UC Berkeley
UC Berkeley Previously Published Works

Title

Apodization Specific Fitting for Improved Resolution, Charge Measurement, and Data Analysis Speed in Charge Detection Mass Spectrometry

Permalink

https://escholarship.org/uc/item/19c2n38m

Journal

Journal of The American Society for Mass Spectrometry, 33(11)

ISSN

1044-0305

Authors

Miller, Zachary M
Harper, Conner C
Lee, Hyuncheol
et al.

Publication Date

2022-11-02

DOI

10.1021/jasms.2c00213

Apodization Specific Fitting for Improved Resolution, Charge Measurement, and Data Analysis Speed in Charge Detection Mass Spectrometry

Zachary M. Miller[#], Conner C. Harper[#], Hyuncheol Lee[‡], Amanda J. Bischoff[#‡†], Matthew B. Francis[#‡†], David V. Schaffer[#‡], and Evan R. Williams[#‡*]

[#]College of Chemistry, University of California, Berkeley, California, 94720-1460, United States
[‡]California Institute for Quantitative Biosciences, University of California, Berkeley, California 94720-1460, United States
[†]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratories, Berkeley, California 94720, United States

[*]Address correspondence to this author.

Email: erw@berkeley.edu

Telephone: (510) 643-7161

Abstract

Short-time Fourier transforms with short segment lengths are typically used to analyze single ion charge detection mass spectrometry (CDMS) data either to overcome effects of frequency shifts that may occur during the trapping period or to more precisely determine the time at which an ion changes mass, charge or enters an unstable orbit. The short segment lengths can lead to scalloping loss unless a large number of zero-fills are used, making computational time a significant factor in real time analysis of data. Apodization specific fitting leads to a 9-fold reduction in computation time compared to zero-filling to a similar extent of accuracy. This makes possible real-time data analysis using a standard desktop computer. Rectangular apodization leads to higher resolution than the more commonly used Gaussian or Hann apodization and makes it possible to separate ions with similar frequencies, a significant advantage for experiments in which the masses of many individual ions are measured simultaneously. Equally important is a >20% increase in S/N obtained with rectangular apodization compared to Gaussian or Hann, which directly translates to a corresponding improvement in accuracy of both charge measurements and ion energy measurements that rely on the amplitudes of the fundamental and harmonic frequencies. Combined with computing the fast Fourier transform in a lower level language, this fitting procedure eliminates computational barriers and should enable real time processing of CDMS data on a laptop computer.

INTRODUCTION

Native mass spectrometry (MS) is widely used to obtain information about the structures, stoichiometries and interactions of molecules and molecular complexes directly from aqueous solution.[1–6] In conventional MS, ensembles of ions are measured to yield m/z spectra, where analyte charge and subsequently, mass, is determined either from the charge-state distributions of individual constituents or from the spacing of isotopic peaks in a single charge state. Analyte mass can become difficult to deduce with increasing molecular size and with increasing sample heterogeneity owing to the inability to resolve individual charge states.[7,8] Mass information has been obtained from highly purified virus capsids with masses ~18 MDa,[9] but often no information can be obtained even for much smaller molecular complexes that are more heterogeneous.

One solution to the problem of sample heterogeneity is to weigh ions individually so that there are no interferences with other components in the sample. Single ion mass measurements have been demonstrated with a number of different types of mass spectrometers, including Fourier-transform ion cyclotron resonance,[10,11] quadrupole ion trap,[12–16] Orbitrap,[17,18] and charge detection mass spectrometers that employ simple electrostatic traps.[19–24] In each of these methods, the frequency of ion motion is related to ion m/z. The charge can be determined either by stripping or adding a single charge[10,11,25] or it can be determined directly from the amplitude of the signal on charge sensitive detectors.[18,23,26–29] Single ion detection has also been demonstrated with time-of-flight mass spectrometry with energy-sensitive superconducting tunnel junction cryodetectors[30] and with nanomechanical resonators.[31,32]

In charge detection mass spectrometry (CDMS) with electrostatic ion traps, the signal amplitude and hence charge of an ion can be determined more accurately by increasing

3

measurement time to improve signal-to-noise (S/N) ratios. This improved accuracy comes at a cost of analysis time. To date, the lowest charge uncertainty that has been reported is 0.174 *e* and was achieved with a transient length of 3 s using a cryogenically cooled charge sensitive preamplifier to reduce noise.[33] Because ion frequencies can evolve with time, short time Fourier-transforms (STFT) are used to analyze CDMS data using step sizes that are sufficiently small so that changes in ion frequency are negligible.[29,34] The FT of an individual time-domain segment results in bins in the frequency domain data that have widths that are inversely related to the length of the transformed data. Shorter time-domain sections can lead to a loss of frequency resolution in the frequency domain and errors in peak amplitude known as scalloping loss. Scalloping loss occurs when there are an insufficient number of points to be able to adequately identify a peak maximum using peak picking routines that represent the centroid of a peak by the highest value point. This leads to an error in peak frequency and a systematic underestimation of peak amplitude. This is especially problematic for CDMS measurements where peak amplitudes are used to determine ion charge[26–28,35] and can also be used to determine ion energy.[26]

Scalloping loss can be reduced by zero-filling, where time domain data is extended by adding values of zero. This increases the number of points in the frequency domain but does not affect peak resolution. Although scalloping loss can be essentially eliminated with a sufficient number of zero-fills, the longer time-domain data sets can increase the computational time significantly. Fitting time-domain data with functions, including quadratics, has also been used to obtain improved estimates of both peak frequency and amplitude in Fourier-transform ion cyclotron resonance MS.[36] Because of the lengthy computation time required in CDMS, data is either processed off-line or has been processed in real time with a 48-core parallelized server.[37,38]

Here, an apodization specific peak fitting method is described which significantly reduces the number of zero-fills necessary to eliminate scalloping losses, making the data analysis significantly more efficient. This also results in improved accuracy in both frequency and amplitude determination and can be done in real time using a modest desktop computer. The application of this method to improve amplitude-based charge state determination and individual ion energy measurements is described.

