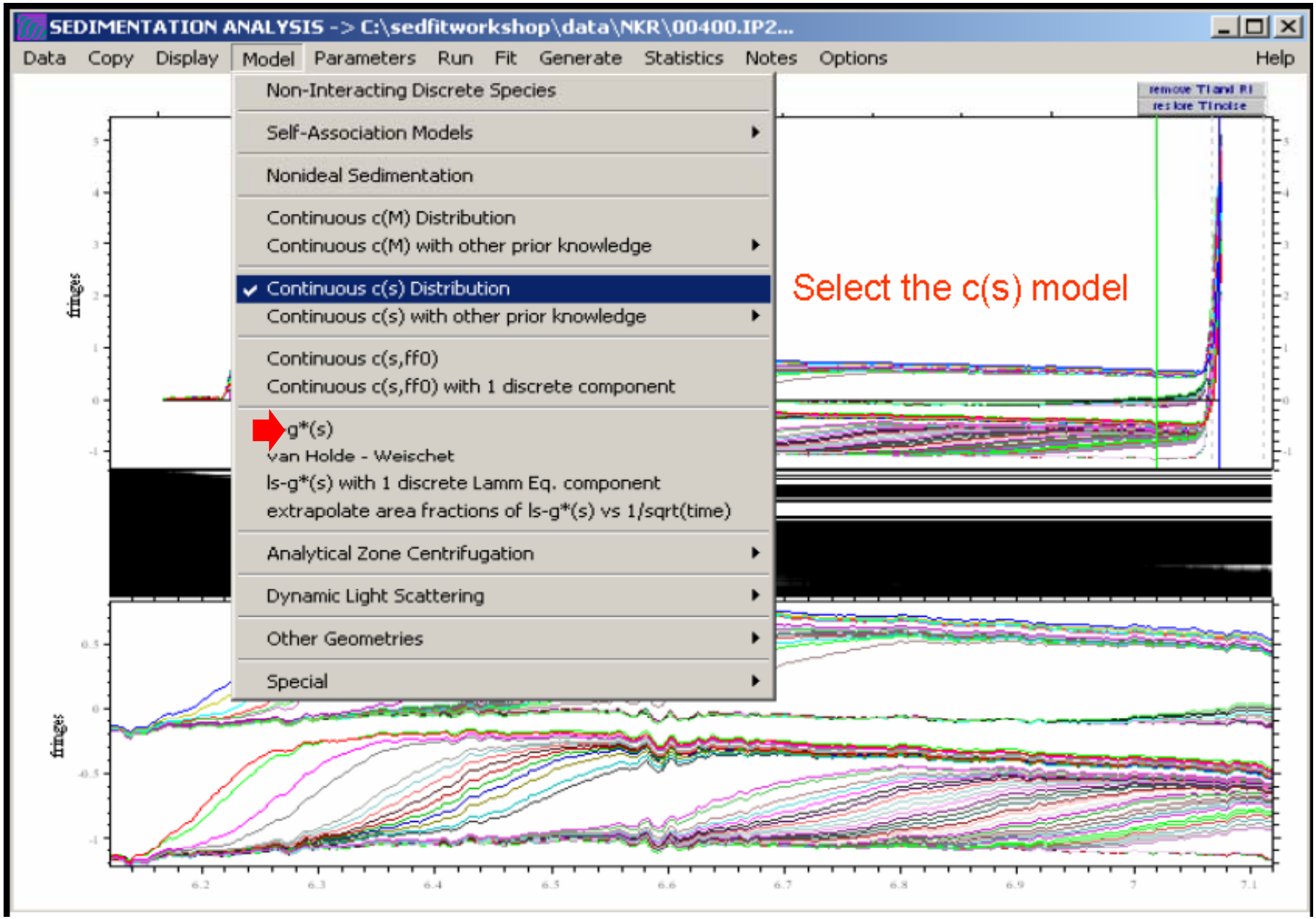


Next, select the meniscus position

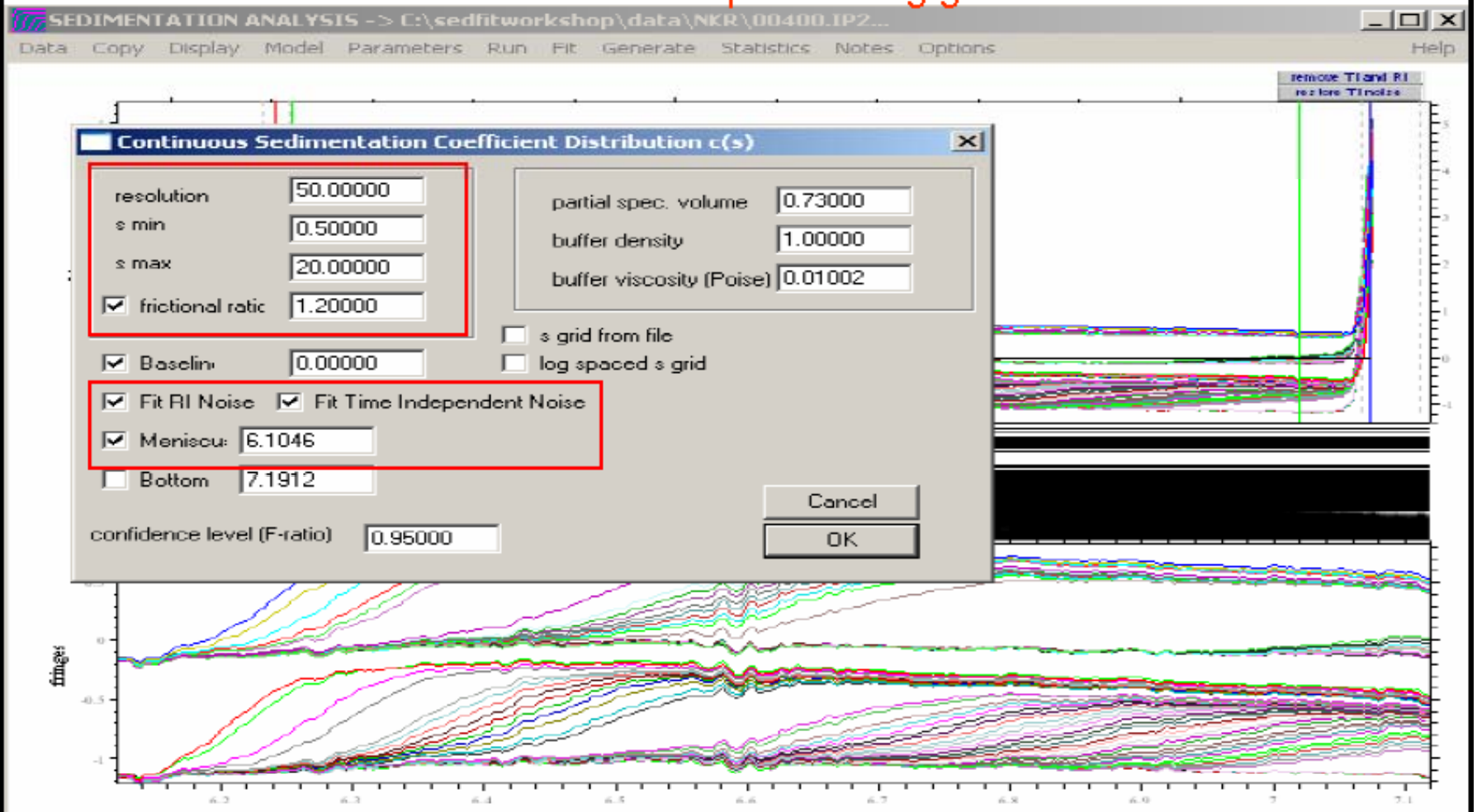

SEDIMENTATION ANALYSIS -> C:\sedfitworkshop\data\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
remove TI and P1
restore TI noise
using left mouse button, slide meniscus to proper position and hit confirm
confirm
cancel
fringes
fringes


SEDIMENTATION ANALYSIS -> C:\sedfitworkshop\data\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
remove TI and RI
restore TInoise
click here and move lower fitting limit to the proper position and confirm
move tlimits, too?
cancel
fringes
fringes
6.5
7


SEDIMENTATION ANALYSIS -> C:\sedfitworkshop\data\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
remove TI and RI
restore TI noise
confirm
cancel
Do the same for the cell bottom and upper fitting limits
fringes
fringes


SEDIMENTATION ANALYSIS -> C:\sedfitworkshop\data\NKR\00400.IP2...
Data
Copy
Display
Model
Parameters
Run
Fit
Generate
Statistics
Notes
Options
Help
Non-Interacting Discrete Species
Self-Association Models
Nonideal Sedimentation
Continuous c(M) Distribution
Continuous c(M) with other prior knowledge
Continuous c(s) Distribution
Continuous c(s) with other prior knowledge
Continuous c(s,ff0)
Continuous c(s,ff0) with 1 discrete component
g*(s)
van Holde - Weischet
ls-g*(s) with 1 discrete Lamm Eq. component
extrapolate area fractions of ls-g*(s) vs 1/sqrt(time)
Analytical Zone Centrifugation
Dynamic Light Scattering
Other Geometries
Special
Select the c(s) model
fringes
fringes

Select PARAMETERS and input starting guesses

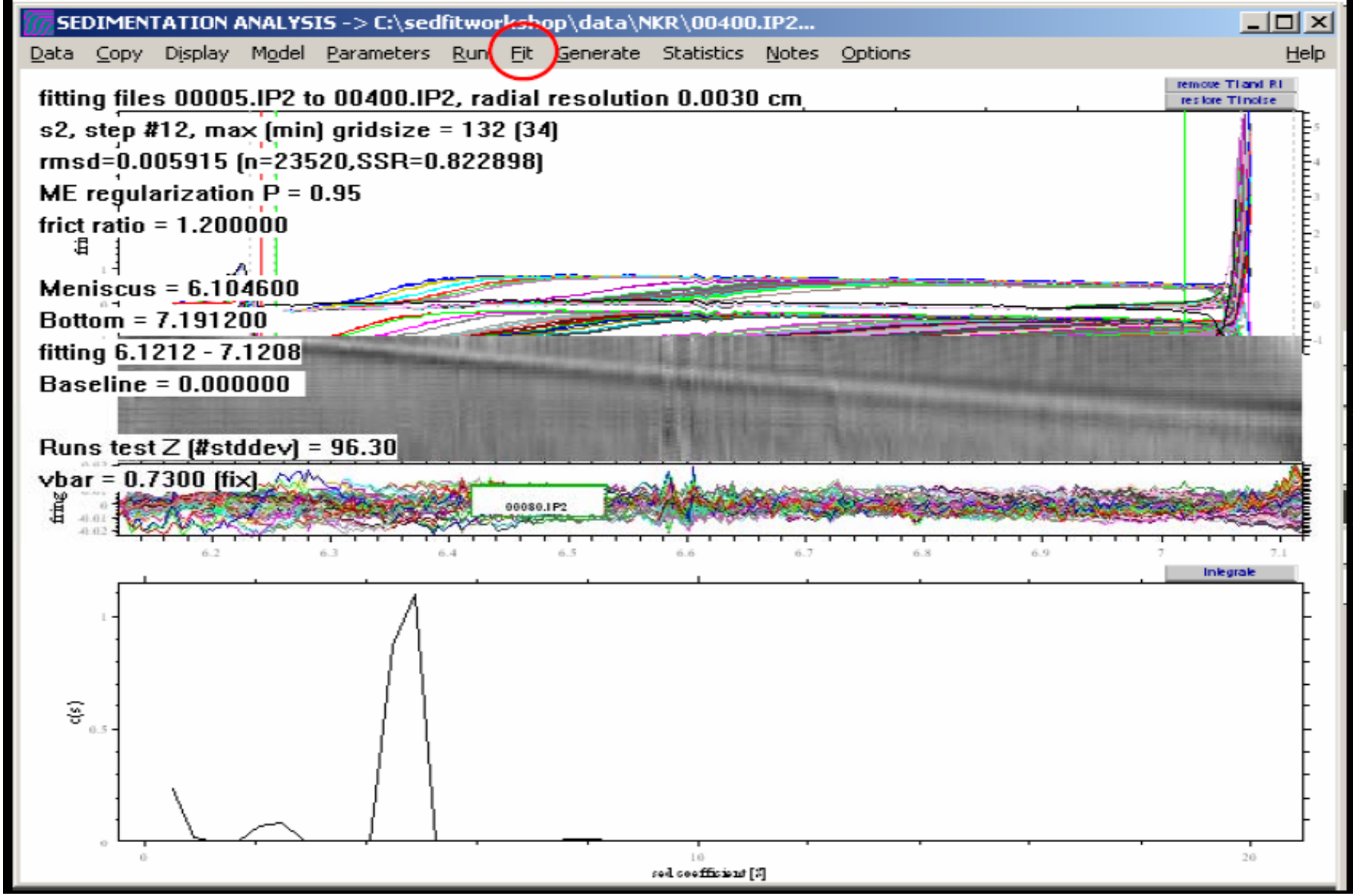


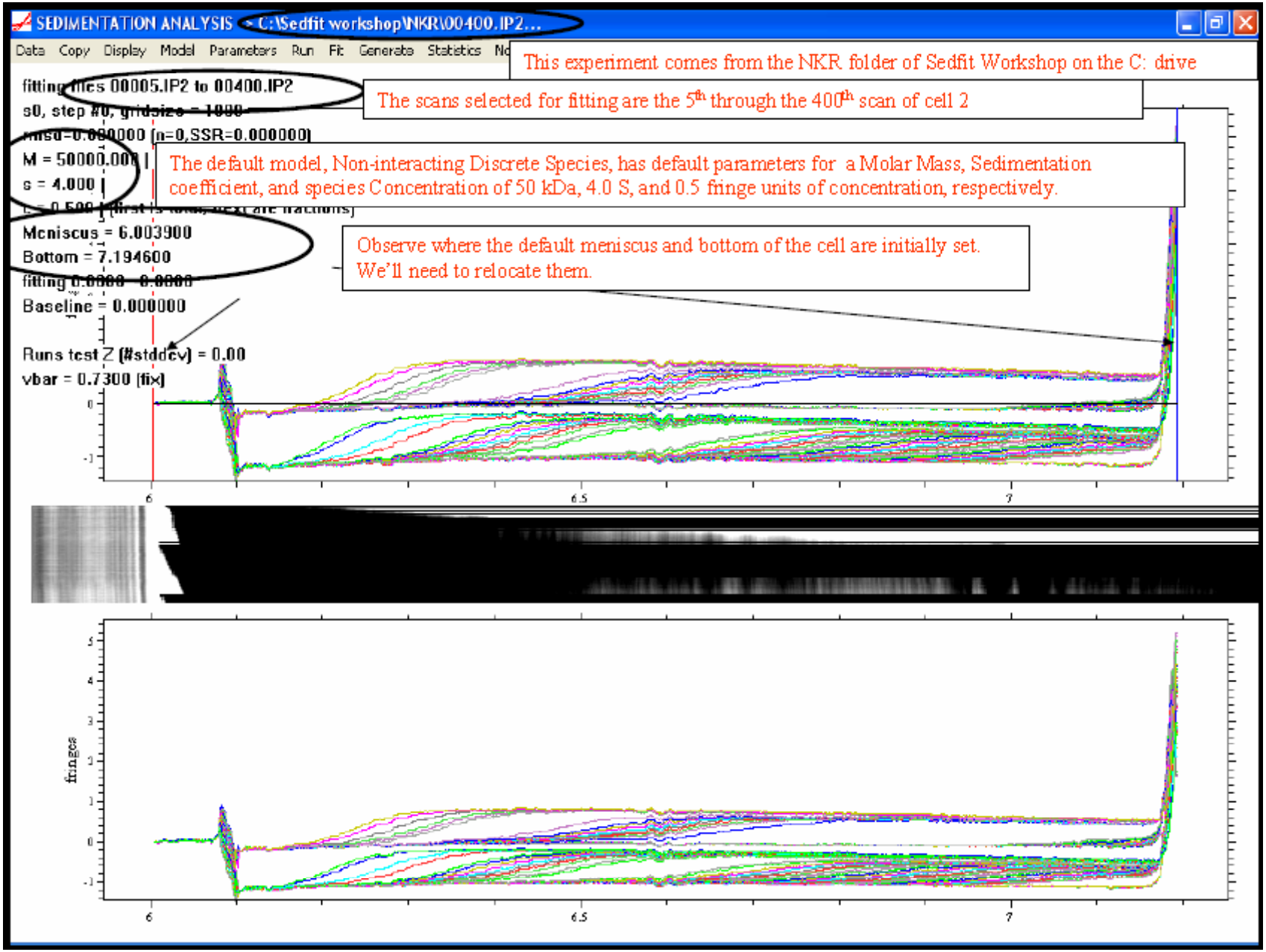
Here we are going to model the experimental data with simulated curves for 50 s-values between 0.5 and 20S.

We will float the frictional ratio Fit for TI and RI noise and Float the meniscus

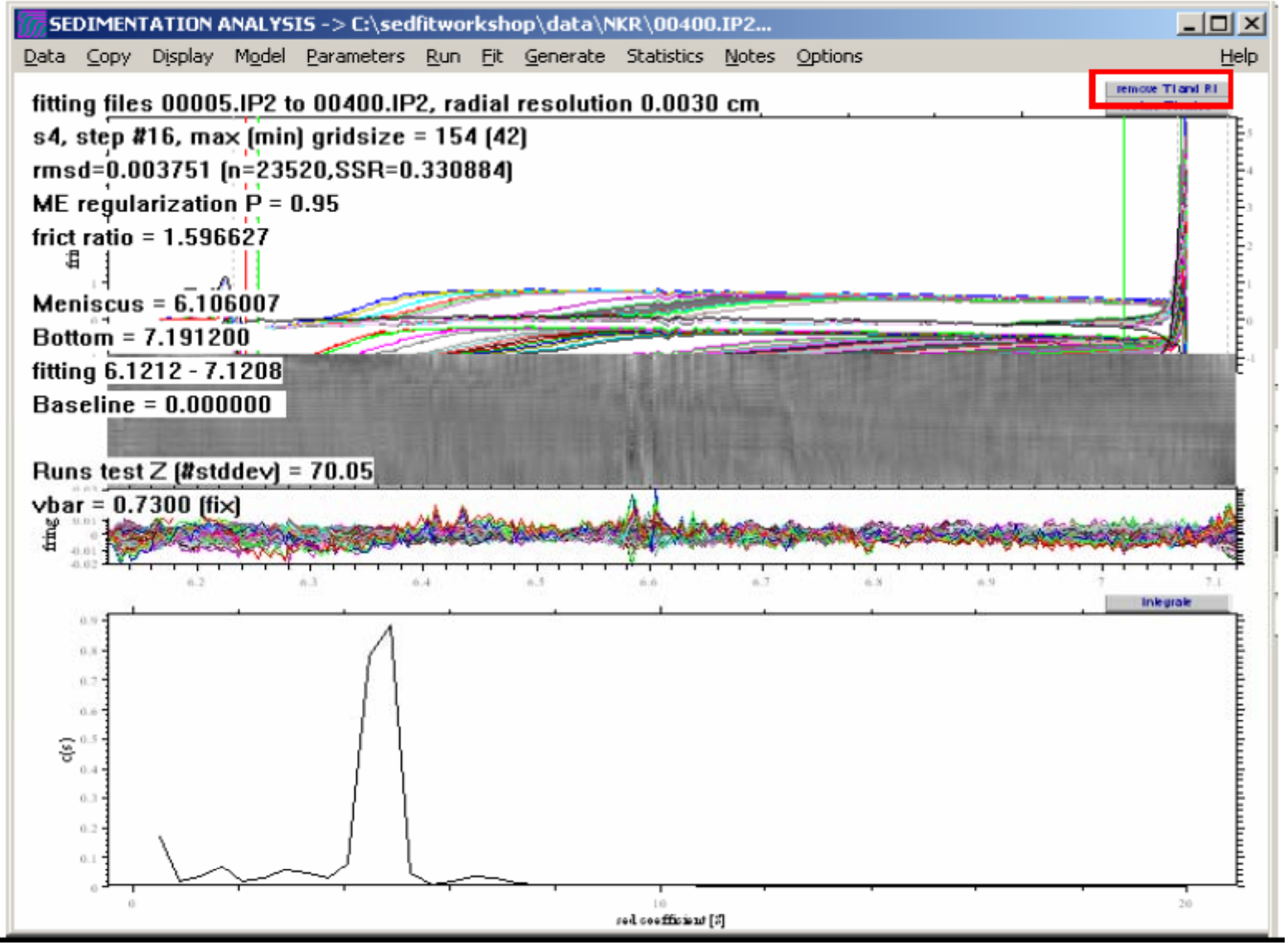
HIT OK

Perform a RUN which optimizes the linear parameters. ……Then FIT

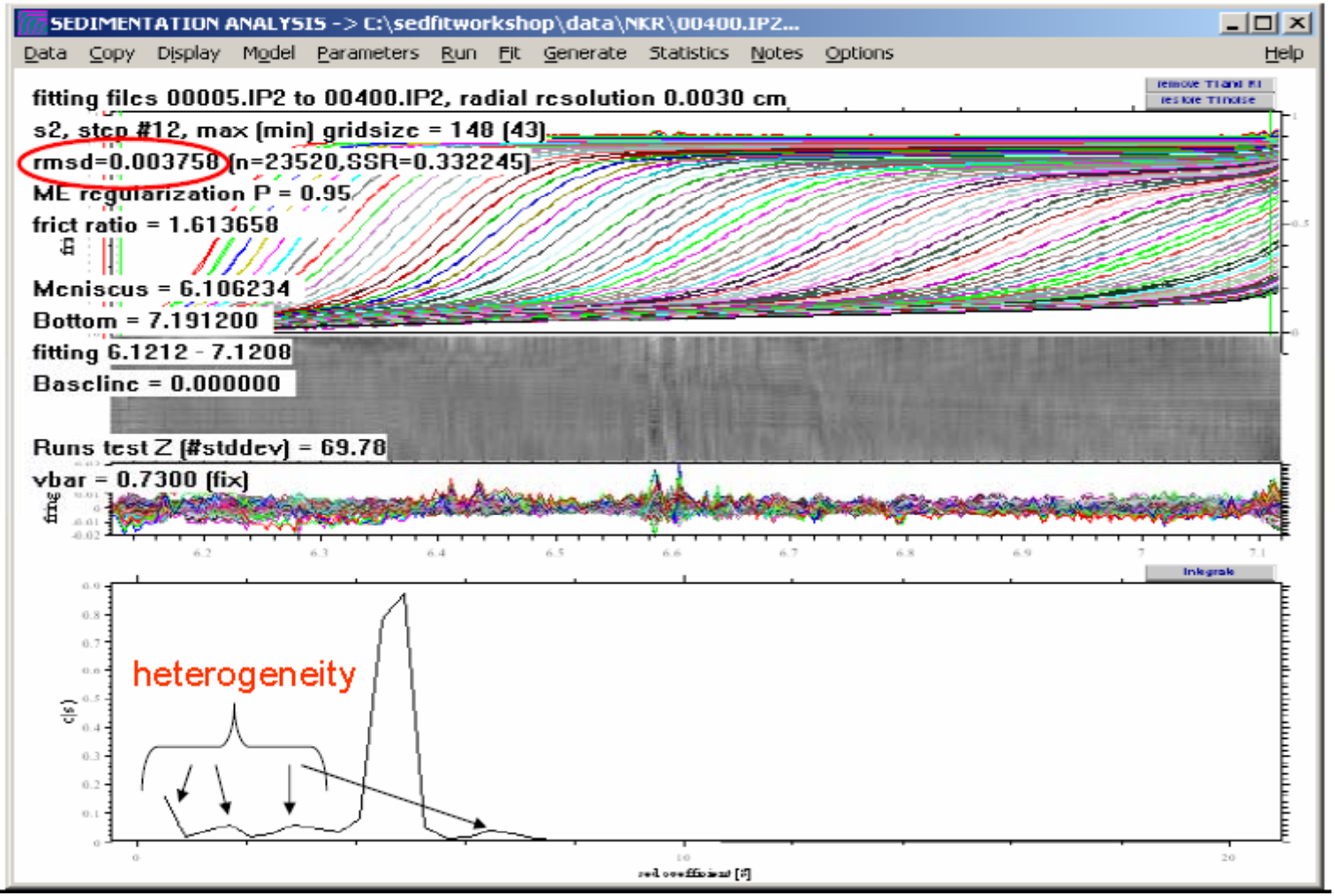


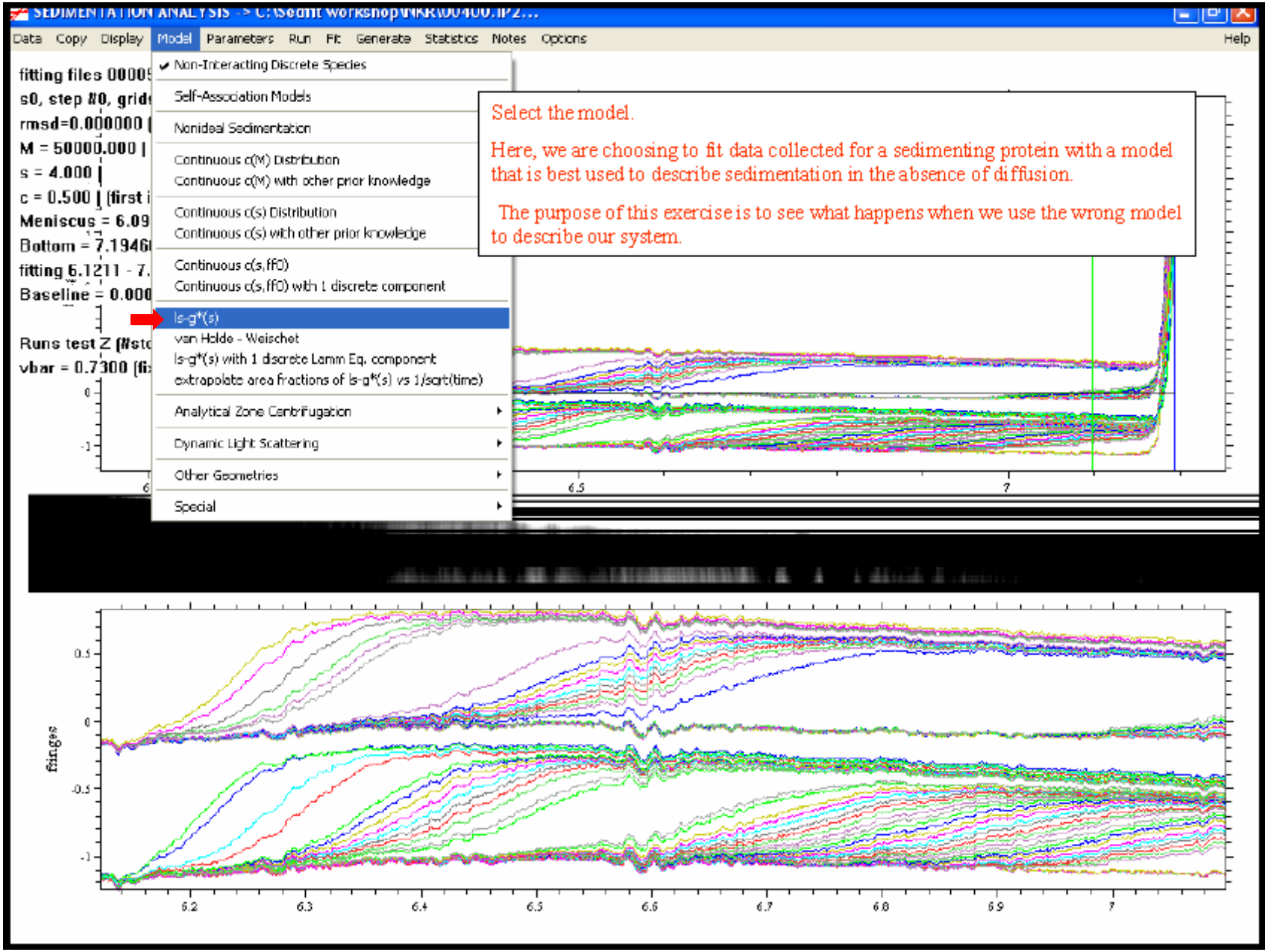

SEDIMENTATION ANALYSIS - C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics
This experiment comes from the NKR folder of Sedfit Workshop on the C: drive
fitting files 00005.IP2 to 00400.IP2
The scans selected for fitting are the 5th through the 400th scan of cell 2
M = 50000.000
s = 4.000
The default model, Non-interacting Discrete Species, has default parameters for a Molar Mass, Sedimentation coefficient, and species Concentration of 50 kDa, 4.0 S, and 0.5 fringe units of concentration, respectively.
Meniscus = 6.003900
Bottom = 7.194600
Observe where the default meniscus and bottom of the cell are initially set.
We'll need to relocate them.
Baseline = 0.000000
Runs test Z (#stddev) = 0.00
vbar = 0.7300 (fix)
fringes
6
6.5
7

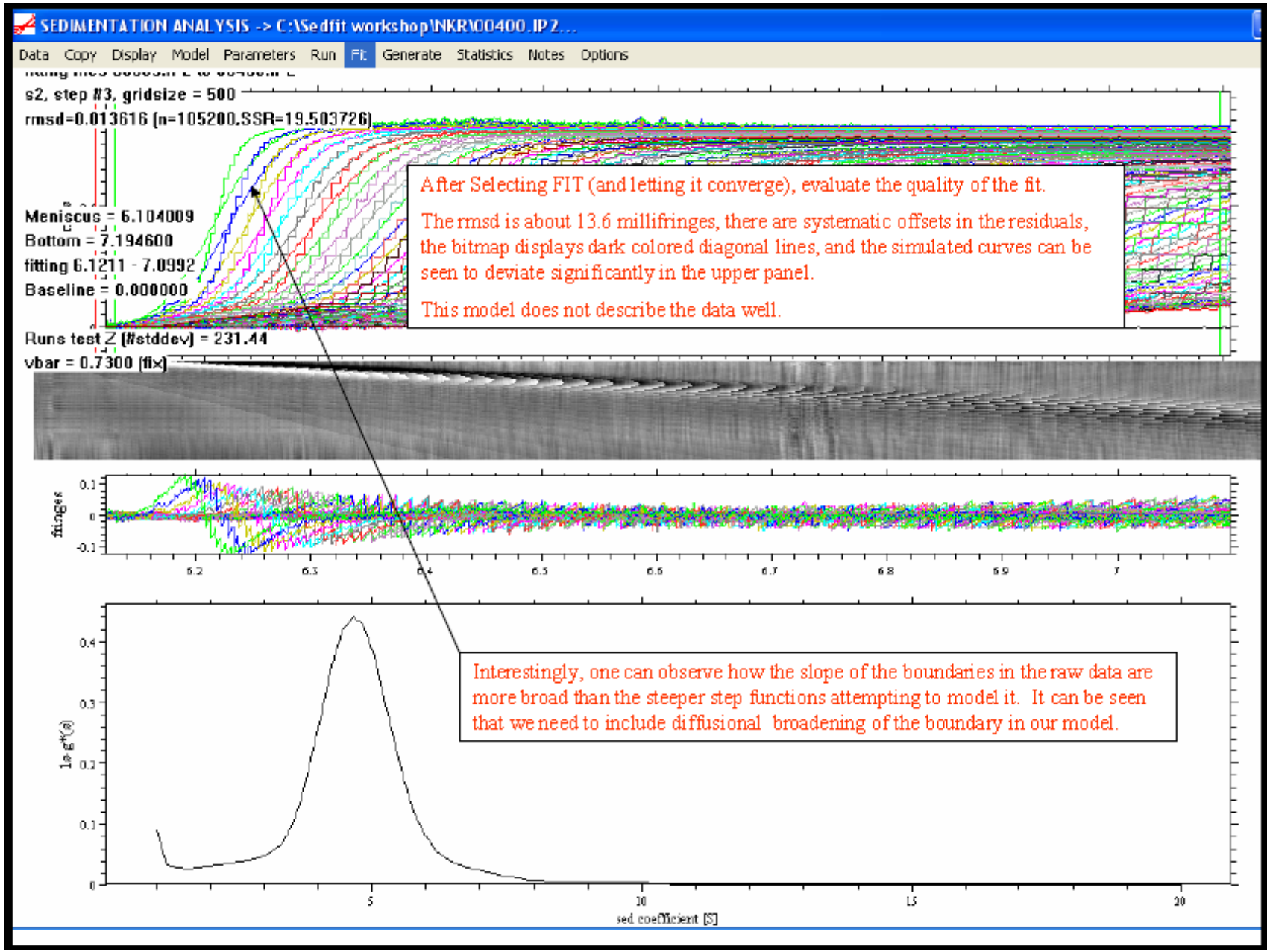
Fitting optimized all parameters. Then subtract the calculated RI and Ti noise.

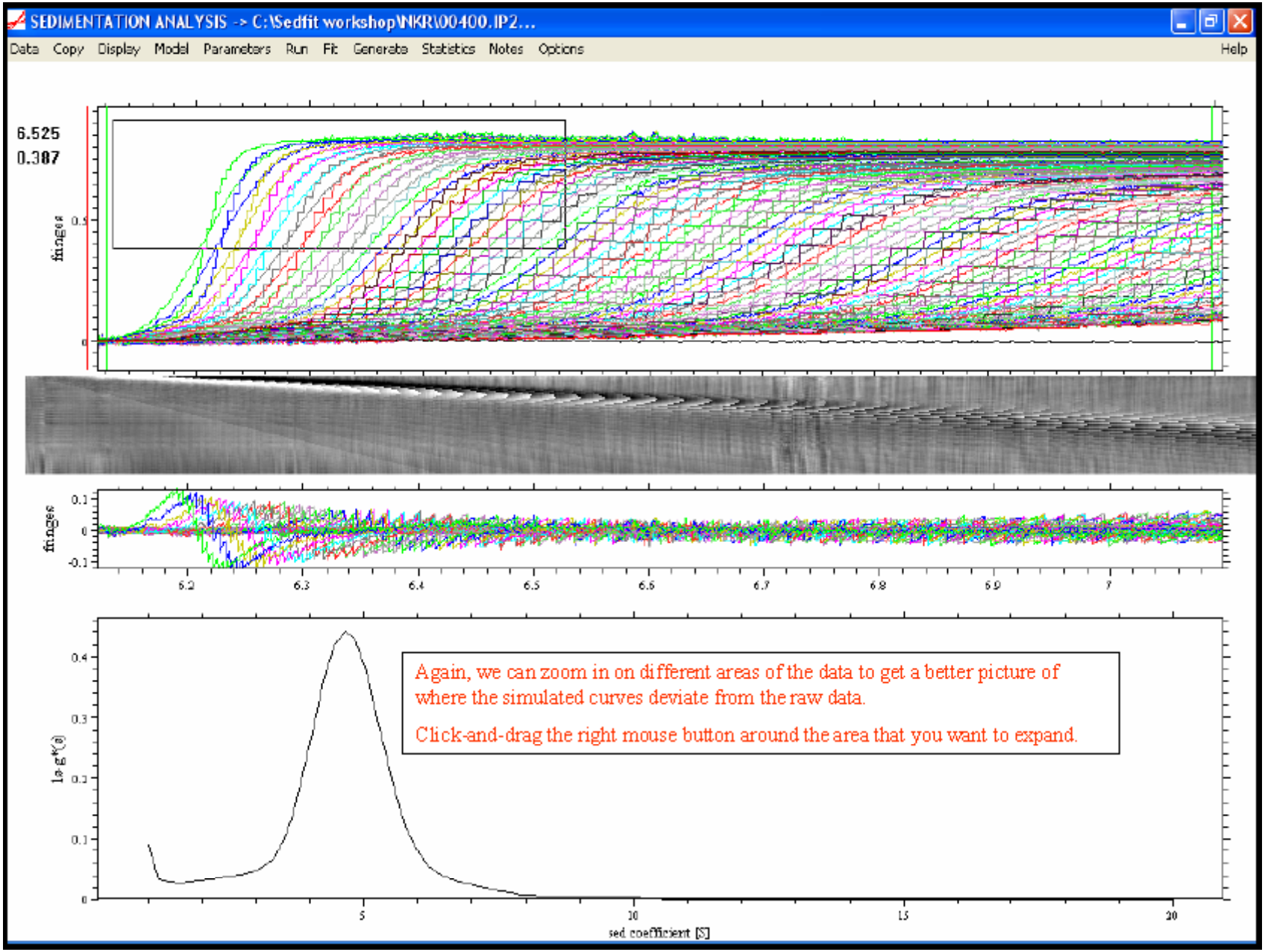


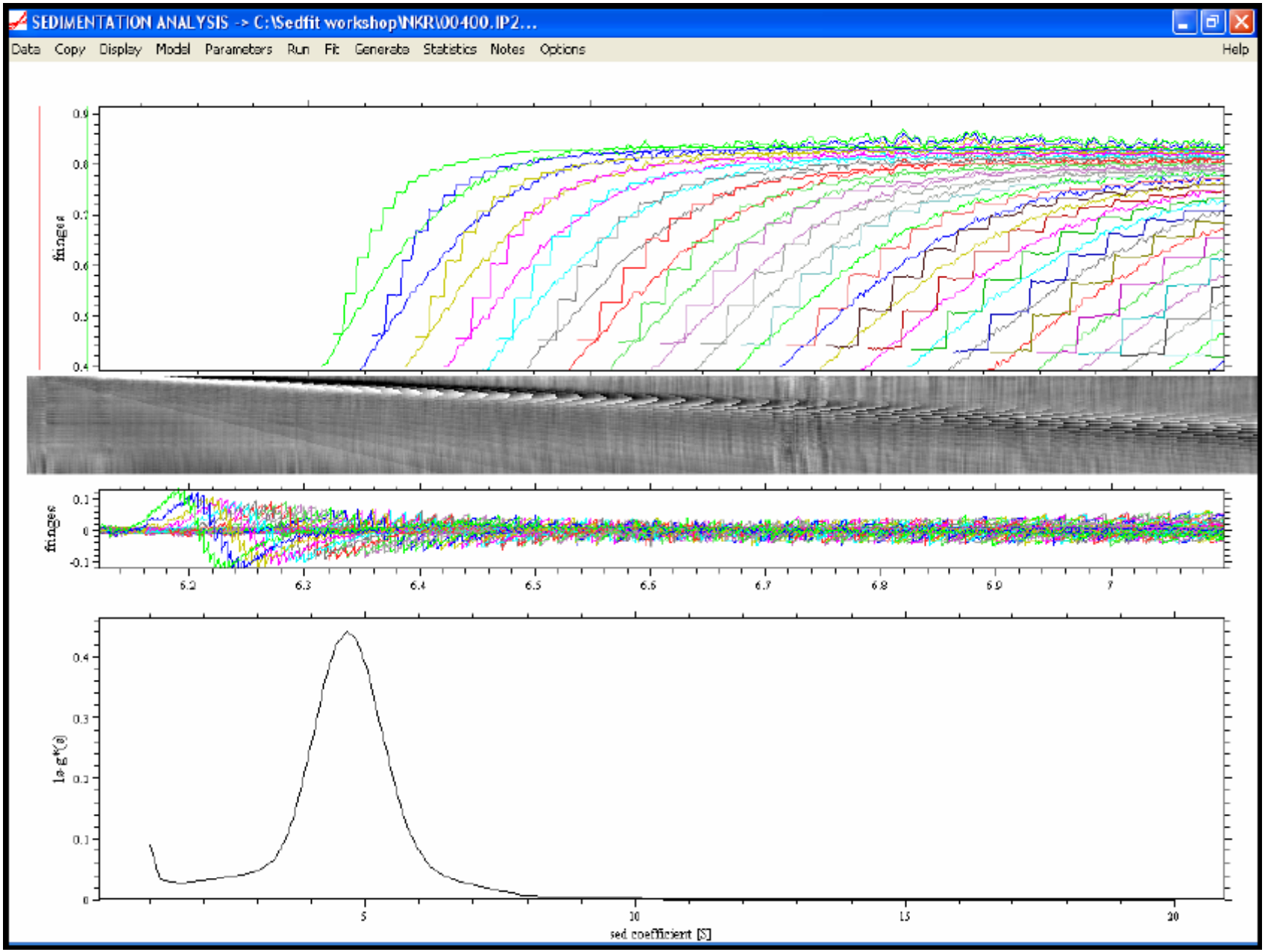
Assess the quality of the fit! Notice root mean square deviation (rmsd) >0.010
Compare to the ls-g*(s) model, and the Non-interacting discrete species model.
Notice the heterogeneity of the sample revealed by the c(s) distribution.

