ELSEVIER

Journal of Neuroscience Methods 155 (2006) 81–91

JOURNAL OF NEUROSCIENCE METHODS

www.elsevier.com/locate/jneumeth

Nyquist interpolation improves neuron yield in multiunit recordings

Timothy J. Blanche *, Nicholas V. Swindale

Department of Ophthalmology and Visual Sciences, University of British Columbia, Canada

Received 27 October 2005; received in revised form 6 December 2005; accepted 22 December 2005

Abstract

Multiunit electrodes, in particular tetrodes and polytrodes, are able to isolate action potentials from many neurons simultaneously. However, inaccuracies in the post-acquisition reconstruction of recorded spike waveforms can affect the reliability of spike detection and sorting. Here we show that bandlimited interpolation with sample-and-hold delay correction reduces waveform variability, leading to improved reliability of threshold-based event detection and improved spike sorting accuracy. Interpolation of continuously acquired data is, however, computationally expensive. A cost-benefit analysis was made of varying sampling rates from 12.5 kHz (no interpolation) to 100 kHz (eight times oversampling, with respect to the Nyquist frequency), taking into consideration the final application of the data. For most purposes, including spike sorting, sample rates below 25 kHz with bandlimited interpolation to 50 kHz were ideal, with negligible gains above this rate. A practical benefit, especially for large electrode arrays, is that the bandwidth and storage requirements can be greatly reduced by using data acquisition rates at or slightly above the Nyquist frequency.
© 2006 Published by Elsevier B.V.

Keywords: Extracellular single-unit recording; Signal processing; Waveform reconstruction; Interpolation; Nyquist criteria; Sampling theorem; Clustering; Tutorial

1. Introduction

Extracellular recordings of neuronal spike activity from single microelectrodes have conventionally been made by passing raw voltage waveforms through a window discriminator and storing time-stamps of triggering events that are assumed to correspond to action potentials of a single neuron. In recordings where only one active neuron lies close to the electrode, confidence about this identification justifies discarding information about the precise shape of the waveform underlying the triggering event, since, for most purposes, knowledge of spike timing is all that is required in later analyses. The advent of multisite, multiunit electrodes such as tetrodes (Gray et al., 1995), multiwire electrode arrays (Nicolelis et al., 2003), and silicon-based polytrodes (Bai et al., 2000; Blanche et al., 2005; Campbell et al., 1991; Csicsvari et al., 2003) has made it possible to record from many neurons simultaneously, but here the identification of spike times from individual neurons depends critically on subtle variations in the amplitude and shape of the waveforms recorded at adjacent electrode sites. Although spike identification with multiunit electrodes can sometimes be done in real time with specialised hardware (Gadicke and Albus, 1995) it is more desirable to store a continuous digital record of the voltage waveform from each of the electrode sites and perform spike identification off-line, after the experiment is completed. This approach requires choices to be made about (i) how to pre-process, i.e. filter, the voltage waveforms prior to analog to digital conversion (ADC), (ii) an appropriate sampling rate, and (iii) how to reconstruct the waveforms from the digital record prior to analysis. These topics are addressed in this paper.

According to sampling theory (Nyquist, 1928; Shannon, 1949; Whittaker, 1935), for signals low-pass filtered with an upper cut-off frequency $W = 5$ kHz, a sample rate of 10 kHz ($2W$, twice the Nyquist frequency) should be sufficient for accurate reconstruction of the original waveform. However, neurophysiologists routinely record spike activity at sample rates between 20 and 40 kHz. Waveforms sampled according to the Nyquist criterion do not appear to provide an accurate representation of the original spike waveforms (Fig. 1), which may explain the common practice of oversampling. Nevertheless, the Shannon–Nyquist theorem states explicitly that if a bandlimited signal $s(t)$ is sampled at a rate $2W$ then $s(t)$ is completely determined, and can be recovered from its samples, s_n, by discrete

* Corresponding author at: Hanse-Wissenschaftskolleg, Brain Research Institute, University of Bremen, Hochschulring 16a, P.O. Box 33 04 40, D-28334 Bremen, Germany. Tel.: +49 421 9160 217

E-mail address: jnm@timblanche.mm.st (T.J. Blanche).

0165-0270/$ – see front matter © 2006 Published by Elsevier B.V.
doi:10.1016/j.jneumeth.2005.12.031

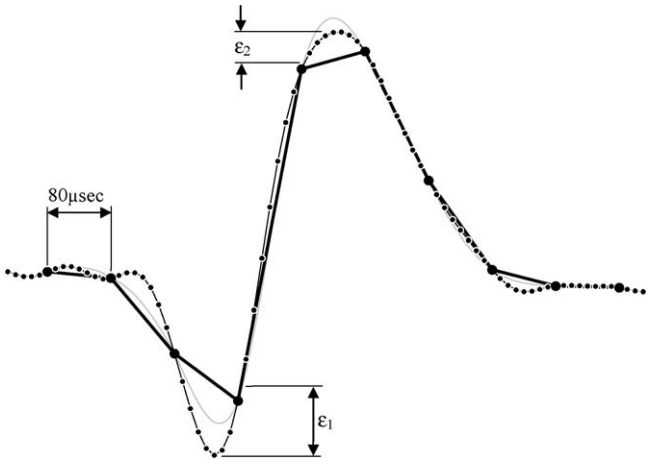

Fig. 1. The case for upsampling. A linearly interpolated spike (~240 μV_{pp} amplitude, thick line) sampled at the Nyquist rate (12.5 kHz, large dots), provides a poor rendition of the original waveform sampled at 100 kHz (fine black line), particularly at the spike onset, peak and valley. Spike amplitudes will be consistently underestimated, and threshold-based spike detectors set in the ranges ε_1 and ε_2 will miss a percentage of spike events due to asynchronous sampling (i.e. spikes occur randomly with respect to the periodic sampling). After interpolation to 100 kHz (small dots) an accurate reconstruction of the underlying waveform is obtained. The cubic spline interpolated waveform (fine grey line) is shown for comparison.

convolution with an appropriate reconstruction function (usually a sinc function, see Section 2).