EXPERIMENTAL SECTION

Instrumentation and Data Processing. Experiments were performed using an in-house built charge detection mass spectrometer that has been described in detail previously.[22,23] In short, ions are generated using electrospray ionization with borosilicate glass capillaries that are held at a potential of 1-3 kV relative to the entrance cone of the mass spectrometer (250 V) at a distance of ~3 mm. Ions pass through a modified Z-spray source (Waters Corporation, Milford, MA, USA), a quadrupole ion guide and are trapped and thermalized for up to 1 s in a linear quadrupole (Ardara Technologies, Ardara, PA, USA) at a potential ~200 V. Ions are pulsed out of the trapping quadrupole and pass through a 90 degree turning quadrupole, which admits ions with a limited spread of kinetic energies into an electrostatic ion trap. The front of the trap is held at 0 V to allow ions to enter the trap and the potential is quickly raised to 330 V to trap and store ions. Measurements are made for trap times of either 1 s or 7.5 s. The charge-induced signals of trapped ions are amplified by a room temperature CoolFET charge-sensitive preamplifier (Amptek, Bedford, MA, USA) and a linear voltage amplifier. The output of the preamplifier is passed through a custom-built Butterworth bandpass filter with low and high cutoff frequencies of 10 kHz, and 300 kHz respectively. The signal is digitized at 1 MHz using an ATS9350 digitizer board (AlazarTech, Pointe Claire, Quebec, CA).

5

Adeno-associated viruses (AAVs) are used as gene delivery agents to treat rare genetic disorders[39,40] and assemblies of the circularly permuted tobacco mosaic viral capsid protein (cpTMVs) are used as models of photosynthetic light harvesting complexes.[41,42] These samples produce ions in the 0.6 – 4.7 MDa range making them well suited to evaluate the performance of data processing methods for single ion CDMS. The preparation of AAV9 samples used in this work is described elsewhere;[41,43–45] a brief description is provided in Supporting Information. A previously unreported mutant of cpTMV containing a non-canonical acid, 3-nitrotyrosine (3NY), at position S65 (cpTMV-S65-3NY) was used in this work due to its assembly homogeneity. Details of the preparation and characterization of cpTMV-S65-3NY are provided in the Supporting Information (Figure S1).

The time-domain data is analyzed using a STFT, where transients are divided into segments of time, the segments are apodized using either a rectangular (i.e., unapodized), Hann or Gaussian ($\sigma = 8$ ms) function, the segments are zero-filled, and subsequently Fourier transformed. In this work, a segment size of 50 ms (50,000 data points) was used, corresponding to 20 Hz bin sizes in frequency space. Zero-filling, where zeros are added to the end of each segment in multiples of the original segment length, is used to reduce the size of frequency domain bins. Herein, adding n equivalents of the window length in zeroes is described as adding n zero-fills, i.e., if a 50 ms segment has two zero-fills, then the zero-filled segment will contain 50 ms of real data, and 100 ms of zeroes (50,000 real data points and 100,000 zeros). Subsequent segments are stepped across the time domain in 5 ms intervals, i.e., 0-50 ms, 5-55 ms, etc., which enables the tracing of ion signals in the frequency domain as ion frequency evolves in time. All data analysis is performed using the Python programming language, version 3.10, on a desktop

computer equipped with a 12 core, 24 thread Ryzen 5900 processor (AMD, Santa Clara, CA, USA), 128 gigabytes of DDR4 RAM (G Skill, Taipei, Taiwan).

The Fourier transform of a rectangular window of arbitrary length yields a sinc function given in equation 1. The sinc function is fit to magnitude Fourier transform data by taking the absolute value, and for signals occurring at any amplitude, width, center frequency and baseline by adding the parameters A, w, f_c, and B, respectively (Equation 2). This process can be repeated for Hann and Gaussian windows to yield an analytical description of their frequency domain peak shapes, but this derivation is not included nor used in this work.

$$\text{Eq. 1 } Sinc(x) = \frac{Sin(\pi x)}{\pi x}$$

$$\text{Eq. 2 } Sinc(f, A, w, f_c, B) = A \left| \frac{Sin(\pi w(f - f_c))}{\pi w(f - f_c)} \right| + B$$

Ion signals are fit to Eq. 2 using non-linear least squares fitting utilizing the Levenberg-Marquardt algorithm and requires initial conditions that are given by the peak height and center frequency from the Fourier transform, a width that is given by the window length in seconds, and a baseline that is initially set to zero. Peak fitting in this analysis is unconstrained, and the final fitted peak height is the sum of parameters A and B after optimization. In the experiments described here, this process of fitting rectangularly apodized frequency domain peaks is referred to as sinc fitting. The process of picking out the highest amplitude and corresponding frequency along a peak in the frequency domain without using the optimized fit equation is referred to as peak picking.

RESULTS AND DISCUSSION

Effects of Data Analysis on Scalloping Loss. In FT-based methods, the resolution or ability to distinguish signals of ions that have similar frequencies depends on several factors, including the length of the time-domain signal. In CDMS, the frequencies of ions evolve with time as a result of small changes in ion trajectories as well as collisions with background gas that reduce ion energy. The latter effect on frequency can be minimized but not entirely eliminated with traps that have harmonic-like potentials.[24] Because ion frequencies can evolve with time, a STFT analysis in which shorter time segments during which ion frequency does not change significantly are used to analyze the time-domain data. A short time window is also advantageous for high time resolution, for example, identifying when an ion fragments or enters an unstable orbit within the trap.[38,46,47] The filter-diagonalization method has been applied to frequency shifts that occur as a result of ion-ion interactions in Fourier-transform ion cyclotron resonance instruments and in principle, this method could also be applied to CDMS data.[48] The ability to accurately determine the frequency and amplitude of individual ions also depends on the duration of the time-domain signal. The length of the time-domain signal can be increased by zero-filling, which does not affect peak width but does increase the accuracy with which a peak centroid and amplitude can be obtained by reducing the size of the corresponding frequency domain bins. This reduces what is referred to as scalloping loss, which is a systematic negative error in the measured peak height when it is determined solely by the highest point along the peak, i.e., when 'peak-picking' is used.

The effect of scalloping loss is illustrated in Figure 1a and 1b, which show measured signal from a single AAV9 ion (4.7 MDa) at two different times over a 1 s trapping period. The signal was processed using 50 ms of data, rectangular apodization, and with two zero-fills (100

8

ms), yielding 6.67 Hz bin widths in frequency space. The peak of the ion signal can fall in-between bins resulting in a maximum value that is offset from the true frequency by up to 3.34 Hz and an amplitude that is lower than the true value (Figure 1a), or it can be centered directly on a bin which results in high accuracy centroid determination, and subsequently, high amplitude accuracy (Figure 1b).

Inaccuracies in peak height measurement that stem from scalloping loss are traditionally mitigated through large numbers of added zero-fills to the measured time-domain data. This is illustrated by the dotted black lines in Figure 1a and 1b which correspond to the same time-domain signals that are zero-filled to total time-domain length of 2000 ms. Although the "true" peak frequency and amplitude can be obtained with a sufficiently large number of zero-fills, the longer time-domain data leads to increased computation time. This necessitates high-performance computing[37] or lower-level (more efficient) software in order to enable real-time processing of CDMS data.