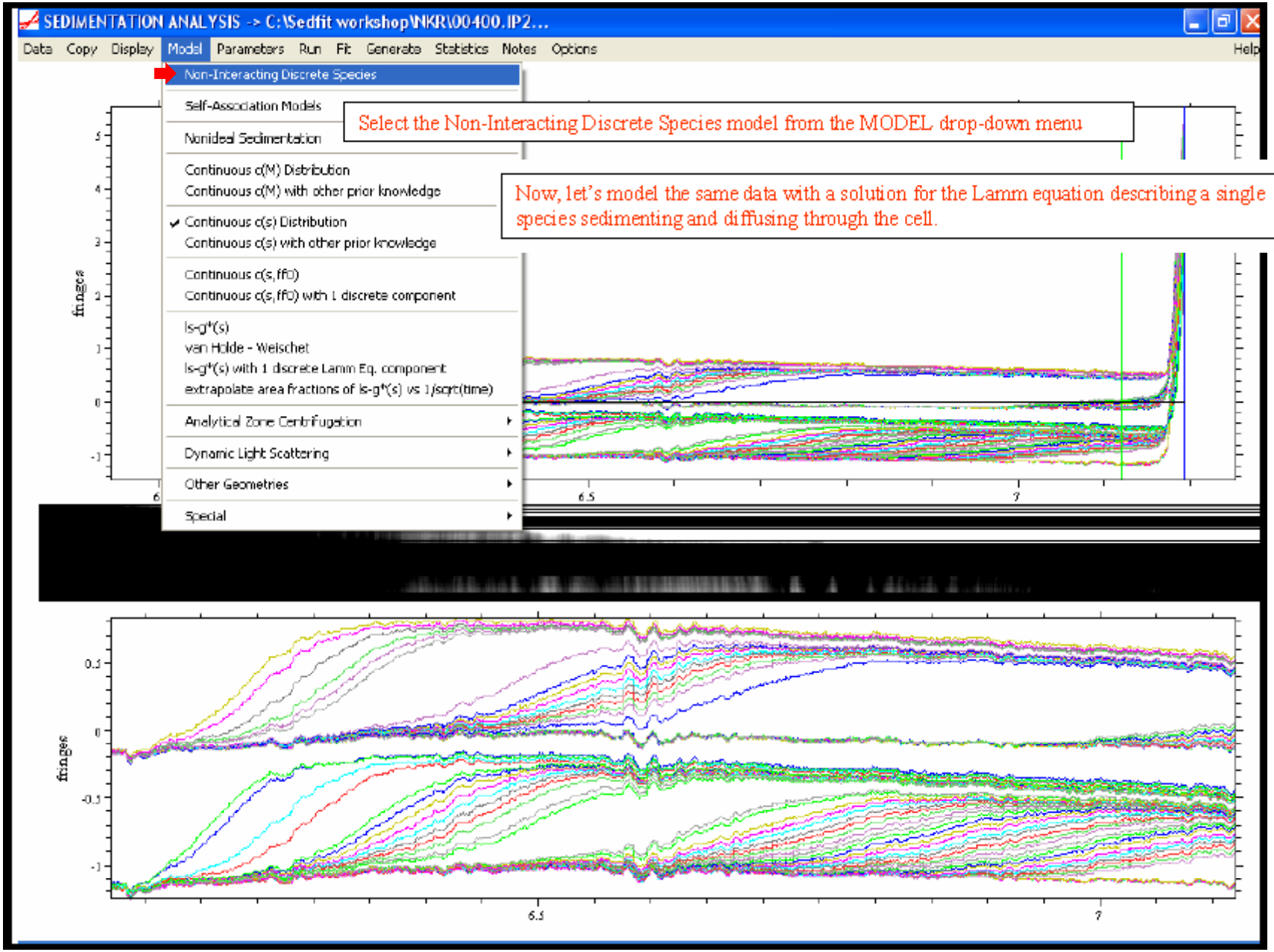


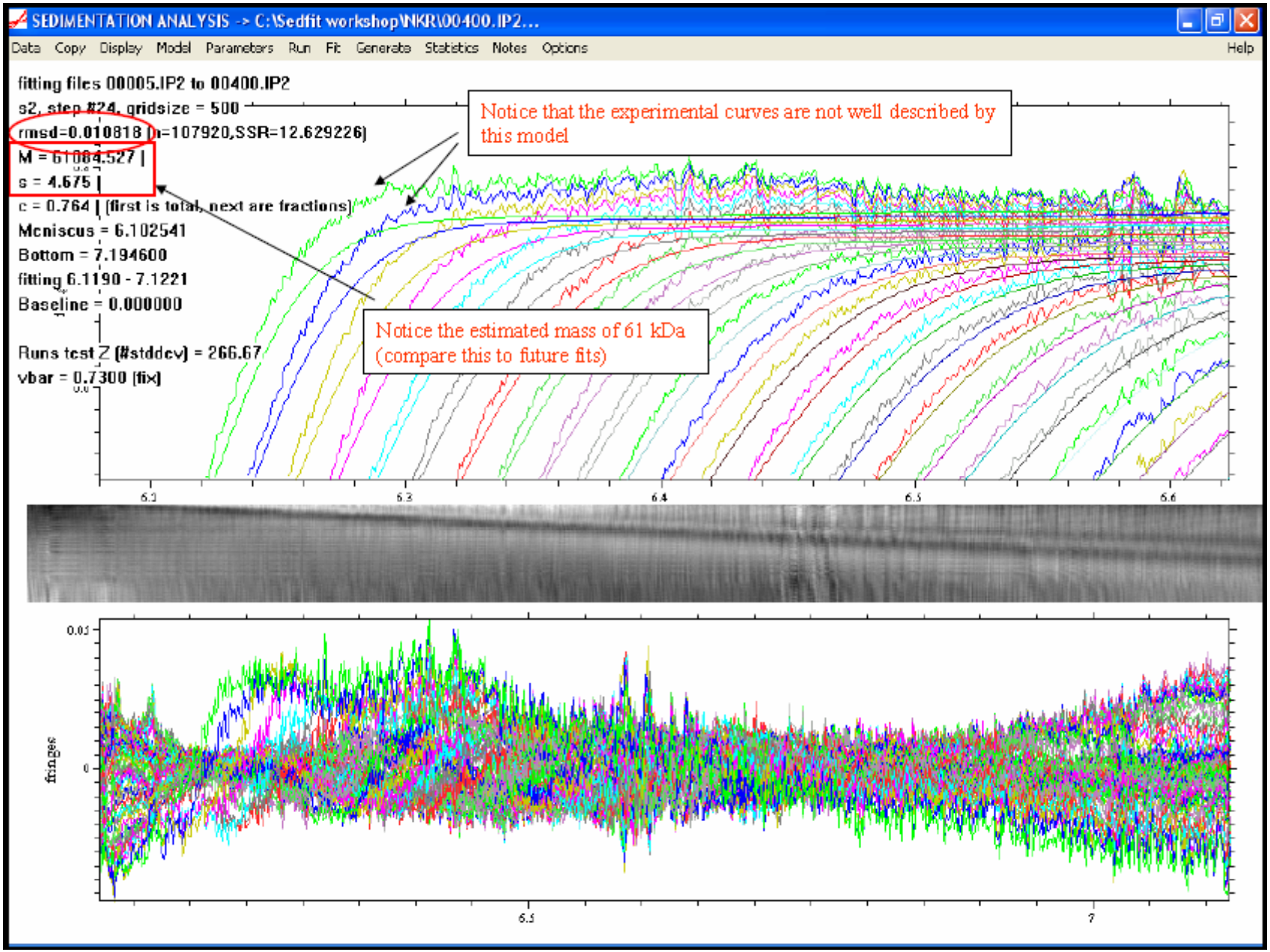

Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
Non-Interacting Discrete Species
Self-Association Models
Nonideal Sedimentation
Continuous c(M) Distribution
Continuous c(M) with other prior knowledge
Continuous c(s) Distribution
Continuous c(s) with other prior knowledge
Continuous c(s,ff0)
Continuous c(s,ff0) with 1 discrete component
ls-g*(s)
van Holde - Weischet
ls-g*(s) with 1 discrete Lamm Eq. component
extrapolate area fractions of ls-g*(s) vs 1/sqrt(time)
Analytical Zone Centrifugation
Dynamic Light Scattering
Other Geometries
Special
Select the model.
Here, we are choosing to fit data collected for a sedimenting protein with a model that is best used to describe sedimentation in the absence of diffusion.
The purpose of this exercise is to see what happens when we use the wrong model to describe our system.
fringes


SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options
s2, step #3, gridsize = 500
rmsd=0.013616 (n=105200,SSR=19.503726)
Meniscus = 6.104009
Bottom = 7.194600
fitting 6.1211 - 7.0992
Baseline = 0.000000
Runs test Z (#stddev) = 231.44
vbar = 0.7300 (fix)
After Selecting FIT (and letting it converge), evaluate the quality of the fit.
The rmsd is about 13.6 millifringes, there are systematic offsets in the residuals, the bitmap displays dark colored diagonal lines, and the simulated curves can be seen to deviate significantly in the upper panel.
This model does not describe the data well.
fringes
Interestingly, one can observe how the slope of the boundaries in the raw data are more broad than the steeper step functions attempting to model it. It can be seen that we need to include diffusional broadening of the boundary in our model.
sed.coefficient [S]

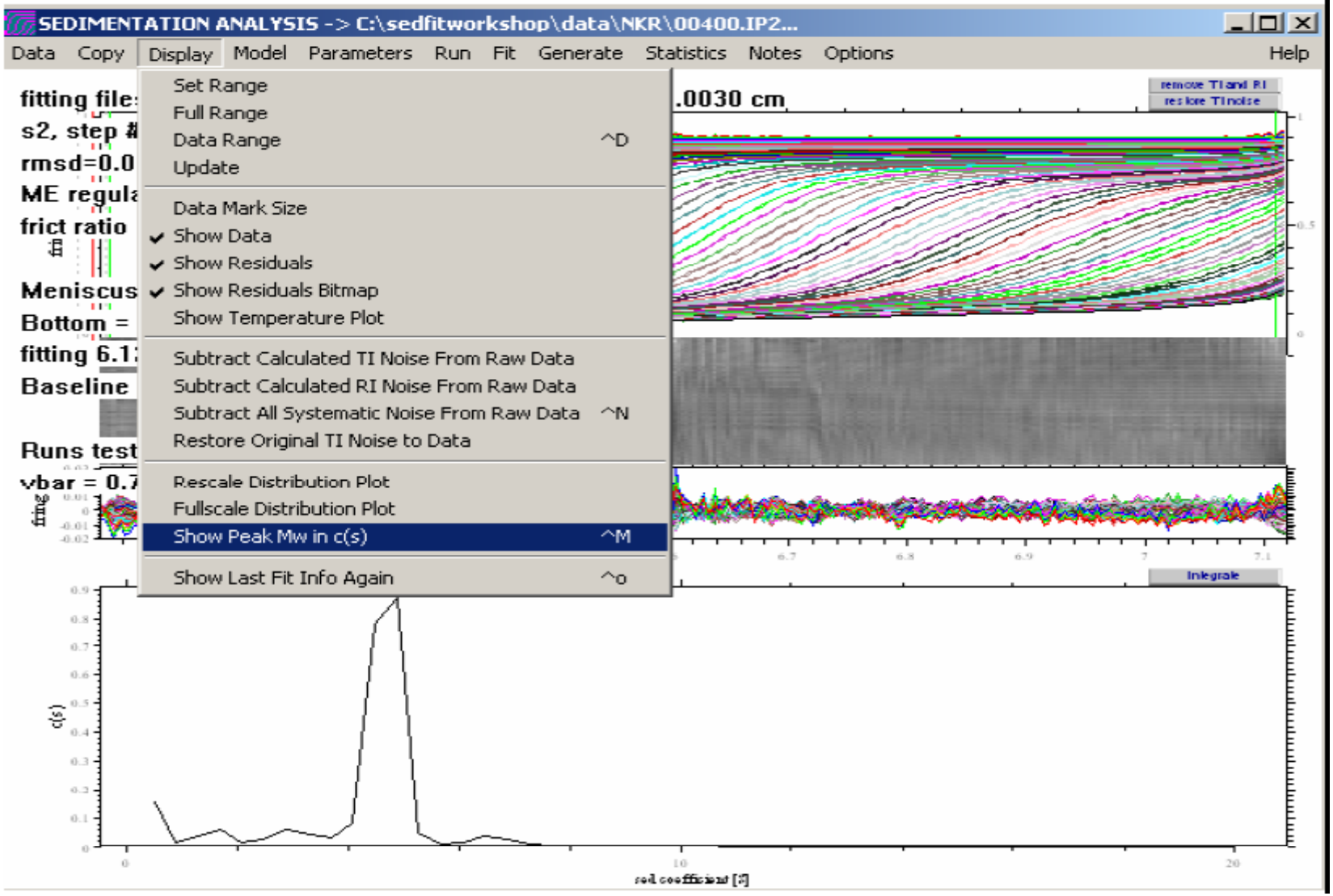

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
6.525
0.387
fringes
0.5
0
fringes
0.1
0
-0.1
6.2
6.3
6.4
6.5
6.6
6.7
6.8
6.9
7
Log(c(s))
0.4
0.3
0.2
0.1
0
5
10
15
20
sed coefficient [S]
Again, we can zoom in on different areas of the data to get a better picture of where the simulated curves deviate from the raw data.
Click-and-drag the right mouse button around the area that you want to expand.


SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
fringes
fringes
ln g*(s)
sed coefficient [S]

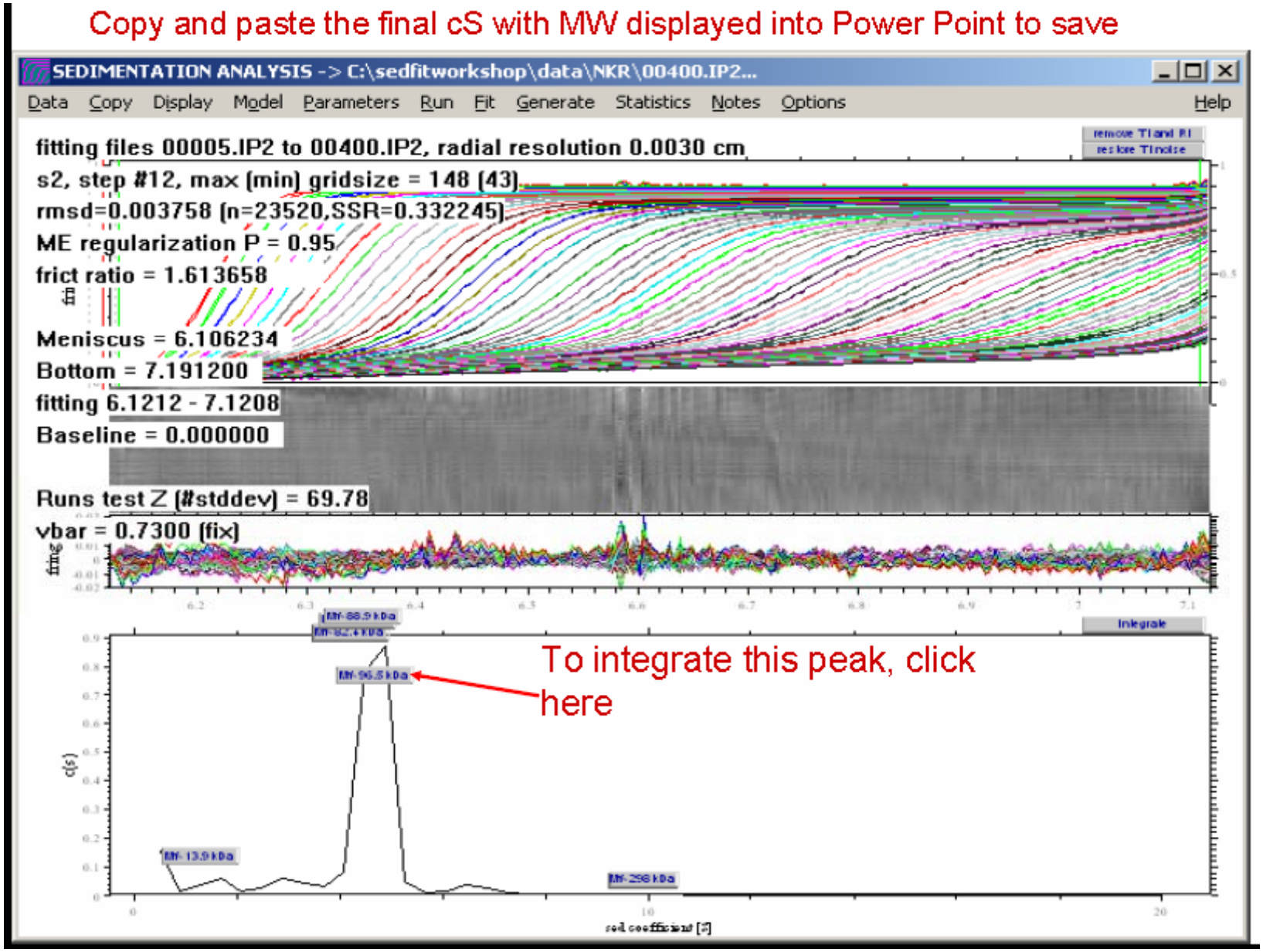

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
Non-Interacting Discrete Species
Self-Association Models
Nonideal Sedimentation
Continuous c(M) Distribution
Continuous c(M) with other prior knowledge
Continuous c(s) Distribution
Continuous c(s) with other prior knowledge
Continuous c(s,ff0)
Continuous c(s,ff0) with 1 discrete component
ls-g*(s)
van Holde - Weischet
ls-g*(s) with 1 discrete Lamm Eq. component
extrapolate area fractions of ls-g*(s) vs 1/sqrt(time)
Analytical Zone Centrifugation
Dynamic Light Scattering
Other Geometries
Special
Select the Non-Interacting Discrete Species model from the MODEL drop-down menu
Now, let's model the same data with a solution for the Lamm equation describing a single species sedimenting and diffusing through the cell.
fringes
6.5
7


SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
fitting files 00005.IP2 to 00400.IP2
s2, step #24, gridsize = 500
rmsd=0.010818 (n=107920,SSR=12.629226)
M = 61064.527 |
s = 4.675 |
c = 0.764 | (first is total, next are fractions)
Mcniscus = 6.102541
Bottom = 7.194600
fitting 6.1190 - 7.1221
Baseline = 0.000000
Runs test Z (#stddev) = 266.67
vbar = 0.7300 (fix)
Notice that the experimental curves are not well described by this model
Notice the estimated mass of 61 kDa (compare this to future fits)
fringes

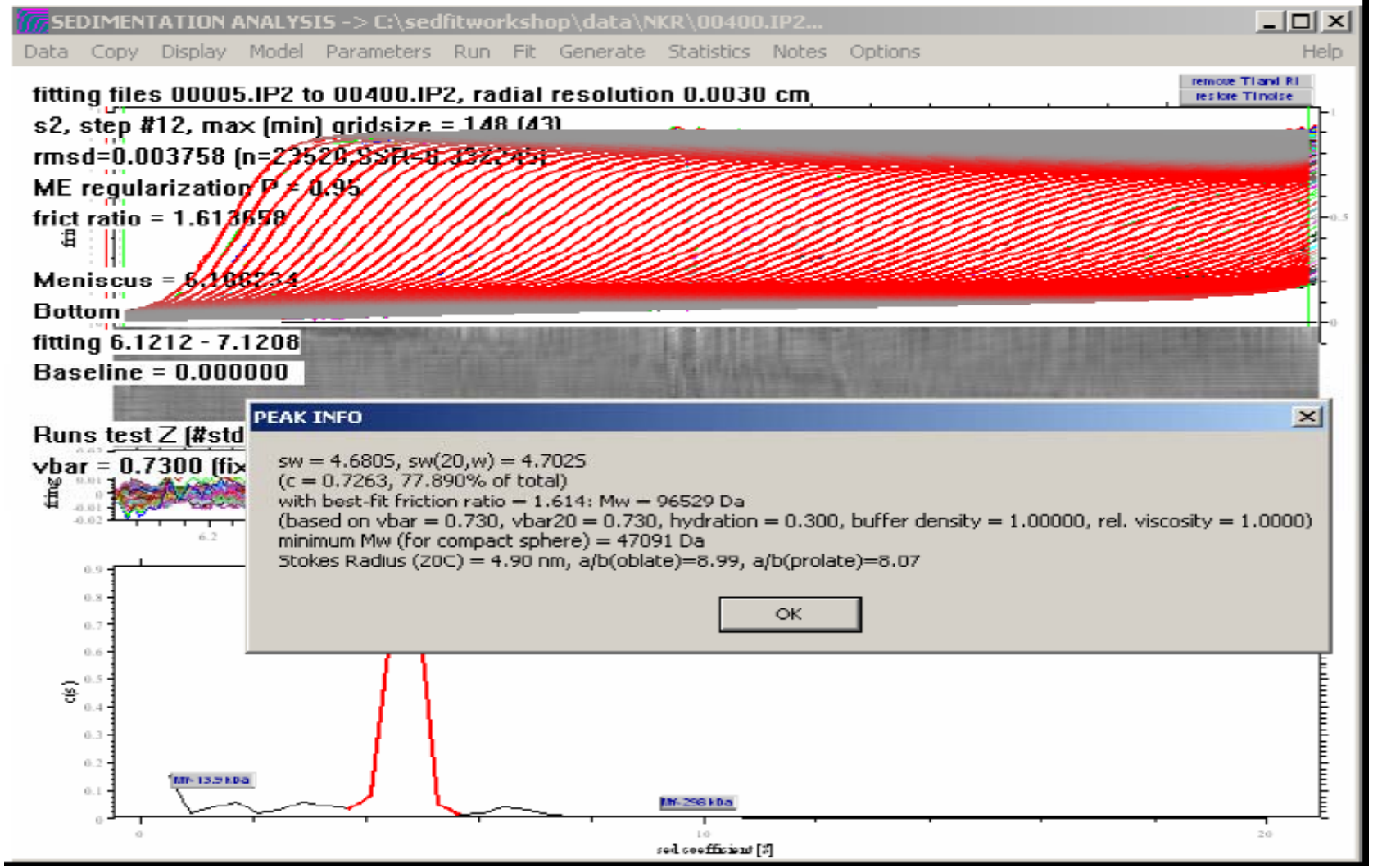
Display the peak MW either from the top menu or by hitting “control M”

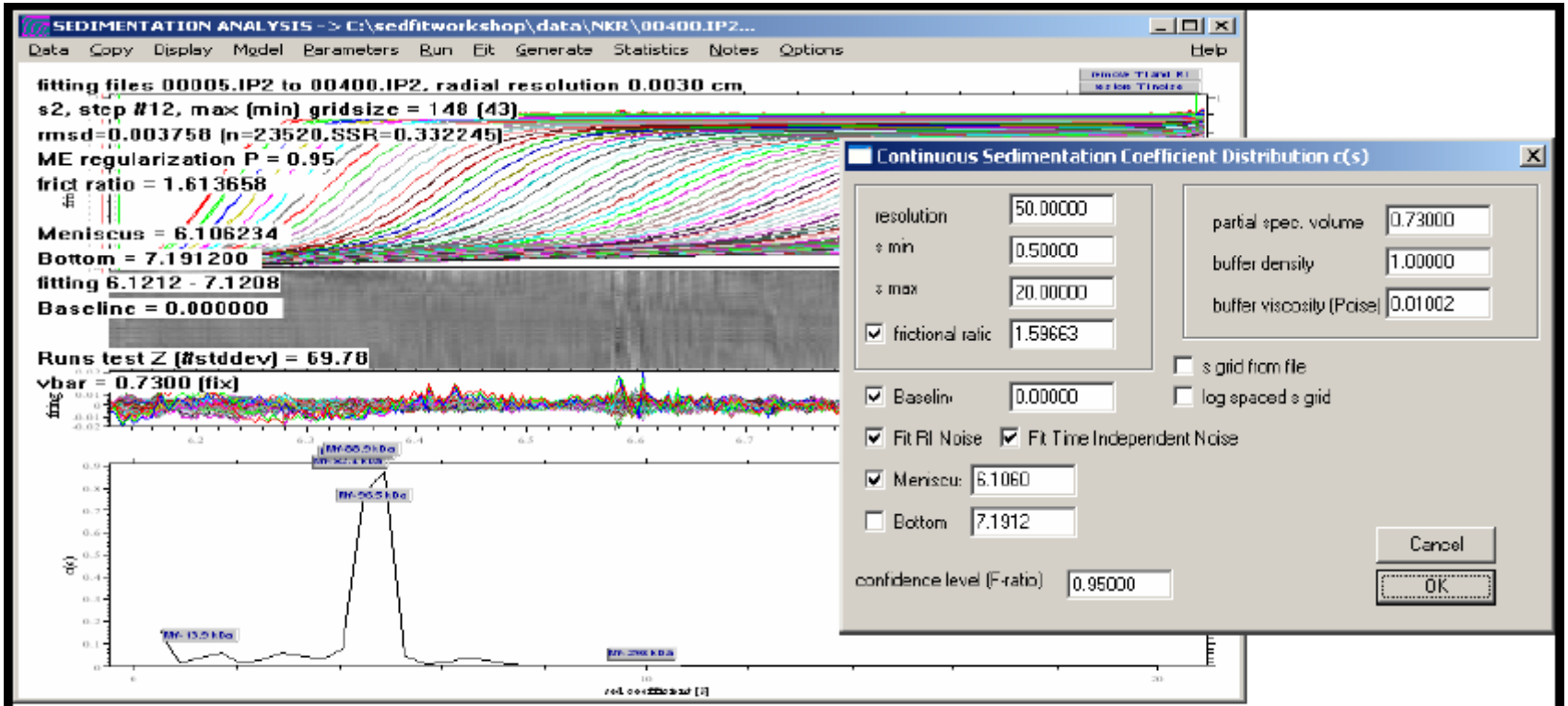


Copy and paste the final cS with MW displayed into Power Point to save

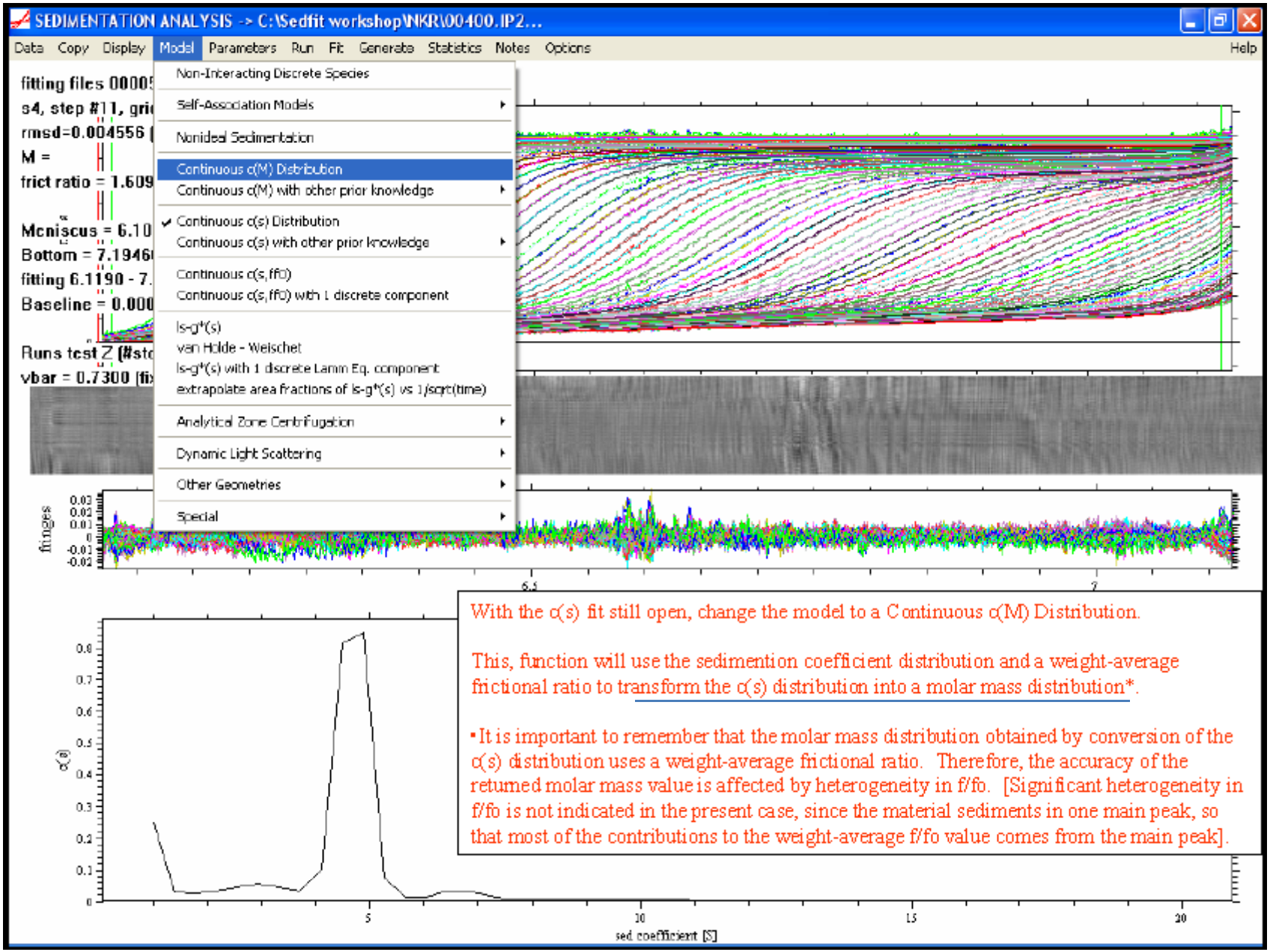


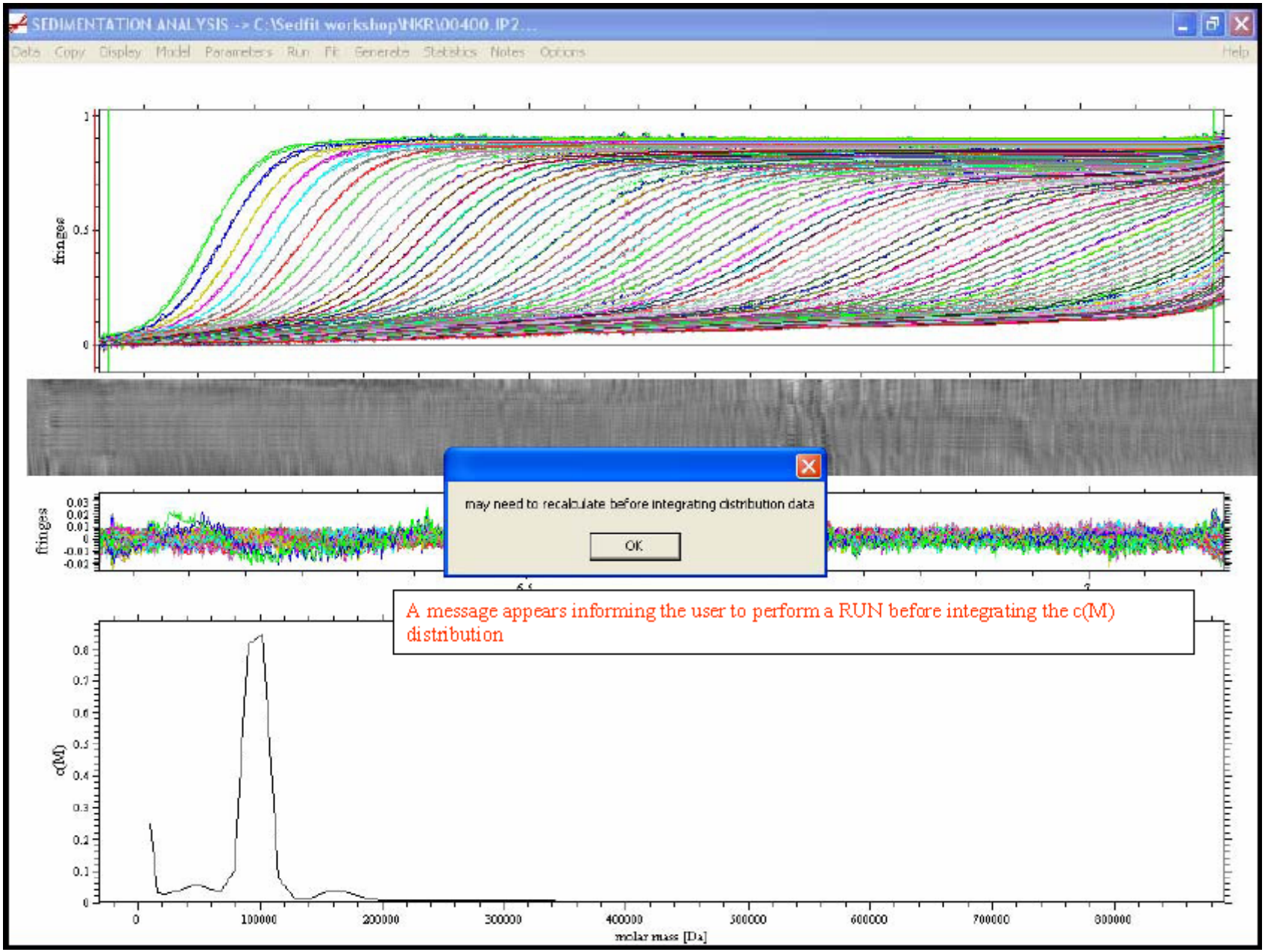
Location of this MW displayed.
Peak information window opens. Copy and paste to ppt.

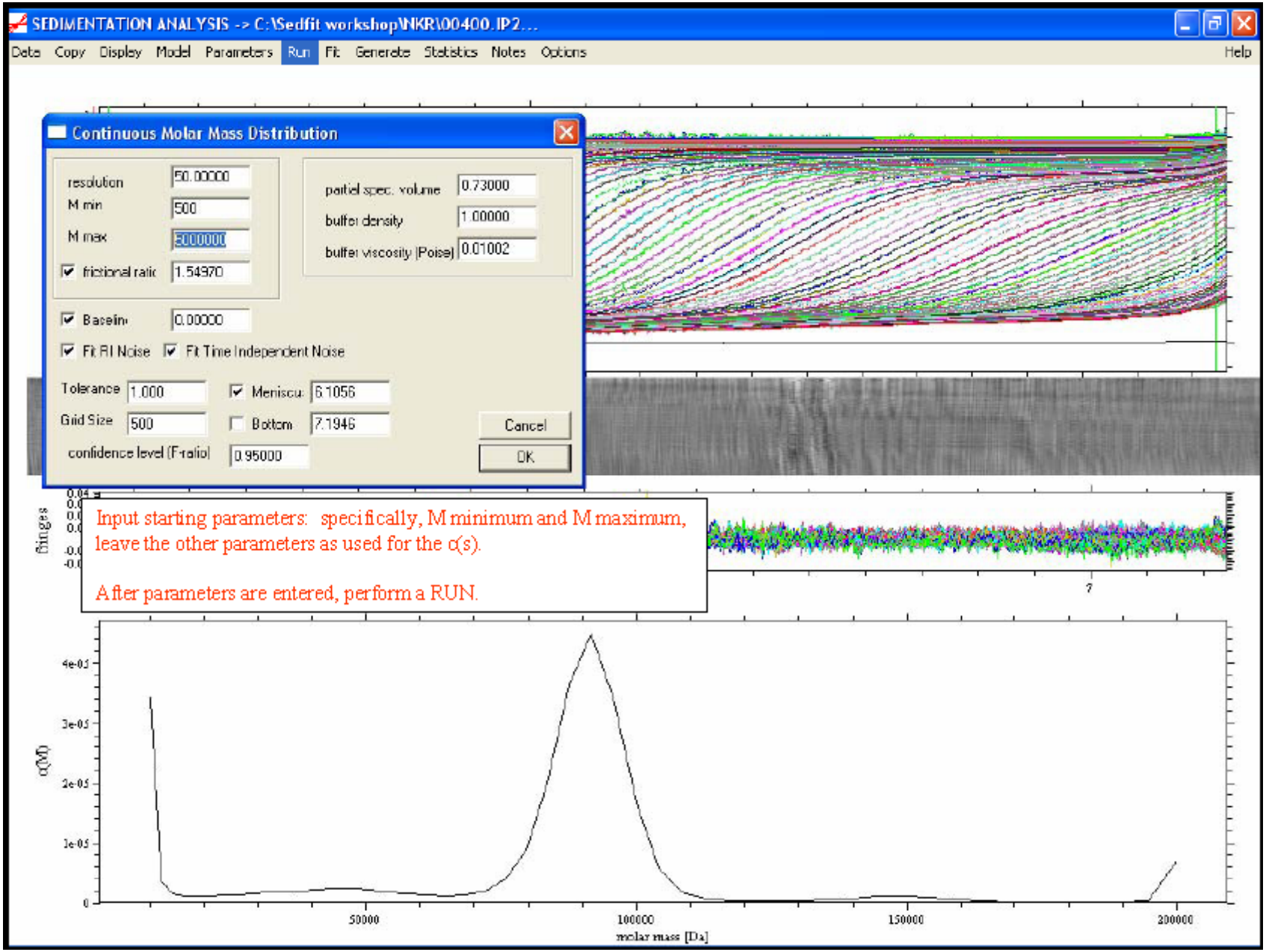


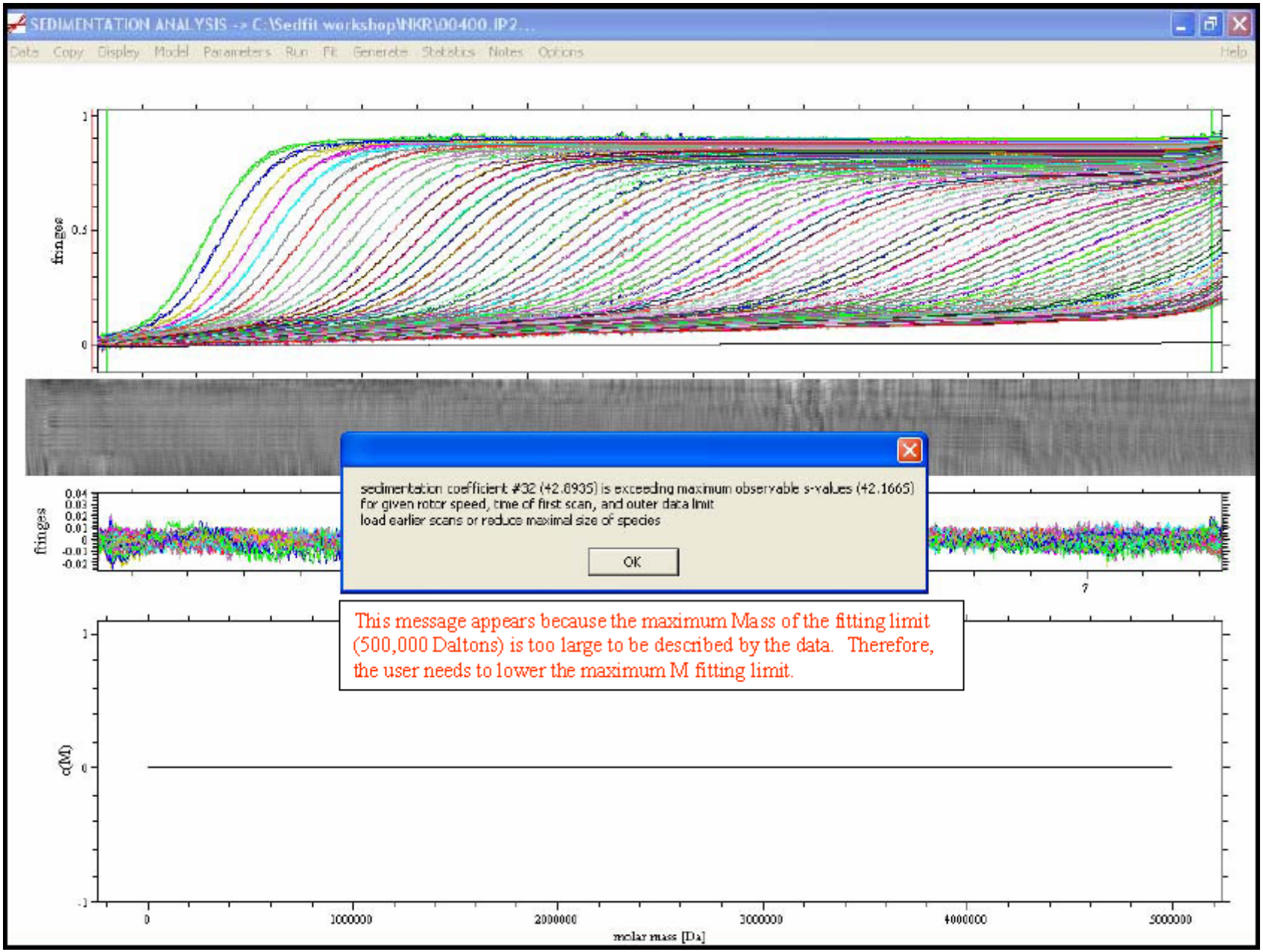


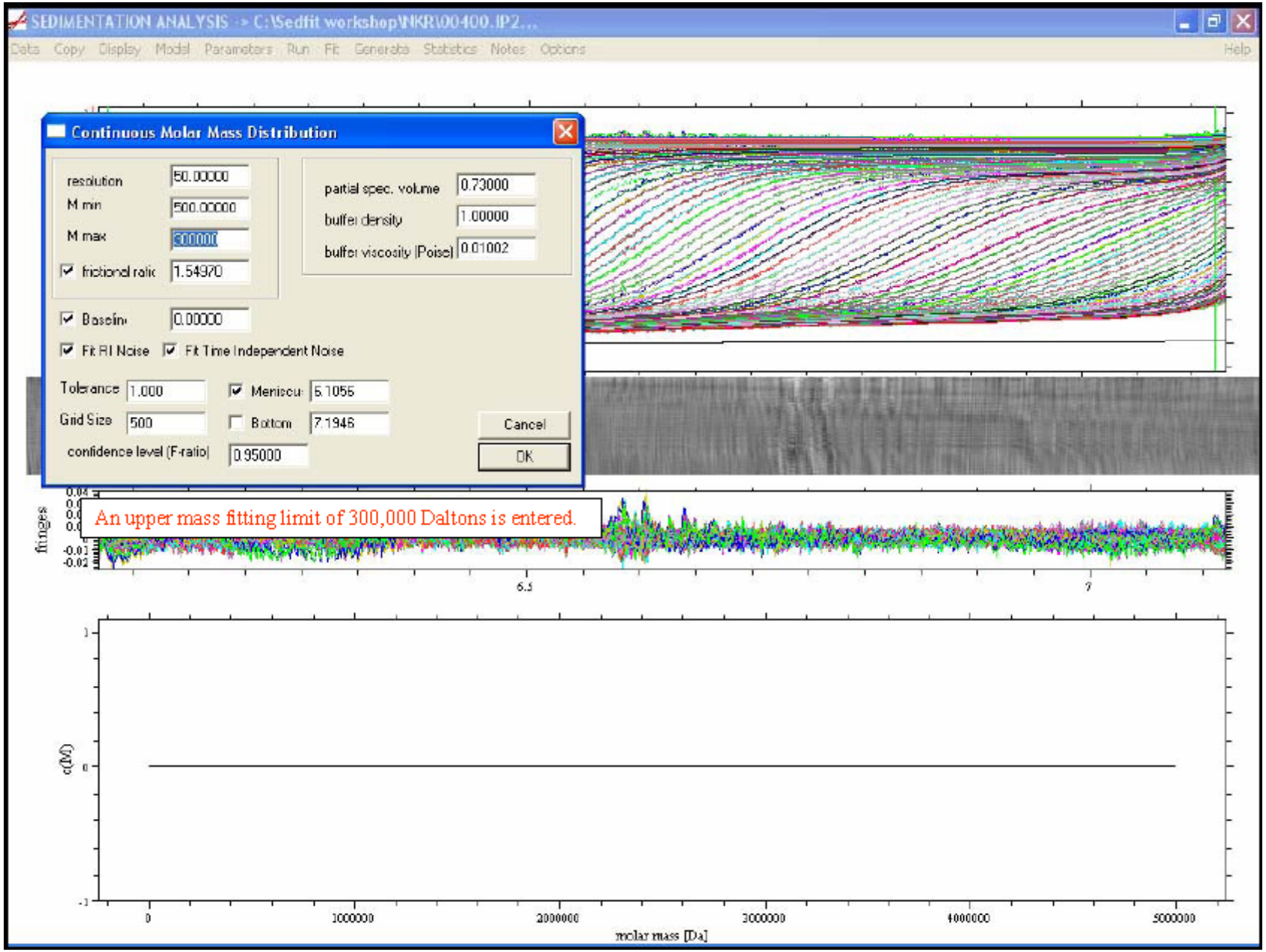
Try the following:
Zoom in the distribution range from 10-20 S (with dragging a rectangle with the right mouse button) – there's not much material out there.
You can further refine the fit by selecting smax = 15, and resolution = 100
Do a RUN – you'll see a smoother distribution.
Then,
Since the distribution goes up at s-min, test what decreasing the lower limit s-min does (for example, set s-min = 0.2). One of the two should be expected: a) a new peak appears and c(s) goes down at the new s-min, and rmsd gets slightly better; b) c(s) goes up further at s-min. What that means is that s-min is low enough to describe all possible sedimentation at the lower end, and is correlated with the baseline (baseline = no sedimentation). This is no problem. In this case, you can simply ignore this partial peak at s-min, and you can zoom into the rest of the distribution to see that better. The option b) is what happens here.