The Nyquist sampling rule is often misconstrued as implying that the raw samples constitute the original signal and do not require further processing. What is usually neglected is the requirement for reconstruction of the voltage waveform using bandlimited interpolation, otherwise known as 'upsampling'. Although this method of signal reconstruction is standard, it is rarely exploited by neurophysiologists (two exceptions include Chandra and Optican (1997) and Pouzat et al. (2002)). In this paper we show that inadequate reconstruction of spike waveforms can result in errors of spike detection, distortion of spike shapes, and compromise the performance of spike sorting algorithms. We also show the need for correction of sample-and-hold delays (SHDs) caused by the serial sampling of channels that is implemented by most multichannel data acquisition boards. These delays, though small, can significantly skew the temporal alignment of multisite recordings made with polytrodes. Upsampling provides tangible benefits with respect to the reliability of spike detection, estimation of spike waveform parameters, accuracy of spike sorting, and ultimately single-unit yield in multiunit recordings.

2. Materials and methods

2.1. Bandlimited interpolation

An analog signal $s(t)$ is bandlimited in accordance with the Nyquist criterion if its Fourier transform:

$$\Im[s(t)] = S(f) = 0 \quad \text{for } |f| > \frac{W}{2} \tag{1}$$

where W is the sampling rate. $W/2$ is commonly referred to as the Nyquist frequency, or Nyquist limit.

Reconstruction of bandlimited digitized signals in the time domain is done by discrete convolution with a sinc function:

$$s(t) = \sum_{n=-\infty}^{n=\infty} s(n)h(t-n) \tag{2}$$

where $s(t)$ is the reconstructed waveform voltage at time t, $s(n)$ is the raw voltage waveform sampled at time intervals $n = 1, 2, 3\ldots$ (for convenience time units are chosen such that the sampling interval $\Delta t = 1$), and $h(t)$ is the sinc reconstruction function:

$$h(t) = \frac{\sin \pi t}{\pi t} \tag{3}$$

The sinc (meaning "cardinal sine") function (Fig. 2B) is the theoretically perfect reconstruction filter for bandlimited signals, since its Fourier transform is a box centered around DC of unitary width (Fig. 2A). The period of the sinc function is twice the sampling period, with zero crossings aligned with the original sample points. It follows that convolution with a zero phase ($t = 0$) sinc kernel leaves these samples unchanged (Fig. 2C, open triangles). The signal at arbitrary time, t, is evaluated by the coefficients of time-shifted sinc kernels (e.g. $t = 0.5$, filled triangles in Fig. 2C).

The sinc is an infinite impulse response (IIR), as any function that has a finite extent in the frequency domain must have an infinite extent in the time domain (and vice versa). The sinc keeps oscillating with ever decreasing amplitude, so for practical purposes it needs to be windowed. Windowing attenuates the sinc function to zero at the extremes, thereby producing a finite impulse response (FIR) filter kernel (Fig. 2C). Tapering the sinc function instead of simply truncating it minimizes bandpass ripple (Gibbs' phenomenon). There are many suitable window functions (Horowitz and Hill, 1989; Smith, 1997, 2004): here a simple Hamming window was used.

Ideally, sinc interpolation should be done with sufficient resolution that interpolating linearly between upsampled points does not introduce error greater than the quantization error of the ADC (Smith, 2004). In practice this is computationally infeasible. For data recorded with a 12-bit ADC (see below), a filter kernel of total length $N = 12$ (i.e. with six zero crossings on either size of the origin at $t = 0$, requiring 13 data points to represent it) was found to provide adequate precision. Combining Eq. (2) with a Hamming window, w_h gives a tapered FIR filter suitable for discrete time series convolution:

$$s(t) = \sum_{n=-6}^{n=6} s(n)w_h(t-n)h(t-n) \tag{4}$$

where

$$w_h(t) = 0.54 - 0.46\cos\pi\left(\frac{2t+N}{N}\right) \tag{5}$$

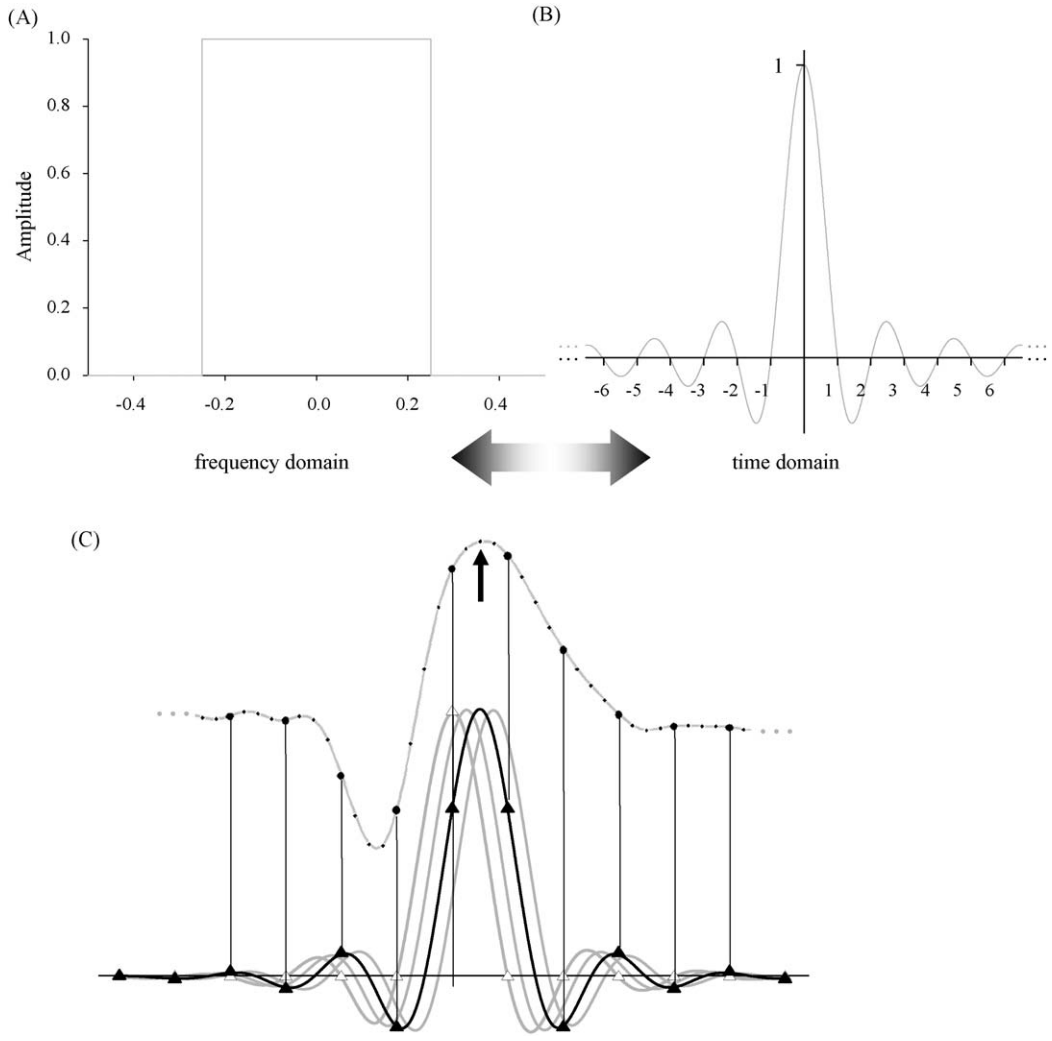

Fig. 2. Bandlimited interpolation using sinc convolution. (A) Frequency response of the ideal low pass filter. Signals below the Nyquist frequency pass unchanged, and those above are excluded. (B) The impulse response of the ideal filter is a sinc function that extends infinitely in time, so for practical purposes it needs to be windowed. (C) Convolution of a bandlimited digitized signal with time-shifted windowed sinc functions can accurately reconstruct the original analog waveform. Depicted are the four 13-point sinc kernels used to upsample a spike shown from 12.5 kHz (large circles) to 50 kHz (small circles). For clarity the filter coefficients are only shown for the zero (open triangles) and 180° (filled triangles) phase kernels. In this example, the interpolant at the peak of the spike (arrow) derives from the 13 surrounding raw sample points convolved with the 180° phase kernel (see text for details).

2.2. Sample-and-hold delay correction

Most digital acquisition cards have a single analog to digital converter and therefore cannot sample multiple channels simultaneously. The usual solution is to combine sample-and-hold circuits with rapid signal multiplexing (so-called "burst-mode" acquisition) to obtain near-simultaneous sampling across all channels. Resultant SHD artefacts can be corrected during the interpolation process by using appropriately phase-shifted sinc kernels, without additional computational overhead:

$$s_i(t) = \sum_{n=-6}^{n=6} s(n) w_h h(t - n + di) \quad (6)$$

where d is the channel–channel SHD interval and i is the ordinal position of the channel in the hold queue. For the cards used in this study operating in 1 MHz burst mode $d = 1$ μs. In more general terms, the channel-specific phase offset is proportional to its ordinal position in the hold queue, typically a small fraction of the sampling period.

Appropriately phase-shifted and windowed sinc kernels, for each combination of interpolant and SHD phase offset, were precalculated and stored in indexed arrays to speed the calculation of Eq. (6).

2.3. Instrumentation and test waveforms

The potential benefits and optimal rate of upsampling for neuronal spike recordings were addressed empirically. Continuously sampled neural activity was recorded with a polytrode – a 54 channel silicon electrode array with closely spaced (46–75 μm) recording sites – inserted into the primary visual

cortex of an anesthetized cat. Full details of the experimental setup are given in Blanche et al. (2005). Signals were amplified 5000× and bandpass filtered (500–6 kHz) with a multichannel amplifier (FA-I 64, Multichannel Systems, Germany), before digital sampling with 12 bit resolution by two 32-channel ADC cards (DT3010s, Data Translation, MA) at 100 kHz/channel. The noise of the system was 3–4 μV_{rms} (~20–30 μV_{pp}), and while some spikes were as large as 1.2 mV_{pp}, more typically they were in the range 100–200 μV_{pp} (Blanche et al., 2005).

Since the original test data were oversampled at 100 kHz, lower sampling rates were obtained by sub-sampling. Improvements in waveform reconstruction, spike detection and sorting were evaluated for Nyquist sampled and two-, four- and eight-times upsampled data, corresponding to effective sampling rates of 12.5, 25, 50 and 100 kHz, respectively.

3. Results

3.1. Ideal bandlimited interpolation

A typical extracellular neural recording made with our system had a power spectral density proportional to $\sim 1/f^2$ beyond a broad peak at ~1.3 kHz (Fig. 3A). Most (92.5%) of the energy was below 6.25 kHz, the Nyquist frequency for data sampled at 12.5 kHz. Only 6.6% of the energy was between 6.25 and 12.5 kHz, and frequencies above 12.5 kHz contributed less than 1% of the total signal energy. Since there was some frequency content above the nominal 6 kHz corner frequency (−3 dB) of the amplifier's anti-alias filter, the "cut-off" frequency can clearly not be equated with the Nyquist frequency (Eq. (1)) and used to determine an adequate sampling rate.