Conversely, apodization specific fitting makes it possible to accurately determine peak centroid frequency and amplitude using fewer zero-fills through interpolation, which results in lower computation time. The results of a rectangular apodization fitting using the same 50 ms of data and 100 ms zero-fills (150 ms FT) are shown as red lines in Figure 1a and 1b. These fits are nearly indistinguishable from the results obtained with the significantly larger number of zero-fills but this fitting requires only a fraction of the computing time.

The error in amplitude due to scalloping loss changes in magnitude over the course of the measurement because the frequency of an ion can shift with time. The shifting ion frequency results in an error in amplitude that is a maximum when it bisects two frequency domain bins (Figure 1a) and a minimum when it occurs on top of a frequency domain bin (Figure 1b). This

variable error is illustrated in Figure 1c, which shows the deviation of peak peaking (blue line) compared to results the 2000 ms zero-filled data (black dashed line) where scalloping loss is negligible for the 1 s duration of the trapping period. The error ranges from -4.46% and -0.00% and has an average value of -1.21%. In contrast, the fitted amplitudes obtained with the same number of zero-fills varies between -0.64% and +0.72% and has an average deviation of just 0.02%.

Other factors affect the variability of the absolute measured amplitude of ion signals, including random noise, collisions with background gas that reduce ion energy and can also result in a small shift in ion trajectories. The latter two phenomenon affect the duty cycle of the ion signal, which changes the amplitudes of the fundamental frequency and the corresponding harmonic frequencies. Thus, the standard deviation in the amplitude of the fundamental frequency does not reflect the standard deviation in charge measurement because the effects of signal duty cycle are taken into account in the determination of ion charge.[27] Moreover, the ratio of the abundances of the second harmonic to the fundamental frequency is used to obtain ion energy throughout the measurement, so an improvement in amplitude measurements correspond to improvements in ion energy resolution. This is essential for a multiplexing scheme that enables multiple ions with the exact same m/z to be analyzed simultaneously if these ions have different kinetic energies.[34,38]

To determine the extent to which sinc fitting improves the absolute amplitude measurements of the fundamental frequency used to determine charge, results from peak picking and apodization specific fitting of 150 ms time-domain data (50 ms experimental data and 100 ms of zero-fills) was compared to 2000 ms time-domain data with the same number of experimental data and the remainder zero-filled (Figure 1d). The signal amplitudes of the

resulting 190 STFT segments that were obtained using peak picking and from sinc fitting have a standard deviations of 3.72 a.u. and 2.68 a.u., respectively. This equates to a ~28% improvement in amplitude standard deviation and, by extension, a ~28% reduction in charge uncertainty. Amplitudes obtained from the 2000 ms time domain standard also have a standard deviation of 2.68 a.u., indicating that fitting approached the maximum amplitude precision that can be achieved through high levels of added zero-fills. Even in cases where the frequency of an individual ion does not change in time and thus scalloping losses are consistent over the trapping period, the amplitudes of the many individual ions required to compile a CDMS mass histogram will have different extents of scalloping losses at their respective frequencies, broadening the overall distribution of amplitudes and increasing charge uncertainty.

In order to determine the improvement in amplitude (charge) accuracy from sinc fitting, data for ~400 cpTMV-S65-3NY ions (602 kDa) trapped for 7.5 seconds were compared to determine the effect on the accuracy of charge-state measurements (Figure 2). Data for these longer trapping times were processed in the same way as the Figure 1 data. Without sinc fitting, individual charge states are not clearly resolved. In contrast, charge states between 55+ and 64+ are individually well-resolved with sinc fitting. Fitting these data to Gaussian peak shapes results in an average charge uncertainty of ~0.254 e. The improved charge-state resolution with sinc fitting is a direct result of reduced scalloping loss and is the first example of amplitude only-based charge state resolution measured using CDMS without the use of a cryogenically cooled preamplifier. It is also important to note that sinc fitting results in a higher median value of charge. This is due to the systematically lower amplitudes that are obtained from peak picking with few zero-fills. In cases where individual charge states are not resolved, the effects of scalloping loss need to be accounted for through calibration.

Comparison of Apodization Functions. Rectangular apodization yields peaks that are the narrowest in the frequency domain compared to other apodization functions.[49] For experiments in which only one or no ions are trapped,[29] higher frequency resolution is not an advantage. However, higher resolution is advantageous in multiplexing experiments in which the individual masses of many ions are measured simultaneous as a result of lower probability that the signals of any two ions will overlap in frequency.[34] Thus, more ion traces in congested regions of frequency space can be resolved leading to an even greater extent of multiplexing that is possible for faster acquisition of mass spectra.[34,38]

The advantage of higher resolution obtained with rectangular apodization in multiplexed CDMS measurements is illustrated in Figure 3a, which shows a limited frequency range over which the frequencies of three AAV9 ions are measured for one second (data from 725 to 975 ms is shown). The frequencies of two of these ions are similar with a minimum frequency separation of ~40 Hz at 885 ms. A one-dimensional FT "slice" of the STFT of data in the 50 ms segment that is centered at 885 ms is shown in Figure 3b using a rectangular (black line), Hann (red line) and a Gaussian (blue line) apodization function. The frequencies of the two ions are clearly resolved as two distinct peaks with a rectangular apodization but are completely unresolved with both the Hann and Gaussian apodization. The higher resolution obtained with rectangular apodization can lead to a significant increase in the number of resolvable ions in multiplexed measurements. Unresolved signals of two ions, such as what occurs with either Hann or Gaussian apodization in this example, would lead to an incorrect assignment of this signal to a single ion with a frequency between those of the two individual ions and an assignment of charge greater than that of either ion. Because ions with similar mass, charge, and energy will have similar frequencies, time-domain data with the signal from many ions may

12

result in the suppression of the most common ions due to higher probability of frequency overlaps.[34] Higher resolution obtained with rectangular apodization will yield the minimum extent of suppression of more abundant ions because narrower peaks reduce the probability of signal overlap.