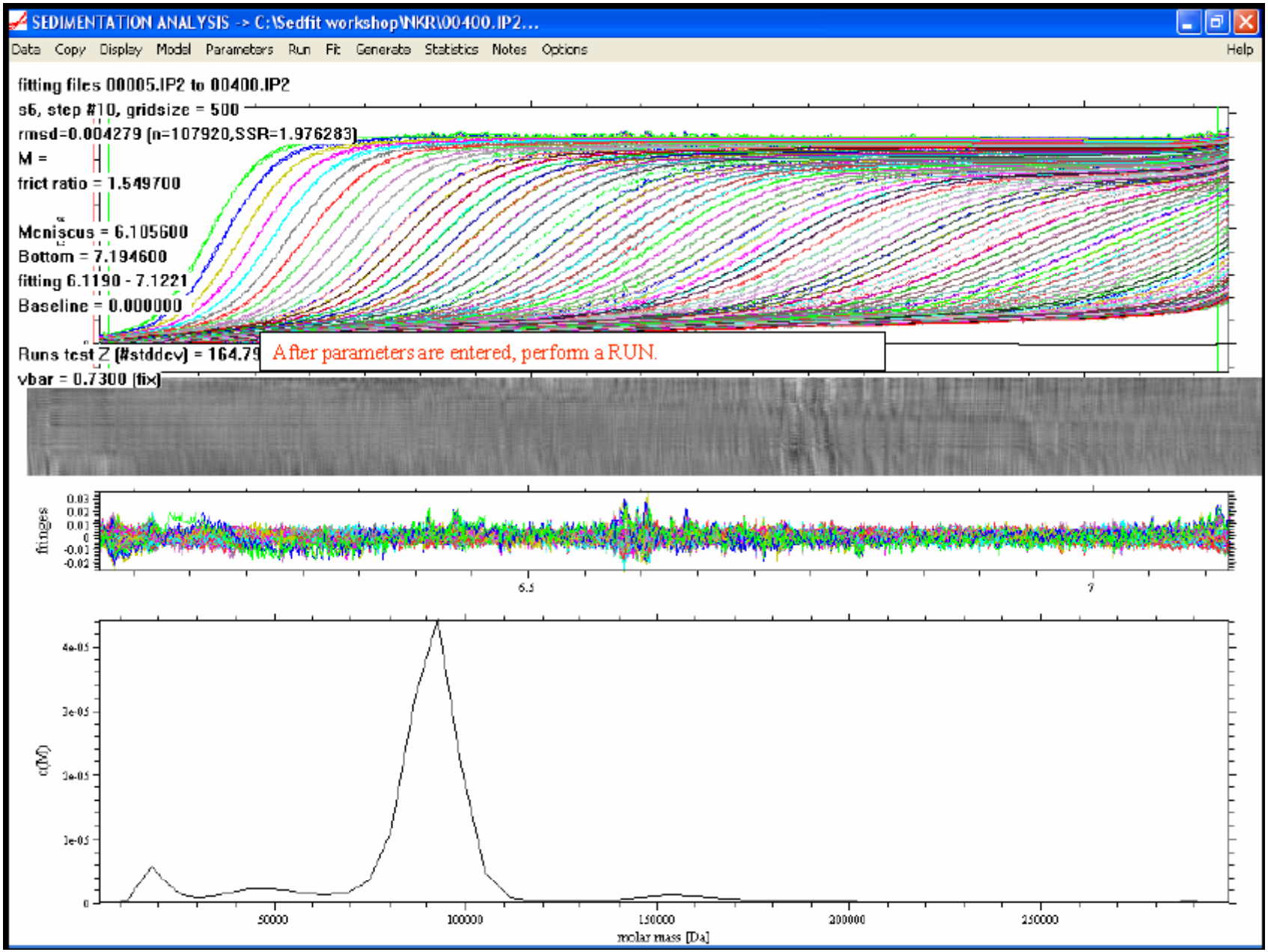

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
Non-Interacting Discrete Species
Self-Association Models
Nonideal Sedimentation
Continuous c(M) Distribution
Continuous c(M) with other prior knowledge
Continuous c(s) Distribution
Continuous c(s) with other prior knowledge
Continuous c(s,ff0)
Continuous c(s,ff0) with 1 discrete component
ls-g*(s)
van Holde - Weischet
ls-g*(s) with 1 discrete Lamm Eq. component
extrapolate area fractions of ls-g*(s) vs 1/sqrt(time)
Analytical Zone Centrifugation
Dynamic Light Scattering
Other Geometries
Special
fitting files 0000
s4, step #11, gri
rmsd=0.004556
M =
frict ratio = 1.609
Meniscus = 6.10
Bottom = 7.1946
fitting 6.1190 - 7.
Baseline = 0.000
Runs test Z (#st
vbar = 0.7300 (fi
fringes
c(s)
sed coefficient [S]
With the c(s) fit still open, change the model to a Continuous c(M) Distribution.
This, function will use the sedimention coefficient distribution and a weight-average frictional ratio to transform the c(s) distribution into a molar mass distribution*.
• It is important to remember that the molar mass distribution obtained by conversion of the c(s) distribution uses a weight-average frictional ratio. Therefore, the accuracy of the returned molar mass value is affected by heterogeneity in f/fo. [Significant heterogeneity in f/fo is not indicated in the present case, since the material sediments in one main peak, so that most of the contributions to the weight-average f/fo value comes from the main peak].

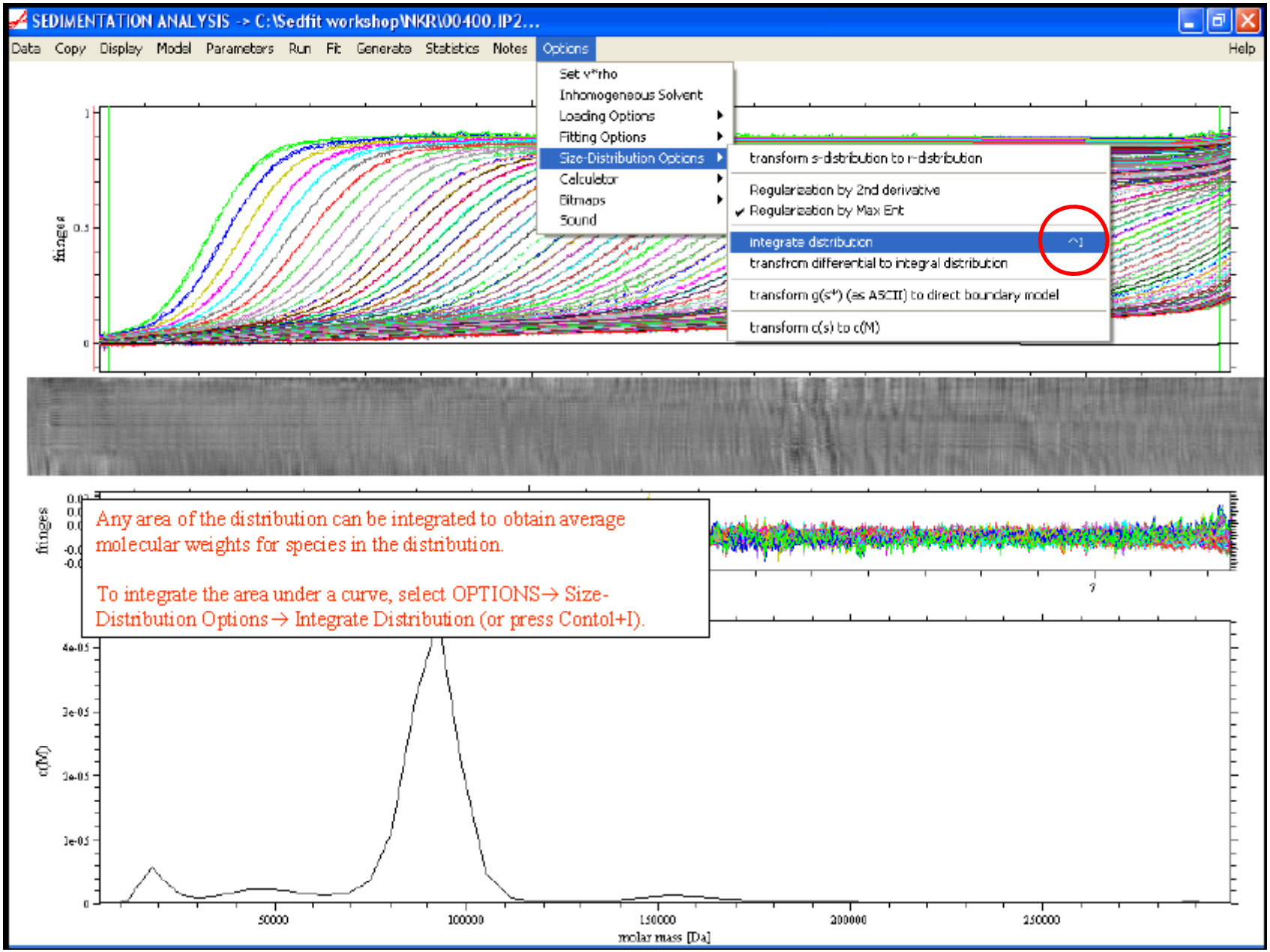

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
fringes
may need to recalculate before integrating distribution data
OK
A message appears informing the user to perform a RUN before integrating the c(M) distribution
c(M)
molar mass [Da]

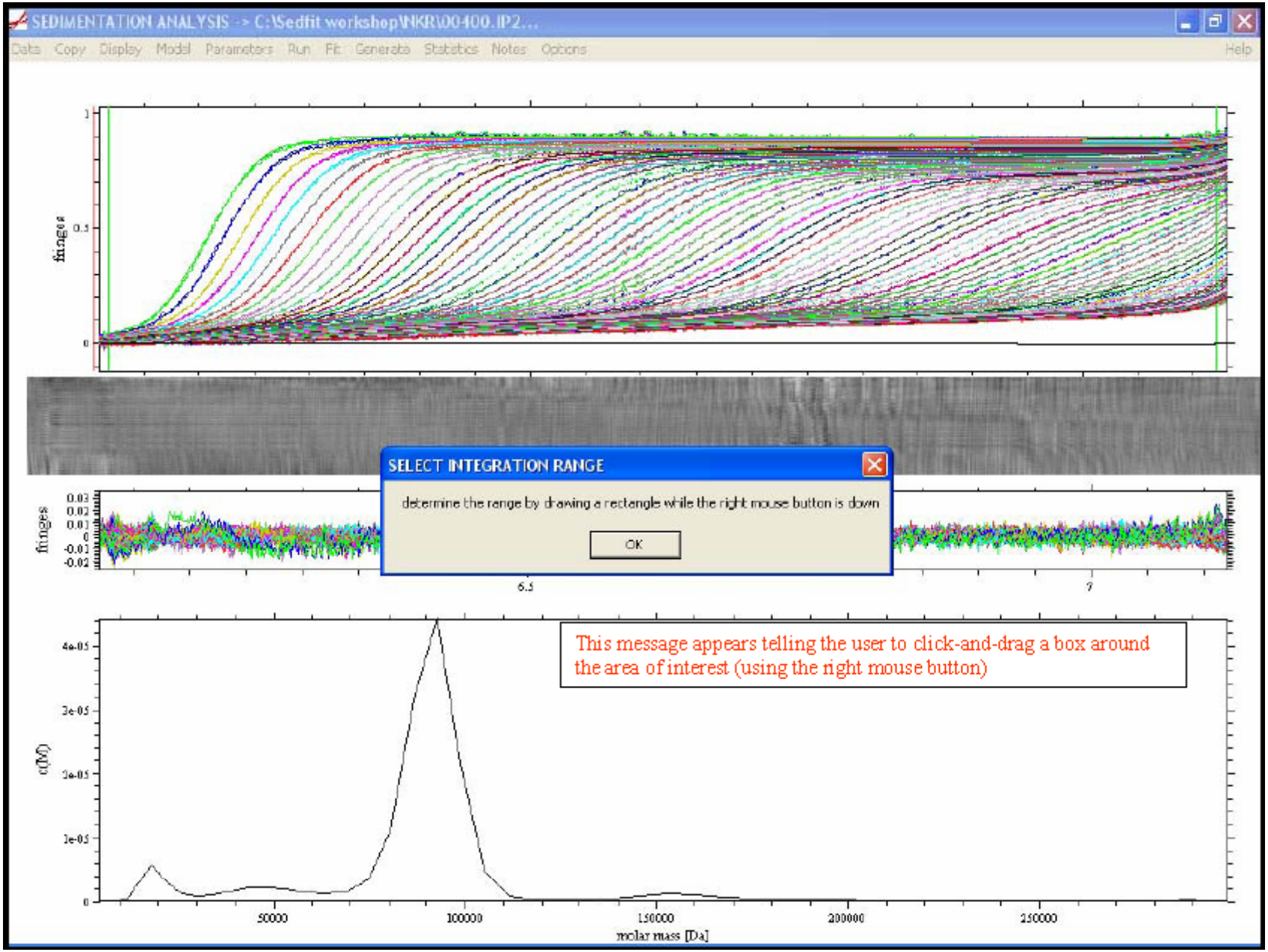

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
Continuous Molar Mass Distribution
resolution 50.00000
M min 500
M max 3000000
frictional ratio 1.54970
partial spec. volume 0.73000
buffer density 1.00000
buffer viscosity (Poise) 0.01002
Baseline 0.00000
Fit RI Noise
Fit Time Independent Noise
Tolerance 1.000
Grid Size 500
confidence level (F-ratio) 0.95000
Meniscus 6.1056
Bottom 7.1946
Cancel
OK
Input starting parameters: specifically, M minimum and M maximum, leave the other parameters as used for the c(s).
After parameters are entered, perform a RUN.
fringes
c(M)
4e-05
3e-05
2e-05
1e-05
0
50000
100000
150000
200000
molar mass [Da]

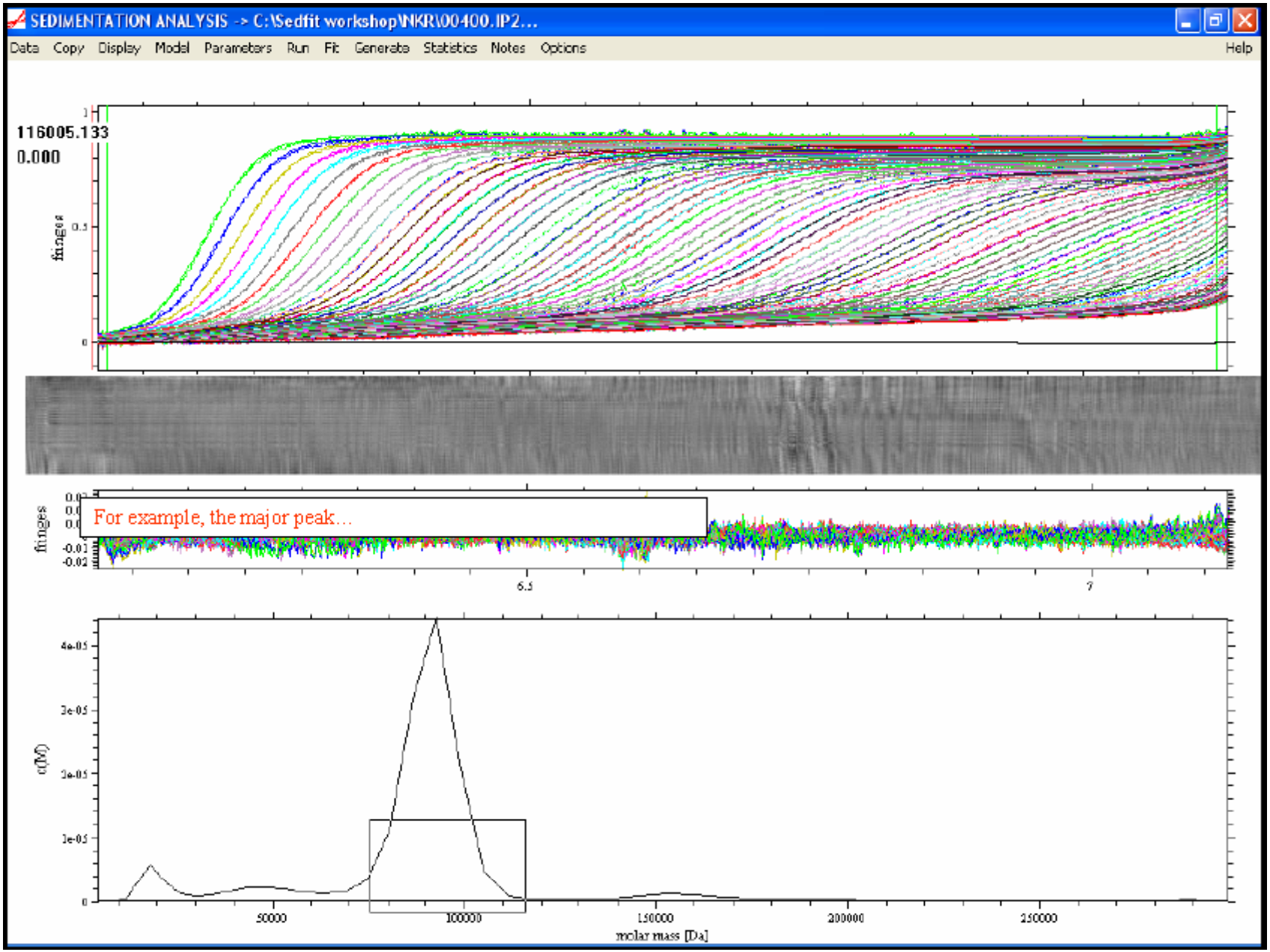

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
fringes
sedimentation coefficient #32 (42.8935) is exceeding maximum observable s-values (42.1665)
for given rotor speed, time of first scan, and outer data limit
load earlier scans or reduce maximal size of species
OK
This message appears because the maximum Mass of the fitting limit (500,000 Daltons) is too large to be described by the data. Therefore, the user needs to lower the maximum M fitting limit.
c(M)
0
1000000
2000000
3000000
4000000
5000000
molar mass [Da]

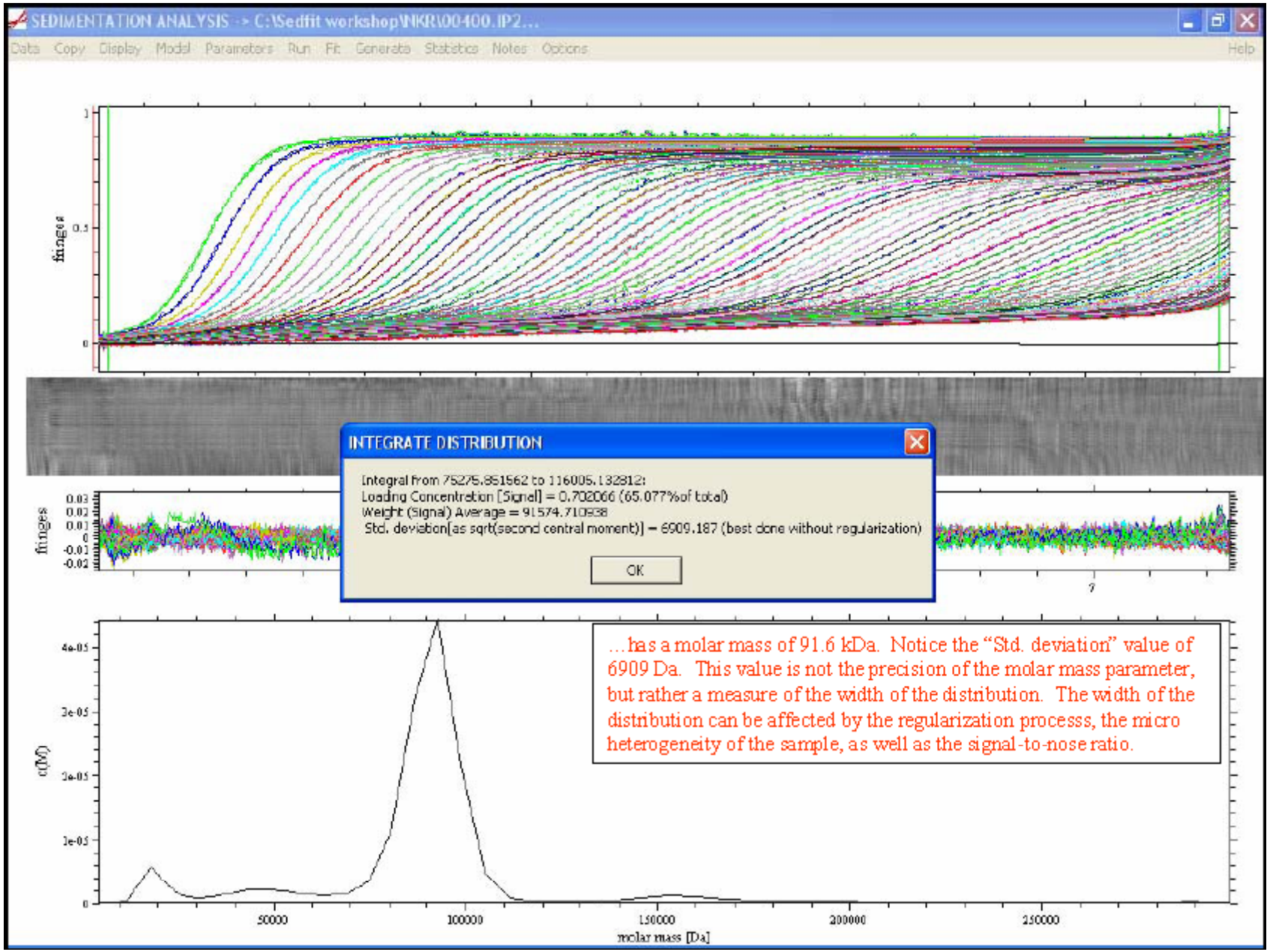

SEDIMENTATION ANALYSIS - » C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
Continuous Molar Mass Distribution
resolution 50.00000
M min 500.00000
M max 300000
frictional ratio 1.54970
partial spec. volume 0.73000
buffer density 1.00000
buffer viscosity [Poise] 0.01002
Baseline 0.00000
Fit RI Noise
Fit Time Independent Noise
Tolerance 1.000
Meniscus 6.1056
Grid Size 500
Bottom 7.1946
confidence level (F-ratio) 0.95000
Cancel
OK
An upper mass fitting limit of 300,000 Daltons is entered.
fringes
6.5
7
c(M)
0
1000000
2000000
3000000
4000000
5000000
molar mass [Da]

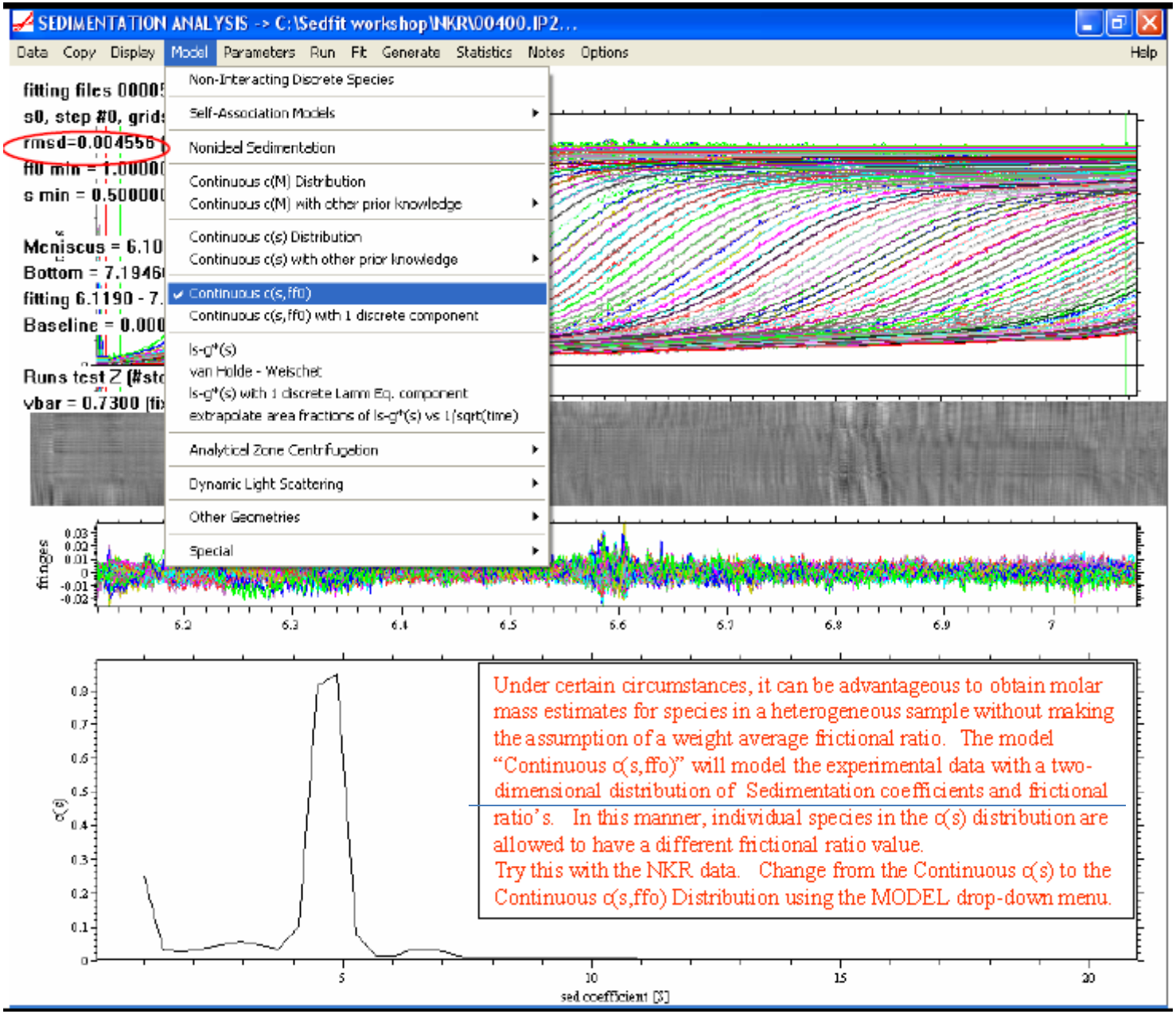

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
fitting files 00005.IP2 to 00400.IP2
s6, step #10, gridsize = 500
rmsd=0.004279 (n=107920,SSR=1.976283)
M =
frict ratio = 1.549700
Meniscus = 6.105600
Bottom = 7.194600
fitting 6.1190 - 7.1221
Baseline = 0.000000
Runs test Z (#stddev) = 164.79
vbar = 0.7300 (fix)
After parameters are entered, perform a RUN.
fringes
c(M)
molar mass [Da]

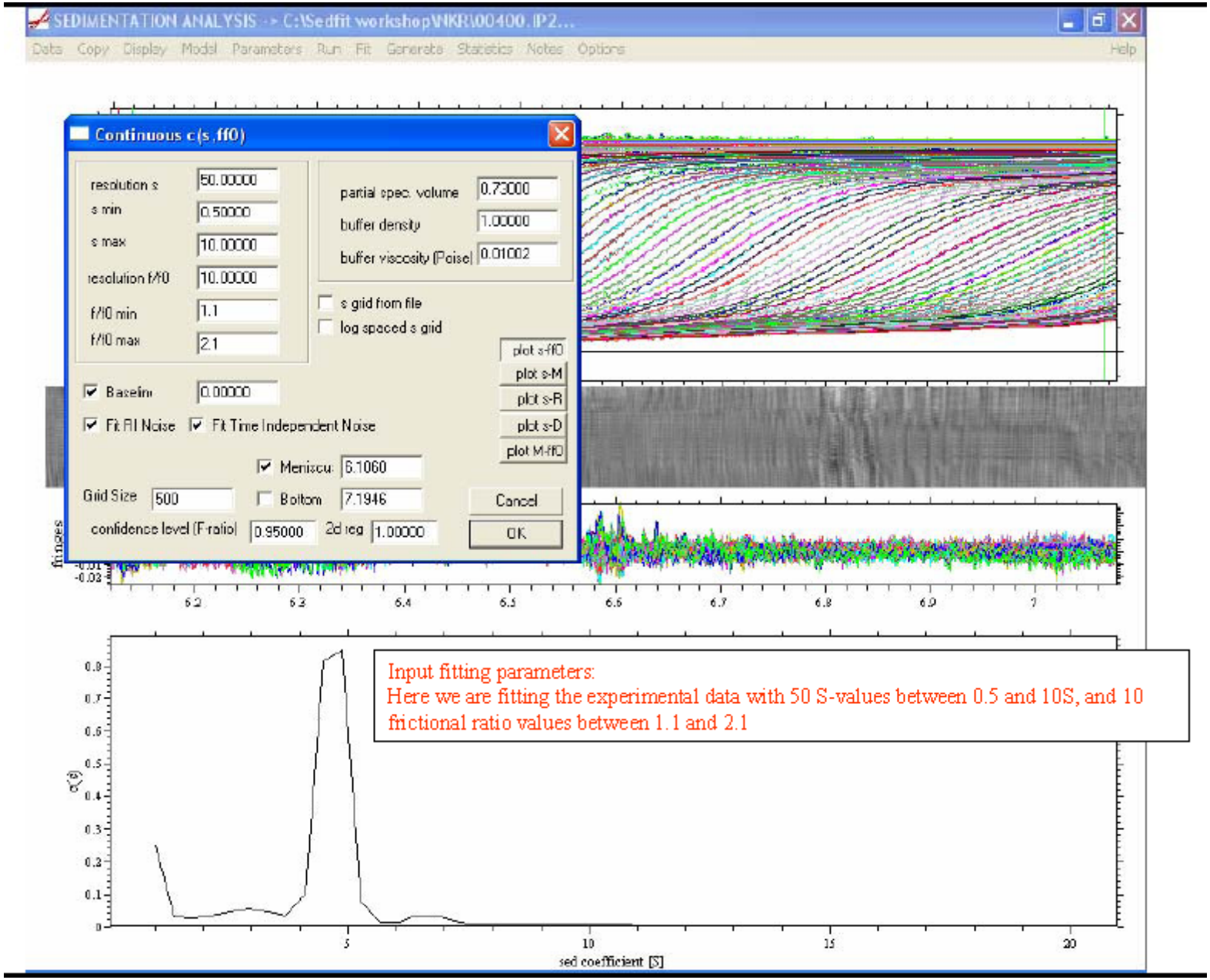

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
Set v*rho
Inhomogeneous Solvent
Loading Options
Fitting Options
Size-Distribution Options
Calculator
Bitmaps
Sound
transform s-distribution to r-distribution
Regularization by 2nd derivative
Regularization by Max Ent
integrate distribution ^I
transfrom differential to integral distribution
transform g(s*) (as ASCII) to direct boundary model
transform c(s) to c(M)
fringes
Any area of the distribution can be integrated to obtain average molecular weights for species in the distribution.
To integrate the area under a curve, select OPTIONS→ Size-Distribution Options→ Integrate Distribution (or press Contol+I).
c(M)
50000
100000
150000
200000
250000
molar mass [Da]

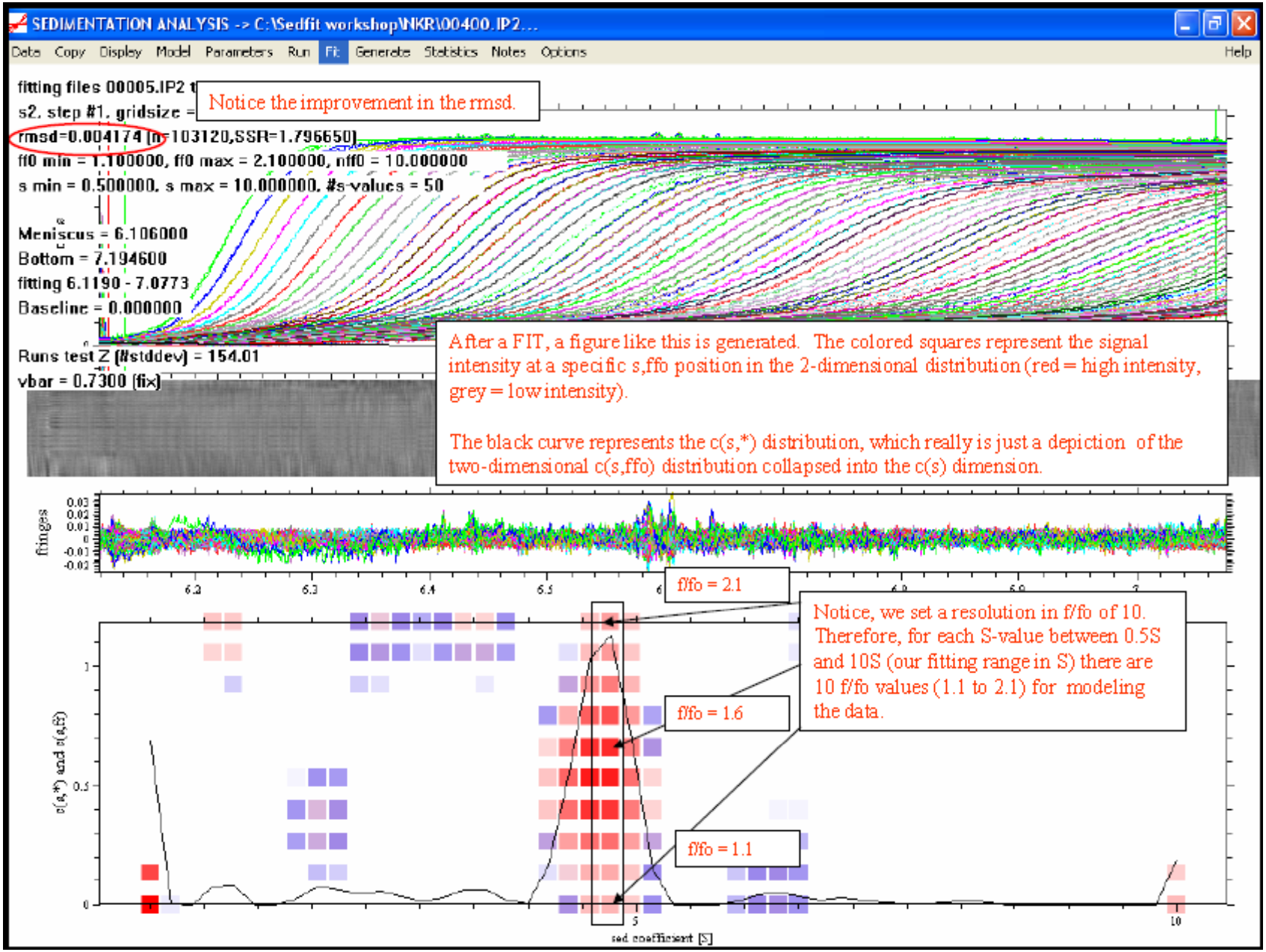

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
fringes
SELECT INTEGRATION RANGE
determine the range by drawing a rectangle while the right mouse button is down
OK
This message appears telling the user to click-and-drag a box around the area of interest (using the right mouse button)
c(M)
50000
100000
150000
200000
250000
molar mass [Da]

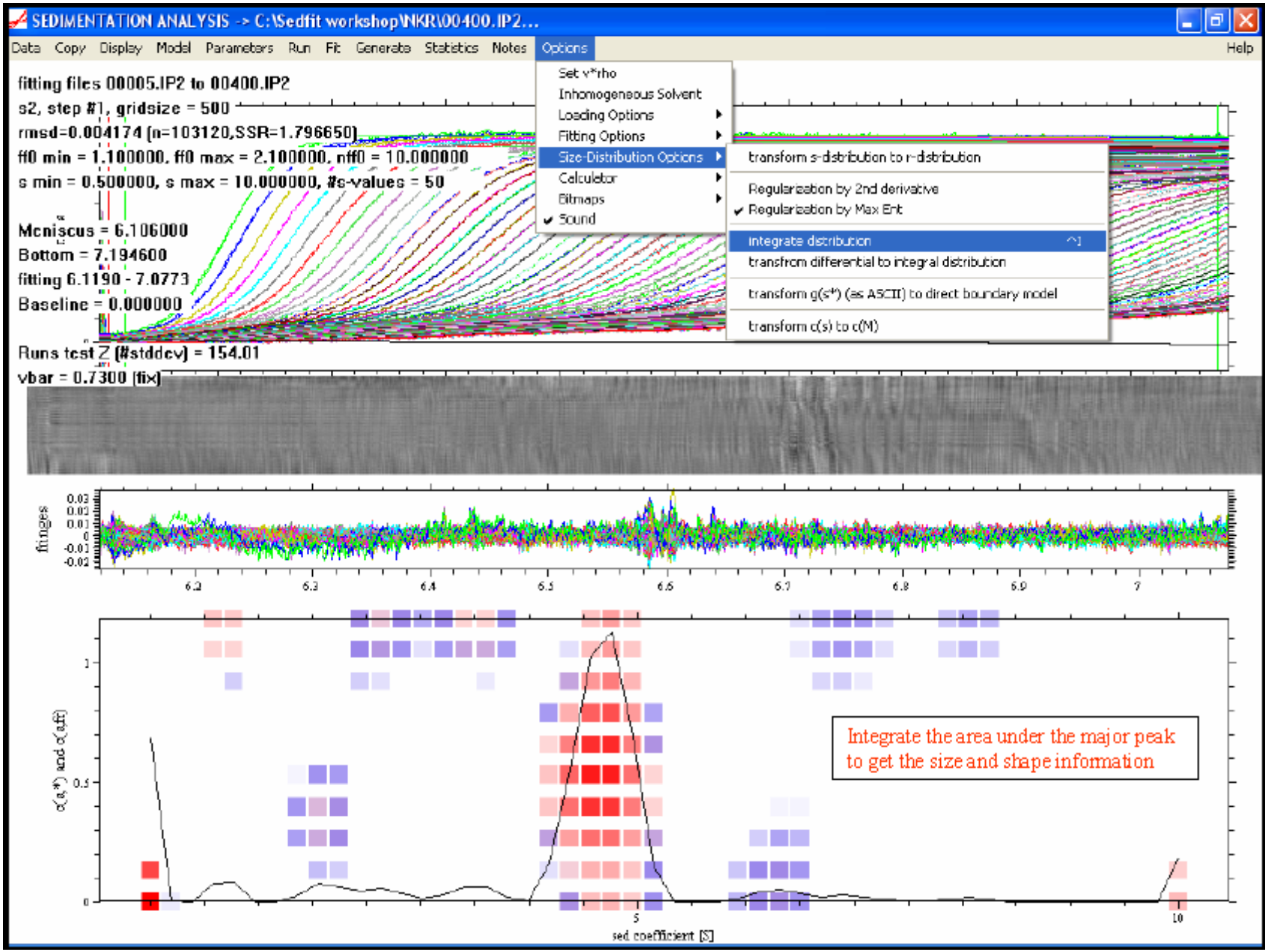

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
116005.133
0.000
fringes
For example, the major peak...
c(M)
molar mass [Da]

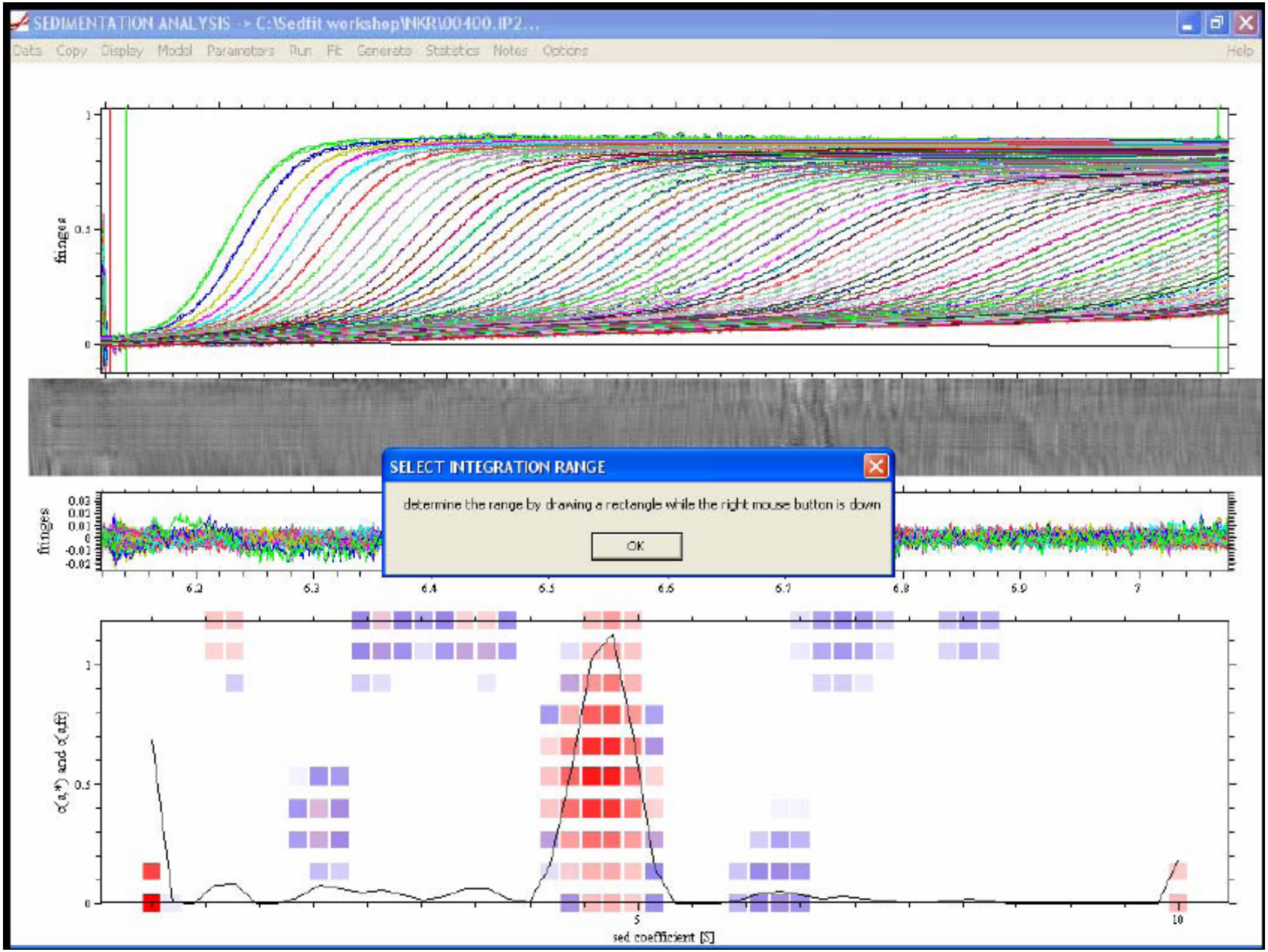

SEDIMENTATION ANALYSIS - > C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
fringes
INTEGRATE DISTRIBUTION
Integral from 75275.851562 to 116005.132812:
Loading Concentration [Signal] = 0.702066 (65.077%of total)
Weight (Signal) Average = 91574.710938
Std. deviation[as sqrt(second central moment)] = 6909.187 (best done without regularization)
OK
c(M)
molar mass [Da]
...has a molar mass of 91.6 kDa. Notice the "Std. deviation" value of 6909 Da. This value is not the precision of the molar mass parameter, but rather a measure of the width of the distribution. The width of the distribution can be affected by the regularization processs, the micro heterogeneity of the sample, as well as the signal-to-nose ratio.