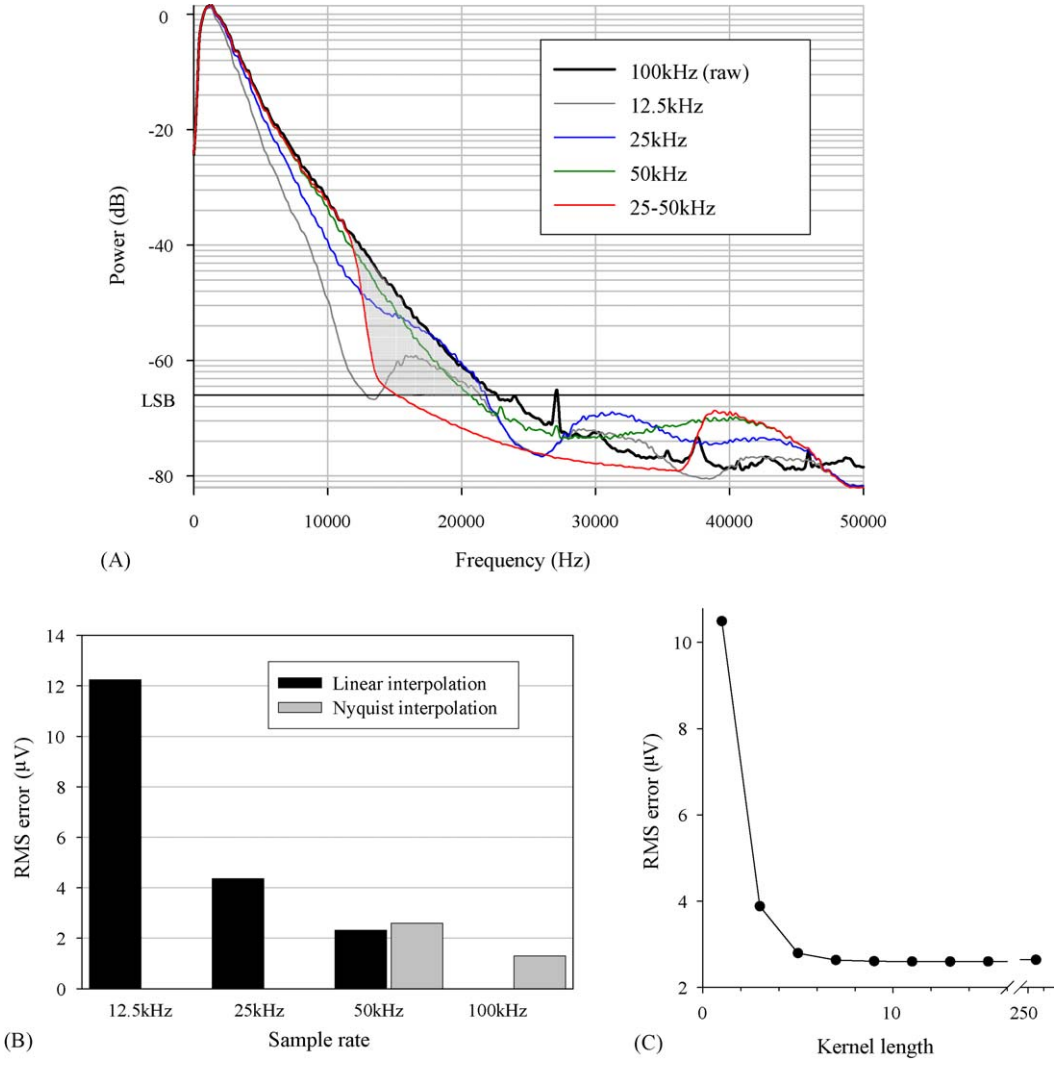

Fig. 3. Signal fidelity at different sample rates. (A) Power spectrum of about 2 min worth of extracellular recording, containing a representative number of spikes, sampled at 100 kHz (solid black line) compared with 12.5, 25, and 50 kHz sampling rates (sub-samples of the 100 kHz record with linear interpolation). Unlike linearly interpolated 25 kHz data (blue line), Nyquist interpolation from 25 to 50 kHz (red line) faithfully restores the frequency content up to 12.5 kHz, the bandpass of interest for spikes. The least significant bit (LSB) line demarcates the precision of the acquisition system (~11.5 usable bits). (B) Interpolation error for spike waveforms is shown as a function of effective sampling rate, for decimated data (black bars) and data upsampled from 25 kHz (grey bars). The residual error at 50 and 100 kHz can be attributed to energy in the transition pass of the anti-alias filter between 12.5 and 50 kHz (shaded region in A). (C) Residual spike waveform error when upsampling from 25 to 50 kHz, as a function of sinc kernel length *N*.

The issue of signal aliasing is considered further in Section 4. For now the key observation is that the frequency content of data sampled at 12.5, 25, or 50 kHz and reconstructed by linear interpolation between samples deviates significantly from the 100 kHz spectrum, whereas data sampled at 25 kHz and upsampled by sinc interpolation to 50 kHz captured the original spectrum almost perfectly up to 12.5 kHz (Fig. 3A).

Notice the strong attenuation of signals above 12.5 kHz (the Nyquist frequency for data upsampled from 25 kHz). Since most neurophysiological signals of interest have frequencies below 12.5 kHz (indeed, all but the fastest spikes have 98% of their energy below 6.25 kHz, see Section 4), the removal of high frequency noise that may exist in the transition bandpass of oversampled raw signals is another side benefit of bandlimited interpolation.

3.2. Spike waveform reconstruction

The root-mean-squared (RMS) error caused by linear interpolation of spike waveforms sampled at 12.5 kHz was three to four times the aggregate noise of the recording system (Fig. 3B). This error decreased rapidly as the original data were progressively oversampled, but the same result was achieved for data upsampled from 25 to 50 or 100 kHz. The small residual error at 50 and 100 kHz is attributable in part to the aforementioned high frequency noise between 12.5 and 50 kHz, and also because linear interpolation of upsampled data is still an approximation of the original waveform. This level of precision was attained with ≥9 point sinc kernels (Fig. 3C), and no further improvements were evident above 13 points (i.e. six zero crossings).