The higher resolution obtained with a rectangular apodization function comes at a cost of a greater extent of spectral leakage, i.e., "ringing". This can be clearly seen in Figure 3b as well as Figure 1a and 1b as the oscillations on either side of the main peaks and as low intensity parallel traces to the main frequencies on Figure 3a. These side bands have the potential to obscure low intensity signals with frequencies on either side of a main peak. Constructive and destructive interference of peak amplitudes also occurs if the ions are close enough in frequency to allow for side lobe overlap with the main peak of an adjacent ion. Whether or not this interference is constructive or destructive in any given segment of time depends on when the ions initially entered the trap and their time-dependent frequencies. The shift between fully constructive and destructive interference is periodic if the ions remain at constant frequencies with a period related to the difference in ion frequency and length of transformed time. Given two ions that are trapped at similar enough frequencies where interference in peak amplitude from adjacent ion side lobes occurs, both ion amplitudes can be accurately determined if the ions are trapped for a sufficiently long time. In this case, the constructive and destructive interference from adjacent ion side lobes will increase the standard deviation of the ion amplitude measurements, but the mean value will be the same as the case where the ions are sufficiently separated in frequency and no interference occurs.

Rectangular apodization has the lowest equivalent noise bandwidth of any apodization function, which translates to a lower noise baseline and higher signal-to-noise ratio (S/N) in

Fourier transforms of sinusoidal signals.[49] To investigate whether this is also true of the unique

periodic ion signals in CDMS, the same AAV9 ion signals from data shown in Figure 1 was

processed with a rectangular, Hann, and Gaussian apodization under otherwise identical

parameters. The S/N resulting from this analysis for the 1 s trap time are shown in Figure 3c

(results and description of S/N analysis is provided in Supporting Information Figure S2). The

S/N with the Hann apodization is slightly greater than that of the Gaussian apodization, but the

rectangular apodization function results in a S/N ratio that is 1.21 and 1.26 times greater than the

Hann and Gaussian functions, respectively. Thus, the S/N advantage of rectangular apodization

also applies to the periodic but not sinusoidal signals of CDMS. This leads to a corresponding

direct improvement in charge measurement accuracy.

Although the rectangular apodization function is not always preferred in applications

such as FT-ICR and other FT-based analytical techniques due to its effect on the dynamic range

of closely spaced signals,[49] it is ideal for CDMS analysis because it makes it possible to analyze

more ions per unit time due to its narrower intrinsic peak width in the frequency domain, and it

decreases charge uncertainty due to its low equivalent noise bandwidth.

Computation Time Savings. Sinc fitting requires significantly fewer zero-fills to obtain

accurate frequency and amplitude (charge) measurements compared to those values obtained

from peak picking. This results in a substantial reduction in computation time that enables

uncompromised real time data processing of CDMS data with relatively limited computing

power. In order to determine the number of zero-fills required to obtain similar accuracy to sinc

fitting that uses two zero-fills, data for 20 individual 1 s trapping periods sampled at 1 MHz was

processed with a varying number of zero-fills. A total of 56 AAV9 ions were identified and

trapped for the duration of these 20 periods. The time required to process these data as a function

of the number of added zero-fills and that for the fitting procedure are shown in Figure 4. The total computation time shown in red (left y-axis) is for processing 20 s of acquired data with peak picking (circles) and peak fitting (triangle) and the percentage difference between the average peak amplitude compared to that with 39 added zero-fills (the equivalent of a 2000 ms time domain sample, creating 0.5 Hz frequency domain bin widths) is shown in blue (right y-axis). The computation time increases only marginally with up to five zero-fills as a result of the ability to parallelize processing these data over multiple processor cores. The error decreases rapidly over this range but is still -0.38% at five zero-fills. Between 5 to 9 zero-fills, the computation time required increases more rapidly, but real time data processing can still be accomplished with up to 9 zero-fills. With 9 zero-fills, the total computation time is 15.4 s and results in an amplitude percent deviation of -0.12%. The computation time increases more linearly with increasing number of zero-fills above 10 due to limitations in internal data transfer speeds in the random access memory of the computer. The accuracy approaches the asymptotic value given by the dashed blue line corresponding to the value with 39 added zero-fills (2000 ms of transformed data).

By comparison, the apodization function specific fitting with two zero-fills results in an average amplitude percent deviation of just +0.03% while requiring only 5 s to analyze 20 individual 1 s transients. Thus, this method enables real time processing of CDMS data with high precision. In contrast, processing with 9 zero-fills without fitting took three times longer and resulted in an amplitude error that is four times higher. In order to obtain a comparable level of accuracy to that obtained by using fitting and 3 zero-fills, 19 zero-fills are necessary (average amplitude error of -0.03%) using only peak picking and 45.6 s of computation time is required

15

for analysis. Thus, the fitting procedure results in more than a 9-fold gain in data analysis speed while providing the same accuracy as zero-filling alone.

Fitting frequency domain peaks introduces an "overhead" cost to the overall computation time compared to the time necessary without fitting when the same number of zero-fills is used. With two zero-fills, the fitting procedure increased the computation time by ~0.6 s. This overhead depends on the number of ions in each file. For these data, there were an average of 2.8 ions per file corresponding to an overhead of 10.7 ms per ion. The overhead time should increase roughly linearly with the number of ions in each file. Based on these data, we expect that files containing signals for up to ~73 ions can still be computed in real time. For files with over ~446 ions, 39 zero-fills becomes more computationally efficient, but ion-ion interactions and frequency overlap would almost certainly preclude this number of simultaneous ion measurements.

The computation time required for this analysis is both computer and software dependent. Python is a higher-level language and has the advantage of being easy to use to develop complex software routines. However, it is not the most efficient for computing Fourier transforms. A lower-level language could substantially reduce these computation times. Initial results in the C++ programming language indicate that a nearly ten-fold reduction in computation time may be achieved, making significantly more zero-fills possible in real time or making real-time processing of data on a laptop computer possible using the sinc fitting method. It should be noted that using lengths of time that correspond to a total number of transformed data points that are factors of two decreases the speed of computing FFTs.[50] However, this performance gain is not realized here because internal data transfer speeds and overhead from the Python programming language are the primary contributors to overall computation time.

CONCLUSIONS

In single ion charge detection mass spectrometry measurements, where the *m/z* of an ion is obtained from the oscillation frequency and ion energy, and the charge of the ion is determined from the amplitude of the induced signal, accurate amplitude measurements are essential. Scalloping loss arising from peak picking of limited time-domain data segments in STFT analysis of CDMS data can occur unless a large number of zero-fills are used for each segment. However, this comes at a significant cost of computation time leading to the use of off-line data analysis on standard desktop computers or a highly efficient computing cluster for real time data analysis. Apodization specific fitting substantially reduces the number of zero-fills needed to reduce scalloping loss to a negligible value and results in a 9-fold reduction in computation time compared to zero-filling to a similar accuracy without fitting. In combination with a lower-level language for the computation of fast Fourier transforms, it should be possible to process these data in real time on a laptop computer.