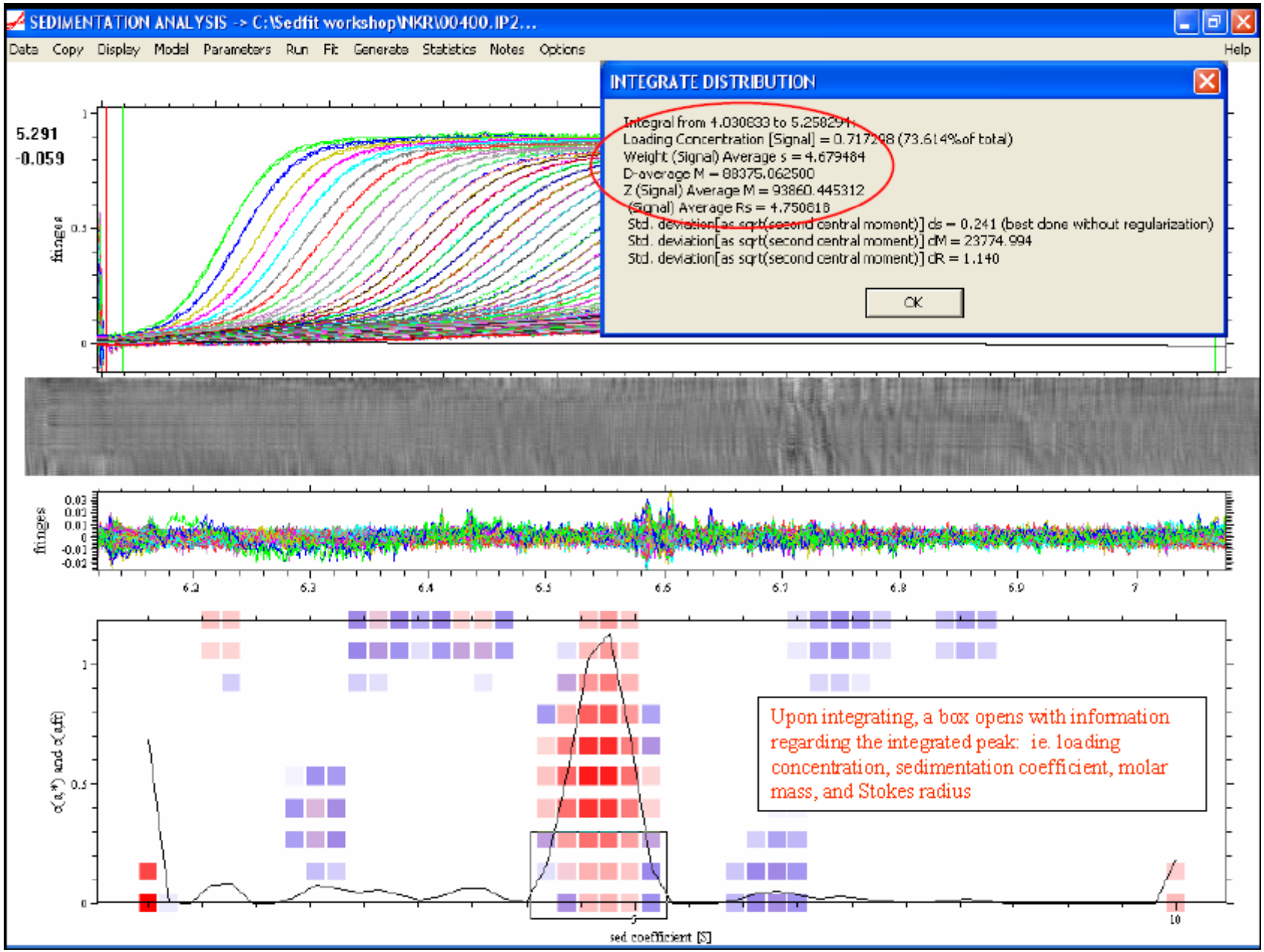

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
Non-Interacting Discrete Species
Self-Association Models
Nonideal Sedimentation
Continuous c(M) Distribution
Continuous c(M) with other prior knowledge
Continuous c(s) Distribution
Continuous c(s) with other prior knowledge
Continuous c(s,ff0)
Continuous c(s,ff0) with 1 discrete component
ls-g*(s)
van Holde - Weischet
ls-g*(s) with 1 discrete Lamm Eq. component
extrapolate area fractions of ls-g*(s) vs 1/sqrt(time)
Analytical Zone Centrifugation
Dynamic Light Scattering
Other Geometries
Special
fitting files 0000
s0, step #0, grid
rmsd=0.004556
ff0 min = 1.0000
s min = 0.50000
Meniscus = 6.10
Bottom = 7.1946
fitting 6.1190 - 7.
Baseline = 0.000
Runs test Z (#st
vbar = 0.7300 (fi
fringes
c(s)
sed coefficient [S]
Under certain circumstances, it can be advantageous to obtain molar mass estimates for species in a heterogeneous sample without making the assumption of a weight average frictional ratio. The model "Continuous c(s,ffo)" will model the experimental data with a two-dimensional distribution of Sedimentation coefficients and frictional ratio's. In this manner, individual species in the c(s) distribution are allowed to have a different frictional ratio value.
Try this with the NKR data. Change from the Continuous c(s) to the Continuous c(s,ffo) Distribution using the MODEL drop-down menu.

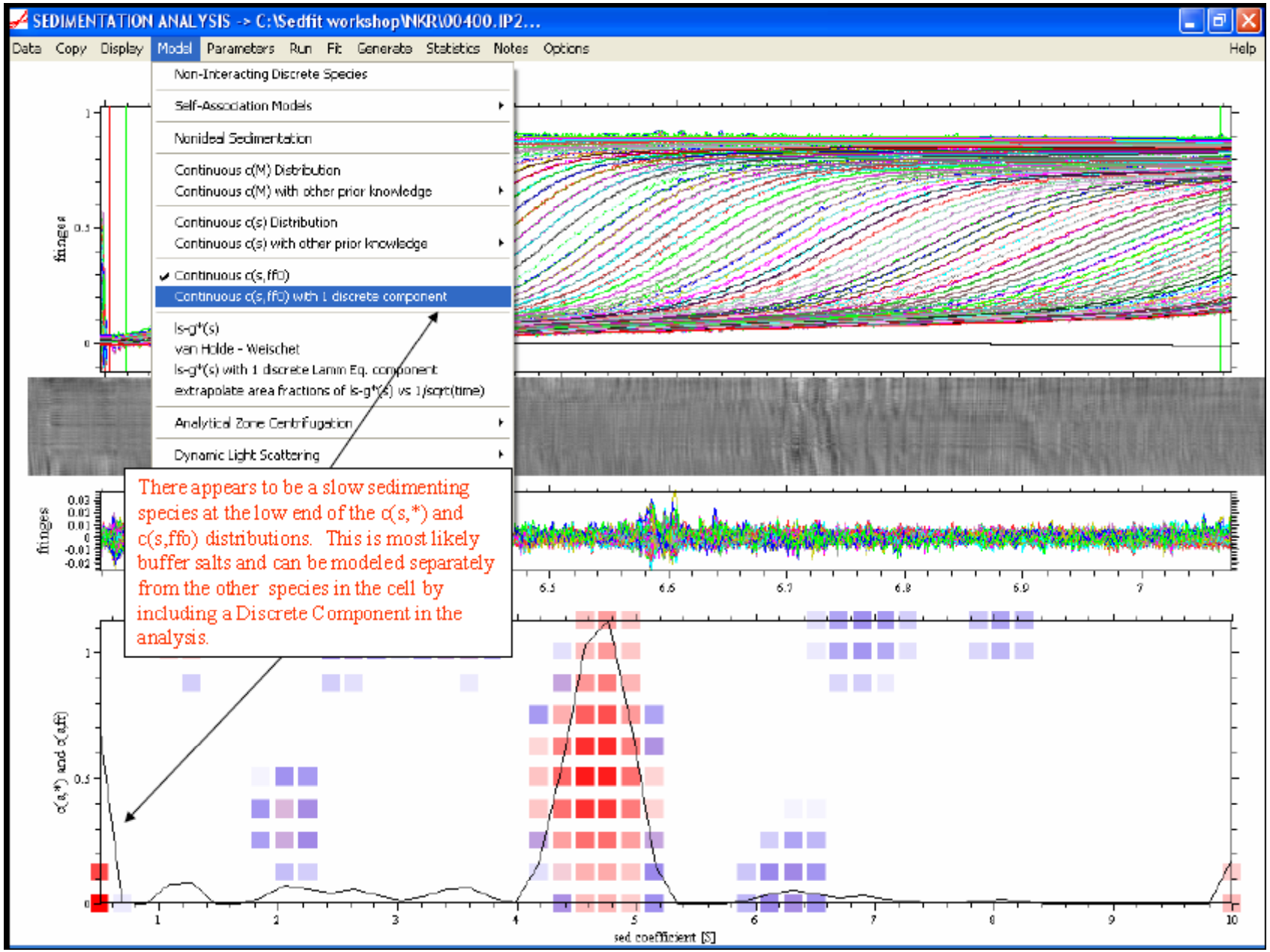

SEDIMENTATION ANALYSIS - C:\Sedfit workshopVNKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
Continuous c(s,ff0)
resolution s
50.00000
s min
0.50000
s max
10.00000
resolution f/f0
10.00000
f/f0 min
1.1
f/f0 max
2.1
partial spec. volume
0.73000
buffer density
1.00000
buffer viscosity (Poise)
0.01002
s grid from file
log spaced s grid
plot s-ff0
plot s-M
plot s-R
plot s-D
plot M-ff0
Baselin
0.00000
Fit RI Noise
Fit Time Independent Noise
Meniscu
6.1060
Grid Size
500
Bottom
7.1946
Cancel
confidence level (F-ratio)
0.95000
2d reg
1.00000
OK
fringes
c(s)
sed coefficient [S]
Input fitting parameters:
Here we are fitting the experimental data with 50 S-values between 0.5 and 10S, and 10 frictional ratio values between 1.1 and 2.1

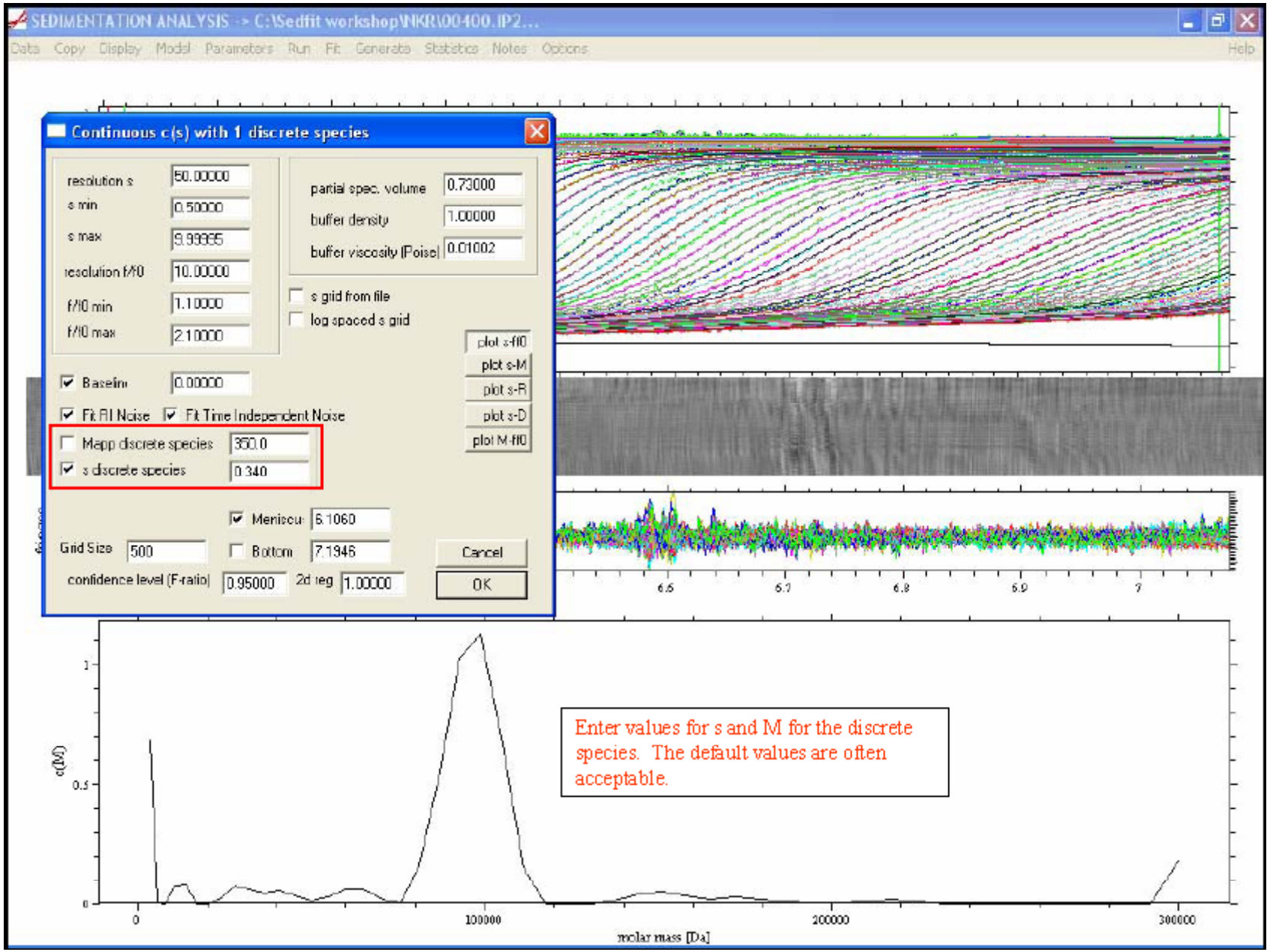

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
fitting files 00005.IP2 t
s2, step #1, gridsize =
rmsd=0.004174 [n=103120,SSR=1.796650]
ff0 min = 1.100000, ff0 max = 2.100000, nff0 = 10.000000
s min = 0.500000, s max = 10.000000, #s-values = 50
Meniscus = 6.106000
Bottom = 7.194600
fitting 6.1190 - 7.0773
Baseline = 0.000000
Runs test Z (#stddev) = 154.01
vbar = 0.7300 (fix)
Notice the improvement in the rmsd.
After a FIT, a figure like this is generated. The colored squares represent the signal intensity at a specific s,ffo position in the 2-dimensional distribution (red = high intensity, grey = low intensity).
The black curve represents the c(s,*) distribution, which really is just a depiction of the two-dimensional c(s,ffo) distribution collapsed into the c(s) dimension.
fringes
f/fo = 2.1
f/fo = 1.6
f/fo = 1.1
Notice, we set a resolution in f/fo of 10. Therefore, for each S-value between 0.5S and 10S (our fitting range in S) there are 10 f/fo values (1.1 to 2.1) for modeling the data.
c(s,*) and c(s,ff)
sed. coefficient [S]

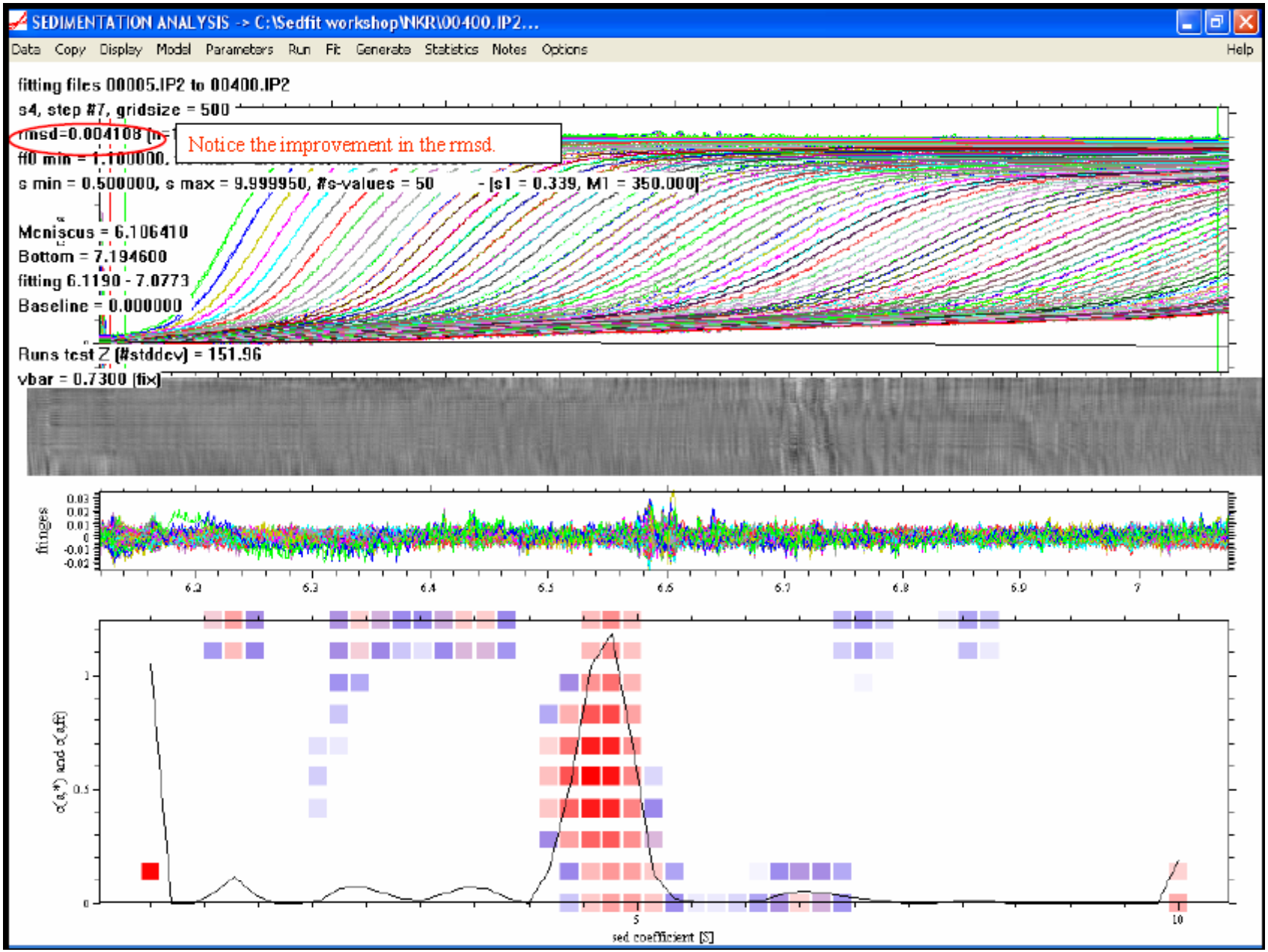

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
Set v*rho
Inhomogeneous Solvent
Loading Options
Fitting Options
Size-Distribution Options
Calculator
Bitmaps
Sound
transform s-distribution to r-distribution
Regularization by 2nd derivative
Regularization by Max Ent
Integrate distribution
transfrom differential to integral distribution
transform g(s*) (as ASCII) to direct boundary model
transform c(s) to c(M)
fitting files 00005.IP2 to 00400.IP2
s2, step #1, gridsize = 500
rmsd=0.004174 (n=103120,SSR=1.796650)
ff0 min = 1.100000, ff0 max = 2.100000, nff0 = 10.000000
s min = 0.500000, s max = 10.000000, #s-values = 50
Meniscus = 6.106000
Bottom = 7.194600
fitting 6.1190 - 7.0773
Baseline = 0.000000
Runs test Z (#stddev) = 154.01
vbar = 0.7300 (fix)
fringes
c(s,ff) and c(ff)
sed coefficient [S]
Integrate the area under the major peak
to get the size and shape information

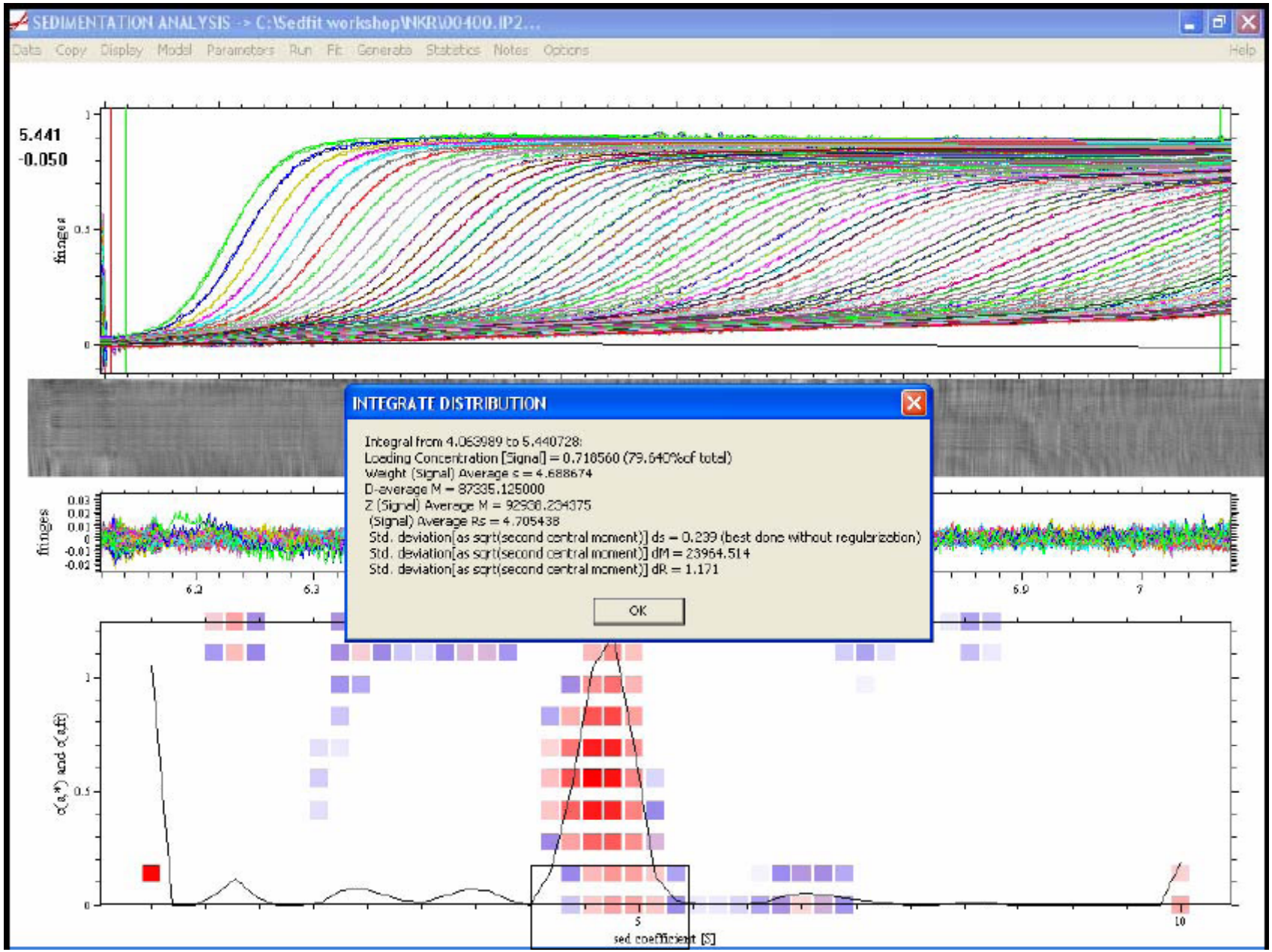

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
fringes
SELECT INTEGRATION RANGE
determine the range by drawing a rectangle while the right mouse button is down
OK
fringes
σ(a,*) and σ(aff)
sed coefficient [S]

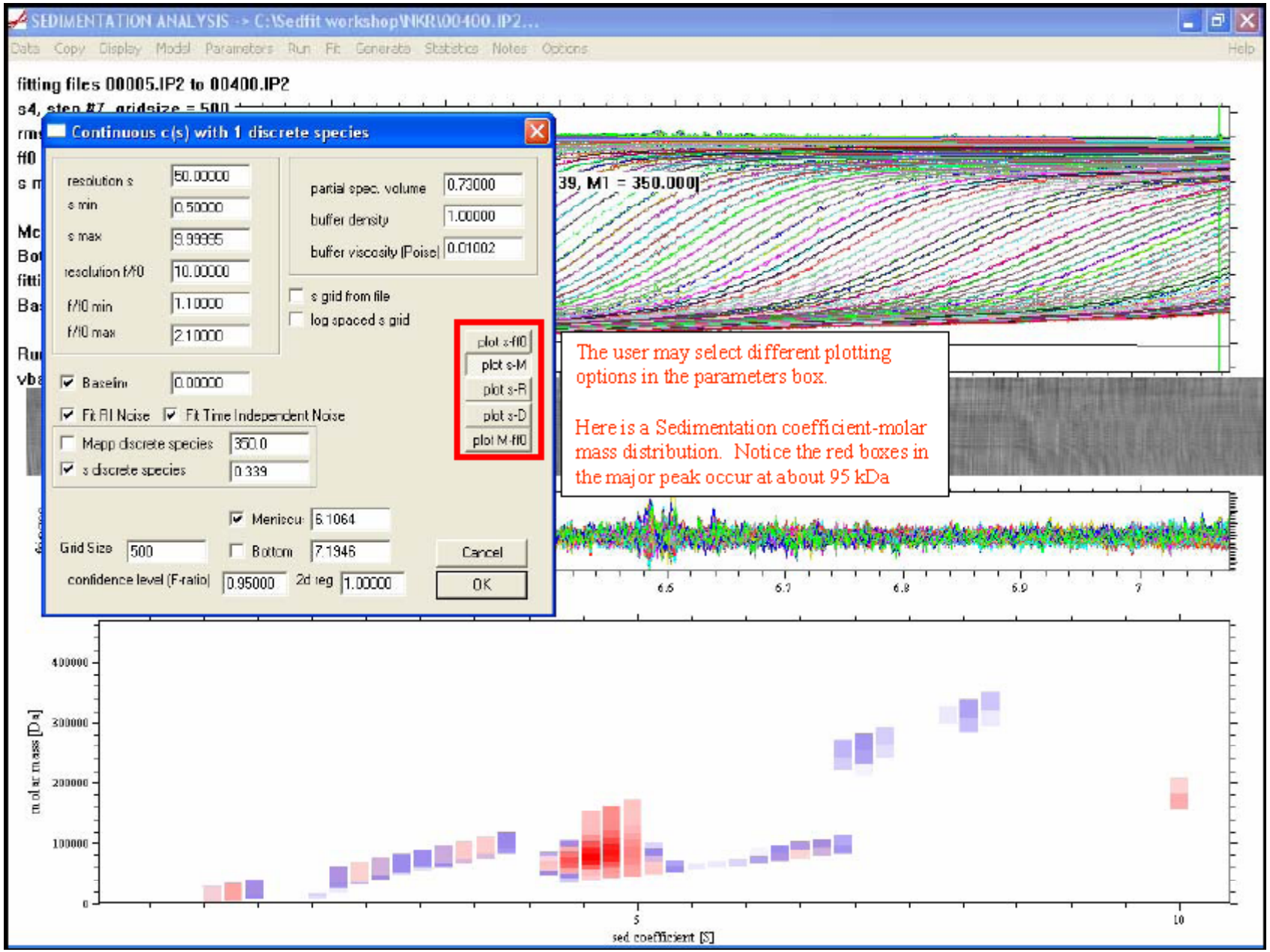

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
INTEGRATE DISTRIBUTION
Integral from 4.030833 to 5.258294
Loading Concentration [Signal] = 0.717298 (73.614%of total)
Weight (Signal) Average s = 4.679484
D-average M = 88375.062500
Z (Signal) Average M = 93860.445312
(Signal) Average Rs = 4.750618
Std. deviation[as sqrt(second central moment)] ds = 0.241 (best done without regularization)
Std. deviation[as sqrt(second central moment)] dM = 23774.994
Std. deviation[as sqrt(second central moment)] dR = 1.140
OK
5.291
-0.059
fringes
sed coefficient [S]
Upon integrating, a box opens with information regarding the integrated peak: ie. loading concentration, sedimentation coefficient, molar mass, and Stokes radius

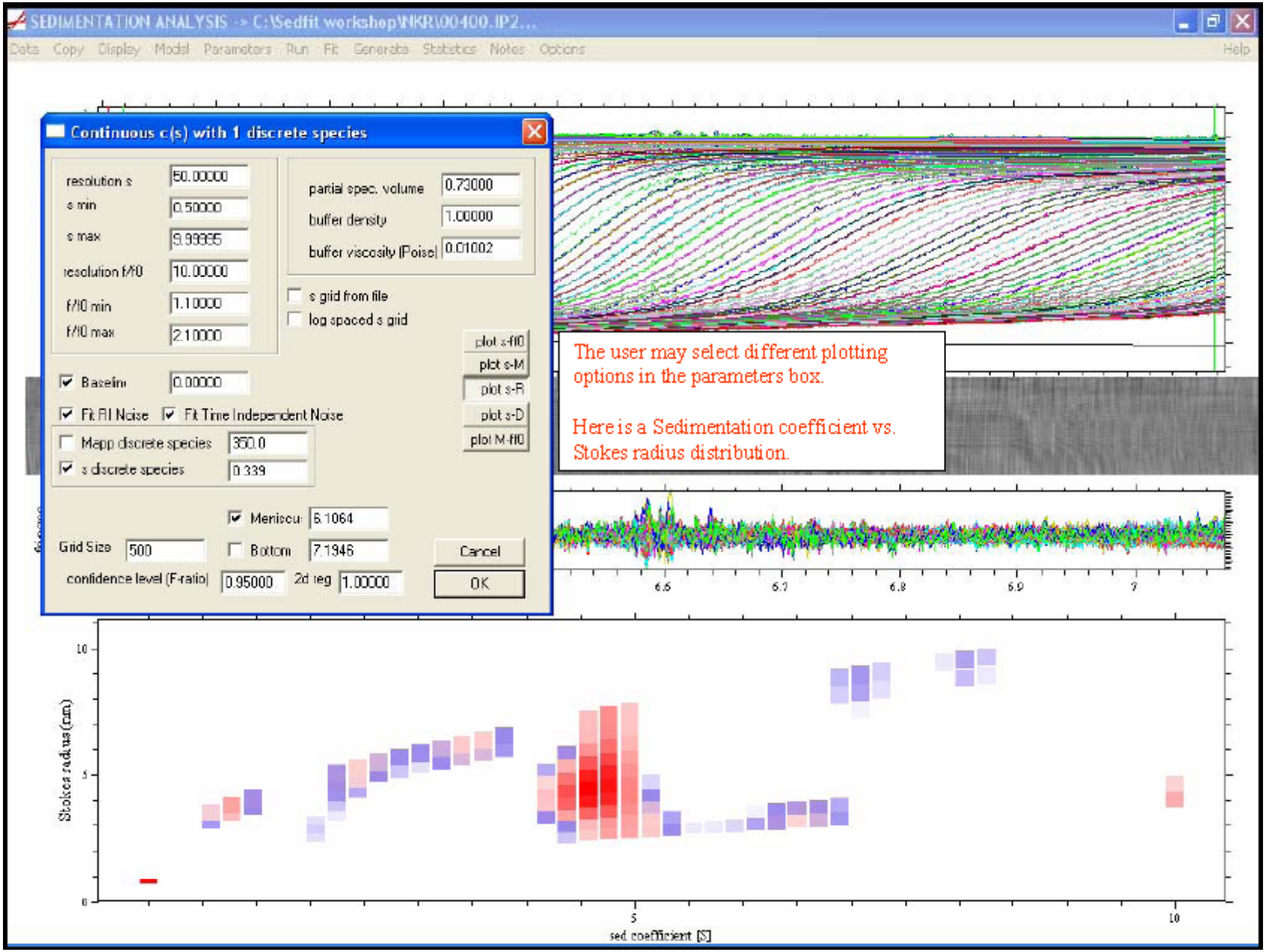

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
Non-Interacting Discrete Species
Self-Association Models
Nonideal Sedimentation
Continuous c(M) Distribution
Continuous c(M) with other prior knowledge
Continuous c(s) Distribution
Continuous c(s) with other prior knowledge
Continuous c(s,ffo)
Continuous c(s,ffo) with 1 discrete component
ls-g*(s)
van Holde - Weischet
ls-g*(s) with 1 discrete Lamm Eq. component
extrapolate area fractions of ls-g*(s) vs 1/sqrt(time)
Analytical Zone Centrifugation
Dynamic Light Scattering
There appears to be a slow sedimenting species at the low end of the c(s,*) and c(s,ffo) distributions. This is most likely buffer salts and can be modeled separately from the other species in the cell by including a Discrete Component in the analysis.
fringes
c(s,*) and c(s,ff)
sed coefficient [S]

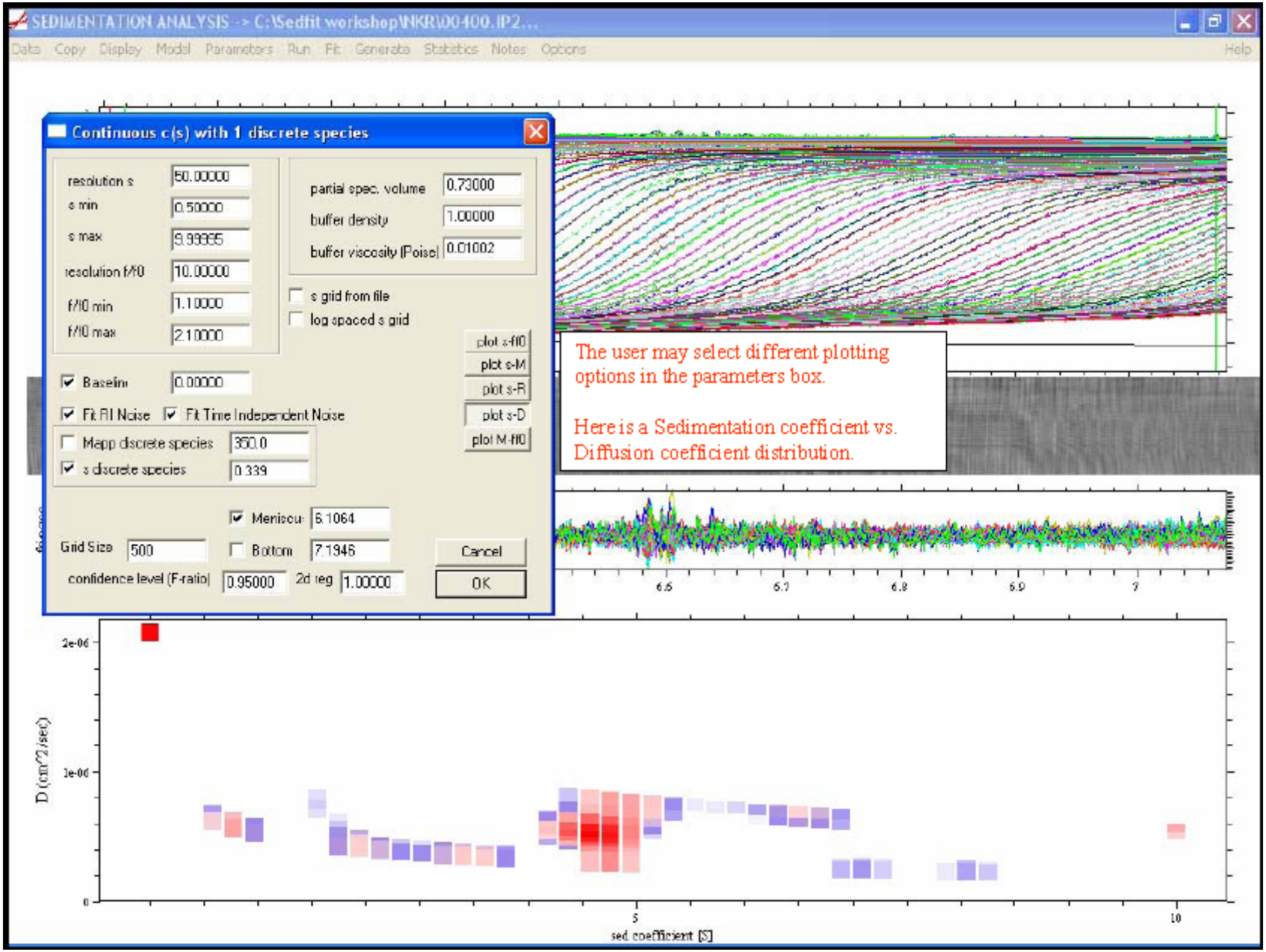

SEDIMENTATION ANALYSIS - > C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
Continuous c(s) with 1 discrete species
resolution s
50.00000
s min
0.50000
s max
9.99995
resolution f/f0
10.00000
f/f0 min
1.10000
f/f0 max
2.10000
partial spec. volume
0.73000
buffer density
1.00000
buffer viscosity [Poise]
0.01002
s grid from file
log spaced s grid
plot s-f/f0
plot s-M
plot s-R
plot s-D
plot M-f/f0
Baseline
0.00000
Fit RI Noise
Fit Time Independent Noise
Mapp discrete species
350.0
s discrete species
0.340
Meniscus
6.1060
Grid Size
500
Bottom
7.1946
Cancel
confidence level (F-ratio)
0.95000
2d reg
1.00000
OK
6.6
6.7
6.8
6.9
7
c(M)
1
0.5
0
0
100000
200000
300000
molar mass [Da]
Enter values for s and M for the discrete species. The default values are often acceptable.

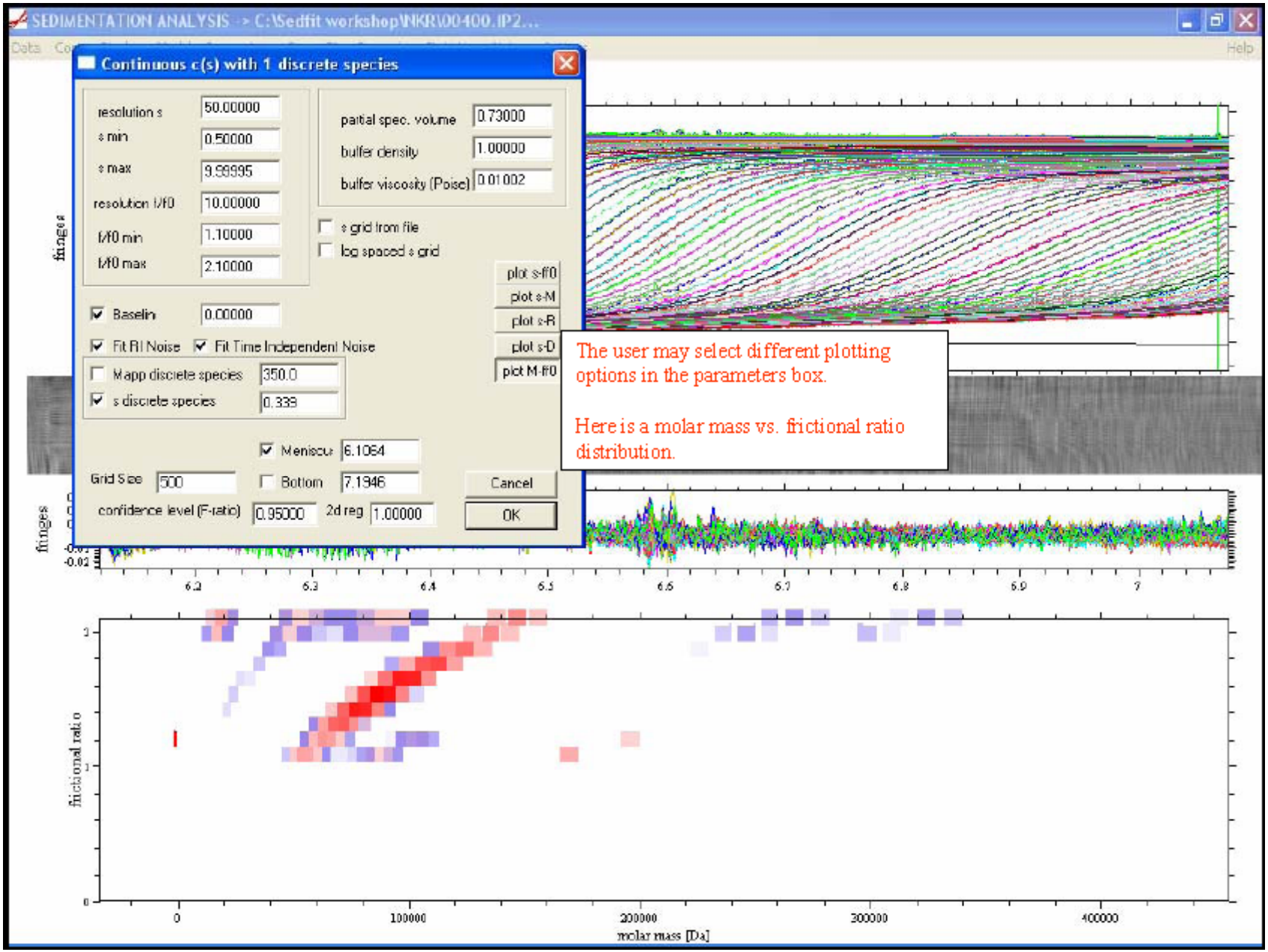

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
fitting files 00005.IP2 to 00400.IP2
s4, step #7, gridsize = 500
rmsd=0.004108
Notice the improvement in the rmsd.
ff0 min = 1.100000,
s min = 0.500000, s max = 9.999950, #s-values = 50 - [s1 = 0.339, M1 = 350.000]
Meniscus = 6.106410
Bottom = 7.194600
fitting 6.1190 - 7.0773
Baseline = 0.000000
Runs test Z (#stddev) = 151.96
vbar = 0.7300 (fix)
fringes
sed coefficient [S]

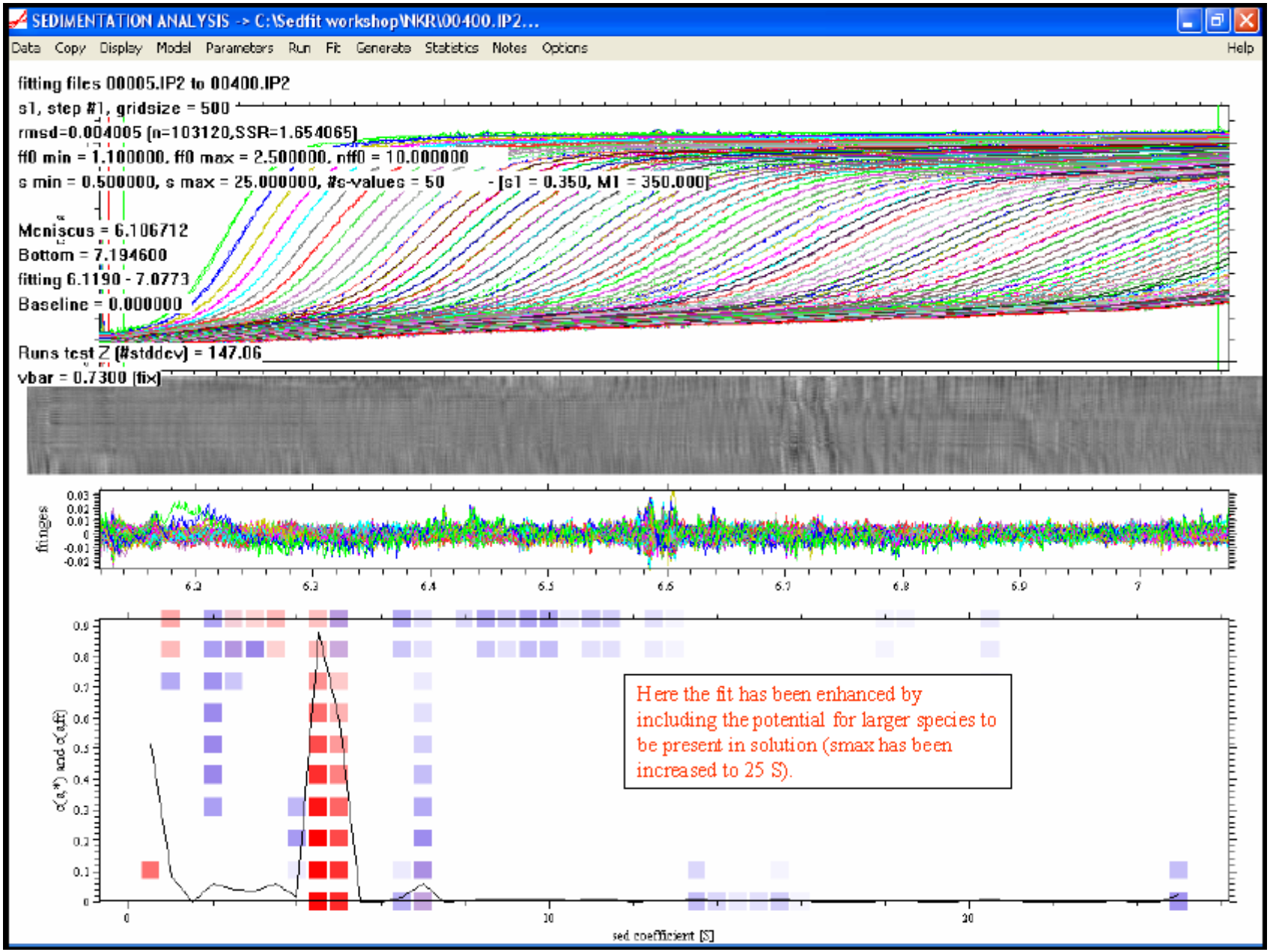

SEDIMENTATION ANALYSIS - > C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
5.441
-0.050
fringes
INTEGRATE DISTRIBUTION
Integral from 4.063989 to 5.440728:
Loading Concentration [Signal] = 0.718560 (79.640%of total)
Weight (Signal) Average s = 4.688674
D-average M = 87335.125000
Z (Signal) Average M = 92938.234375
(Signal) Average Rs = 4.705438
Std. deviation[as sqrt(second central moment)] ds = 0.239 (best done without regularization)
Std. deviation[as sqrt(second central moment)] dM = 23964.514
Std. deviation[as sqrt(second central moment)] dR = 1.171
OK
σ(s,*) and σ(diff)
sed coefficient [S]