Interpolation using windowed sinc convolution accurately reconstructed spike waveforms (Fig. 4A and C) whereas linear

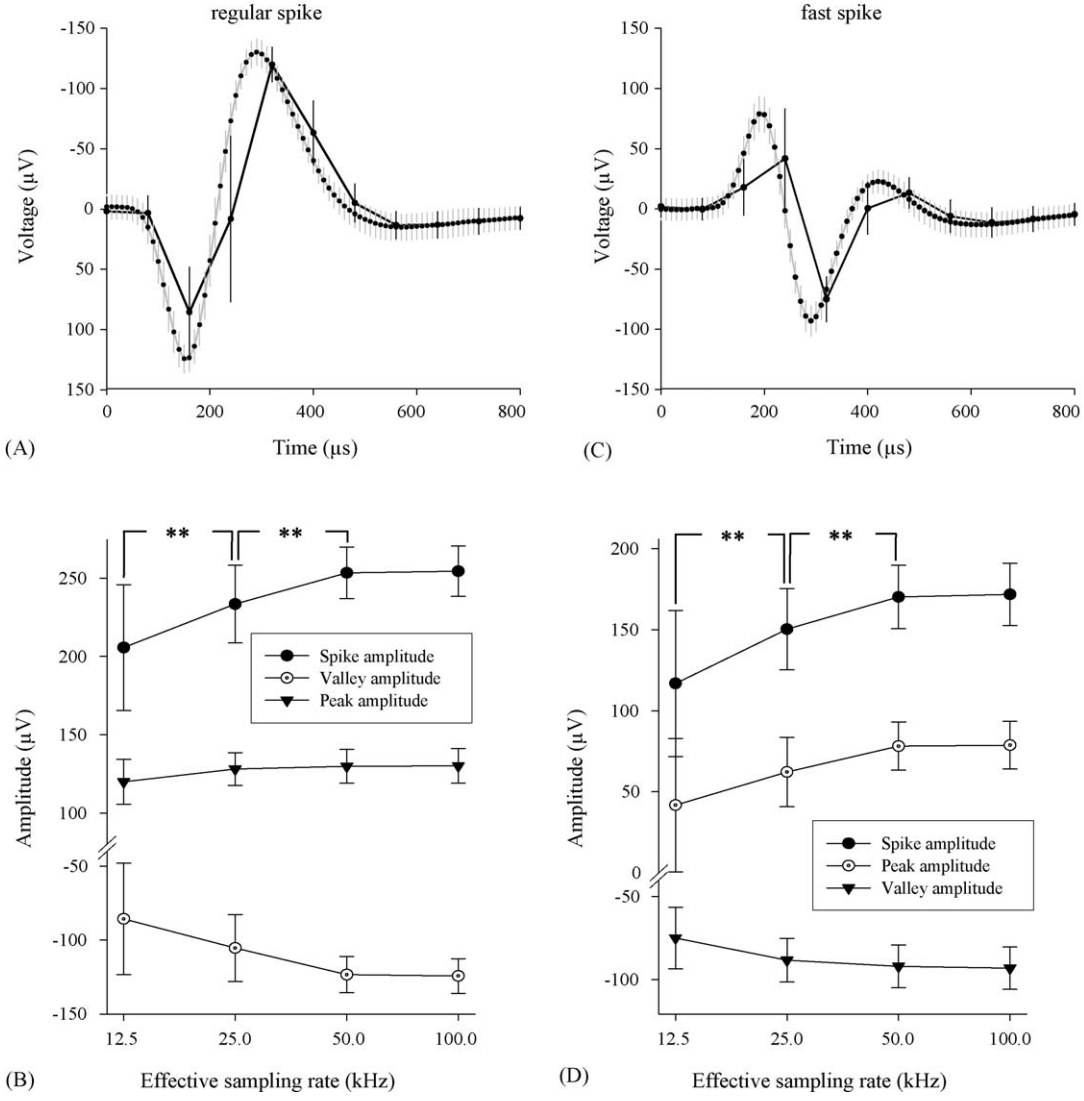

Fig. 4. Interpolation improves spike waveform reconstruction. Sinc interpolation of Nyquist-sampled data faithfully reconstructs spike waveforms. (A) An averaged ($n = 313$) medium amplitude spike was poorly represented at 12.5 kHz (dots), whereas upsampling to 100 kHz (small dots) accurately matched the underlying spike shape (grey line, acquired at 100 kHz). (B) Without upsampling, waveform-derived spike parameters showed increased variance and underestimation of peak and valley amplitudes (nb: spike amplitude = peak − valley amplitude). (C) A fast-spiking neuron ($n = 1640$) had even greater waveform distortion and variability without adequate interpolation. Consequently (D) spike amplitudes were miscalculated. Error bars are the mean ± S.D. $^{**}p < 0.00001$.

interpolation gave a poor rendition of the underlying spike shape. Accordingly, estimates of spike amplitude (Fig. 4B and D) were consistently low. For example, the regular spiking neuron in Fig. 4 had an average estimated spike amplitude of 172 μVpp at 100 kHz, but only 117 μVpp at 12.5 kHz, 68% of the true amplitude. Averaging across many spikes is not a solution, because the asynchronous nature of the sample times with respect to the spike occurences gives rise to the variability shown in Fig. 4A and C. Without interpolation, estimates of spike width based on the average waveform were relatively accurate (108 μs at 12.5 kHz versus 96 μs at 100 kHz for the fast-spiking neuron), however, spike–spike variability caused by temporal aliasing (80 μs at 12.5 kHz) was unacceptably high and manifest as increased noise in the *y*-axis values. Spike width standard deviation was proportional to roughly half the sampling period (i.e. 44.8 μs at 12.5 kHz compared with just 5.3 μs at 100 kHz). Therefore, even if spike-triggered averages are used for the generation of spike sorting templates, or to derive waveform metrics, spike amplitude and shape variability is still affected by the temporal misalignment of individual spike events.

Sinc interpolation increased the effective sampling rate, with concomitant improvement in waveform fidelity. There was a gradual improvement in spike shape with increased sampling rate, most noticeably for fast spiking neurons and in regions of the waveform with high slew rates, such as the depolarisation phase (Fig. 4A and C). Slower regions of the spike waveform, including the repolarisation and after-hyperpolarisation phases, were in a sense already oversampled because they were comprised of lower frequencies. The high variability and poor accuracy of linearly interpolated Nyquist-sampled data is not only undesirable for spike sorting (see below) but is also inadequate for other applications, such as modeling of neuronal field potentials and classification of cell type based on spike amplitude and width (Bartho et al., 2004; Blanche et al., 2003).