Although Gaussian, Hann, and other apodization functions are often used in other FTMS methods, as well as in some CDMS measurements,[28] we demonstrate that rectangular apodization significantly improves frequency resolution for the individual ion signals measured in CDMS. This is advantageous for ion multiplexing measurements where ions with similar frequencies can be more easily resolved. Equally important is the >20% S/N gain obtained with rectangular apodization compared to Gaussian and Hann apodization. This improvement in S/N translates into a corresponding improvement in charge measurement accuracy and in ion energy measurements that are obtained from signal amplitudes of the fundamental and harmonic frequencies.

Supporting Information

Supplementary experimental section, including cpTMV-S65-3NY preparation and monomer characterization by LC-MS (Figure S1) and AAV9 preparation, and noise characterization methods and frequency domain plots (Figure S2)

References

(1)   Loo, J. A. Electrospray Ionization Mass Spectrometry: A Technology for Studying Noncovalent Macromolecular Complexes. *Int. J. Mass Spectrom.* **2000**, *200* (1–3), 175–186.

(2)   Heck, A. J. R.; Van Den Heuvel, R. H. H. Investigation of Intact Protein Complexes by Mass Spectrometry. *Mass Spectrom. Rev.* **2004**, *23* (5), 368–389.

(3)   Hernández, H.; Robinson, C. V. Determining the Stoichiometry and Interactions of Macromolecular Assemblies from Mass Spectrometry. *Nat. Protoc.* **2007**, *2* (3), 715–726.

(4)   Painter, A. J.; Jaya, N.; Basha, E.; Vierling, E.; Robinson, C. V.; Benesch, J. L. P. Real-Time Monitoring of Protein Complexes Reveals Their Quaternary Organization and Dynamics. *Chem. Biol.* **2008**, *15* (3), 246–253.

(5)   Han, L.; Hyung, S. J.; Mayers, J. J. S.; Ruotolo, B. T. Bound Anions Differentially Stabilize Multiprotein Complexes in the Absence of Bulk Solvent. *J. Am. Chem. Soc.* **2011**, *133* (29), 11358–11367.

(6)   Kintzer, A. F.; Sterling, H. J.; Tang, I. I.; Abdul-Gader, A.; Miles, A. J.; Wallace, B. A.; Williams, E. R.; Krantz, B. A. Role of the Protective Antigen Octamer in the Molecular Mechanism of Anthrax Lethal Toxin Stabilization in Plasma. *J. Mol. Biol.* **2010**, *399* (5), 741–758.

(7)   Weiss, V. U.; Bereszcazk, J. Z.; Havlik, M.; Kallinger, P.; Gösler, I.; Kumar, M.; Blaas, D.; Marchetti-Deschmann, M.; Heck, A. J. R.; Szymanski, W. W.; Allmaier, G. Analysis of a Common Cold Virus and Its Subviral Particles by Gas-Phase Electrophoretic Mobility Molecular Analysis and Native Mass Spectrometry. *Anal. Chem.* **2015**, *87* (17), 8709–8717.

(8)   McKay, A. R.; Ruotolo, B. T.; Ilag, L. L.; Robinson, C. V. Mass Measurements of
        Increased Accuracy Resolve Heterogeneous Populations of Intact Ribosomes. *J. Am.
        Chem. Soc.* **2006**, *128* (35), 11433–11442.

(9)   Snijder, J.; Rose, R. J.; Veesler, D.; Johnson, J. E.; Heck, A. J. R. Studying 18 MDa Virus
        Assemblies with Native Mass Spectrometry. *Angew. Chemie - Int. Ed.* **2013**, *52* (14),
        4020–4023.

(10)   Smith, R. D.; Cheng, X.; Brace, J. E.; Hofstadler, S. A.; Anderson, G. A. Trapping,
        Detection and Reaction of Very Large Single Molecular Ions by Mass Spectrometry.
        Nature **1994**, *369* (6476), 137–139.

(11)   Bruce, J. E.; Cheng, X.; Bakhtiar, R.; Wu, Q.; Hofstadler, S. A.; Anderson, G. A.; Smith,
        R. D. Trapping, Detection, and Mass Measurement of Individual Ions in a Fourier
        Transform Ion Cyclotron Resonance Mass Spectrometer. *J. Am. Chem. Soc.* **1994**, *116*
        (17), 7839–7847.

(12)   Wuerker, R. F.; Shelton, H.; Langmuir, R. V. Electrodynamic Containment of Charged
        Particles. *J. Appl. Phys.* **1959**, *30* (3), 342–349.

(13)   Philip, M. A.; Gelbard, F.; Arnold, S. An Absolute Method for Aerosol Particle Mass and
        Charge Measurement. *J. Colloid Interface Sci.* **1983**, *91* (2), 507–515.

(14)   Hars, G.; Tass, Z. Application of Quadrupole Ion Trap for the Accurate Mass
        Determination of Submicron Size Charged Particles. *J. Appl. Phys.* **1995**, *77* (9), 4245–
        4250.

(15)   Schlemmer, S.; Illemann, J.; Wellert, S.; Gerlich, D. Nondestructive High-Resolution and
        Absolute Mass Determination of Single Charged Particles in a Three-Dimensional
        Quadrupole Trap. *J. Appl. Phys.* **2001**, *90* (10), 5410–5418.

(16)   Bell, D. M.; Howder, C. R.; Johnson, R. C.; Anderson, S. L. Single CdSe/ZnS Nanocrystals in an Ion Trap: Charge and Mass Determination and Photophysics Evolution with Changing Mass, Charge, and Temperature. *ACS Nano* **2014**, *8* (3), 2387–2398.

(17)   Kafader, J. O.; Melani, R. D.; Senko, M. W.; Makarov, A. A.; Kelleher, N. L.; Compton, P. D. Measurement of Individual Ions Sharply Increases the Resolution of Orbitrap Mass Spectra of Proteins. *Anal. Chem.* **2019**, *91* (4), 2776–2783.

(18)   Wörner, T. P.; Snijder, J.; Bennett, A.; Agbandje-McKenna, M.; Makarov, A. A.; Heck, A. J. R. Resolving Heterogeneous Macromolecular Assemblies by Orbitrap-Based Single-Particle Charge Detection Mass Spectrometry. *Nat. Methods* **2020**, *17* (4), 395–398.

(19)   Benner, W. H. A Gated Electrostatic Ion Trap to Repetitiously Measure the Charge and m/z of Large Electrospray Ions. *Anal. Chem.* **1997**, *69* (20), 4162–4168.