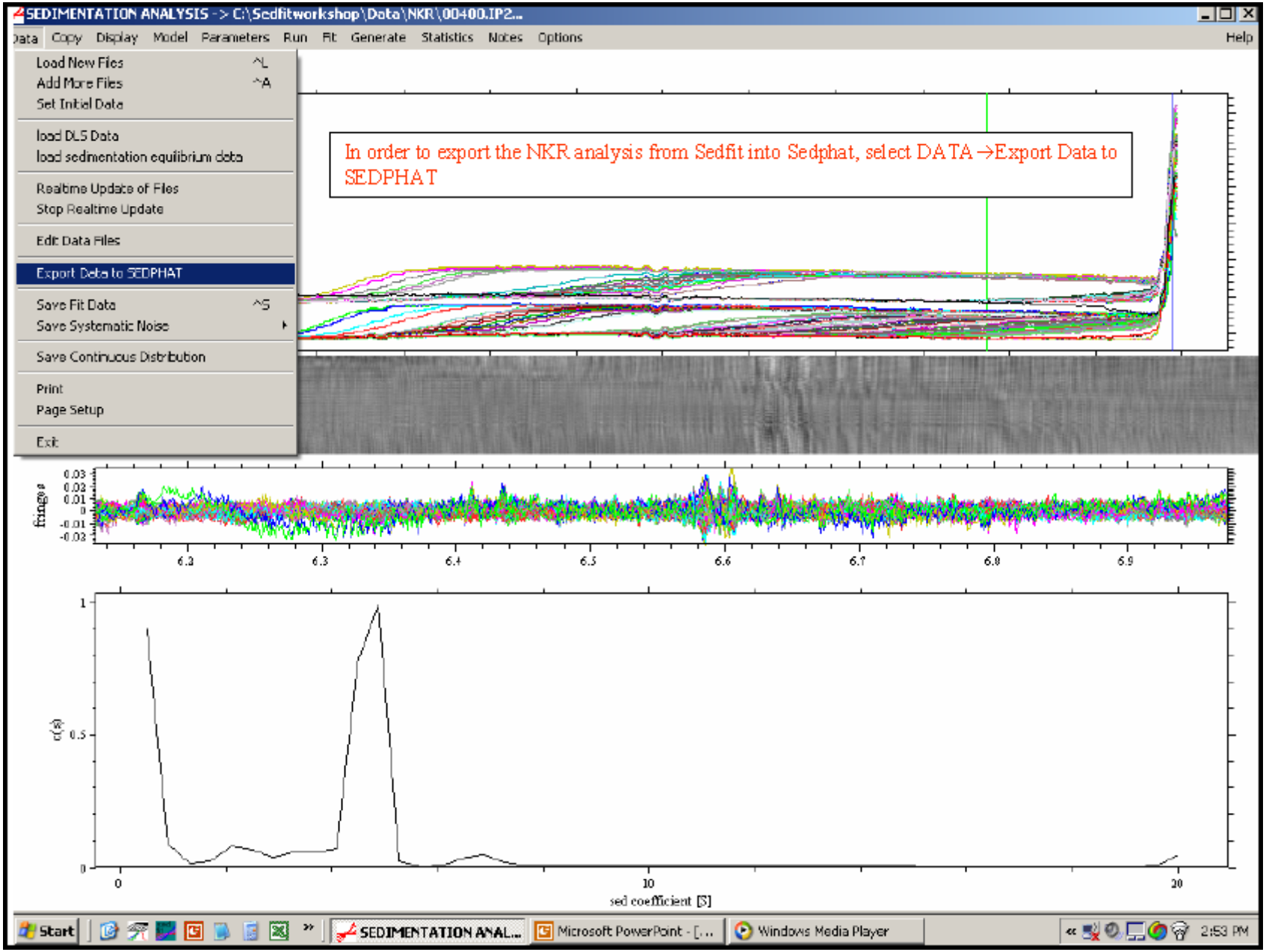

SEDIMENTATION ANALYSIS - > C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
fitting files 00005.IP2 to 00400.IP2
Continuous c(s) with 1 discrete species
resolution s 50.00000
s min 0.50000
s max 9.99995
resolution f/f0 10.00000
f/f0 min 1.10000
f/f0 max 2.10000
partial spec. volume 0.73000
buffer density 1.00000
buffer viscosity [Poise] 0.01002
s grid from file
log spaced s grid
plot s-ff0
plot s-M
plot s-R
plot s-D
plot M-ff0
Baselin 0.00000
Fit RI Noise
Fit Time Independent Noise
Mapp discrete species 350.0
s discrete species 0.339
Meniscu 6.1064
Grid Size 500
Bottom 7.1946
Cancel
confidence level (F-ratio) 0.95000
2d reg 1.00000
OK
39, M1 = 350.000
The user may select different plotting options in the parameters box.
Here is a Sedimentation coefficient-molar mass distribution. Notice the red boxes in the major peak occur at about 95 kDa
6.6
6.7
6.8
6.9
7
400000
300000
200000
100000
0
molar mass [Da]
5
10
sed. coefficient [S]

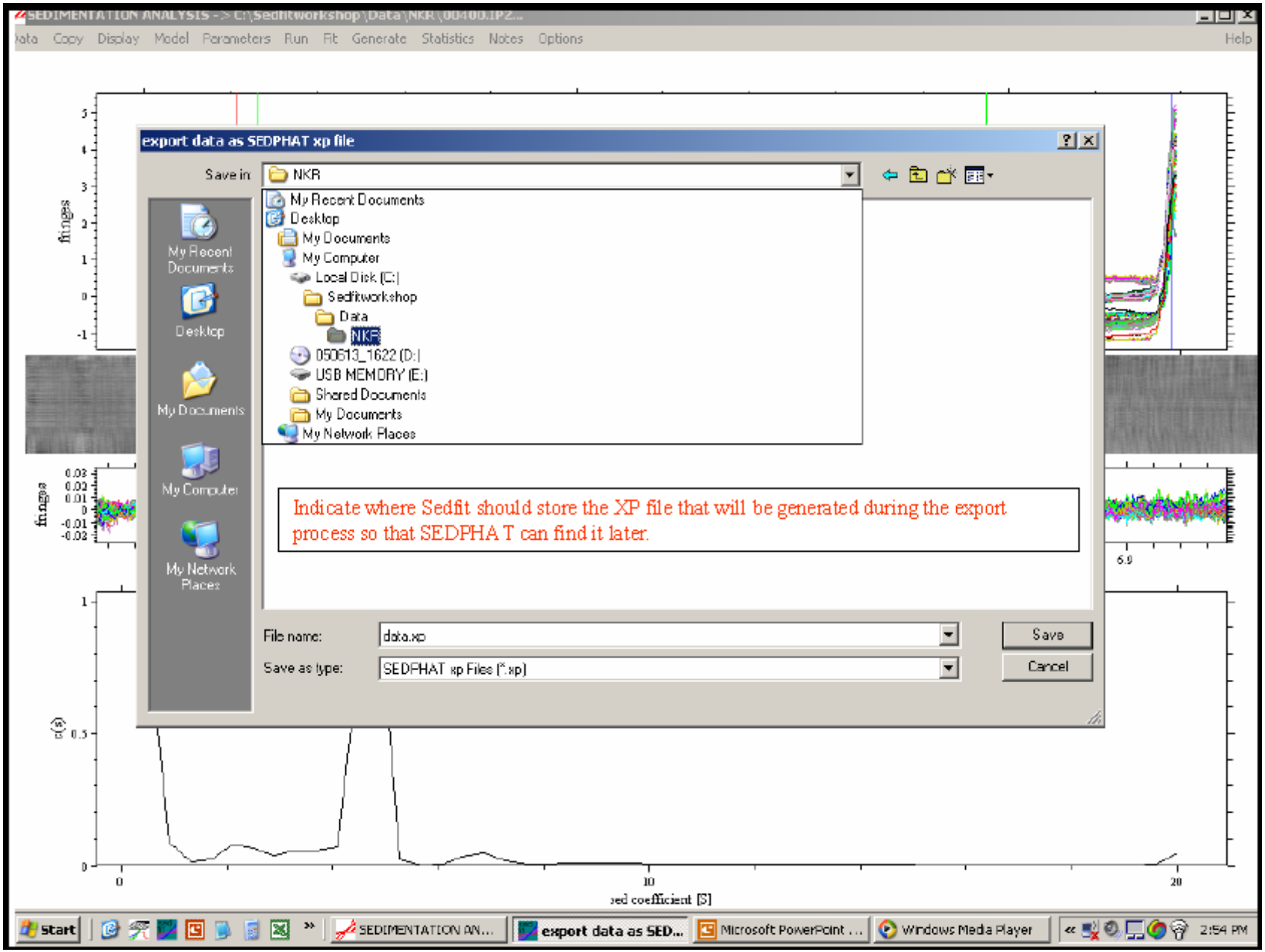

SEDIMENTATION ANALYSIS - > C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
Continuous c(s) with 1 discrete species
resolution s
50.00000
s min
0.50000
s max
9.99995
resolution f/f0
10.00000
f/f0 min
1.10000
f/f0 max
2.10000
partial spec. volume
0.73000
buffer density
1.00000
buffer viscosity [Poise]
0.01002
s grid from file
log spaced s grid
plot s-f/f0
plot s-M
plot s-R
plot s-D
plot M-f/f0
Baseline
0.00000
Fit RI Noise
Fit Time Independent Noise
Mapp discrete species
350.0
s discrete species
0.339
Meniscus
6.1064
Grid Size
500
Bottom
7.1946
Cancel
confidence level (F-ratio)
0.95000
2d reg
1.00000
OK
The user may select different plotting options in the parameters box.
Here is a Sedimentation coefficient vs. Stokes radius distribution.
6.6
6.7
6.8
6.9
7
10
5
0
Stokes radius (nm)
5
10
sed. coefficient [S]

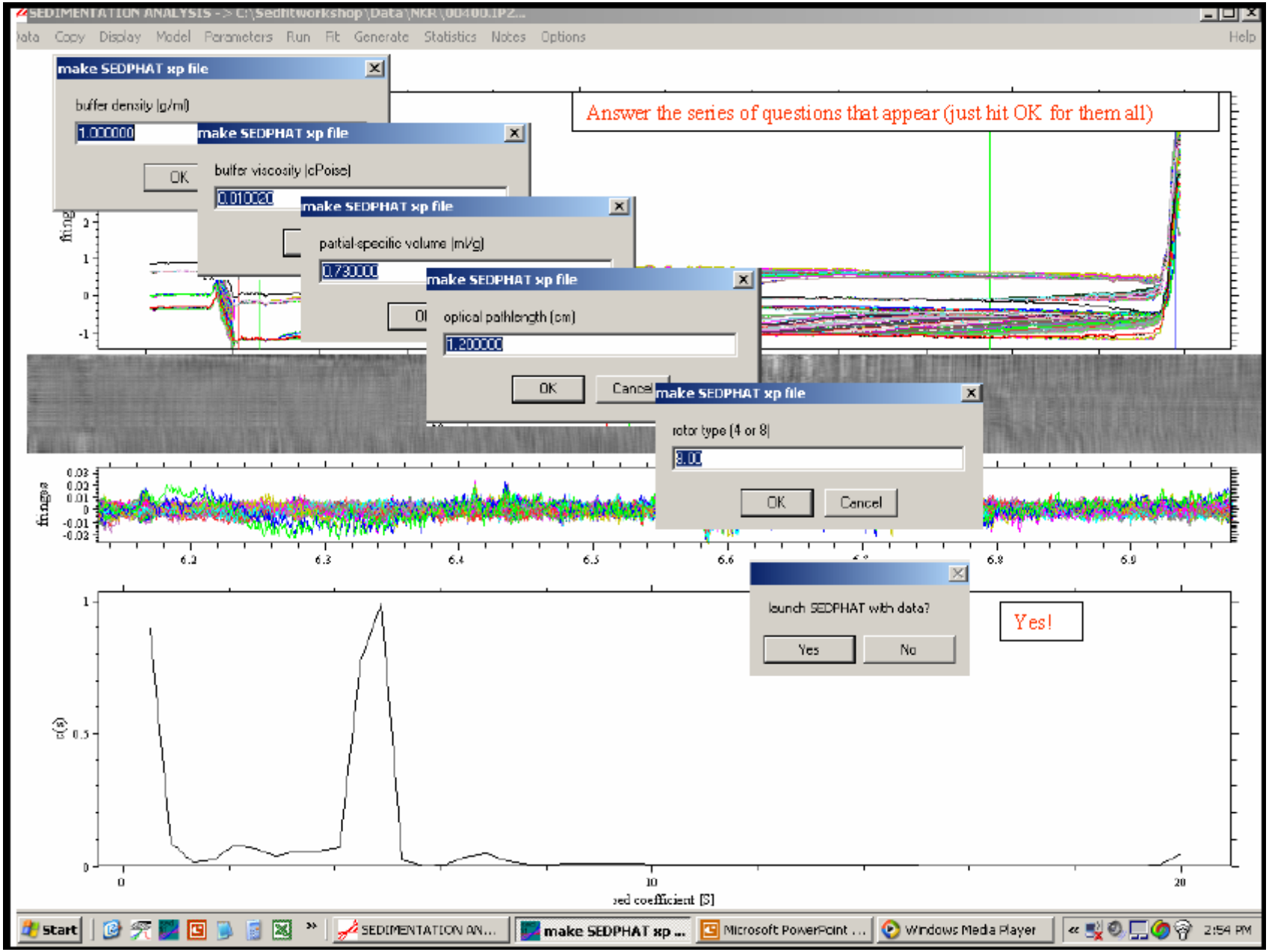

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
Continuous c(s) with 1 discrete species
resolution s
50.00000
s min
0.50000
s max
9.99995
resolution f/f0
10.00000
f/f0 min
1.10000
f/f0 max
2.10000
partial spec. volume
0.73000
buffer density
1.00000
buffer viscosity (Poise)
0.01002
s grid from file
log spaced s grid
plot s-f/f0
plot s-M
plot s-R
plot s-D
plot M-f/f0
Baseline
0.00000
Fit RI Noise
Fit Time Independent Noise
Mapp discrete species
350.0
s discrete species
0.339
Meniscus
5.1064
Grid Size
500
Bottom
7.1946
Cancel
confidence level (F-ratio)
0.95000
2d reg
1.00000
OK
The user may select different plotting options in the parameters box.
Here is a Sedimentation coefficient vs. Diffusion coefficient distribution.
6.5
6.7
6.8
6.9
7
2e-06
1e-06
0
D (cm^2/sec)
5
10
sed coefficient [S]

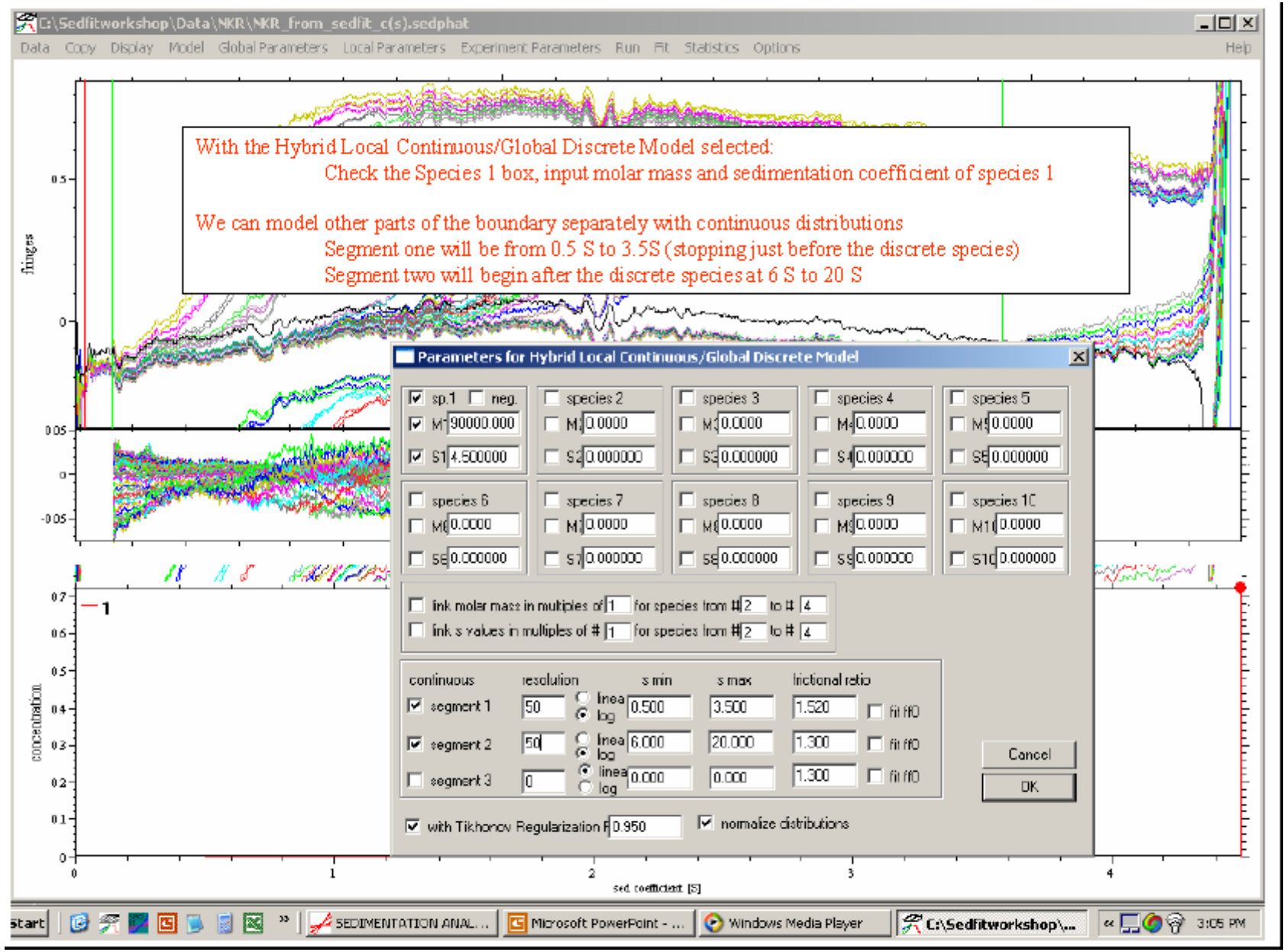

SEDIMENTATION ANALYSIS - > C:\Sedfit workshop\NKR\00400.IP2...
Continuous c(s) with 1 discrete species
resolution s
50.00000
s min
0.50000
s max
9.99995
resolution f/f0
10.00000
f/f0 min
1.10000
f/f0 max
2.10000
partial spec. volume
0.73000
buffer density
1.00000
buffer viscosity (Poise)
0.01002
s grid from file
log spaced s grid
plot s-ff0
plot s-M
plot s-R
plot s-D
plot M-ff0
Baselin
0.00000
Fit RI Noise
Fit Time Independent Noise
Mapp discrete species
350.0
s discrete species
0.339
Meniscu
6.1064
Grid Size
500
Bottom
7.1946
confidence level (F-ratio)
0.95000
2d reg
1.00000
Cancel
OK
The user may select different plotting options in the parameters box.
Here is a molar mass vs. frictional ratio distribution.
fringes
frictional ratio
molar mass [Da]

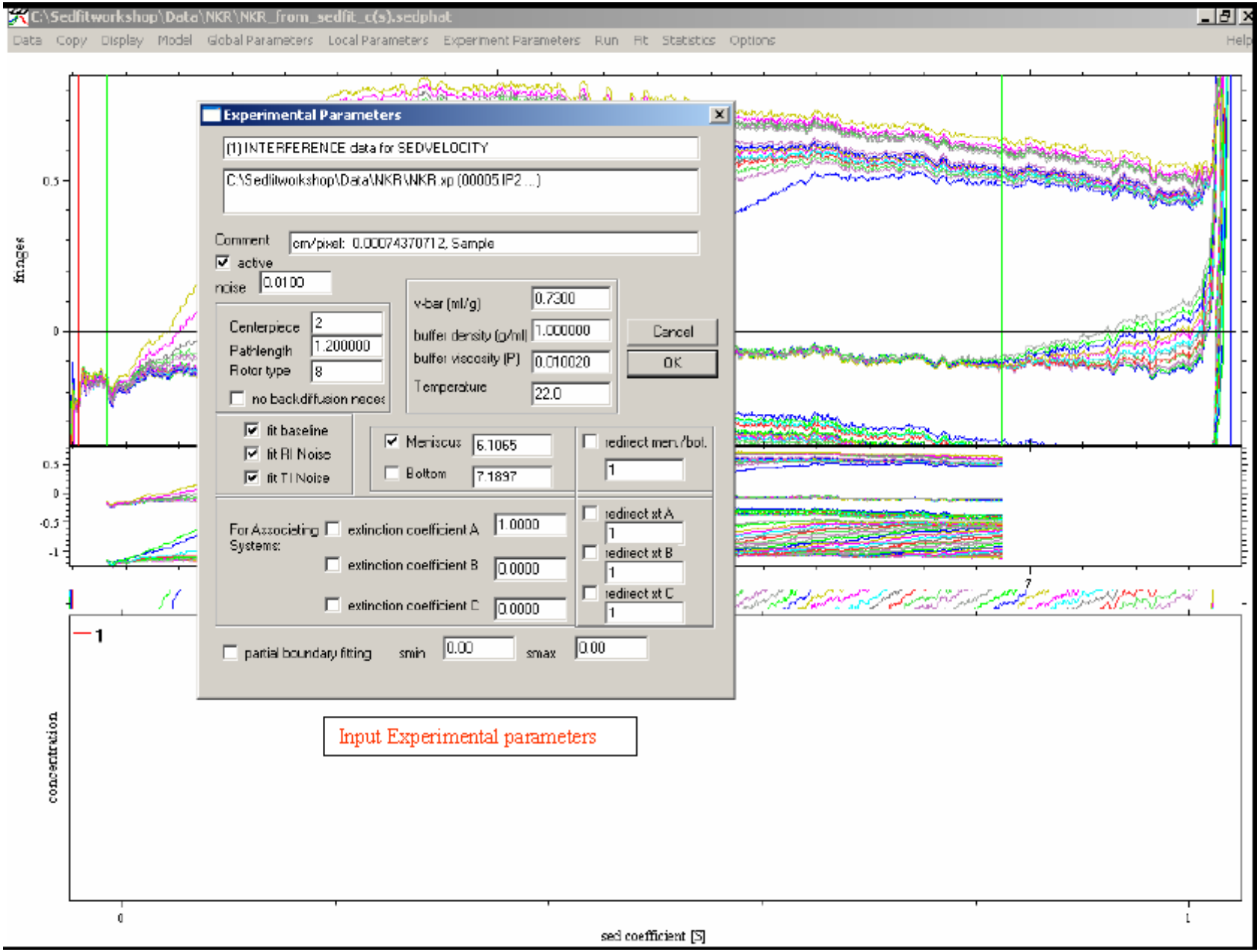

SEDIMENTATION ANALYSIS -> C:\Sedfit workshop\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
fitting files 00005.IP2 to 00400.IP2
s1, step #1, gridsize = 500
rmsd=0.004005 (n=103120,SSR=1.654065)
ff0 min = 1.100000, ff0 max = 2.500000, nff0 = 10.000000
s min = 0.500000, s max = 25.000000, #s-values = 50
- [s1 = 0.350, M1 = 350.000]
Meniscus = 6.106712
Bottom = 7.194600
fitting 6.1190 - 7.0773
Baseline = 0.000000
Runs test Z (#stddev) = 147.06
vbar = 0.7300 [fix]
fringes
c(s,ff0) and c(s,aff)
sed coefficient [S]
Here the fit has been enhanced by including the potential for larger species to be present in solution (smax has been increased to 25 S).

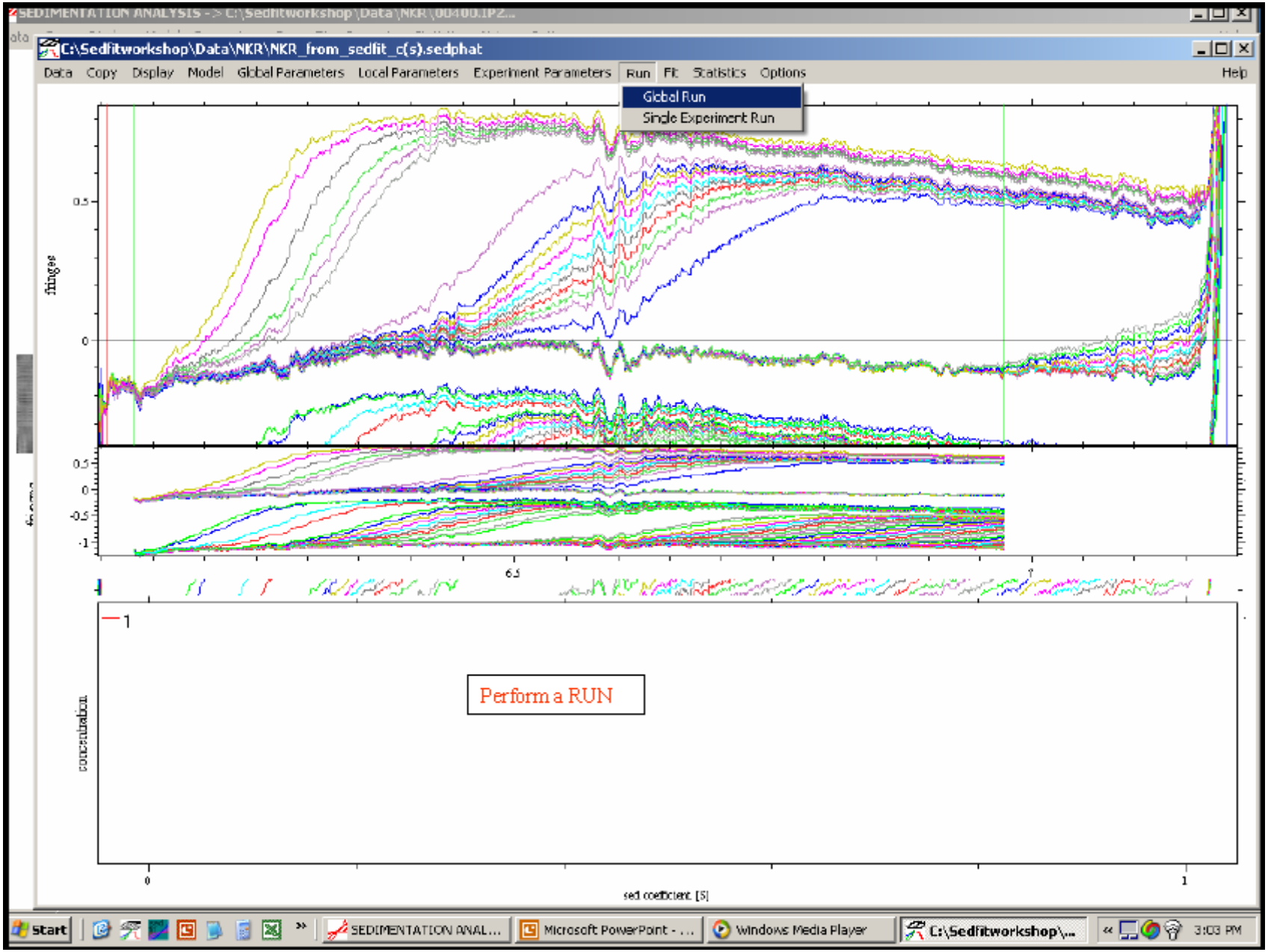

SEDIMENTATION ANALYSIS -> C:\Sedfitworkshop\Data\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
Load New Files ^L
Add More Files ^A
Set Initial Data
load DLS Data
load sedimentation equilibrium data
Realtime Update of Files
Stop Realtime Update
Edit Data Files
Export Data to SEDPHAT
Save Fit Data ^S
Save Systematic Noise
Save Continuous Distribution
Print
Page Setup
Exit
In order to export the NKR analysis from Sedfit into Sedphat, select DATA →Export Data to SEDPHAT
fringes
sed coefficient [S]
Start
SEDIMENTATION ANAL...
Microsoft PowerPoint - [...
Windows Media Player
2:53 PM

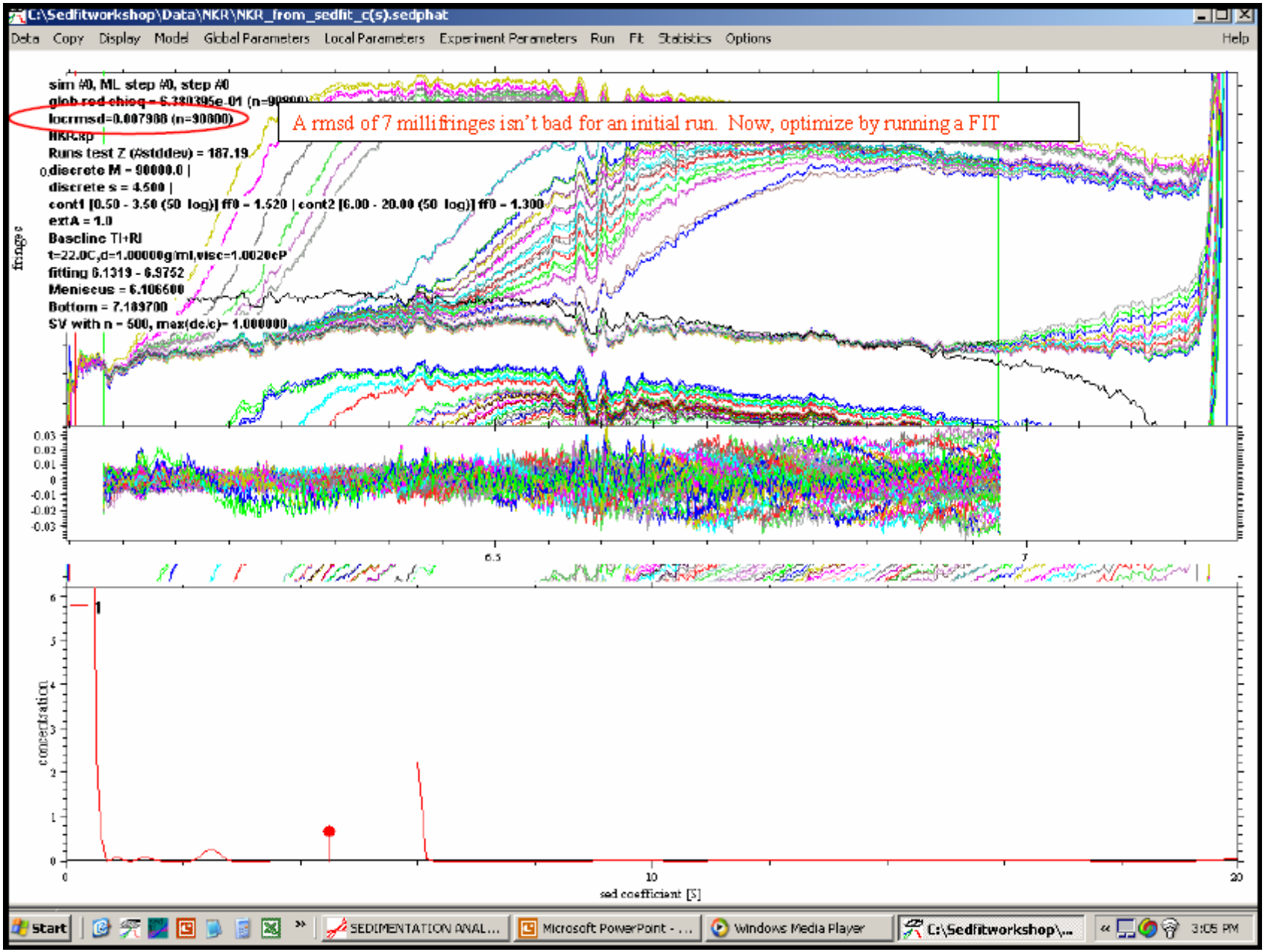

SEDIMENTATION ANALYSIS -> C:\Sedfitworkshop\Data\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
export data as SEDPHAT xp file
Save in: NKR
My Recent Documents
Desktop
My Documents
My Computer
Local Disk (C:)
Sedfitworkshop
Data
NKR
050613_1622 (D:)
USB MEMORY (E:)
Shared Documents
My Documents
My Network Places
Indicate where Sedfit should store the XP file that will be generated during the export process so that SEDPHAT can find it later.
File name: data.xp
Save as type: SEDPHAT xp Files (*.xp)
Save
Cancel
fringes
c(s)
sed coefficient [S]
start
SEDIMENTATION AN...
export data as SED...
Microsoft PowerPoint ...
Windows Media Player
2:54 PM

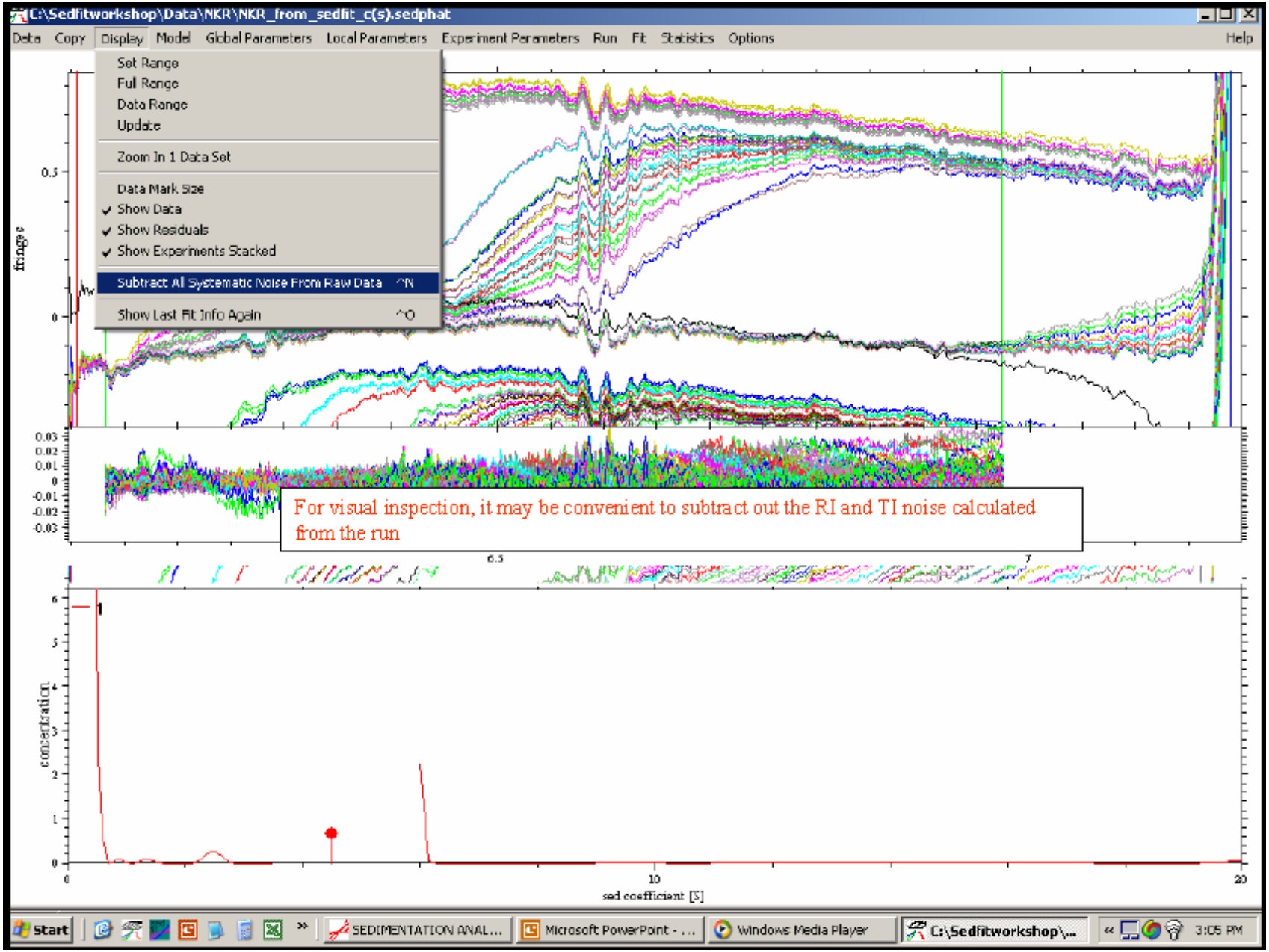

SEDIMENTATION ANALYSIS -> C:\SedNtworkshop\Data\NKR\00400.IP2...
Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
make SEDPHAT xp file
buffer density (g/ml)
1.000000
OK
make SEDPHAT xp file
buffer viscosity (cPoise)
0.010020
make SEDPHAT xp file
partial-specific volume (ml/g)
0.730000
make SEDPHAT xp file
optical pathlength (cm)
1.200000
OK
Cancel
make SEDPHAT xp file
rotor type (4 or 8)
8.00
OK
Cancel
Answer the series of questions that appear (just hit OK for them all)
launch SEDPHAT with data?
Yes
No
Yes!
fringes
6.2
6.3
6.4
6.5
6.6
6.8
6.9
c(s)
sed coefficient [S]
start
SEDIMENTATION AN...
make SEDPHAT xp ...
Microsoft PowerPoint ...
Windows Media Player
2:54 PM

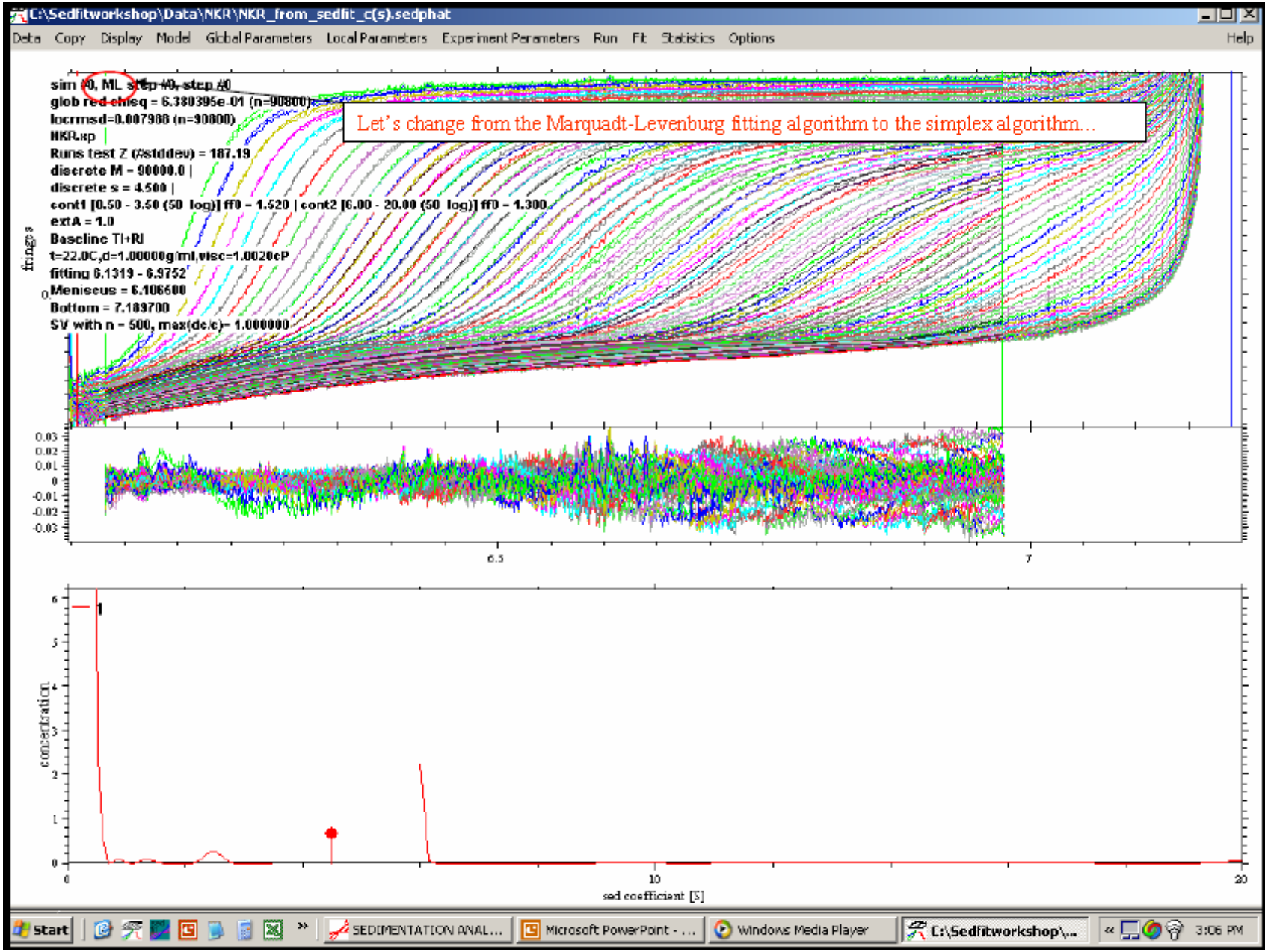

C:\Sedfitworkshop\Data\NKR\NKR_from_sedfit_c(s).sedphat
Data Copy Display Model Global Parameters Local Parameters Experiment Parameters Run Fit Statistics Options Help
With the Hybrid Local Continuous/Global Discrete Model selected:
Check the Species 1 box, input molar mass and sedimentation coefficient of species 1
We can model other parts of the boundary separately with continuous distributions
Segment one will be from 0.5 S to 3.5S (stopping just before the discrete species)
Segment two will begin after the discrete species at 6 S to 20 S
Parameters for Hybrid Local Continuous/Global Discrete Model
sp.1 neg. M1 90000.000 S1 4.500000
species 2 M2 0.0000 S2 0.000000
species 3 M3 0.0000 S3 0.000000
species 4 M4 0.0000 S4 0.000000
species 5 M5 0.0000 S5 0.000000
species 6 M6 0.0000 S6 0.000000
species 7 M7 0.0000 S7 0.000000
species 8 M8 0.0000 S8 0.000000
species 9 M9 0.0000 S9 0.000000
species 10 M10 0.0000 S10 0.000000
link molar mass in multiples of 1 for species from # 2 to # 4
link s values in multiples of # 1 for species from # 2 to # 4
continuous resolution s min s max frictional ratio
segment 1 50 linear log 0.500 3.500 1.520 fit ff0
segment 2 50 linear log 6.000 20.000 1.300 fit ff0
segment 3 0 linear log 0.000 0.000 1.300 fit ff0
Cancel
OK
with Tikhonov Regularization P 0.950
normalize distributions
fringes
concentration
sed. coefficient [S]
SEDIMENTATION ANAL...
Microsoft PowerPoint - ...
Windows Media Player
C:\Sedfitworkshop\...
3:05 PM