3.3. Reliability of event detection

Variability in the amplitude of spike peaks caused by asynchronous digital sampling can result in missed spike events (threshold ranges ε_1 and ε_2 in Fig. 1). The magnitude of this error was quantified by counting the number of spikes from a single neuron interpolated to different effective sampling rates over a range of trigger thresholds. In the first of two representative examples (Fig. 5A) a large-amplitude neuron fired 303 spikes, all of which were detected at a threshold level of +250 μV or lower, irrespective of the sample rate. As the threshold was increased, the number of triggers decreased, so that by +450 μV no spikes were detected. The relationship between the number of detected events and the threshold level was well described by the sigmoidal function $y = a/(1 + e^{(x-x_0)/b})$ where a is the maximum height of the function, x_0 the position at half-height, and b is the inverse of the slope (i.e. large values of b mean a small slope). Lateral shifts in the sigmoid result from changes in the mean spike amplitude at different sampling rates. Broader distributions of spike amplitude produce shallower slopes, and vice versa. Factors affecting spike amplitude include intrinsic variability of the action potential, biological and recording system noise, and the discrete, asynchronous nature of digital sampling. Changes in the slope of the sigmoid (and hence the range over which the same spikes were detected) at different sampling rates indicate differences in spike amplitude variance attributable entirely to sampling. For the large amplitude spike shown in Fig. 5A the shallower slope at 12.5 kHz ($b = 23$) compared with that for 25 and 50 kHz ($b = 16$) reflects a small increase in sampling related spike amplitude variability. However, for this unit there was little benefit of higher sampling rates in terms of detection reliability, because simply lowering the threshold to +250 μV gave perfect detection without an increase in spurious triggers from noise or other spike events. This was not the case in the second example (Fig. 5B), where the broader spread

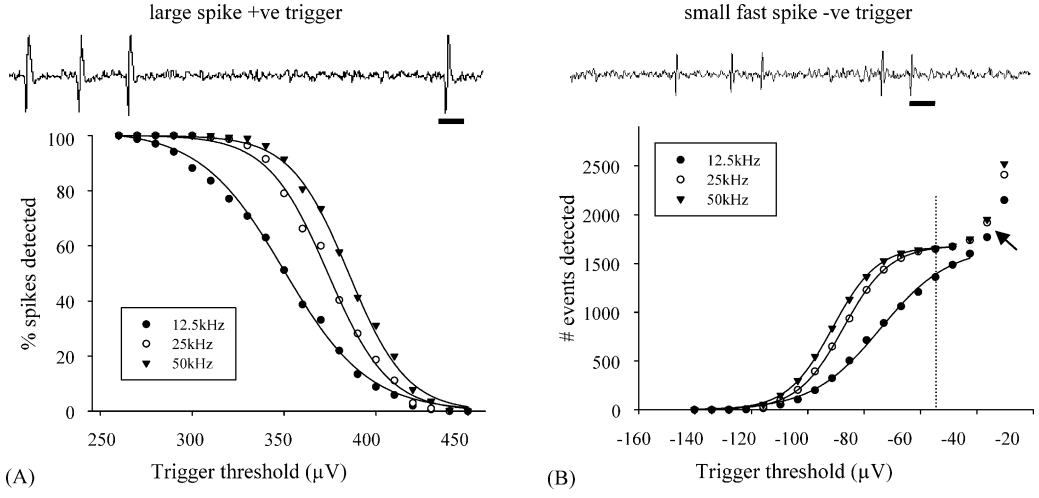

Fig. 5. Interpolation improves spike detection reliability. Different trigger thresholds were required to detect the same spike events from a given neuron at different sampling rates. (A) For large amplitude solitary spikes, upsampling afforded no benefit in terms of detection reliability. All spikes from this neuron could be distinguished from other spikes (and the background noise) at any sample rate simply by adjusting the threshold. (B) Fast spiking, lower amplitude spikes, however, were often only reliably detected after upsampling to 25 kHz or more. At the ideal threshold for higher sampling rates (−55 μV, dotted line) many spikes were missed at 12.5 kHz, but lowering the absolute value of the threshold dramatically increased the number of false triggers (arrow). In both examples the fitted curves are sigmoidal functions ($r^2 \geq 0.99$). As the 100 kHz plots were indistinguishable from those at 50 kHz, they were removed for clarity. Scalebars = 2 ms.

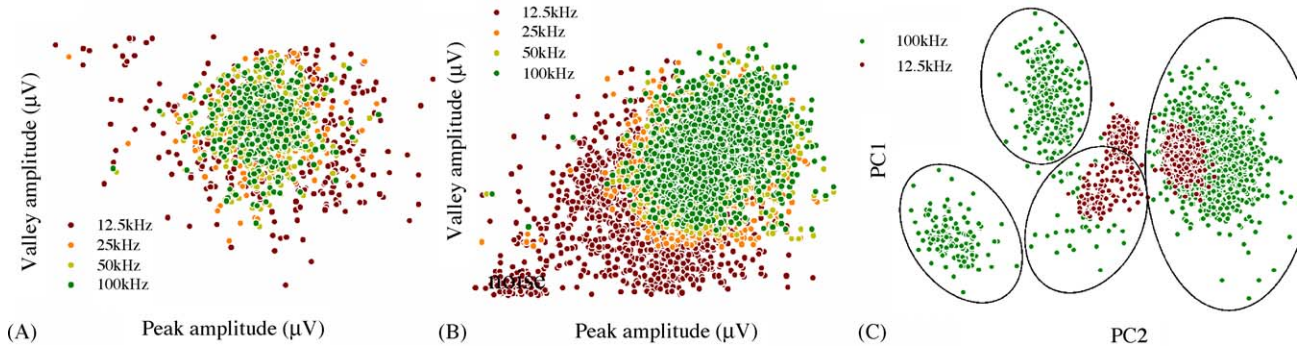

Fig. 6. Interpolation reduces cluster variability. Regardless of the features used for spike sorting, higher sampling rates yielded lower cluster variability. (A) Spike peak vs. valley amplitudes from a single neuron ($n = 316$), mean-subtracted to show the cluster spreads at each of the four sampling rates tested. (B) Uncentred spike amplitude clusters from another neuron ($n = 1253$). Underestimation of spike amplitudes produced the drift in cluster centres at low sample rates. High cluster variance at 12.5 kHz, made separation of spikes from noise very difficult. (C) PCA-based cluster plots of spikes upsampled to 100 kHz produced three neuron clusters (small ellipses), each distinct from each other and the noise cluster (large ellipse). In contrast, clusters derived from the same spike events sampled at the Nyquist rate were mixed, and less well separated from the noise cluster.