(20)   Doussineau, T.; Yu Bao, C.; Clavier, C.; Dagany, X.; Kerleroux, M.; Antoine, R.; Dugourd, P. Infrared Multiphoton Dissociation Tandem Charge Detection-Mass Spectrometry of Single Megadalton Electrosprayed Ions. *Rev. Sci. Instrum.* **2011**, *82* (8), 084104.

(21)   Contino, N. C.; Jarrold, M. F. Charge Detection Mass Spectrometry for Single Ions with a Limit of Detection of 30 Charges. *Int. J. Mass Spectrom.* **2013**, *345–347*, 153–159.

(22)   Elliott, A. G.; Merenbloom, S. I.; Chakrabarty, S.; Williams, E. R. Single Particle Analyzer of Mass: A Charge Detection Mass Spectrometer with a Multi-Detector Electrostatic Ion Trap. *Int. J. Mass Spectrom.* **2017**, *414*, 45–55.

(23)   Elliott, A. G.; Harper, C. C.; Lin, H. W.; Williams, E. R. Mass, Mobility and MS[N] Measurements of Single Ions Using Charge Detection Mass Spectrometry. *Analyst* **2017**, *142* (15), 2760–2769.

(24) Hogan, J. A.; Jarrold, M. F. Optimized Electrostatic Linear Ion Trap for Charge Detection Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (10), 2086–2095.

(25) Cheng, X.; Bakhtiar, R.; Van Orden, S.; Smith, R. D. Charge-State Shifting of Individual Multiply-Charged Ions of Bovine Albumin Dimer and Molecular Weight Determination Using an Individual-Ion Approach. *Anal. Chem.* **1994**, *66* (13), 2084–2087.

(26) Harper, C. C.; Elliott, A. G.; Lin, H. W.; Williams, E. R. Determining Energies and Cross Sections of Individual Ions Using Higher-Order Harmonics in Fourier Transform Charge Detection Mass Spectrometry (FT-CDMS). *J. Am. Soc. Mass Spectrom.* **2018**, *29* (9), 1861–1869.

(27) Elliott, A. G.; Harper, C. C.; Lin, H. W.; Williams, E. R. Effects of Individual Ion Energies on Charge Measurements in Fourier Transform Charge Detection Mass Spectrometry (FT-CDMS). *J. Am. Soc. Mass Spectrom.* **2019**, *30* (6), 946–955.

(28) Keifer, D. Z.; Shinholt, D. L.; Jarrold, M. F. Charge Detection Mass Spectrometry with Almost Perfect Charge Accuracy. *Anal. Chem.* **2015**, *87* (20), 10330–10337.

(29) Pierson, E. E.; Contino, N. C.; Keifer, D. Z.; Jarrold, M. F. Charge Detection Mass Spectrometry for Single Ions with an Uncertainty in the Charge Measurement of 0.65 E. *J. Am. Soc. Mass Spectrom.* **2015**, *26* (7), 1213–1220.

(30) Sipe, D. M.; Plath, L. D.; Aksenov, A. A.; Feldman, J. S.; Bier, M. E. Characterization of Mega-Dalton-Sized Nanoparticles by Superconducting Tunnel Junction Cryodetection Mass Spectrometry. *ACS Nano* **2018**, *12* (3), 2591–2602.

(31) Sage, E.; Brenac, A.; Alava, T.; Morel, R.; Dupré, C.; Hanay, M. S.; Roukes, M. L.; Duraffourg, L.; Masselon, C.; Hentz, S. Neutral Particle Mass Spectrometry with Nanomechanical Systems. *Nat. Commun.* **2015**, *6* (1), 1–5.

(32)   Erdogan, R. T.; Alkhaled, M.; Kaynak, B. E.; Alhmoud, H.; Pisheh, H. S.; Kelleci, M.;
Karakurt, I.; Yanik, C.; Şen, Z. B.; Sari, B.; Yagci, A. M.; Özkul, A.; Hanay, M. S.
Atmospheric Pressure Mass Spectrometry of Single Viruses and Nanoparticles by
Nanoelectromechanical Systems. *ACS Nano* **2022**, *16* (3), 3821–3833.

(33)   Todd, A. R.; Alexander, A. W.; Jarrold, M. F. Implementation of a Charge-Sensitive
Amplifier without a Feedback Resistor for Charge Detection Mass Spectrometry Reduces
Noise and Enables Detection of Individual Ions Carrying a Single Charge. *J. Am. Soc.*
Mass Spectrom. **2020**, *31* (1), 146–154.

(34)   Harper, C. C.; Williams, E. R. Enhanced Multiplexing in Fourier Transform Charge
Detection Mass Spectrometry by Decoupling Ion Frequency from Mass to Charge Ratio.
J. Am. Soc. Mass Spectrom. **2019**, *30* (12), 2637–2645.

(35)   Elliott, A. G.; Harper, C. C.; Lin, H. W.; Susa, A. C.; Xia, Z.; Williams, E. R.
Simultaneous Measurements of Mass and Collisional Cross-Section of Single Ions with
Charge Detection Mass Spectrometry. *Anal. Chem.* **2017**, *89* (14), 7701–7708.

(36)   Blakney, G. T.; Hendrickson, C. L.; Marshall, A. G. Predator Data Station: A Fast Data
Acquisition System for Advanced FT-ICR MS Experiments. *Int. J. Mass Spectrom.* **2011**,
306 (2–3), 246–252.

(37)   Draper, B. E.; Jarrold, M. F. Real-Time Analysis and Signal Optimization for Charge
Detection Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (6), 898–904.

(38)   Harper, C. C.; Elliott, A. G.; Oltrogge, L. M.; Savage, D. F.; Williams, E. R. Multiplexed
Charge Detection Mass Spectrometry for High-Throughput Single Ion Analysis of Large
Molecules. *Anal. Chem.* **2019**, *91* (11), 7458–7468.

(39)   Mendell, J. R.; Al-Zaidy, S. A.; Rodino-Klapac, L. R.; Goodspeed, K.; Gray, S. J.; Kay,

C. N.; Boye, S. L.; Boye, S. E.; George, L. A.; Salabarria, S.; Corti, M.; Byrne, B. J.; Tremblay, J. P. Current Clinical Applications of In Vivo Gene Therapy with AAVs. *Mol. Ther.* **2021**, *29* (2), 464–488.

(40) Kotterman, M. A.; Chalberg, T. W.; Schaffer, D. V. Viral Vectors for Gene Therapy: Translational and Clinical Outlook. *Annu. Rev. Biomed. Eng.* **2015**, *17*, 63–89.