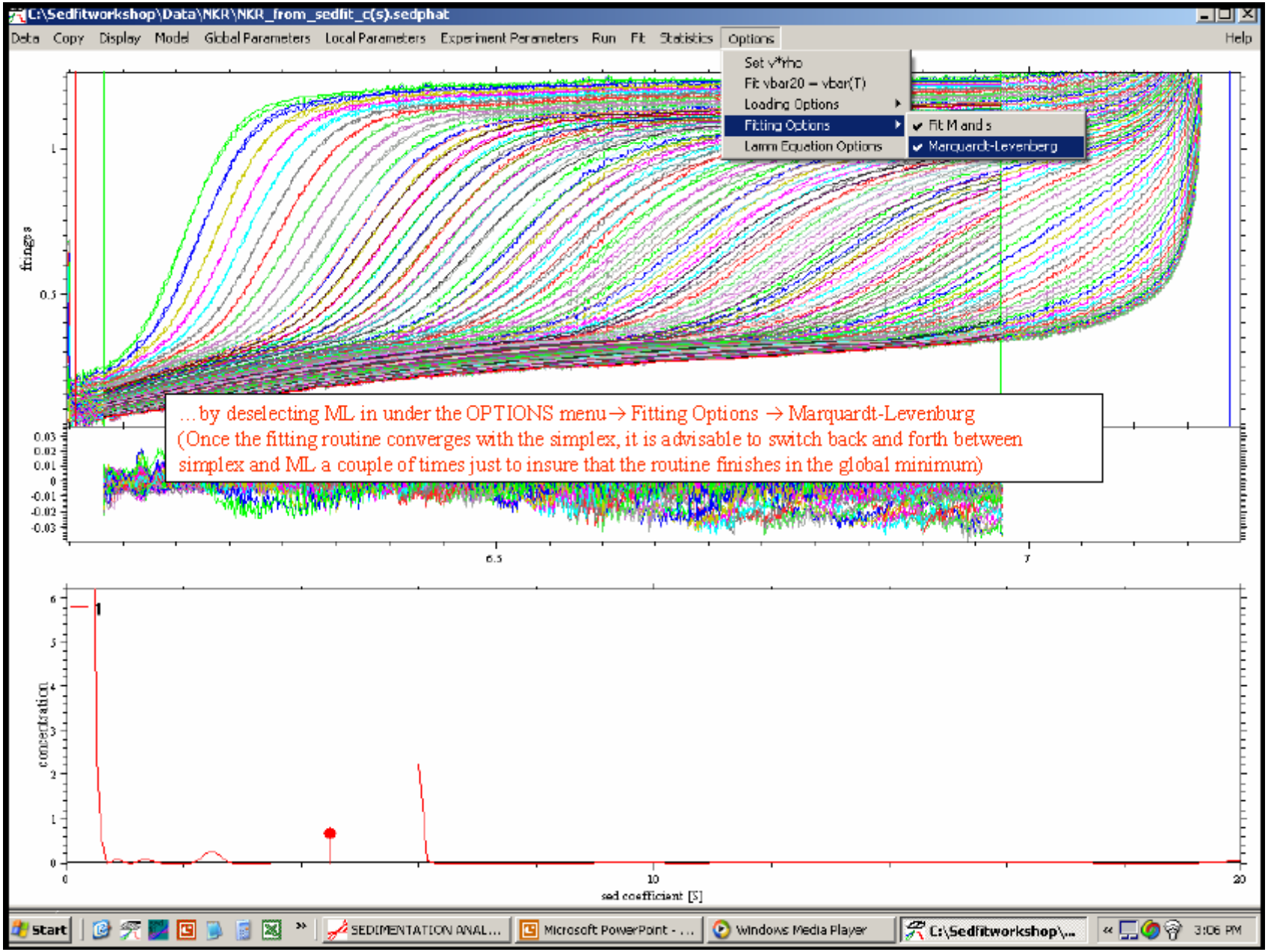

C:\Sedfitworkshop\Data\NKR\NKR_from_sedfit_c(s).sedphat
Data Copy Display Model Global Parameters Local Parameters Experiment Parameters Run Fit Statistics Options Help
Experimental Parameters
(1) INTERFERENCE data for SEDVELOCITY
C:\Sedfitworkshop\Data\NKR\NKR.xp (00005.IP2 ...)
Comment
cm/pixel: 0.0007437071 2, Sample
active
noise
0.0100
Centerpiece
2
Pathlength
1.200000
Rotor type
8
no backdiffusion neces
v-bar (ml/g)
0.7300
buffer density (g/ml)
1.000000
buffer viscosity (P)
0.010020
Temperature
22.0
Cancel
OK
fit baseline
fit RI Noise
fit TI Noise
Meniscus
6.1065
Bottom
7.1897
redirect men./bot.
1
For Associating Systems:
extinction coefficient A
1.0000
extinction coefficient B
0.0000
extinction coefficient C
0.0000
redirect xt A
redirect xt B
redirect xt C
partial boundary fitting
smin
0.00
smax
0.00
Input Experimental parameters
fringes
concentration
sed coefficient [S]

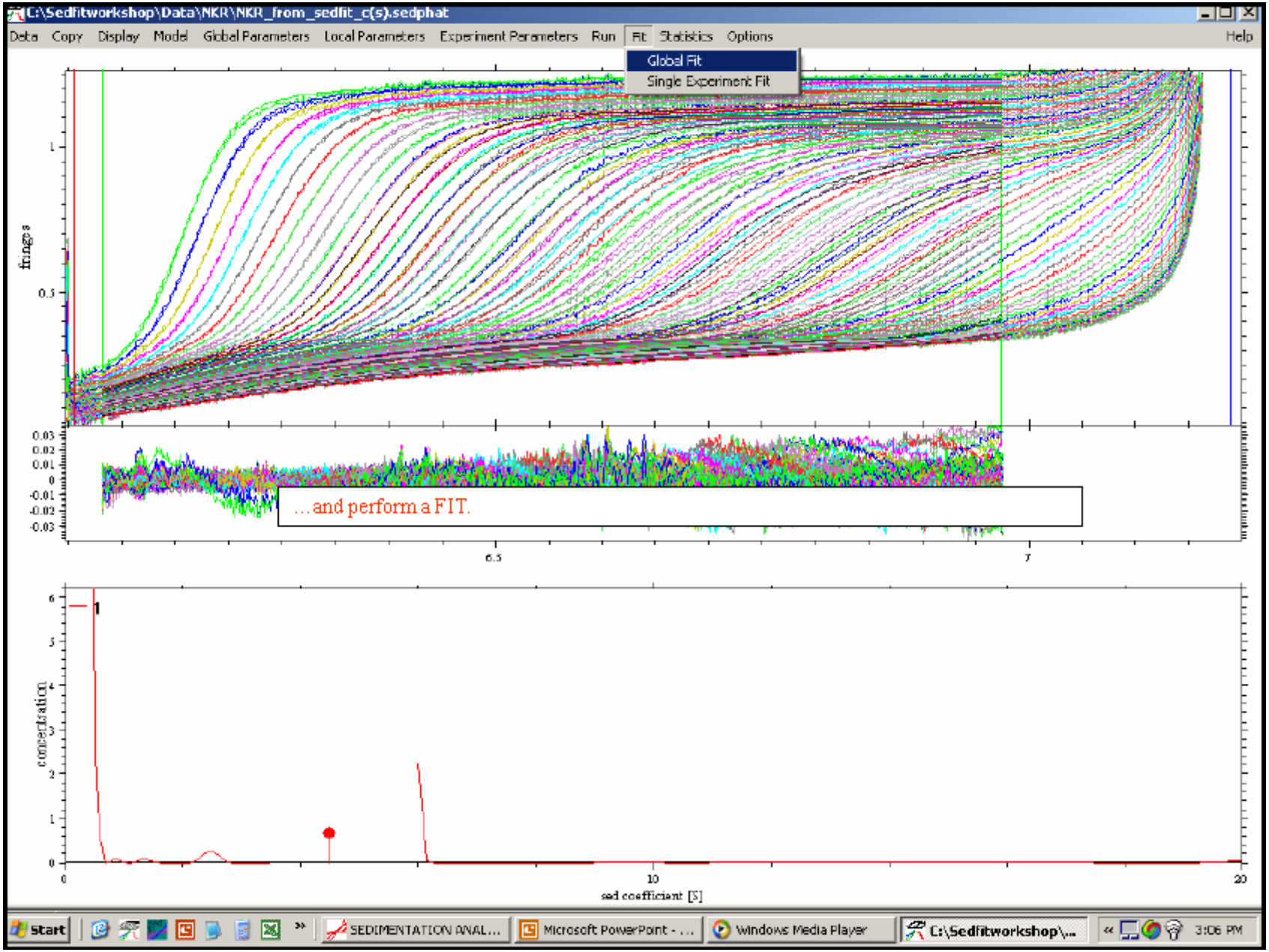

SEDIMENTATION ANALYSIS -> C:\Sedfitworkshop\Data\NKR\00400.IP2...
C:\Sedfitworkshop\Data\NKR\NKR_from_sedfit_c(s).sedphat
Data Copy Display Model Global Parameters Local Parameters Experiment Parameters Run Fit Statistics Options Help
Global Run
Single Experiment Run
fringes
0.5
0
-0.5
-1
6.5
1
Perform a RUN
concentration
0
1
sed coefficient [S]
start
SEDIMENTATION ANAL...
Microsoft PowerPoint - ...
Windows Media Player
C:\Sedfitworkshop\...
3:03 PM

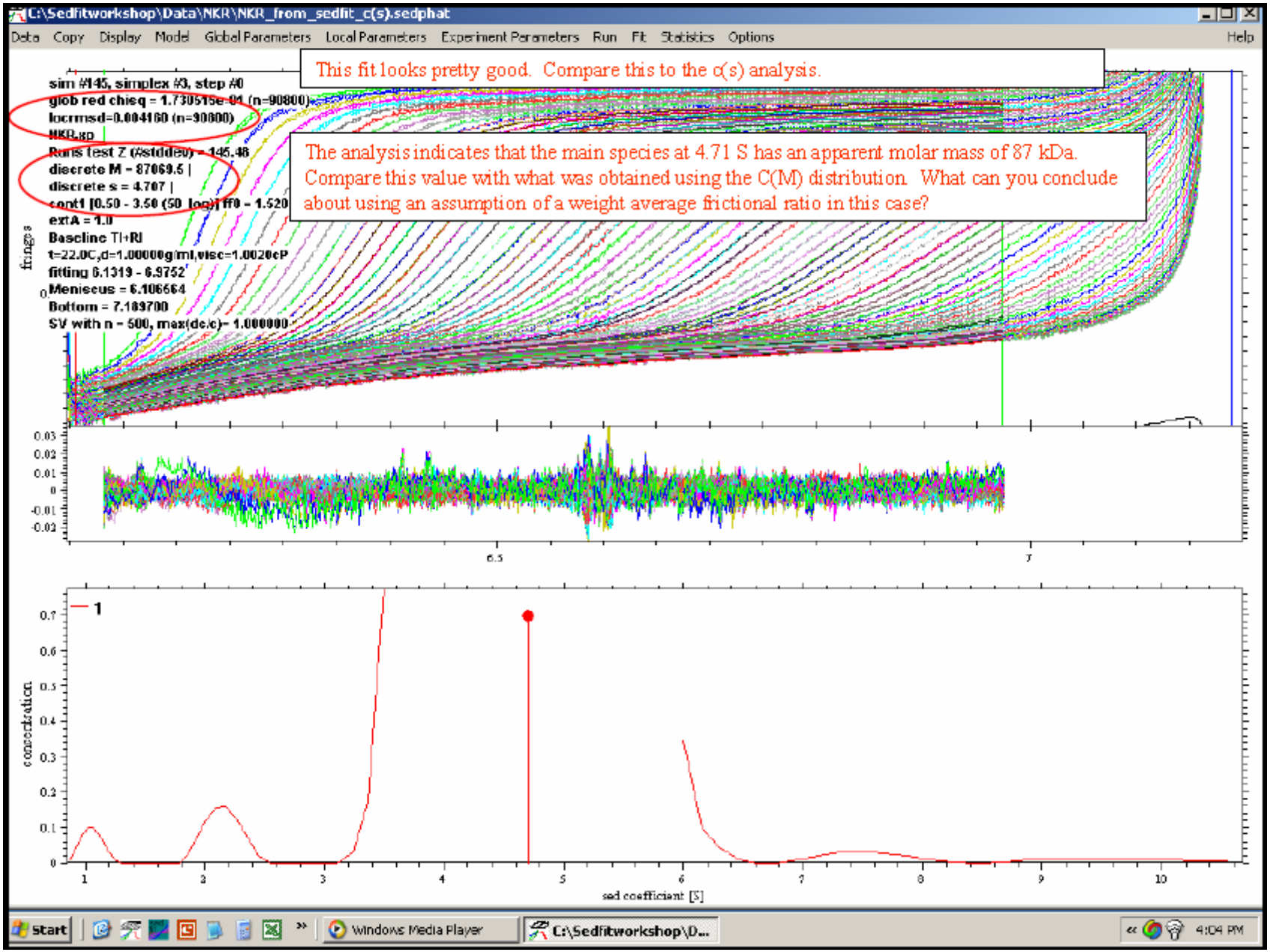

C:\Sedfitworkshop\Data\NKR\NKR_from_sedfit_c(s).sedphat
Data Copy Display Model Global Parameters Local Parameters Experiment Parameters Run Fit Statistics Options Help
sim #0, ML step #0, step #0
glob red chisq = 6.380395e-01 (n=90000)
locrmsd=0.007988 (n=90000)
NKR.sp
Runs test Z (#stddev) = 187.19
discrete M = 90000.0 |
discrete s = 4.500 |
cont1 [0.50 - 3.50 (50 log)] ff0 = 1.520 | cont2 [6.00 - 20.00 (50 log)] ff0 = 1.300
extA = 1.0
Baseline TI+RI
t=22.0C,d=1.00000g/ml,visc=1.0020cP
fitting 6.1319 - 6.9752
Meniscus = 6.106600
Bottom = 7.189700
SV with n = 500, max(dc/c)= 1.000000
A rmsd of 7 millifringes isn't bad for an initial run. Now, optimize by running a FIT
fringes
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
6.5
7
1
concentration
sed coefficient [S]
start
SEDIMENTATION ANAL...
Microsoft PowerPoint - ...
Windows Media Player
C:\Sedfitworkshop\...
3:05 PM

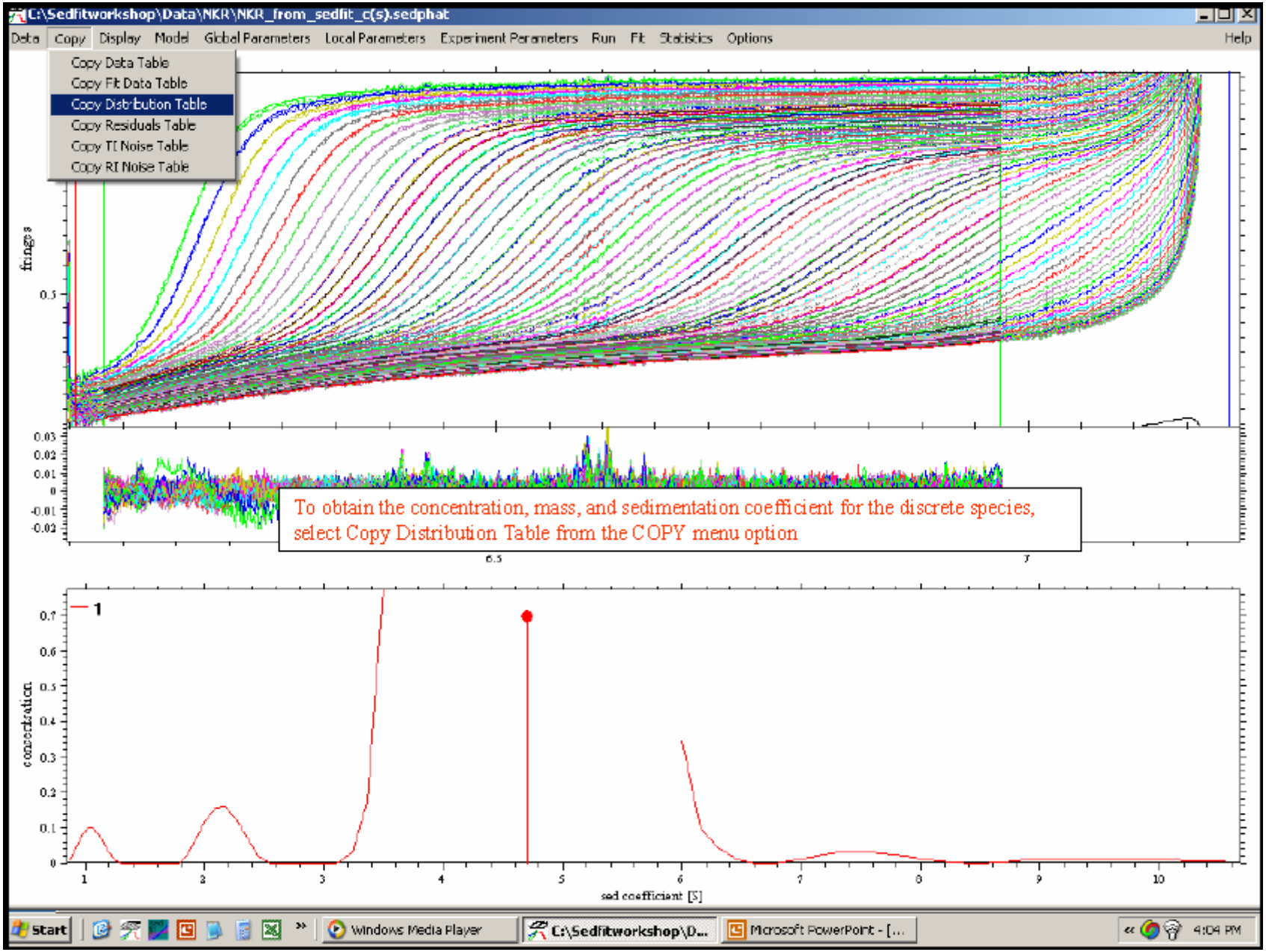

C:\Sedfitworkshop\Data\NKR\NKR_from_sedfit_c(s).sedphat
Data Copy Display Model Global Parameters Local Parameters Experiment Parameters Run Fit Statistics Options Help
Set Range
Full Range
Data Range
Update
Zoom In 1 Data Set
Data Mark Size
Show Data
Show Residuals
Show Experiments Stacked
Subtract All Systematic Noise From Raw Data ^N
Show Last Fit Info Again ^O
fringe
For visual inspection, it may be convenient to subtract out the RI and TI noise calculated from the run
concentration
sed coefficient [S]
start
SEDIMENTATION ANAL...
Microsoft PowerPoint - ...
Windows Media Player
C:\Sedfitworkshop\...
3:05 PM

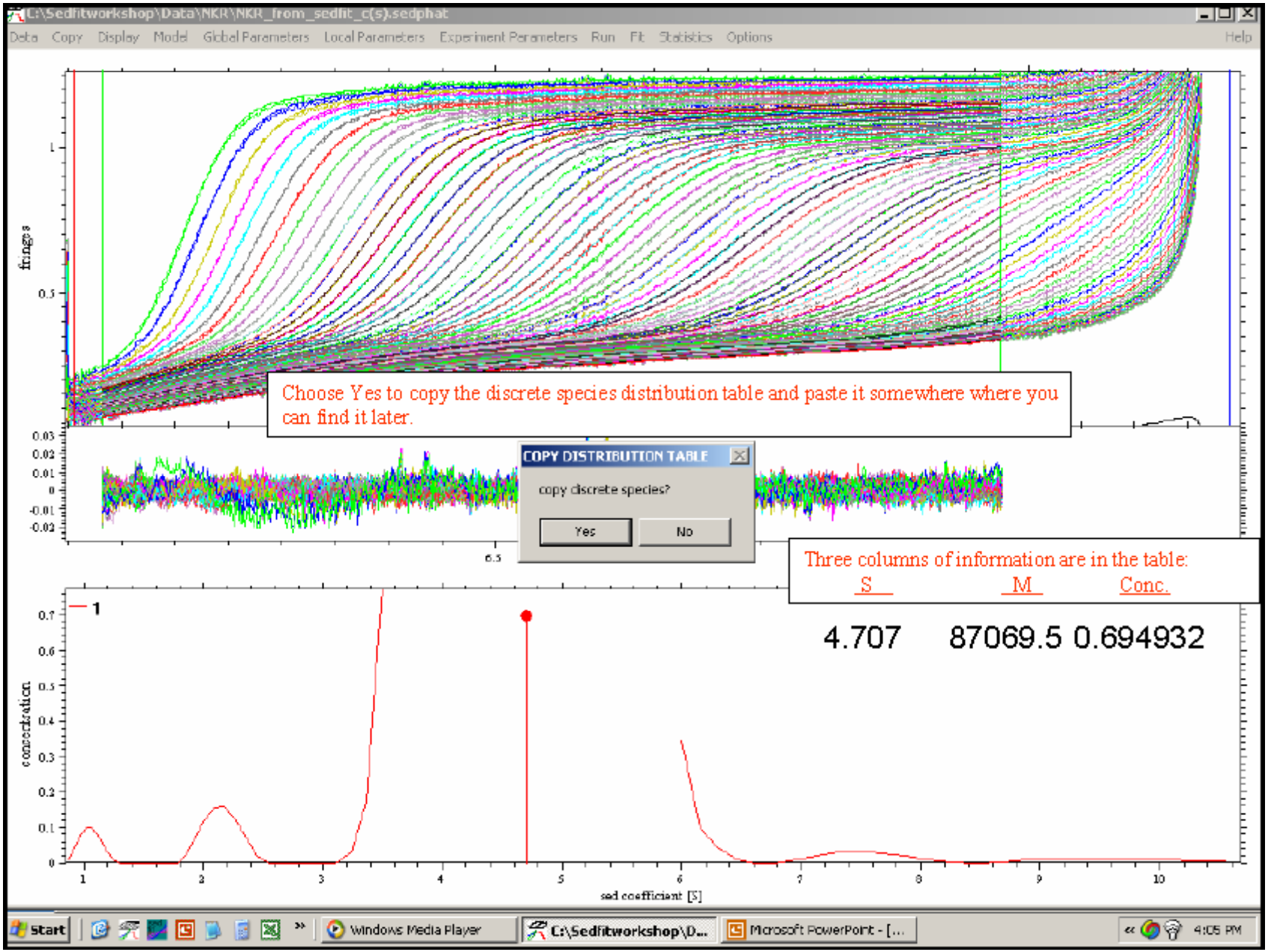

C:\Sedfitworkshop\Data\NKR\NKR_from_sedfit_c(s).sedphat
Data Copy Display Model Global Parameters Local Parameters Experiment Parameters Run Fit Statistics Options Help
sim #0, ML step #0, step #0
glob red-chisq = 6.380395e-01 (n=90800)
locrmsd=0.007988 (n=90800)
NKR.sp
Runs test Z (#stddev) = 187.19
discrete M = 90000.0
discrete s = 4.500
cont1 [0.50 - 3.50 (50 log)] ff0 = 1.520 | cont2 [6.00 - 20.00 (50 log)] ff0 = 1.300
extA = 1.0
Baseline TI+RI
t=22.0C,d=1.00000g/ml,visc=1.0020cP
fitting 6.1319 - 6.9752
Meniscus = 6.106600
Bottom = 7.189700
SV with n = 500, max(dc/c)= 1.000000
Let's change from the Marquadt-Levenburg fitting algorithm to the simplex algorithm…
fringes
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
6.5
7
1
concentration
sed coefficient [S]
0
10
20
start
SEDIMENTATION ANAL…
Microsoft PowerPoint - …
Windows Media Player
C:\Sedfitworkshop\…
3:06 PM


C:\Sedfitworkshop\Data\NKR\NKR_from_sedfit_c(s).sedphat
Data Copy Display Model Global Parameters Local Parameters Experiment Parameters Run Fit Statistics Options Help
Set v*rho
Fit vbar20 = vbar(T)
Loading Options
Fitting Options
Lamm Equation Options
Fit M and s
Marquardt-Levenberg
fringes
…by deselecting ML in under the OPTIONS menu→ Fitting Options → Marquardt-Levenburg
(Once the fitting routine converges with the simplex, it is advisable to switch back and forth between simplex and ML a couple of times just to insure that the routine finishes in the global minimum)
concentration
sed coefficient [S]
start
SEDIMENTATION ANAL...
Microsoft PowerPoint - ...
Windows Media Player
C:\Sedfitworkshop\...
3:06 PM


C:\Sedfitworkshop\Data\NKR\NKR_from_sedfit_c(s).sedphat
Data Copy Display Model Global Parameters Local Parameters Experiment Parameters Run Fit Statistics Options Help
Global Fit
Single Experiment Fit
fringes
...and perform a FIT.
concentration
sed coefficient [S]
start
SEDIMENTATION ANAL...
Microsoft PowerPoint - ...
Windows Media Player
C:\Sedfitworkshop\...
3:06 PM


C:\Sedfitworkshop\Data\NKR\NKR_from_sedfit_c(s).sedphat
Data Copy Display Model Global Parameters Local Parameters Experiment Parameters Run Fit Statistics Options Help
This fit looks pretty good. Compare this to the c(s) analysis.
The analysis indicates that the main species at 4.71 S has an apparent molar mass of 87 kDa. Compare this value with what was obtained using the C(M) distribution. What can you conclude about using an assumption of a weight average frictional ratio in this case?
sim #145, simplex #3, step #0
glob red chisq = 1.730515e-04 (n=90800)
locrmsd=0.004160 (n=90000)
NKR.sp
Runs test Z (#stddev) = -145.48
discrete M = 87069.5 |
discrete s = 4.707 |
cont1 [0.50 - 3.50 (50 log)] ff0 = 1.520
extA = 1.0
Baseline TI+RI
t=22.0C,d=1.00000g/ml,visc=1.0020cP
fitting 6.1319 - 6.9752
Meniscus = 6.106664
Bottom = 7.189700
SV with n = 500, max(dc/c)= 1.000000
fringes
0.03
0.02
0.01
0
-0.01
-0.02
6.5
7
1
concentration
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
6
7
8
9
10
sed coefficient [S]
start
Windows Media Player
C:\Sedfitworkshop\D...
4:04 PM

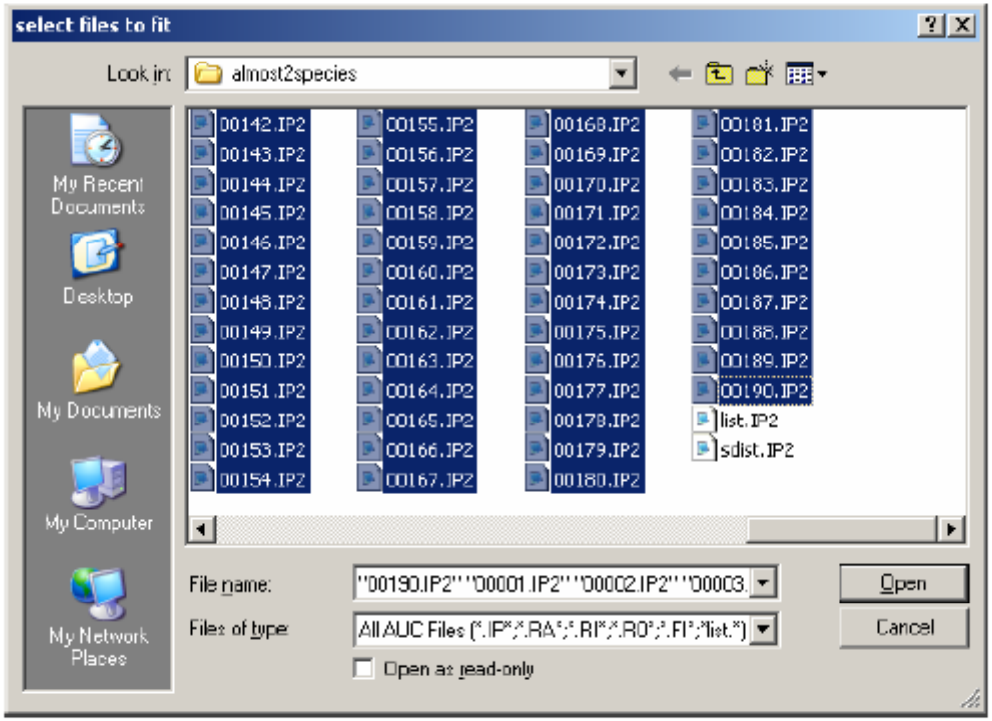

C:\Sedfitworkshop\Data\NKR\NKR_from_sedfit_c(s).sedphat
Data
Copy
Display
Model
Global Parameters
Local Parameters
Experiment Parameters
Run
Fit
Statistics
Options
Help
Copy Data Table
Copy Fit Data Table
Copy Distribution Table
Copy Residuals Table
Copy TI Noise Table
Copy RI Noise Table
fringes
To obtain the concentration, mass, and sedimentation coefficient for the discrete species, select Copy Distribution Table from the COPY menu option
1
concentration
sed coefficient [S]
start
Windows Media Player
C:\Sedfitworkshop\D...
Microsoft PowerPoint - [...
4:04 PM

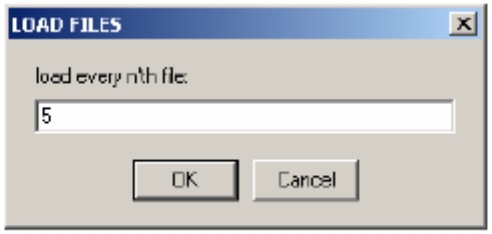

C:\Sedfitworkshop\Data\NKR\NKR_from_sedfit_c(s).sedphat
Data Copy Display Model Global Parameters Local Parameters Experiment Parameters Run Fit Statistics Options Help
fringes
Choose Yes to copy the discrete species distribution table and paste it somewhere where you can find it later.
COPY DISTRIBUTION TABLE
copy discrete species?
Yes
No
Three columns of information are in the table:
S M Conc.
4.707 87069.5 0.694932
concentration
sed coefficient [S]
start
Windows Media Player
C:\Sedfitworkshop\D...
Microsoft PowerPoint - [...
4:05 PM

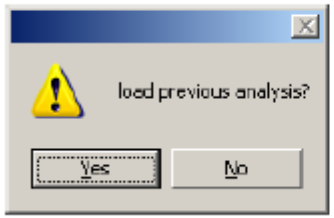
IV-1: Lamm equation modeling with a monomer-dimer system

Let's suppose we have a protein ~ 100 kDa that, based on c(s) seems to form dimers. There's also a smaller Mw impurity. The goal of this exercise is to make the transition from c(s) in SEDFIT to set up an analysis of the same data as a interacting monomer-dimer system in SEDPHAT.

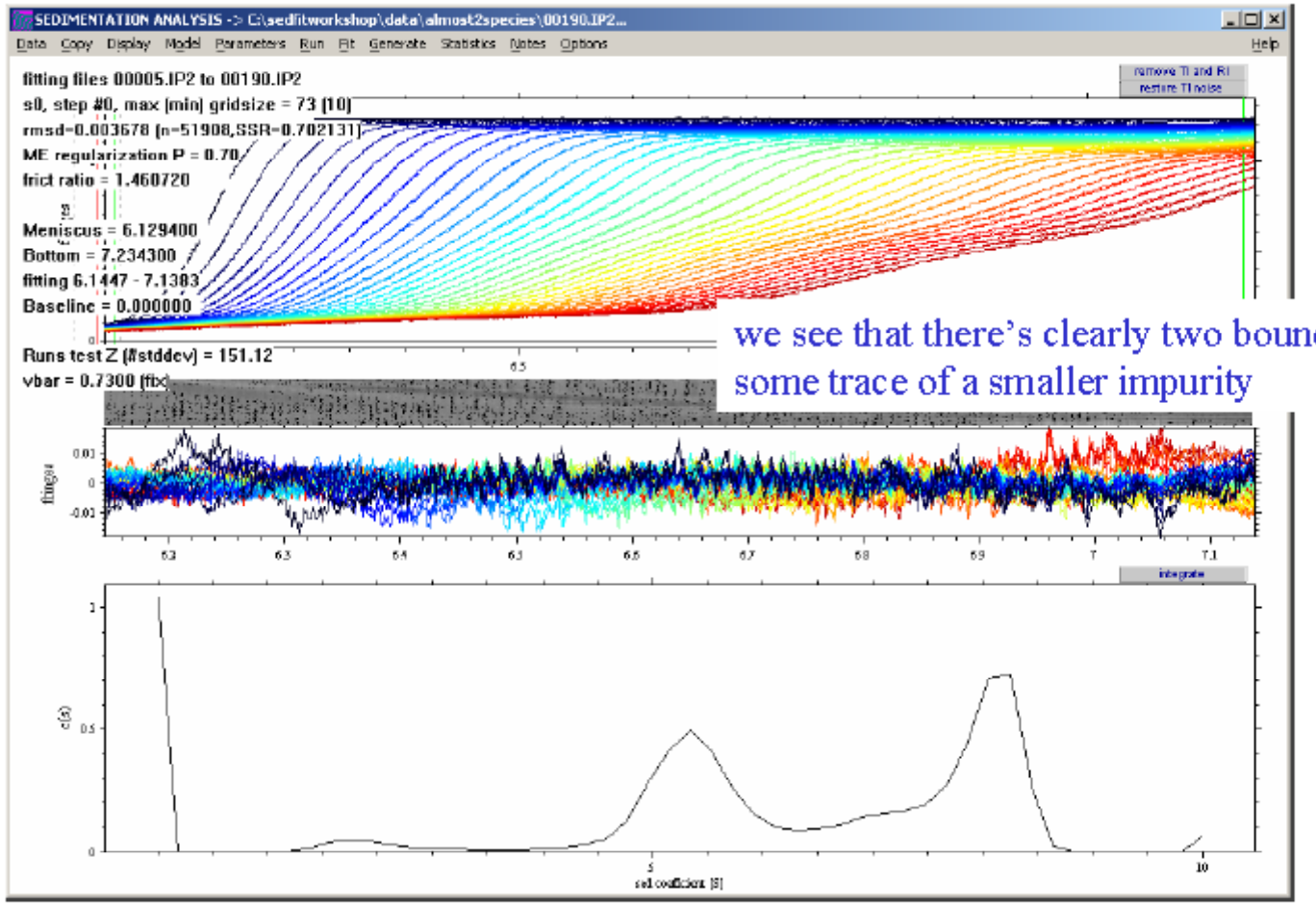
open SEDFIT, and load all files in the example “almost2species”





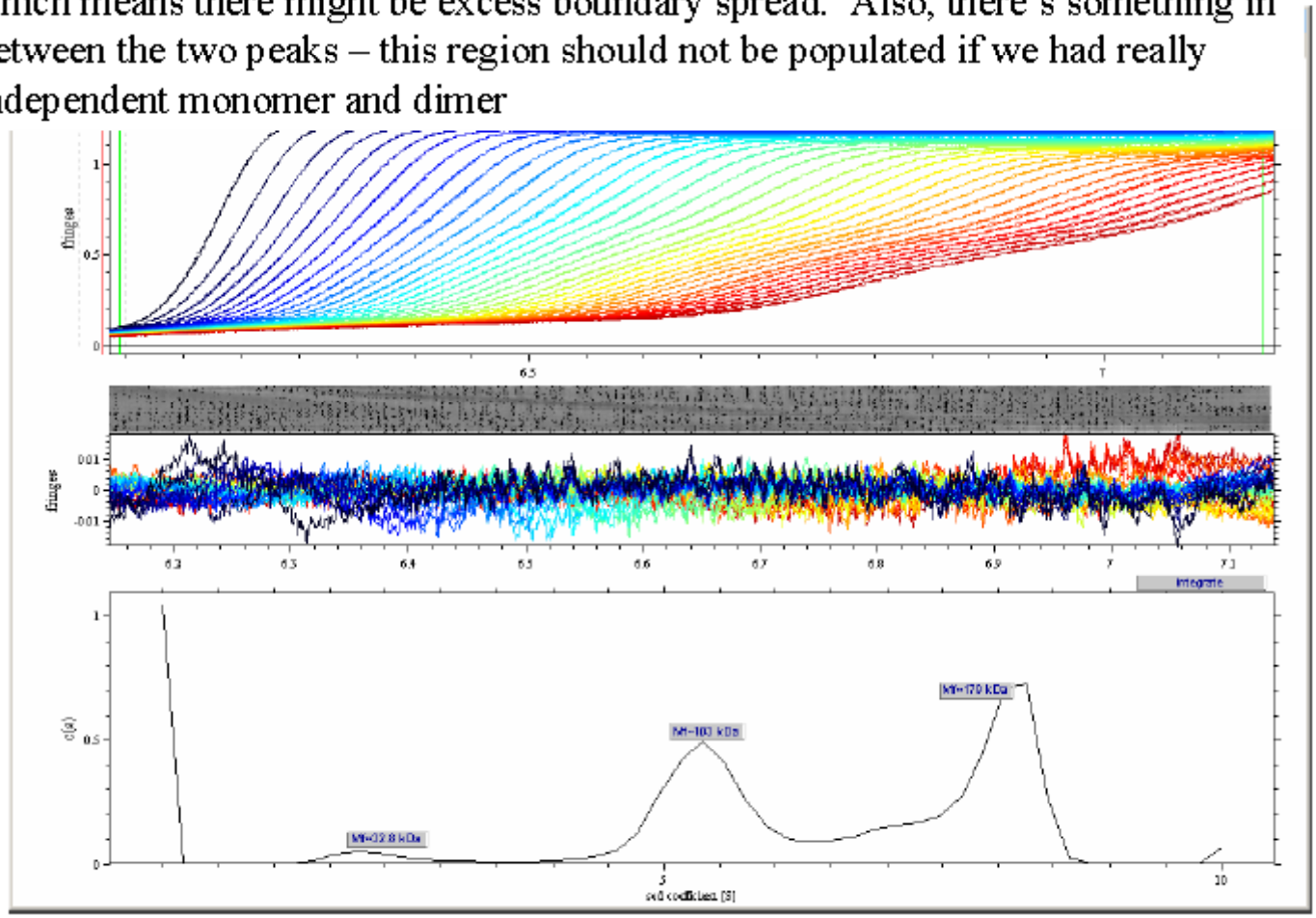


YES, then to a RUN command
then do control-D [data range] and control-N [subtract TI+RI noise]
This will reproduce a previously made standard c(s)
[if you don't get to restore a previously saved fit, see tutorials on c(s)]

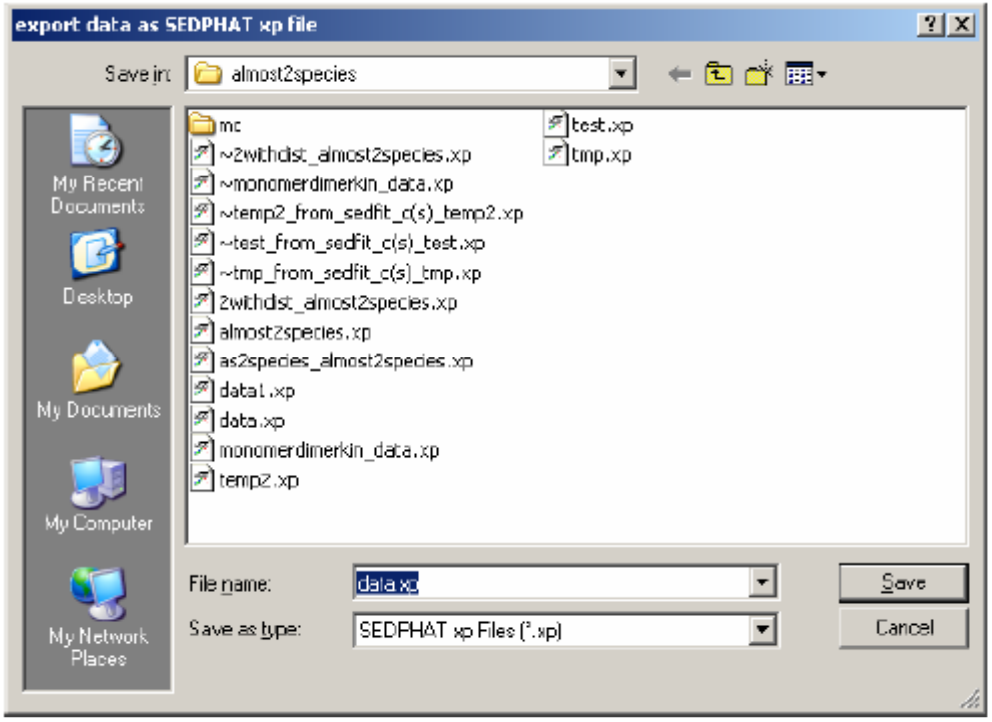


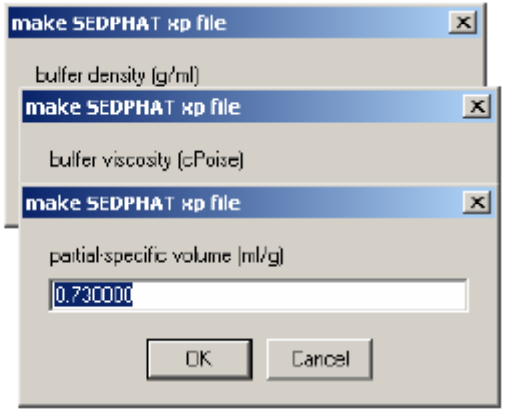
we see that there's clearly two boundaries, and some trace of a smaller impurity

control-M shows the approximate Mw estimates: The 103 kDa should be the monomer, but the 179 kDa for what should be the dimer is too low a number, which means there might be excess boundary spread. Also, there's something in between the two peaks – this region should not be populated if we had really independent monomer and dimer

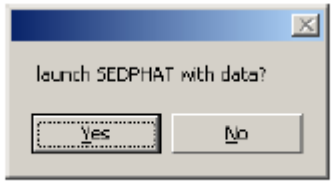


let's export this. Use the function Data → Export Data to SEDPHAT
Give the xp-file to be generated some filename, and default through all the parameter boxes.

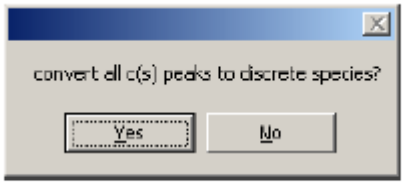




...