of amplitudes at 12.5 kHz ($b = 12$ versus 8 at 25 kHz) meant that all spikes from this neuron ($n = 1650$) could only be delineated from the background noise after interpolation. Even at the most favourable threshold (Fig. 5B, dotted line), just 80% of the Nyquist-sampled spike events were detectable, and raising the threshold produced an exponential increase in the number of false positives triggered by other small amplitude neurons and noise. For all spikes examined ($n = 15$), the proportionally largest improvement in detection reliability was from 12.5 to 25 kHz, with no further gains beyond 50 kHz.

3.4. Cluster variability

In order for spike clustering methods to effectively isolate single units, clusters from different neurons need to have minimal or no overlap with each other and the noise cluster. One of the consequences of sampling related variability is increased cluster scatter, which, as before, can be minimised with Nyquist interpolation (Fig. 6). Two-, four- and eight-times upsampling each approximately halved the standard deviation of spike amplitude clusters compared to those generated from Nyquist sampled data (e.g. 28–15 μV for the spikes shown in Fig. 6A, 36–21 μV for Fig. 6B). For spike sorting based on dimension-reduced waveform features, for example principal component analysis (PCA, Abeles and Goldstein, 1977), Nyquist interpolation with precisely aligned spikes produced clusters that were delineated from each other and the noise (Fig. 6C). In this example, average cluster separation measures based on the Mahalanobis distance (Blanche et al., 2005; Mahalanobis, 1936) were 1.8 for the mixed raw clusters, and 9.2 for the 100 kHz upsampled clusters. Improvements in cluster separation translate into higher neuron yields with better quality single unit isolation.

3.5. SHD correction

While SHDs of a few microseconds are of no consequence for tetrode recordings, with polytrodes they accumulate to produce significant phase disparities in waveforms recorded across sites (Fig. 7A). With our recording system the delay of a given site was 1 μs/channel. The maximum phase advance caused by the SHDs of the 32 channel DT3010 data acquisition cards was thus only 31 μs, however,this would be sufficient to corrupt instantaneous spike amplitude measurements (Fig. 7B) needed for modeling the spatial distribution of spike potentials across channels under quasistatic conditions (Blanche et al., 2003). Likewise, studies of microsecond precision spike timing (Brand et al., 2002; Kawasaki et al., 1988; Wagner et al., 2005) would be confounded by SHDs of this magnitude. For these sorts of analyses, SHD artefacts should be compensated (Eq. (6)) during the interpolation process.

4. Discussion

The task of isolating single units in multiunit recordings is challenging enough without the added variability caused by inaccurate waveform reconstruction. Linear interpolation of analog signals sampled at the Nyquist frequency cannot capture the true shape of a spike waveform. Accordingly, spike events may be missed, and spike metrics that are needed for later analyses such as spike sorting will be compromised. While the usual solution is to oversample, this places an extra burden on the acquisition hardware and data storage requirements. The development of high-speed high-channel acquisition boards has not kept pace with recent advances in multisite electrode arrays that have hundreds (Litke et al., 2004) or thousands (Bai et al., 2000) of recording sites. These devices generate gigabytes of data per minute sampling at 25 kHz per channel, so in spite of the remarkable storage capacities of modern desktop computers, there are real incentives for avoiding oversampling. In this report we have demonstrated that windowed sinc interpolation can achieve the same end as oversampling (as it should on theoretical grounds). In addition, we have shown that SHD correction removes systematic sampling delays that can have a deleterious effect on cross-channel spike waveform analyses that depend on truly simultaneous cross-channel measurements.

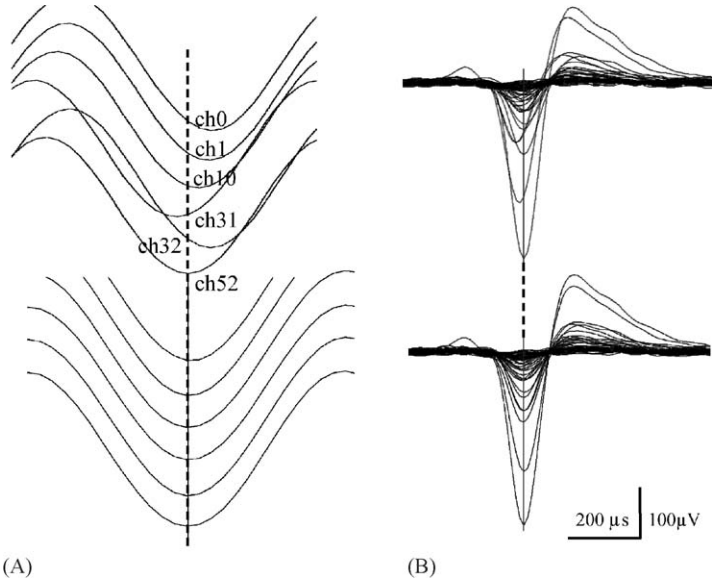

Fig. 7. Sample-and-hold delay correction. (A) A 1 kHz sine-wave signal sampled "near simultaneously" on different channels of the two 32-channel data acquisition cards (sharing a common sampling clock). Accumulation of SHDs produced artefactual phase misalignments (upper panel). The same waveforms interpolated with SHD correction (lower panel) were perfectly aligned (note that ch0/ch32, ch1/ch33, etc. were the corresponding pairs of channels on each acquisition card, and therefore exhibit the same delay). (B) A similar demonstration of the effect of SHDs on a spike recorded by 16 neighbouring polytrode recording sites, before and after SHD correction.