(41) Delor, M.; Dai, J.; Roberts, T. D.; Rogers, J. R.; Hamed, S. M.; Neaton, J. B.; Geissler, P. L.; Francis, M. B.; Ginsberg, N. S. Exploiting Chromophore-Protein Interactions through Linker Engineering to Tune Photoinduced Dynamics in a Biomimetic Light-Harvesting Platform. *J. Am. Chem. Soc.* **2018**, *140* (20), 6278–6287.

(42) Dedeo, M. T.; Duderstadt, K. E.; Berger, J. M.; Francis, M. B. Nanoscale Protein Assemblies from a Circular Permutant of the Tobacco Mosaic Virus. *Nano Lett.* **2010**, *10* (1), 181–186.

(43) Ramsey, A. V.; Bischoff, A. J.; Francis, M. B. Enzyme Activated Gold Nanoparticles for Versatile Site-Selective Bioconjugation. *J. Am. Chem. Soc.* **2021**, *143* (19), 7342–7350.

(44) Dias Florencio, G.; Precigout, G.; Beley, C.; Buclez, P. O.; Garcia, L.; Benchaouir, R. Simple Downstream Process Based on Detergent Treatment Improves Yield and in Vivo Transduction Efficacy of Adeno-Associated Virus Vectors. *Mol. Ther. - Methods Clin. Dev.* **2015**, *2*, 15024.

(45) Barnes, C. R.; Lee, H.; Ojala, D. S.; Lewis, K. K.; Limsirichai, P.; Schaffer, D. V. Genome-Wide Activation Screens to Increase Adeno-Associated Virus Production. *Mol. Ther. - Nucleic Acids* **2021**, *26* (S2), 94–103.

(46) Harper, C. C.; Brauer, D. D.; Francis, M. B.; Williams, E. R. Direct Observation of Ion Emission from Charged Aqueous Nanodrops: Effects on Gaseous Macromolecular

Charging. *Chem. Sci.* **2021**, *12* (14), 5185–5195.

(47) Keifer, D. Z.; Alexander, A. W.; Jarrold, M. F. Spontaneous Mass and Charge Losses from Single Multi-Megadalton Ions Studied by Charge Detection Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (3), 498–506.

(48) Leach, F. E.; Kharchenko, A.; Vladimirov, G.; Aizikov, K.; O'Connor, P. B.; Nikolaev, E.; Heeren, R. M. A.; Amster, I. J. Analysis of Phase Dependent Frequency Shifts in Simulated FTMS Transients Using the Filter Diagonalization Method. *Int. J. Mass Spectrom.* **2012**, *325–327*, 19–24.

(49) Harris, F. J. On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform. *Proc. IEEE* **1978**, *66* (1), 51–83.