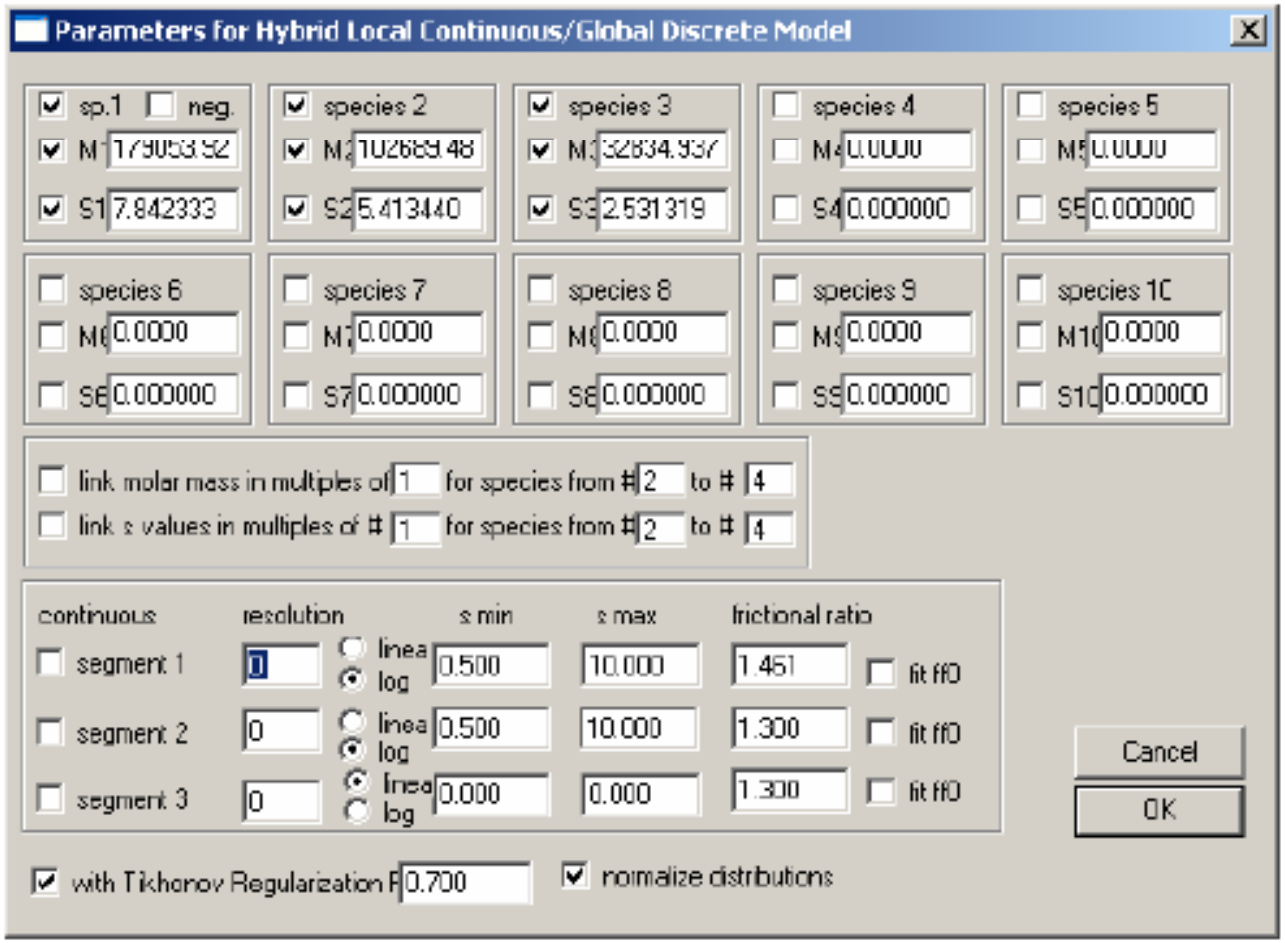


YES



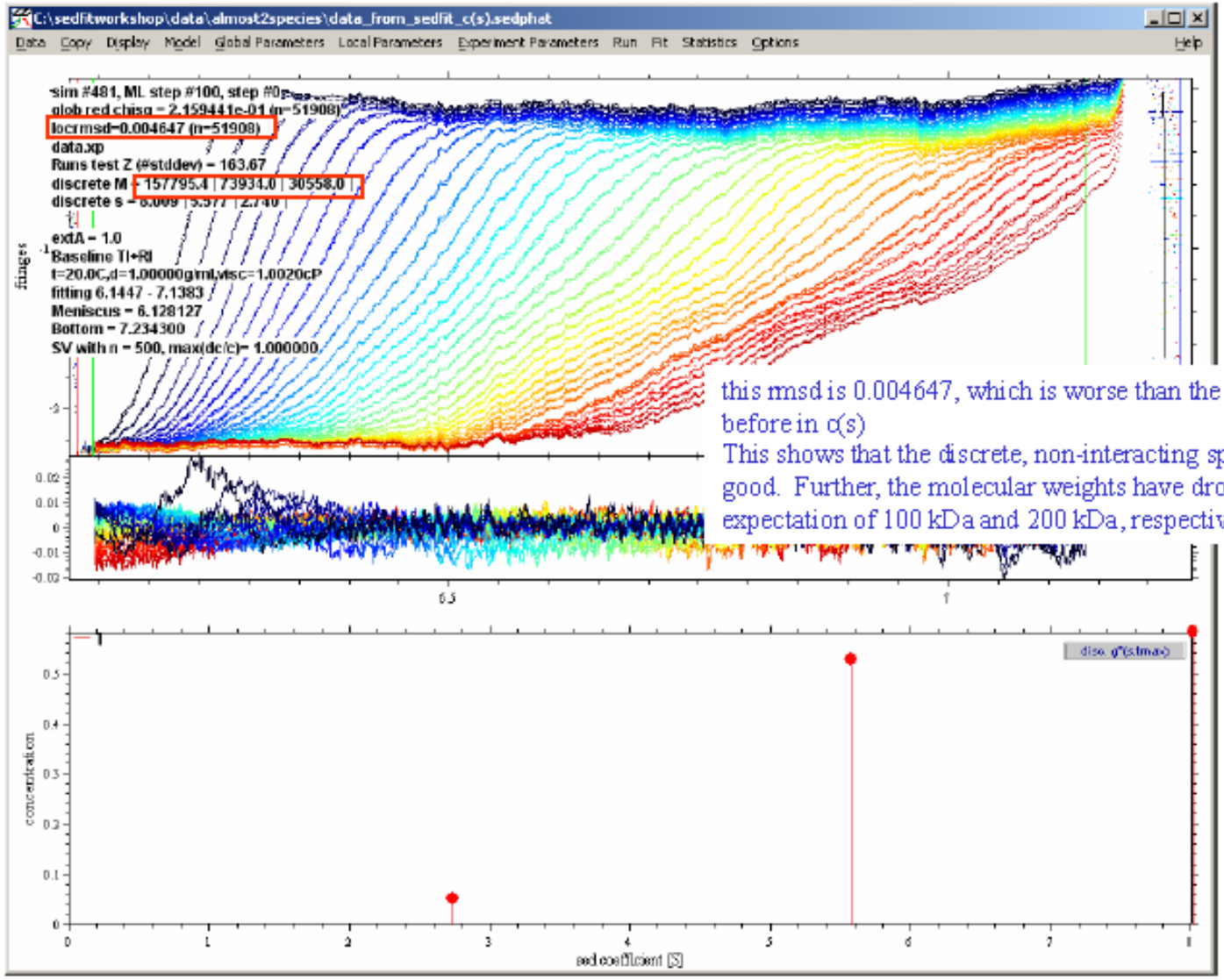
YES

this will use the automatic integration in SEDFIT to generate starting guesses for SEDPHAT model treating all species as discrete species.



accept, do RUN and then FIT

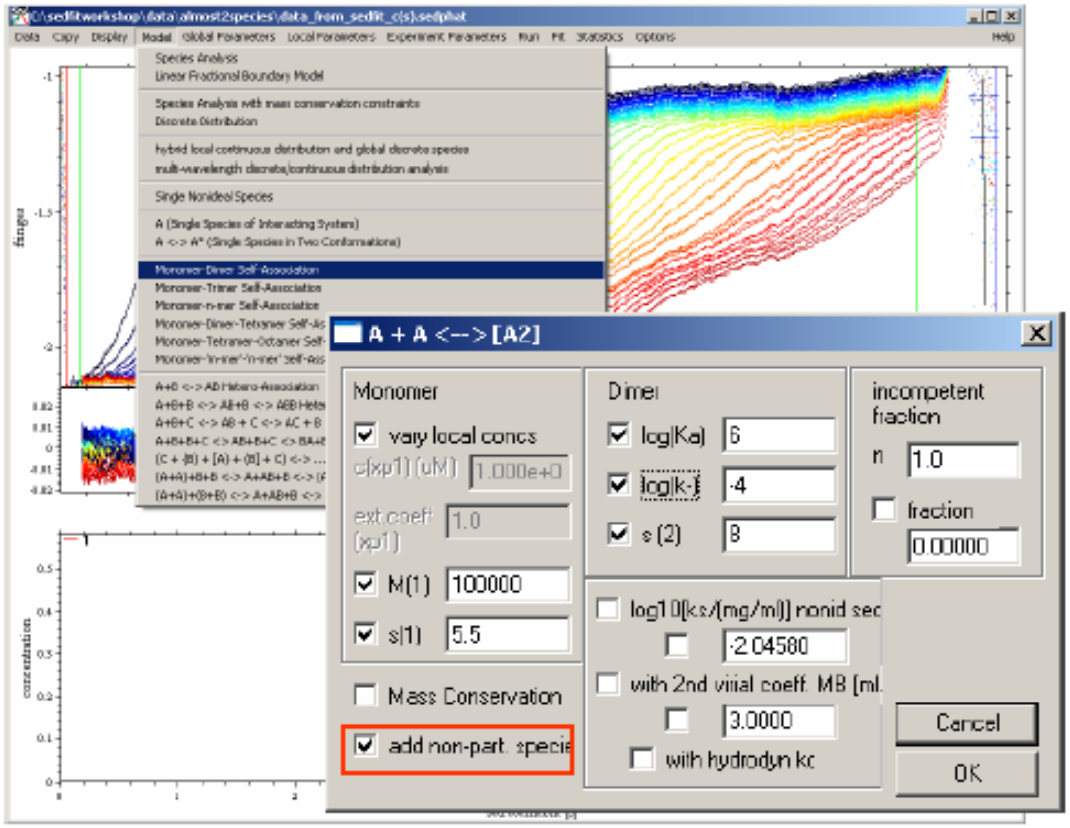
then fine-adjust the display by using control-D [set data range] and right-double click in the distribution window [full range distribution plot] and control-O [show last fit info]



this rmsd is 0.004647, which is worse than the 0.003678 found before in c(s)

This shows that the discrete, non-interacting species fit is not good. Further, the molecular weights have dropped below the expectation of 100 kDa and 200 kDa, respectively.

Let's try the monomer-dimer model in order to account for finite reaction kinetics.



also check 'add non-part. species' box, since we will need to account for the slow sedimenting degradation product when fitting this data.

Then hit OK

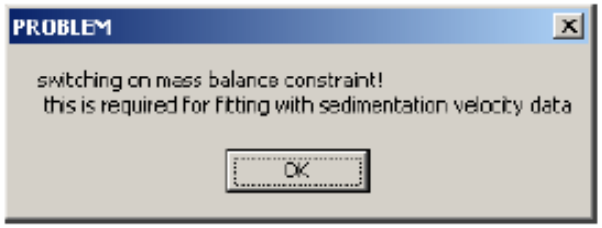
we use the following initial estimates:

M1 = 100kDa [that should be the Mw from sequence, but we don't know the exact vbar, therefore we'll float it here]

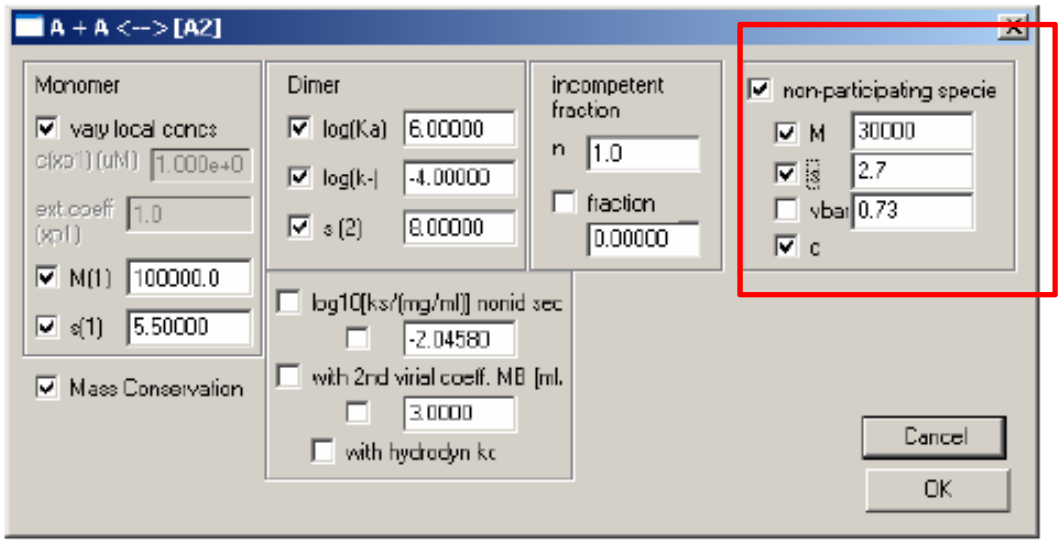
s1 = 5.5, s2 = 8 [these are the s-values from the discrete model, which should be excellent starting guesses and we'll float]

log(Ka) = 6 [since we loaded ~ uM concentration, and see roughly equal height of monomer and dimer peaks in c(s), we should be close to Kd at these uM concs]

log(k-) = -4 [the reaction must be fairly slow (i.e. –4 or slower) for us to resolve peaks in c(s), but we want to put the starting guess at a value close to the range of –3 to – 4 where the SV experiment is sensitive to the value of log(koff). Otherwise the optimization of this parameter would become difficult.]



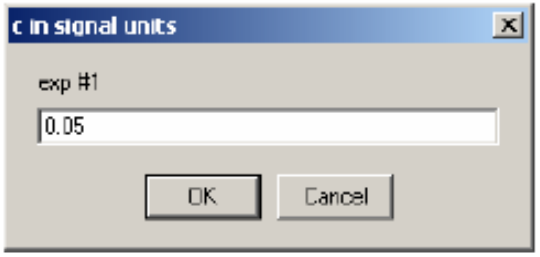
this is no problem, all SV models obey mass balance.



because we had checked the 'add non-part. species' box, now an extended box shows up with a field where we enter the parameters of the non-participating species.

This will be 30kDa and 2.7 S, which are the estimates taken from the discrete species fit

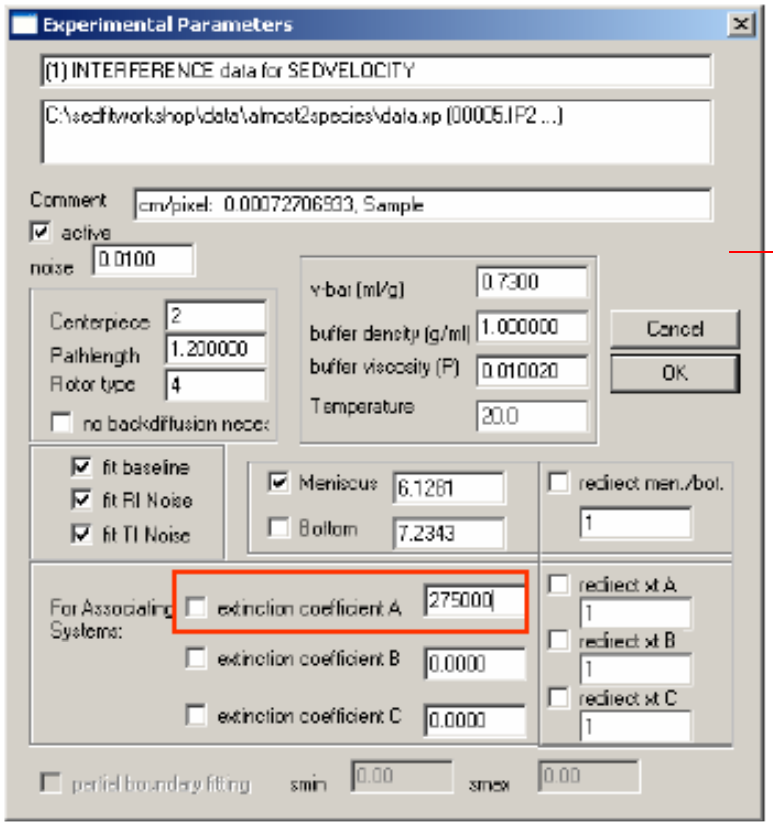
Then hit OK



Now a input box pops up asking us how large the concentration of the non-participating species is. This is in signal units, and from the discrete species fit we can estimate a value of 0.05. (Or this can be obtained also from integrating the previous c(s) around this contaminant peak).

One more set of parameters is required, due to the fact that in the interacting species models we want to deal with molar binding constants. Therefore, we need starting concentrations in molar units, as well as extinction coefficients to relate the molar concentration to signal.

Open the Experimental Parameters:



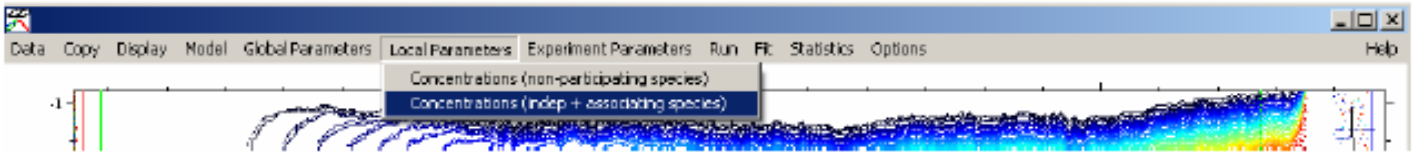
For the 'extinction coefficient A', enter a value of 275000. This is interference data, and based on the sensitivity for proteins of 3.3 fringes/mg/ml, we have the 'interference extinction coefficient' 2.75 x Mw.

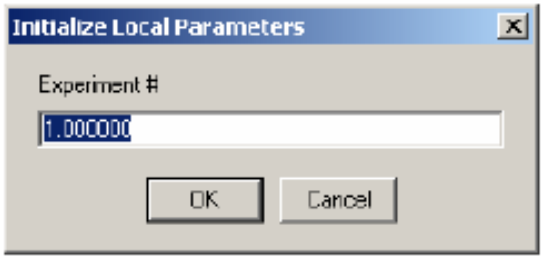
Note we also have a pathlength of 1.2 cm, which will allow SEDPHAT to correct for the 12 mm thickness of the centerpiece used.

If this was absorbance data, we would enter here the conventional molar extinction coefficient (OD/M/cm), just as taken, for example, from SEDNTERP.

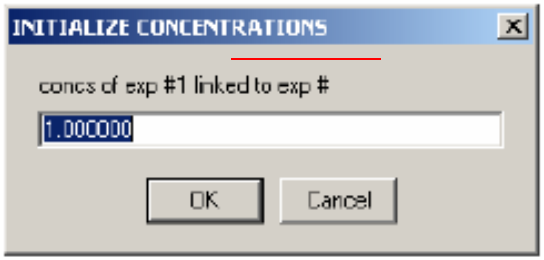
Click on OK

Next, choose the "concentrations (indep + associating species)" from the menu.

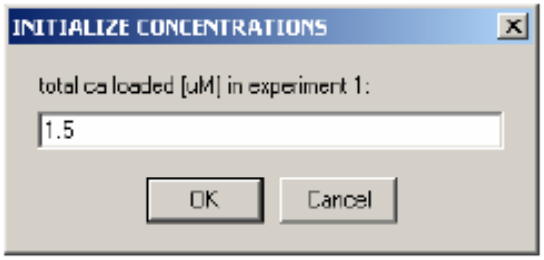




we're referring to xp #1 (have only 1 loaded anyway…)

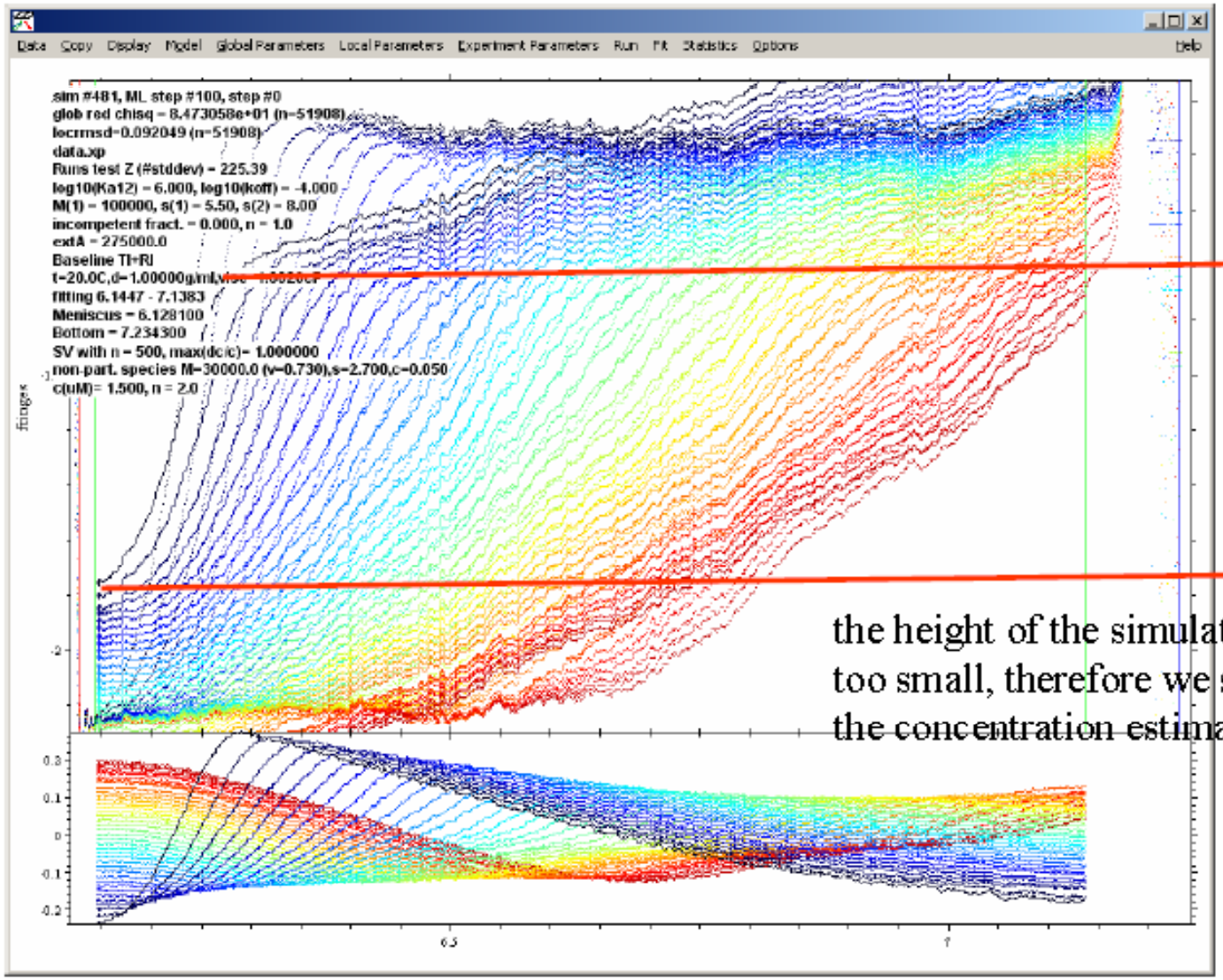


we'll not redirect this xp's concentration to another xp
(again, since we only have one loaded – this function would make sense, e.g., if we had absorbance data from the same cell loaded as xp2 such that we should constrain the concentrations to be the same – 'redirected'. See the sedimentation equilibrium tutorials, where heavy use of this option is made.)



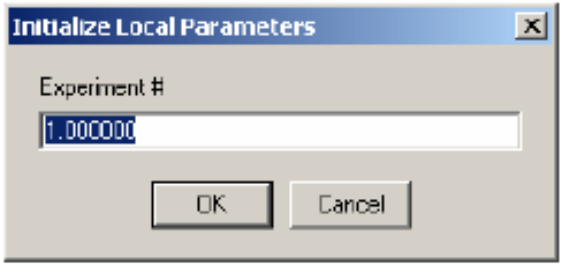
I believe we loaded a concentration of ~1.5 uM in this cell, therefore we enter the value 1.5.

Do a RUN to test the starting guesses.

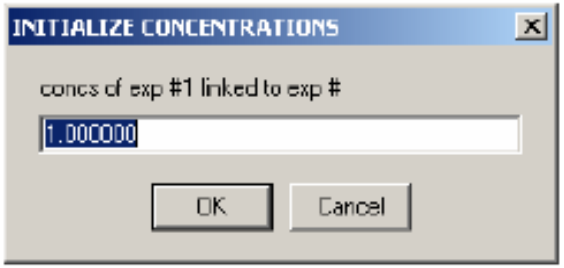


the height of the simulated boundary is too small, therefore we should increase the concentration estimate.

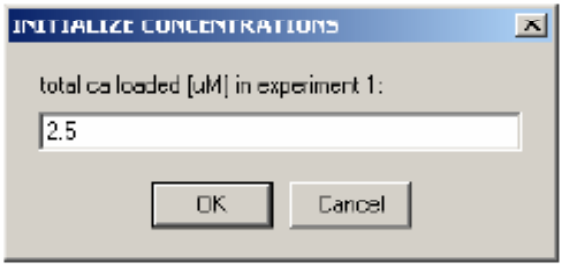
Again, choose the "concentrations (indep + associating species)" from the menu.



we're referring to xp #1 (have only 1 loaded anyway…)

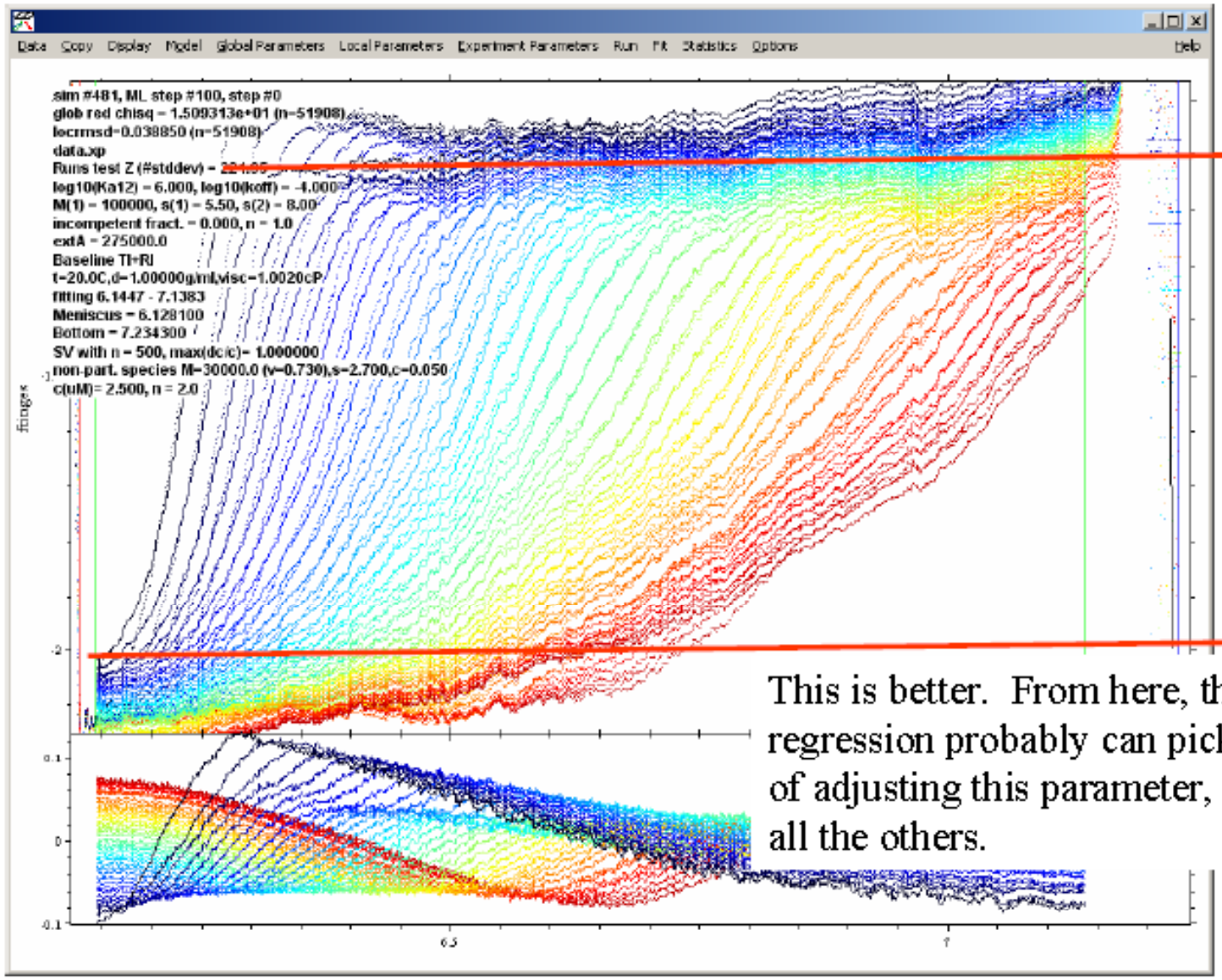


we'll not redirect this xp's concentration to another xp



Let's try a value of 2.5 uM.

Do a RUN to test the starting guesses.



This is better. From here, the non-linear regression probably can pick up the job of adjusting this parameter, along with all the others.

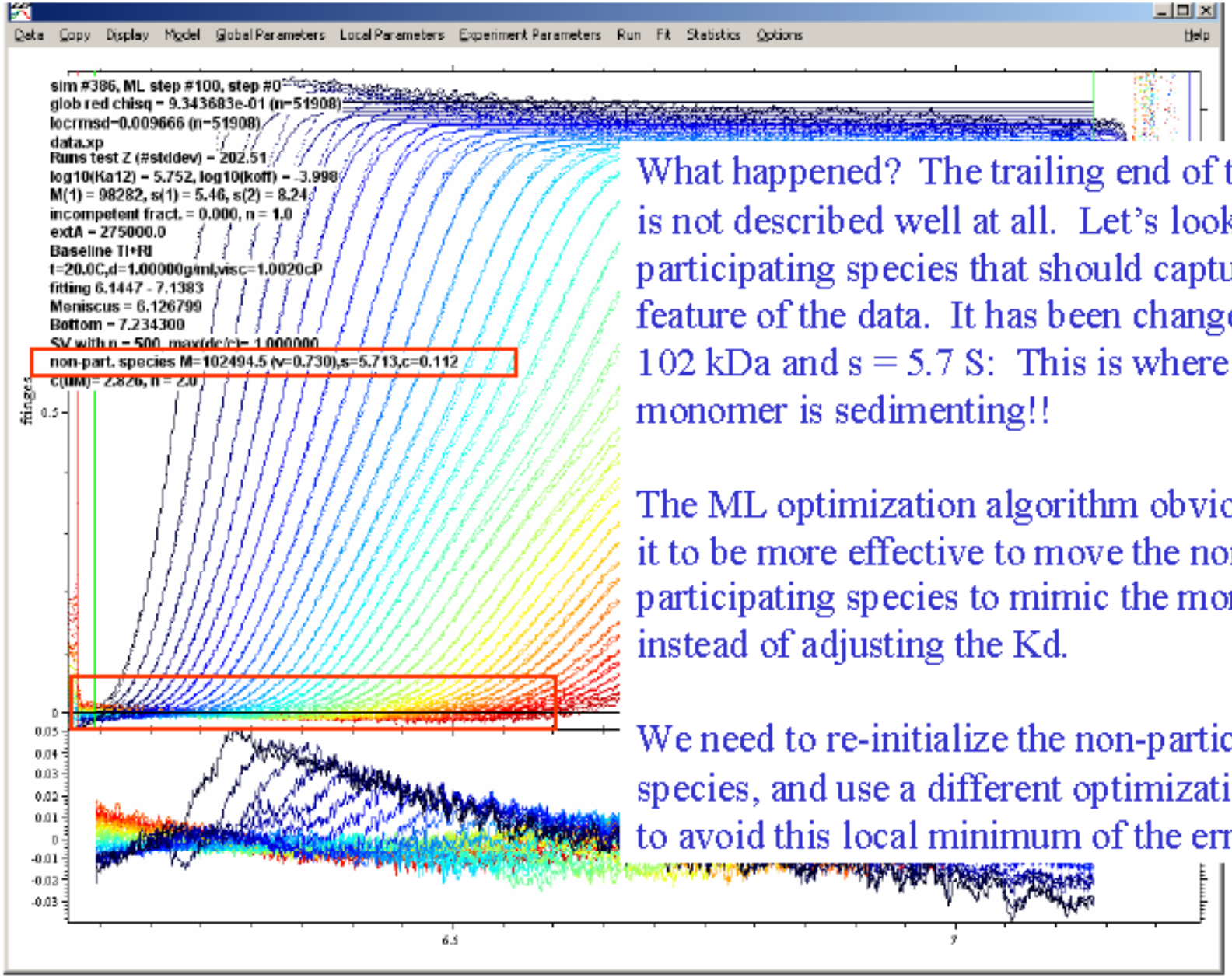
Now click on FIT. Note, the 'ML' in the first line indicates we are currently using Marquardt-Levenberg optimization, which will need to do a large number of simulations to initialize the gradient information and seems not to do much before updating the screen… nevertheless, this is a very efficient algorithm overall.

If your copy of SEDPHAT is configured to use the Simplex or the Simulated Annealing algorithm for minimization, the following sequence of events will be different. There are many different equally valid options how to do the fitting.

As it turns out, the following documentation will not be of a straightforward fit, but one where we have to avoid getting trapped into local minima. We use a combination of fitting algorithms to get to the best fit in a shallow minimum. (Remember the error surfaces of the banana function and the mountain range.) The details of the following are probably not reproducible, because the Simplex routine involves seed points generated randomly, and therefore it will slightly differ each time. However, it should be straightforward to apply the same strategy to your particular situation, as well as to other fits.

From here on, this is more an example on switching between fitting algorithms.

After the fit has stopped, hit control-N [subtract noise] and control-O [show best-fit info again]. This is where I ended up with after the fit. This does not look very good, yet.



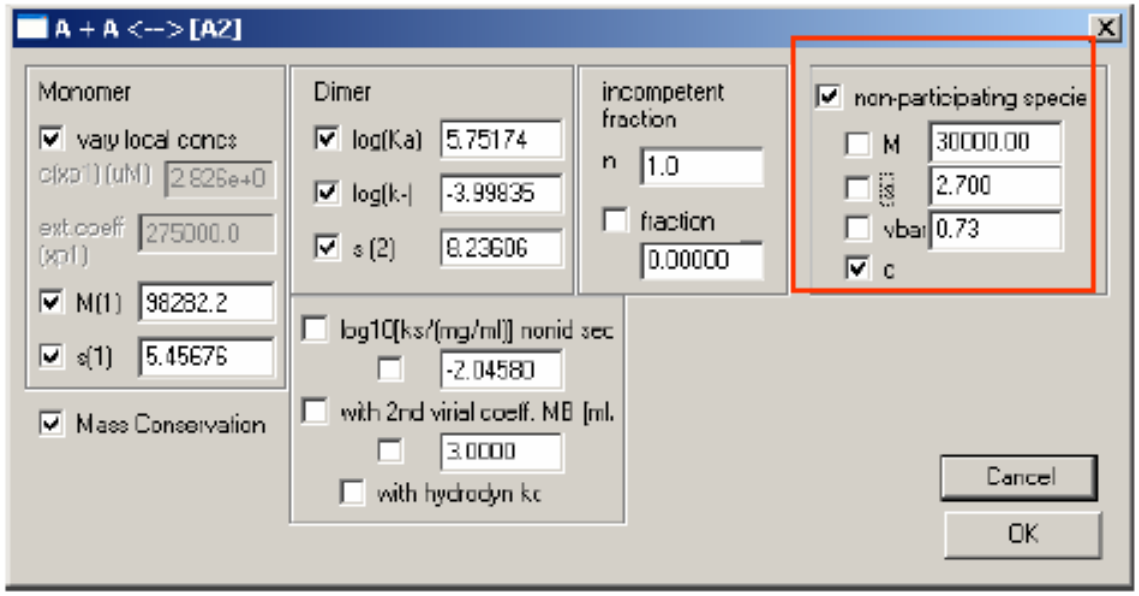
What happened? The trailing end of the boundary is not described well at all. Let's look at the non-participating species that should capture this feature of the data. It has been changed to M = 102 kDa and s = 5.7 S: This is where the monomer is sedimenting!!

The ML optimization algorithm obviously found it to be more effective to move the non-participating species to mimic the monomer, instead of adjusting the Kd.

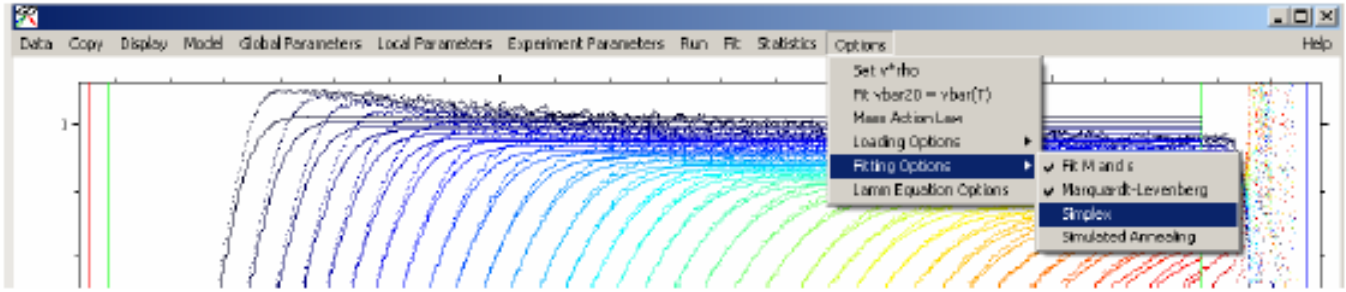
We need to re-initialize the non-participating species, and use a different optimization strategy to avoid this local minimum of the error surface.

in the Global parameters box, enter M = 30000 and s = 2.7

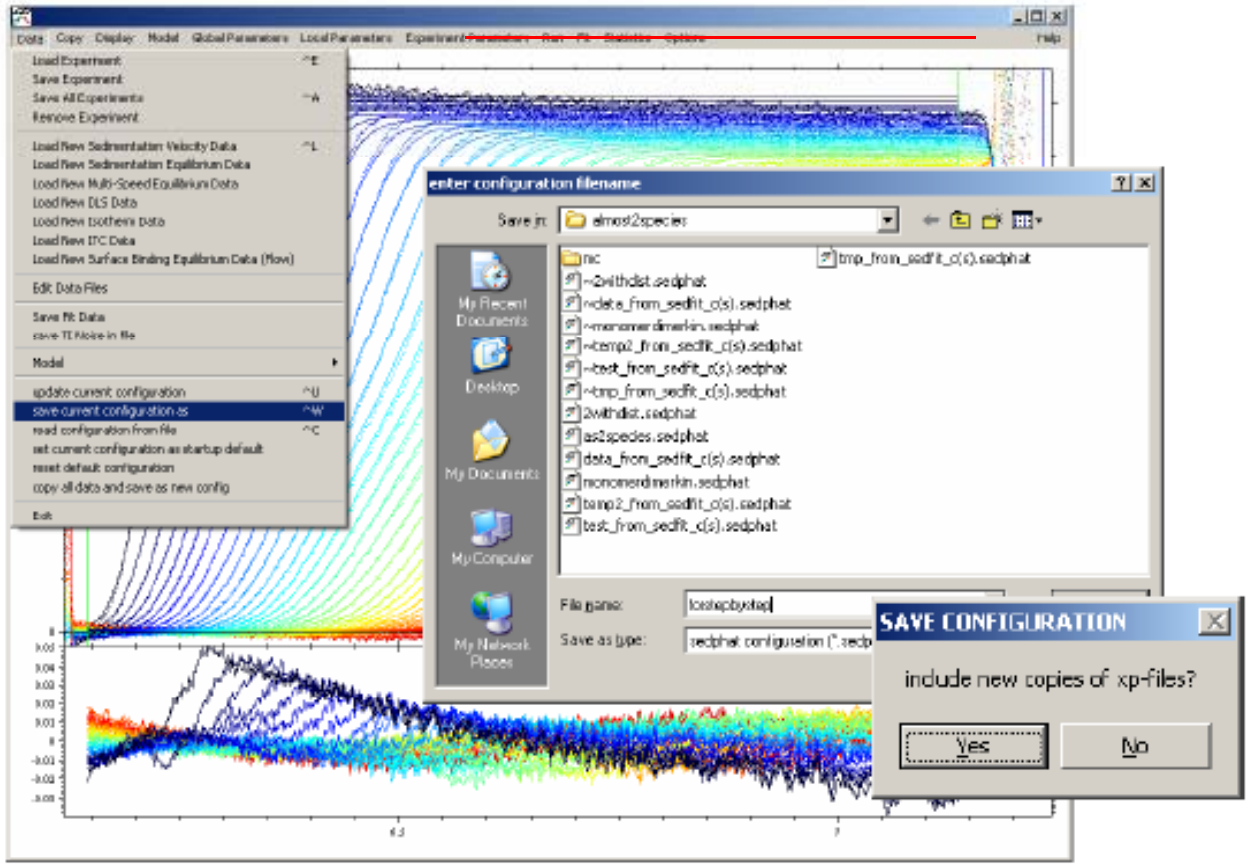
Also, we want to keep these parameters fixed now (we'll set them free later).



click on the "Simplex" menu function. This will toggle off Marquardt-Levenberg and switch on the Simplex algorithm



Also, just in case we need to go back, let us save this state of the analysis.
In the Data menu, click on "save current configuration as", and enter a filename.

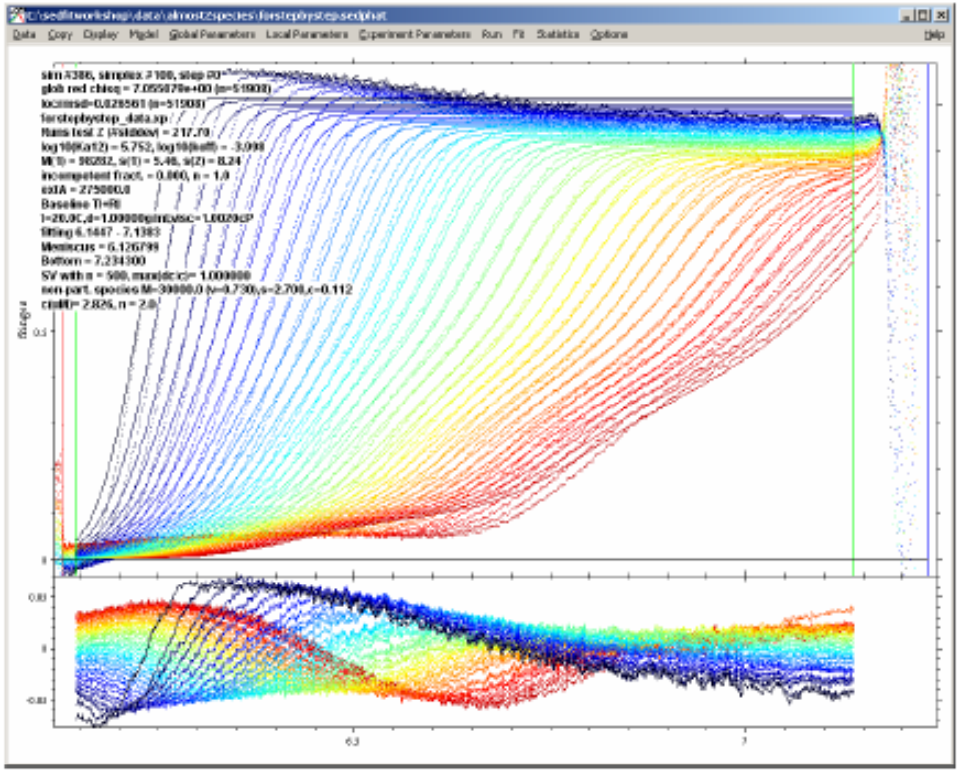


Yes, let's also save the xp file.

This way, we can recreate this analysis later, and we'll be able to tell from the title bar the filename.

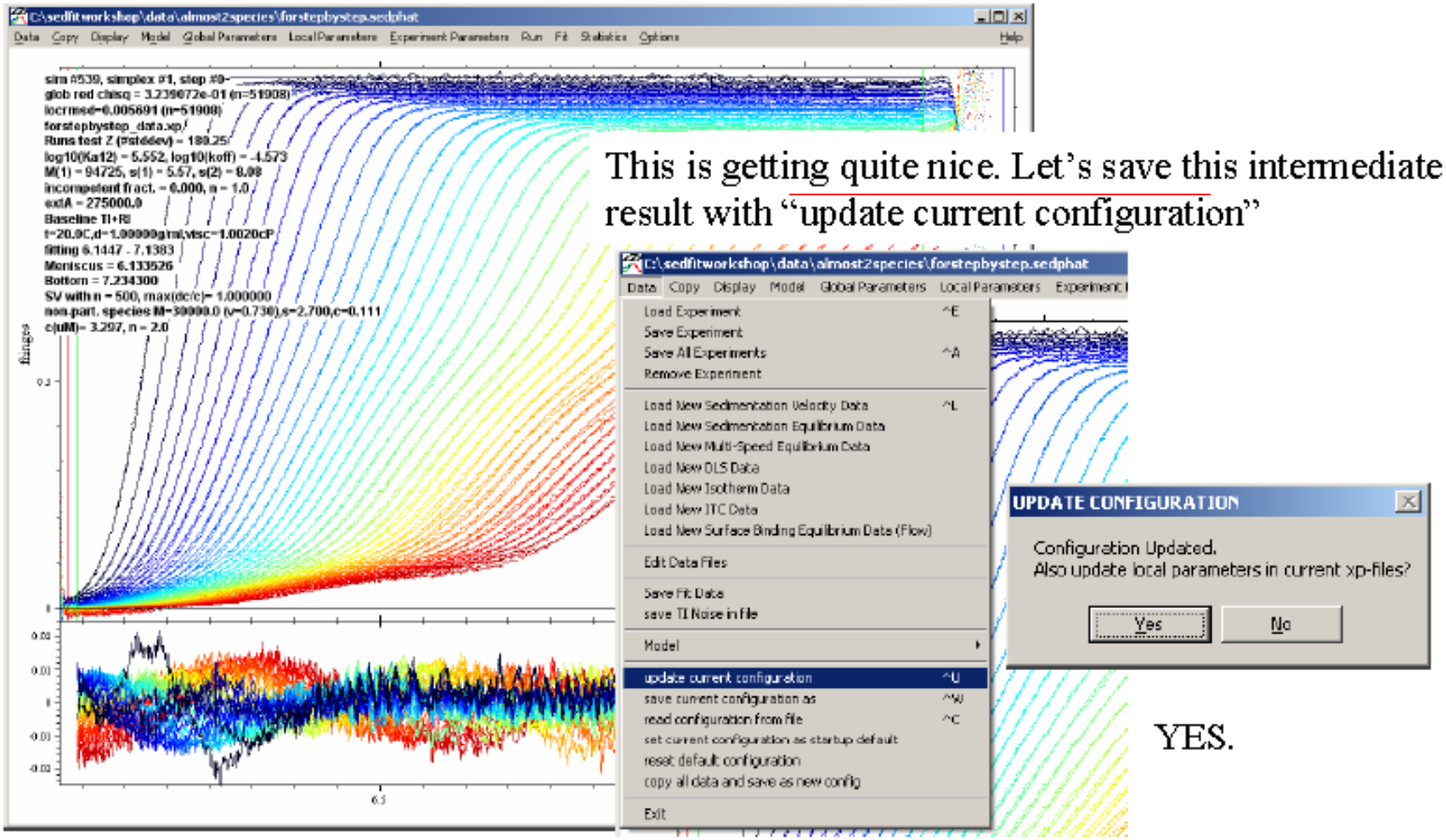
back to the fitting problem…

After just doing a RUN and control-N, we find this to be a good starting point. Not surprisingly, it seems to be worse then after the previous fit, but it will become hopefully much better again after we do a FIT.



Click on FIT. You'll notice the different behavior of the display during a Simplex fit as compared to the previous Marquardt-Levenberg fit. Now, we see things changing and gradually improving earlier. However, at some point the convergence seems to become quite slow.

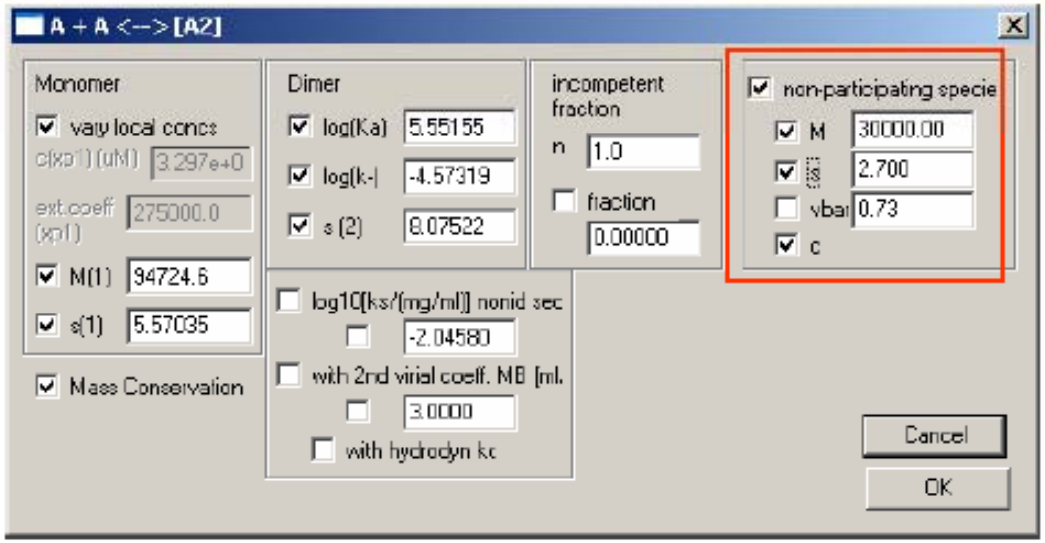
We could just let this go until it stops. I recommend that. However, being impatient, after the rmsd of the fit has dropped below 0.006, I interrupted this by hitting the space bar *once*. After control-N and control-O to get a good display, I find:



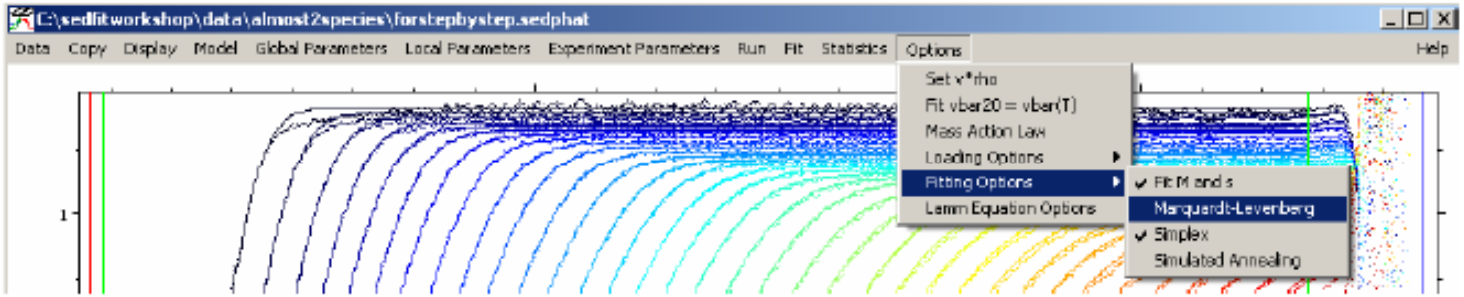
This is getting quite nice. Let's save this intermediate result with "update current configuration"

YES.

go back to the Global parameters box, and **float M and s of the non-participating species**

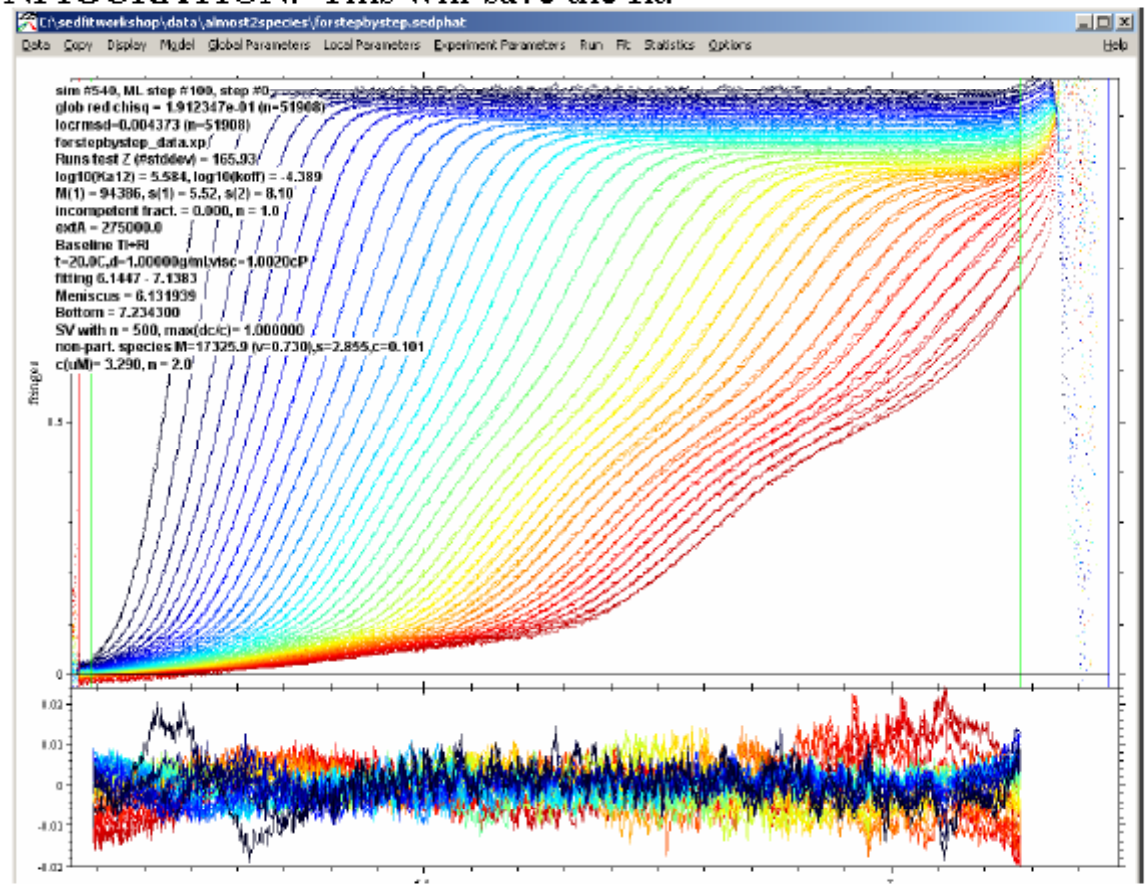


Also, we change back to Marquardt-Levenberg, to take advantage of its better 'homing-in' capability. Click on the "Marquardt-Levenberg" menu function to toggle.



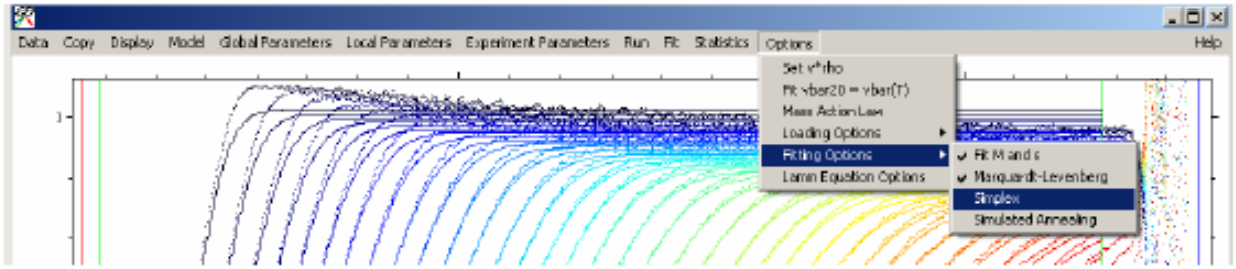
Then do a FIT

after a minute or two, we find the following best-fit (I used control-N and control-O to optimize display). use control-U or the menu function to UPDATE CURRENT CONFIGURATION. This will save the fit.



this fit does not look optimal, yet, as judged by the systematic residuals and misfit in the faster boundary. Therefore, we switch back to Simplex optimization and try to improve further.

click on the “Simplex” menu function. This will toggle off Marquardt-Levenberg and switch on the Simplex algorithm

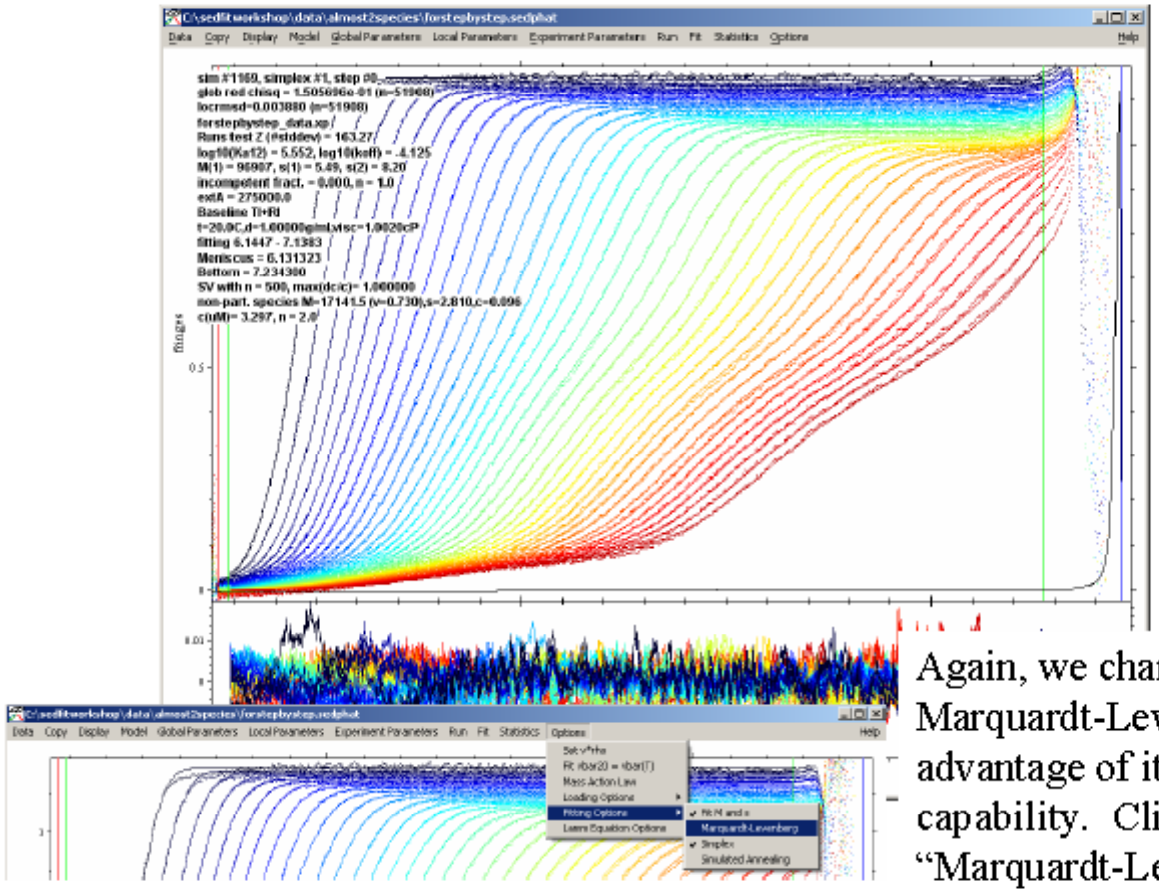


Then click on **FIT**

I saw initially a big improvement, indicating that the Marquardt-Levenberg again had been trapped in a small local minimum, from which the Simplex helped us escape. But after a while, convergence seemed to slow down very much, such that I again hit the space bar *once* to interrupt the fit.

[Hitting it once just stops the fit. It will then do a last simulation using the so-far best fitting parameters. If you inadvertently hit it twice, you are aborting this last simulation, and the fitted lines will be far off. However, you can rescue this by using a RUN command to re-establish the best-fit curves.]

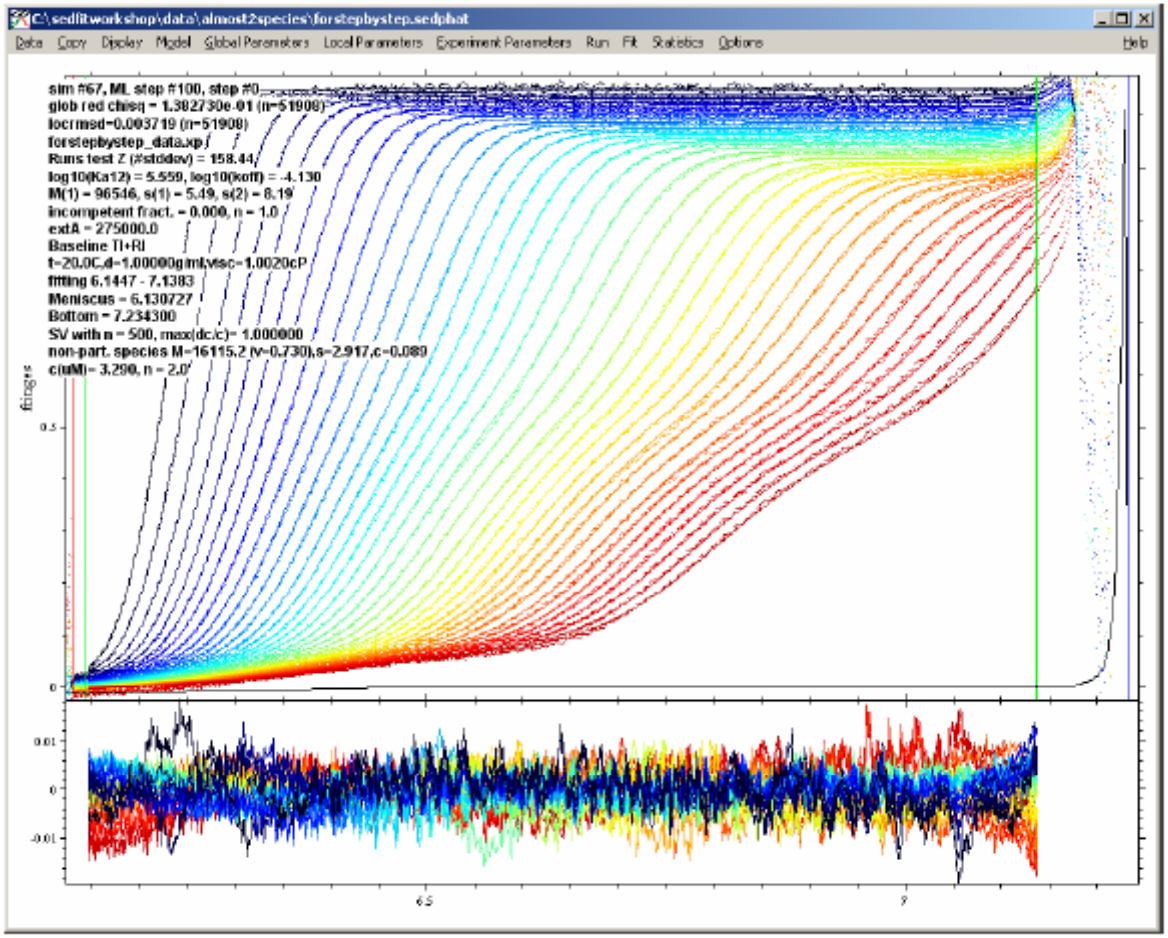
This is what I found. Use control-U or the menu function to UPDATE CURRENT CONFIGURATION. This will save the fit. Always also save the xp-file with it.



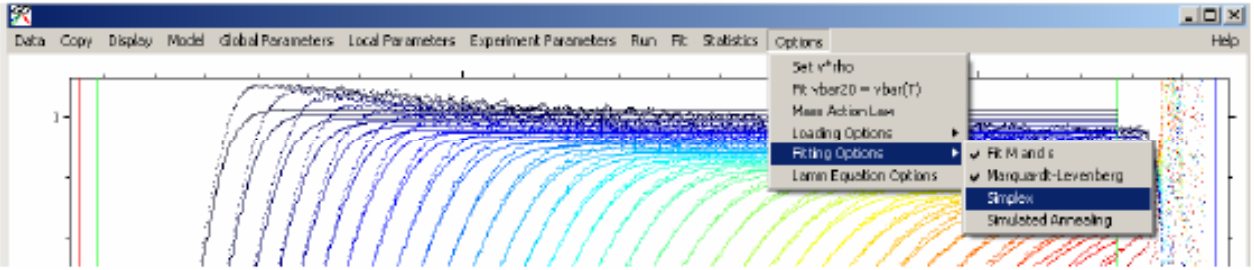
Again, we change back to Marquardt-Levenberg, to take advantage of its better 'homing-in' capability. Click on the "Marquardt-Levenberg" menu function to toggle.

Then do a FIT

This improved it slightly (judged by rmsd). Use control-U or the menu function to UPDATE CURRENT CONFIGURATION. This will save the fit.



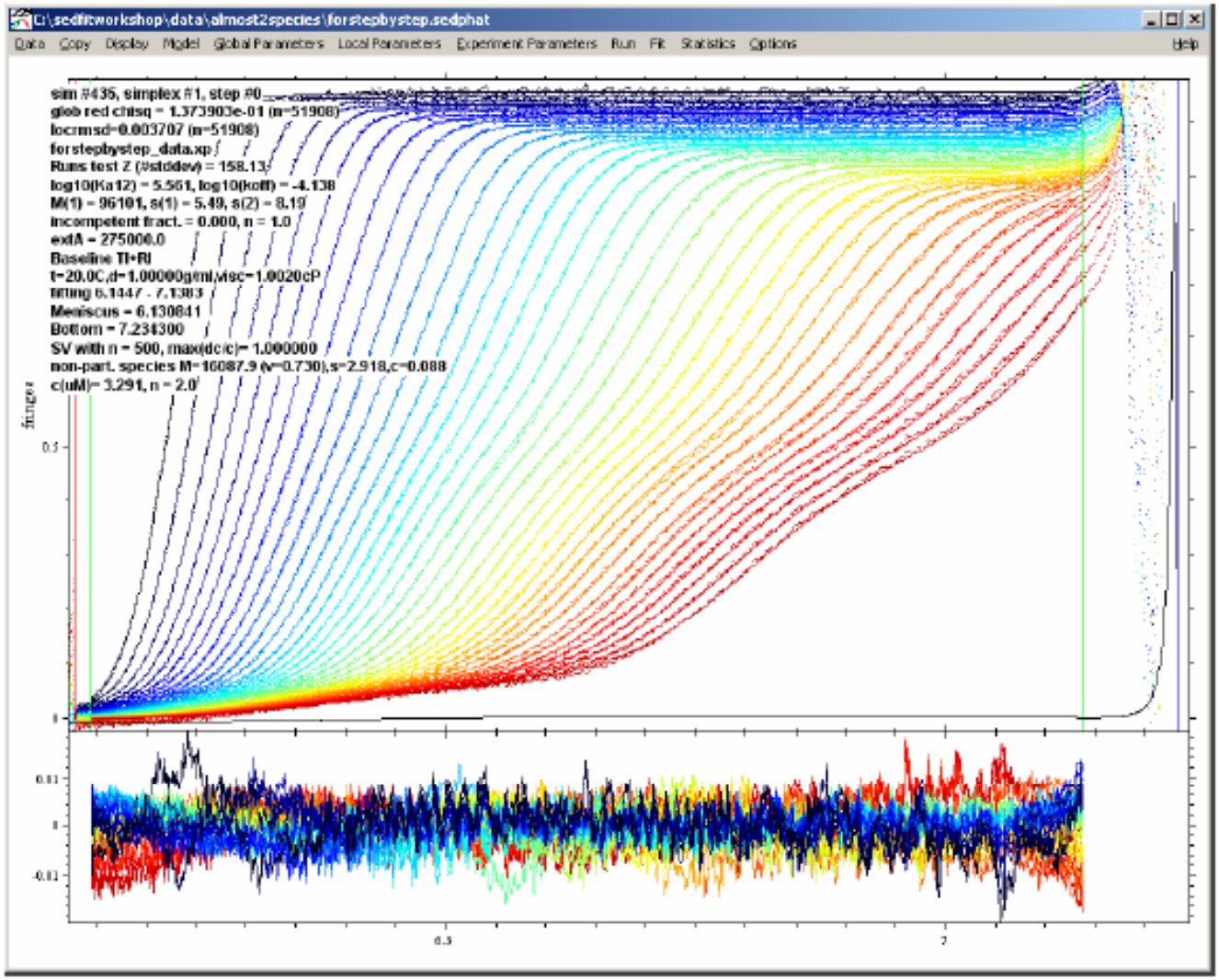
Again, click on the "Simplex" menu function. This will toggle off Marquardt-Levenberg and switch on the Simplex algorithm



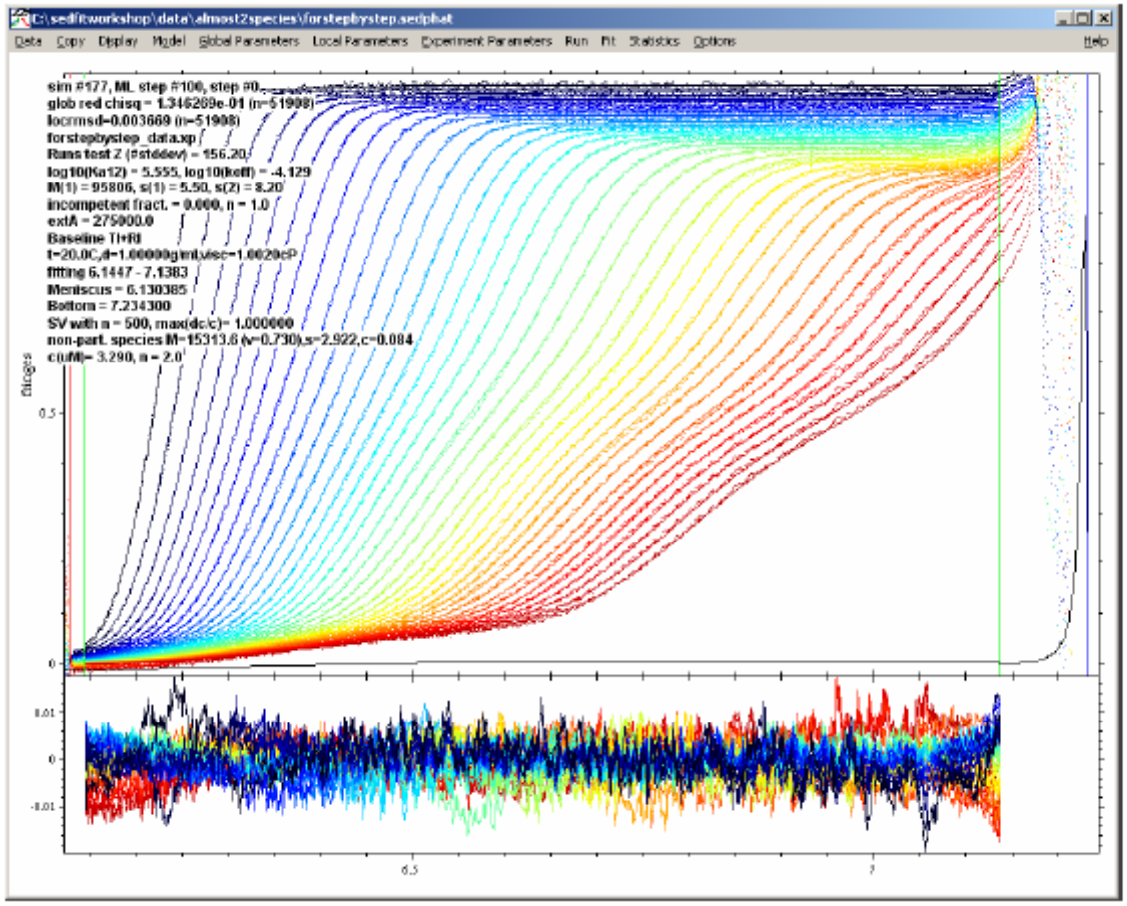
Then do a FIT

I find this still improves the fit slightly. After convergence gets very slow, I stopped it again and restarted the FIT. The reason for this is to take advantage of the random seeding of points each time you start a Simplex fit. This helped to get the rmsd further down. I repeat this procedure, but the second time it didn't seem to help.

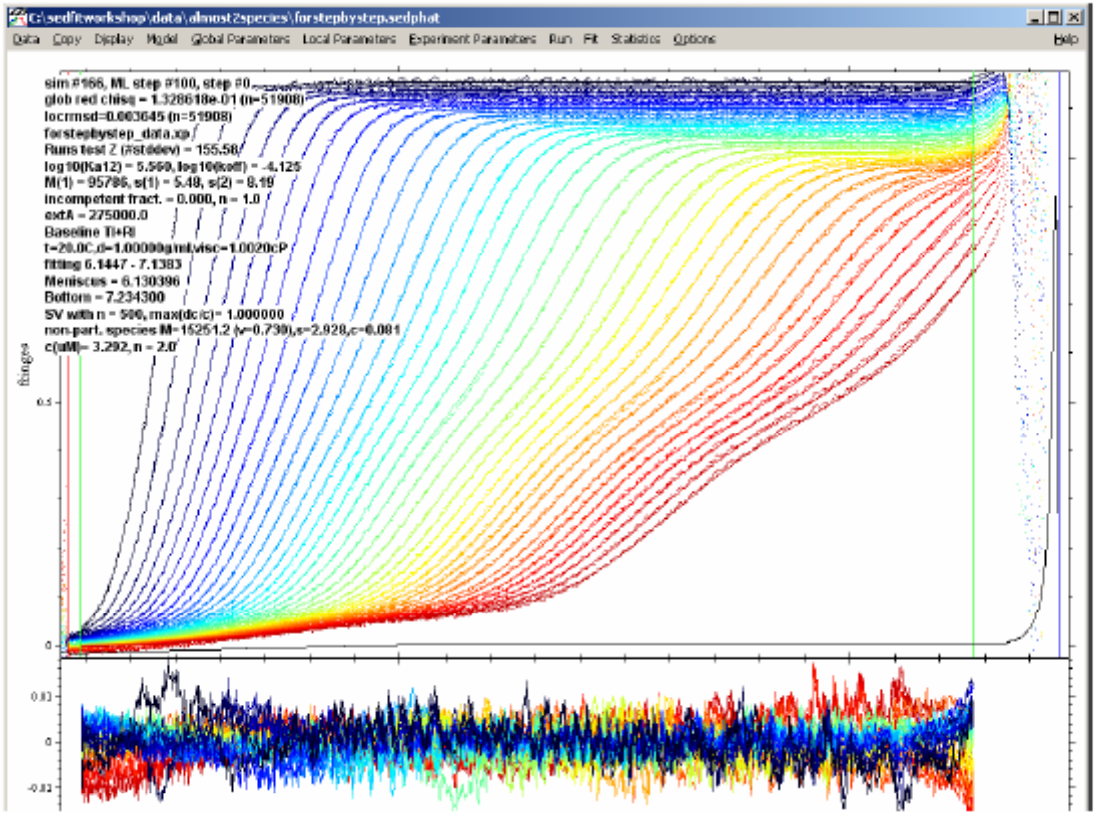
Save the configuration (control-U).



Then I switched back to the Marquardt-Levenberg method and tried again that way. This is the result: ***Further Improvement!!*** This is obviously a very badly shaped error surface, with a rather shallow valley that has many small ripples.



Same thing, again, after another round of alternating Simplex and Marquardt-Levenberg fits.



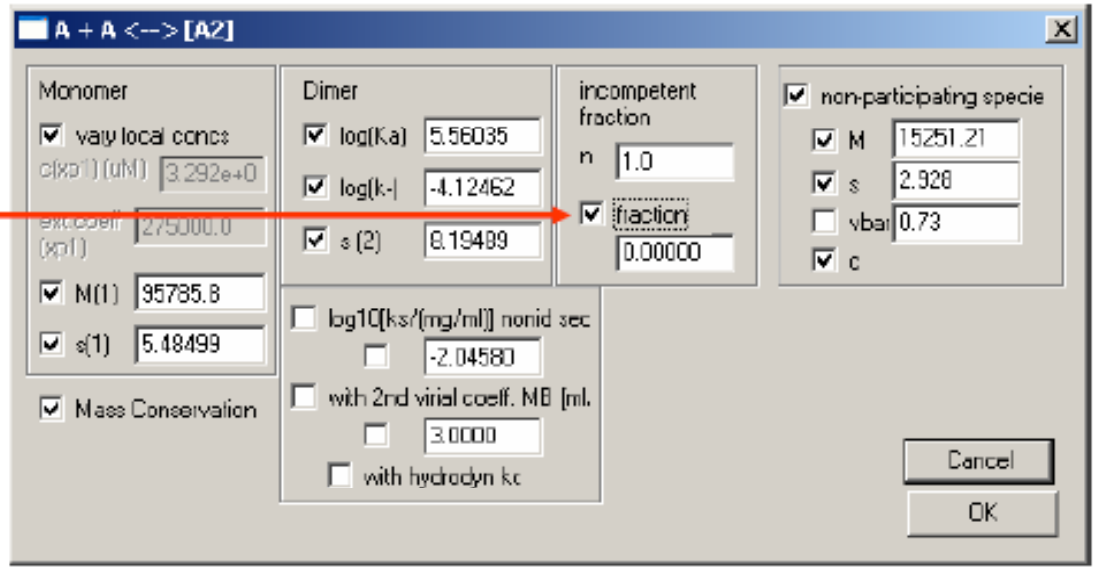
We could proceed in this way until we don't see any further improvement with any method. Another option would be to use the Simulated Annealing method, which I want to demonstrate in the following.

Simulated Annealing is better for error surfaces with many local minima, but more time-consuming. We could have used the Simulated Annealing method right from the beginning, but I didn't know yet that the error surface was so badly shaped.

At this point, let's float two more parameters that potentially could play a role in this fit:

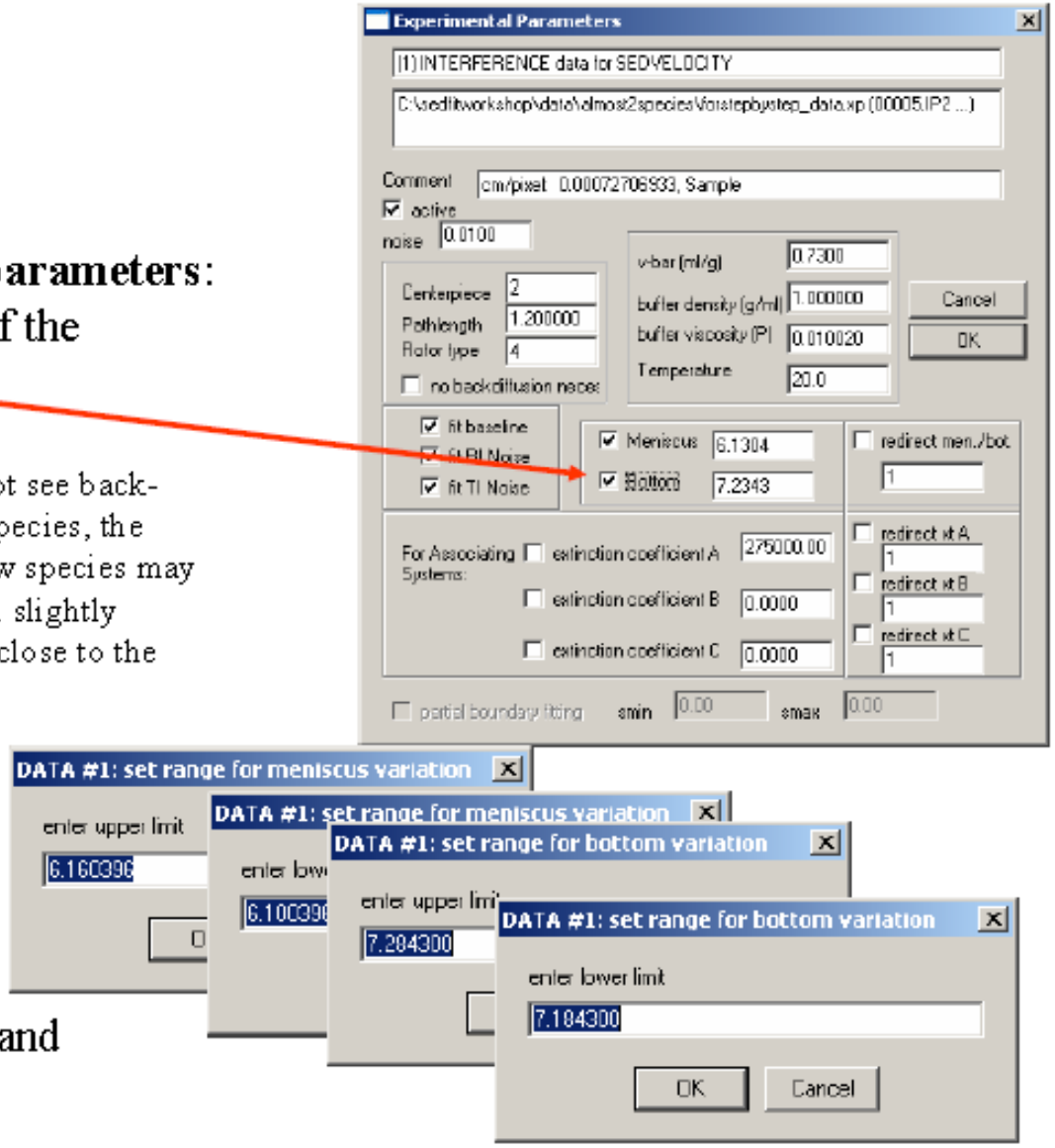
in the **global parameters**: switch on the incompetent fraction

using only one experiment from a single concentration is not really very good to determine this parameter (because it likely will be correlated with the Kd), but the fine-points of the boundary shape could carry some information about it.



in the **experiment parameters**: switch on floating of the bottom.

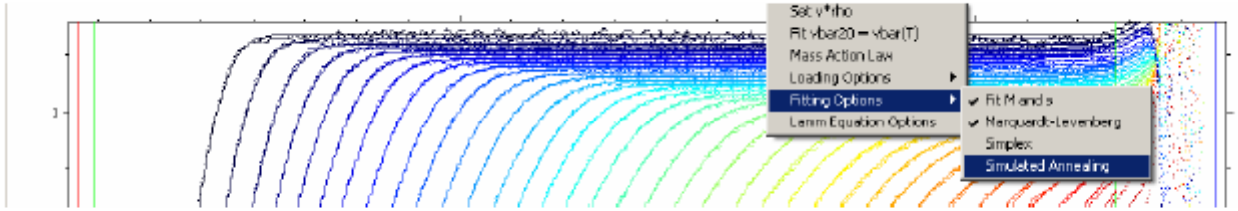
even though we should not see back-diffusion from the main species, the contaminating smaller Mw species may experience back-diffusion slightly influencing our residuals close to the right fitting limit.



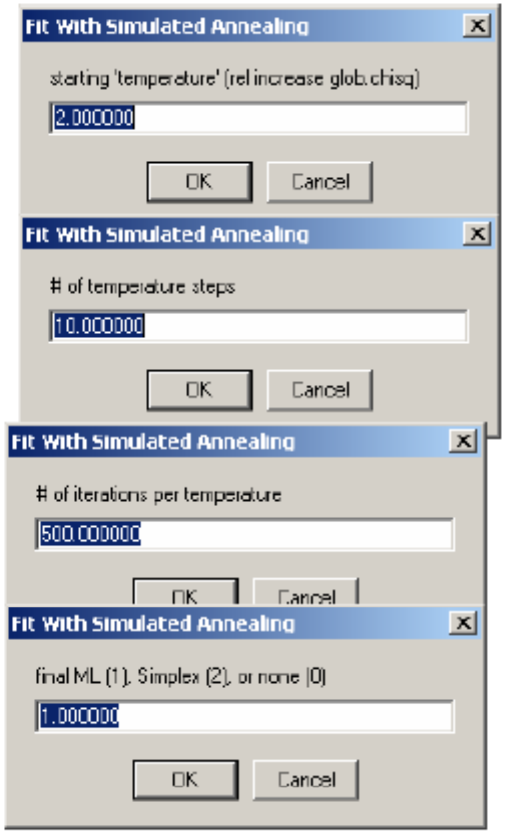
accept the default limits for meniscus and bottom variation.

before doing anything else, use SAVE CURRENT CONFIGURATION AS and save under a different filename (I used ‘forstepbystepSA’).

Click on the 'Simulated Annealing' function. This will toggle it on (and the other algorithms off)



The following boxes will show up. Accept the default for all.



This will set the starting tolerance for accepting worse points during the procedure. A value of 2 means that points within a chisquare of twofold the starting chisquare will be tolerated.

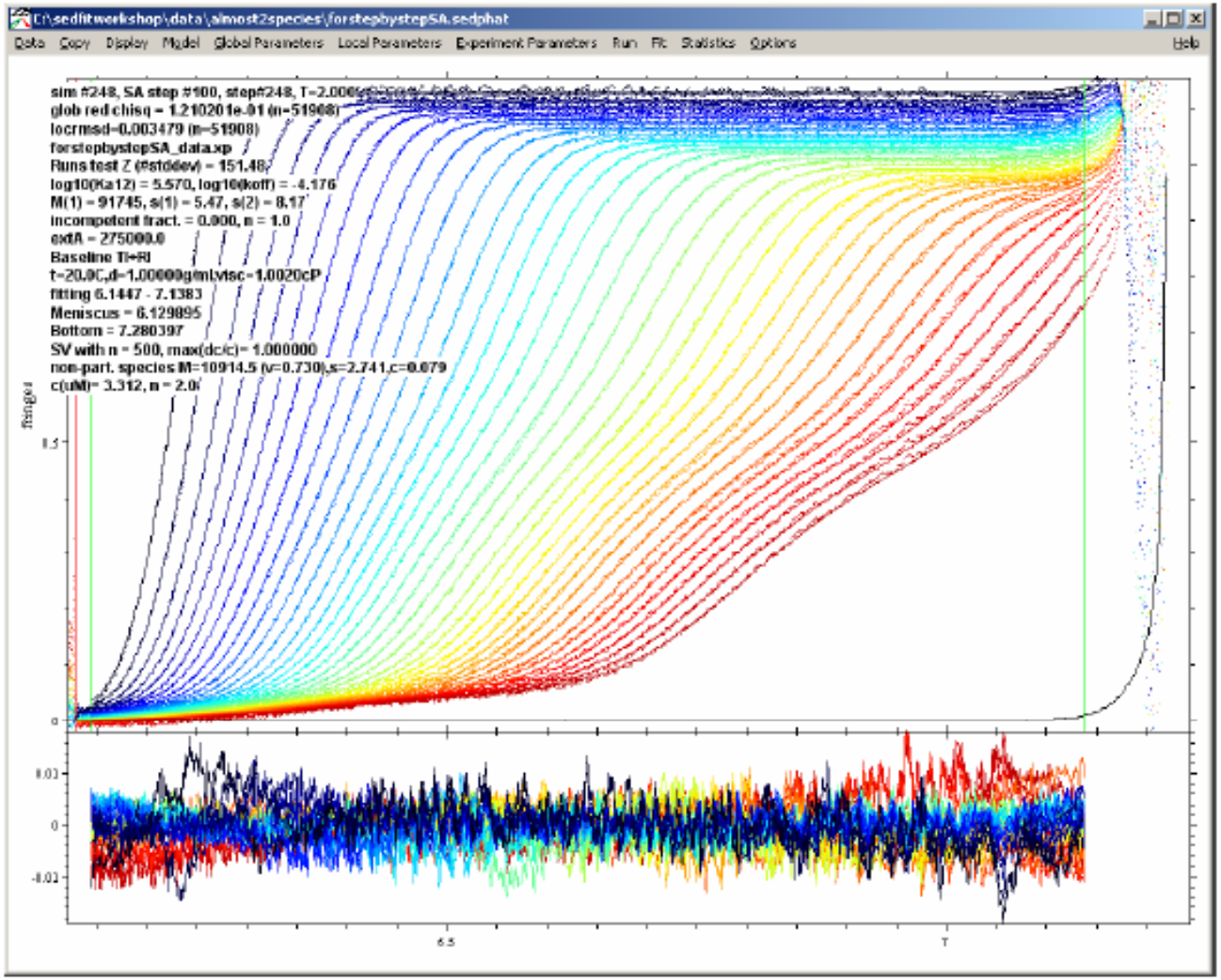
This determines the 'temperature' schedule, i.e. the number of times when we reduce the threshold chisquare for rejecting worse points.

Then do a **FIT**
This will take a while.

Since we will not expect any form of convergence, we have to predefine the number of iterations to be done at each temperature. You can get a feel for a ball-park number by looking at how many iterations the Simplex algorithm used previously.

We know that we likely will need another algorithm to home-in. The combination with a final round of Marquardt-Levenberg optimization seems to be reasonable.

This is where it ended up with SA – quite a bit better still. In fact, this is now better than the original c(s) fit we started with. We should continue searching with alternate Simplex, Simulated Annealing and/or Marquardt-Levenberg steps for better fits.



Finally, after a second trial with Simulated Annealing, no improvement was possible with Simplex and Marquardt-Levenberg. This is the best fit I found. Unfortunately, there's no guarantee that this is really the global optimum.