(50) Cooley, J. W.; Tukey, J. W. An Algorithm for the Machine Calculation of Complex Fourier Series. *Math. Comput.* **1965**, *19* (90), 297.
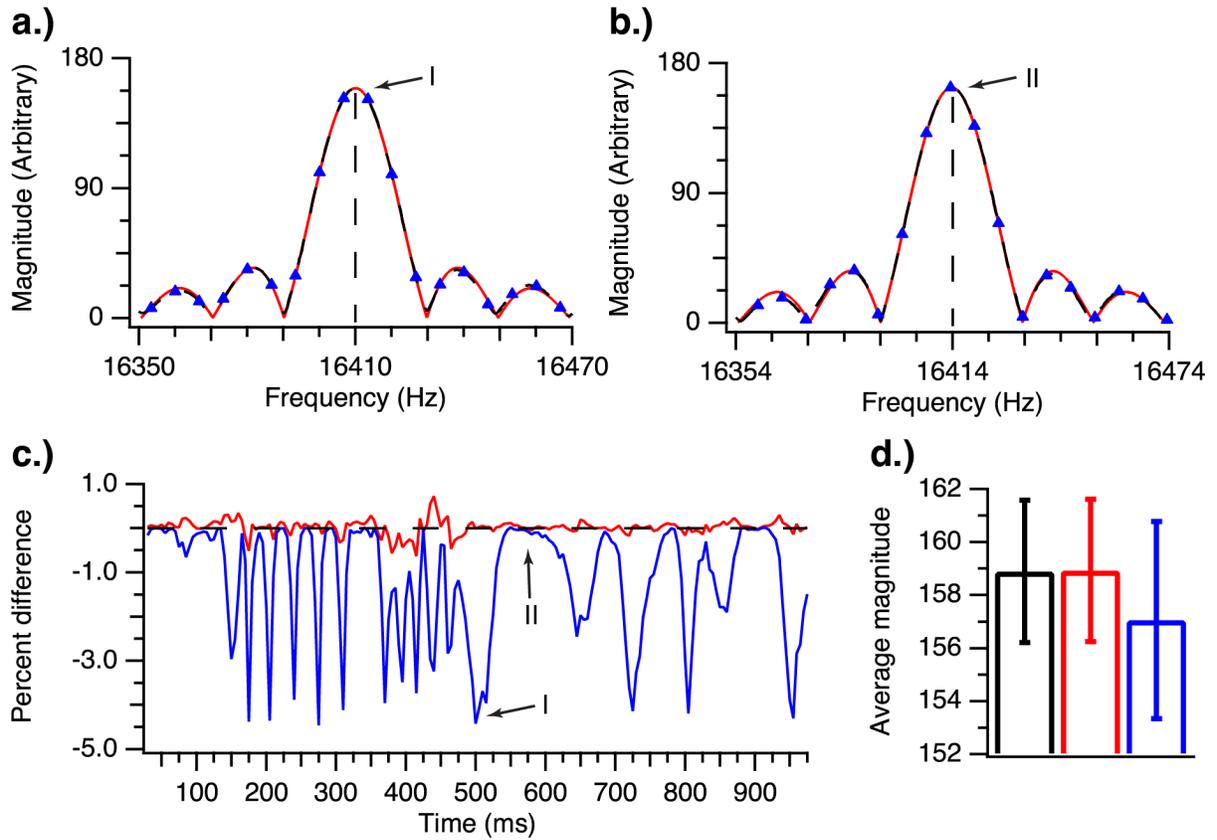
Figure 1: Comparison of peak picking and sinc fitting for a single AAV9 ion using 50 ms of time-domain data and two zero-fills (150 ms of time-domain data) compared to 39 zero-fills (2000 ms of time domain data); time-domain data starting at a) 500 ms shows significant scalloping loss with peak picking (blue points) compared to sinc fitting function (red line) and zero-filled to 2000 ms (black dashed line) whereas at b) 570 ms, scalloping loss is minimal; c) deviations in peak amplitude from peak picking (blue line) and sinc fitting (red line) compared to data that was zerofilled to 2000 ms (black dashed line) over the course of the 1 s trapping period; d) average peak amplitude and standard deviation of a single AAV9 ion obtained using peak picking and 39 zero-fills (black), sinc fitting and two zero-fills (red), and peak picking and two zero-fills (blue).
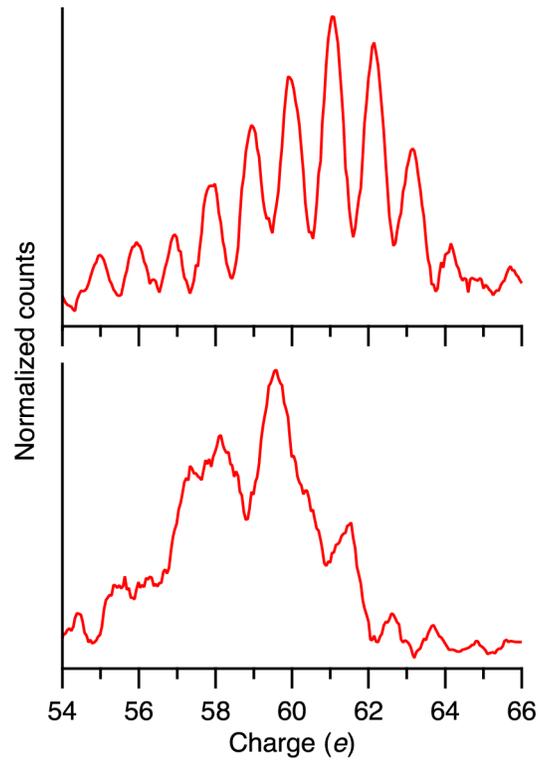
Figure 2: Charge histograms showing normalized counts as a function of peak amplitude (charge) for ~400 cpTMV-S65-3NY ions using sinc fitting (top) and peak picking (bottom) showing improved charge state resolution with sinc fitting. Data bin widths are 0.05 e and are smoothed using a Savitsky-Golay filter.
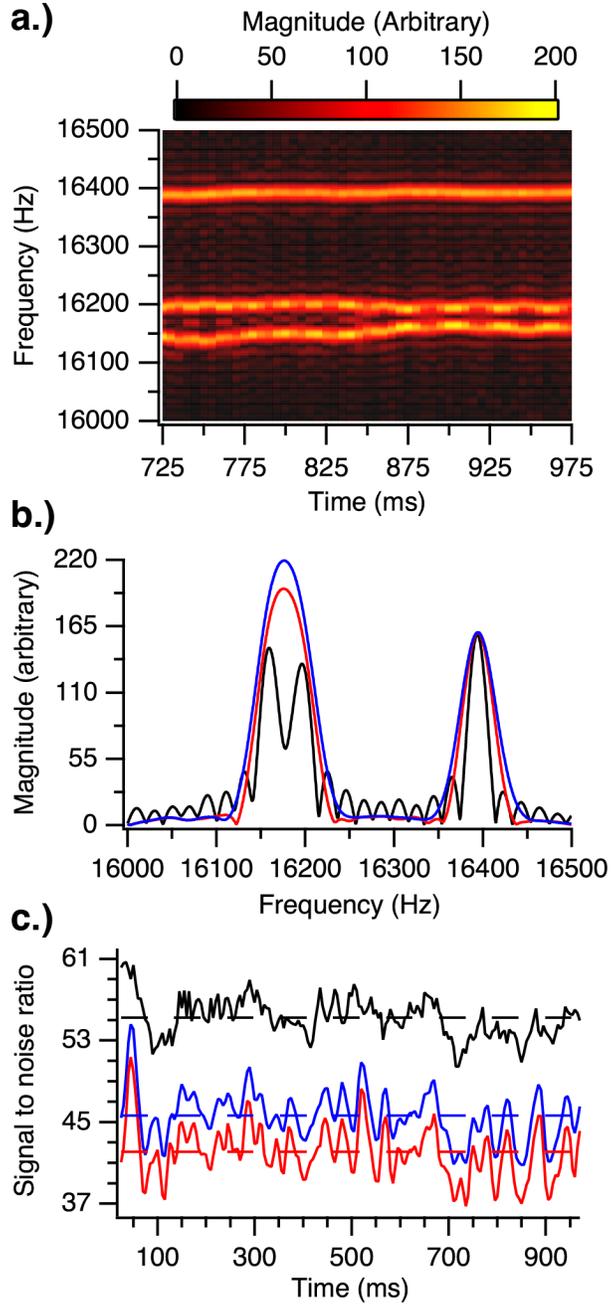
Figure 3: a) STFT data between 725 and 975 ms with rectangular apodization showing

frequency vs. time for three AAV9 ions that have similar oscillation frequencies, b) FT data in a

window centered at 885 ms (a single "slice" of the STFT) obtained using a rectangular (black

line), Hann (red line) and Gaussian (blue) apodization function showing that two ions not

resolved by the latter two functions are clearly resolved with rectangular apodization and c) the

S/N of a single AAV9 ion over each "slice" of the STFT for the 1 s trapping period for the three apodization functions.
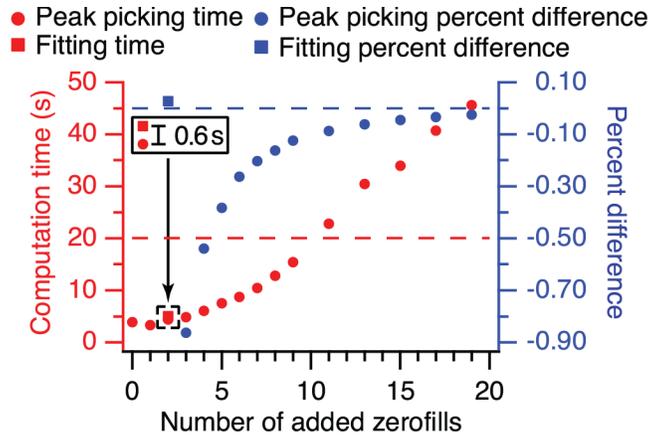
Figure 4: Computation time (left *y*-axis) required to obtain the frequencies and signal amplitudes (charge) of 20 one second trapping events for 56 total ions as a function of number of added zero-fills (*x*-axis) for peak picking (red circles) and sinc fitting (red square). The red dashed line indicates 20 seconds of computation time and represents the threshold for real time analysis. The average percent difference in peak amplitudes (right *y*-axis) for peak picking (blue circles) and sinc fitting (blue square) compared to amplitudes obtained using 39 zero-fills (2000 ms time-domain data length). The blue dashed line indicates zero percent deviation.