That windowed sinc interpolation restores the original spike waveform does not imply that information over and above that contained in the original samples has been added, since irrespective of whether the data were sampled according to the Nyquist criterion, these are either already filtered out of the original data or are aliased into frequencies below the Nyquist limit. Thus, upsampling 25 kHz sampled data to 100 kHz does not produce waveforms containing frequencies over 12.5 kHz (Fig. 3A, red line). It follows that if only Fourier methods are used in the subsequent analysis of the raw data (e.g. Kayser et al., 2003; McGill and Dorfman, 1984) there is no advantage to interpolating in the time domain. However, most electrophysiological analyses such as spike detection, slope and amplitude estimation, peri-stimulus time histograms, burst analysis, reverse correlation and so forth are done in the time domain, and it is here that the benefits of correct interpolation will be realised.

Supplementary digital filters can be incorporated into the interpolation process, just as the Hamming window was convolved with the IIR filter to give the FIR kernel. As convolution is associative, repetitious convolution of the time series for each filter pass can be avoided by convolving the actual sinc kernel with the filter or filters of interest. Kernels for a multitude of purposes, including high-pass, low-pass, and notch-filters, kernels for smoothing or removal of dc signals, differentiation, integration and so forth (Smith, 1997), can thus be applied without additional computational overhead.

4.1. An optimal level of upsampling?

Interpolation by convolution is computationally intensive. Both the quantity of data that must be managed and the time it takes to be processed is proportional to the rate of interpolation (Table 1). For multiunit spike data all significant improvements in accuracy and precision afforded by interpolation had been attained by 50 kHz. Upsampling to 100 kHz does not, therefore, justify the doubling in computation time or bandwidth. There is no single optimal rate of oversampling; it depends on how the data will ultimately be used. For analyses exclusively in the frequency domain, sample at the Nyquist frequency. If the goal is to generate tuning response curves from average spike counts, then for large amplitude spikes that are already well isolated it could be argued that undersampling is sufficient. Any missed spike triggers would be randomly distributed across trials, and the shape of the tuning curve would still be accurate. At the other extreme, attempting to infer intracellular state from multisite extracellular waveform distributions (Henze et al., 2000), or studies of spike timing reliability and precision on the submillisecond (DeWeese et al., 2003; Mainen and Sejnowski, 1995) and microsecond (Brand et al., 2002; Kawasaki et al., 1988; Wagner et al., 2005) timescale, may benefit considerably from upsampling.

Interpolation may also be applied to specific stages in the data processing stream. For example, the quality of PCA-based spike sorting relies on precise spike alignment (Lewicki, 1998), however, the realism of the spike waveform is unimportant so long as the shape is consistent. By upsampling to 50 kHz to ensure reliable spike detection and peak alignment, followed by subsampling back to 10 kHz to speed the calculation of the principal eigenvalues, processing of large scale recordings from polytrodes can be made more efficient while at the same time improving the quality of sorting (data not shown). Similarly, waveform subtraction algorithms that attempt to decompose overlapping spikes (Atiya, 1992; Lewicki, 1994) perform poorly if the canonical spike templates are inaccurate. Misalignment of the spike

Table 1
Cost-benefit analysis of upsampling

Interpolation factor	Effective sampling rate (kHz)	Spike detection	Spike amplitude	Spike shape	Spike width[a]	CPU load (ms/chan/s)[b]	Bandwidth (KB/chan/s)[c]
1×	12.5	+	+	+	+	≤1[d]	25
2×	25	+++	+++	++	+++	4.0 ± 0.1	50
4×	50	++++	++++	+++	++++	8.1 ± 0.3	100
8×	100	++++	++++	++++	++++	16.2 ± 0.4	200

The benefits of spike waveform interpolation at different rates of upsampling, balanced against computational load and bandwidth requirements. Rating scale + to ++++, relative to raw data sampled at the Nyquist rate (more pluses denote higher accuracy and precision).

[a] Applies to both single spikes and the alignment precision of multiple, averaged spikes.

[b] Average ± S.D. of 100 measurements for 13 pt sinc kernels, computed on an Athlon 1800+ processor.

[c] For 12–16 bit data.

[d] Overhead for (sequential) file streaming from the hard disk.

template and the measured waveform can introduce spurious spike-like artefacts. To keep these residual artefacts below the RMS noise level, Lewicki (1994) estimated the required alignment precision to be in the range of 2–10 times the Nyquist sampling rate, similar to that suggested here.

4.2. Other interpolation methods

Meijering (2002) provides a comprehensive review of the many available methods of interpolation, but only Fourier-based techniques such as bandlimited interpolation have a solid theoretical grounding. Alternative methods, such as cosine, Gaussian, polynomial and cubic spline interpolation (Schafer and Rabiner, 1973) are useful for display purposes but do not give *bona fide* reconstructions. Cubic spline interpolation (based upon smooth, continuous first and second derivatives) can be used to accurately locate waveform peak and valley times (Wheeler and Smith, 1987) and requires only a few samples around the raw peak to be interpolated. In turn PCA-based spike sorting can be made more accurate. Aside from this specific purpose, however, cubic splines are not useful for quantitatively accurate waveform reconstruction (Fig. 1). Moreover, it is computationally prohibitive to process continuously acquired waveforms with cubic splines, making it unsuitable for improving spike detection reliability. In any case there is no reason for not using the theoretically correct Fourier methods of interpolation given that standard desktop computers are capable of processing hundreds of channels at a rate faster than real time (Table 1).

Computing interpolated sample points by applying suitable phase rotations in the frequency domain is equivalent to sinc convolution in the time domain. But since the raw waveforms are usually stored as time series and the length of the sinc kernels are short, converting the data into the frequency domain, upsampling it, and converting it back to the time domain for further processing is slower than direct convolution in the time domain, even if the fast Fourier transform (Press et al., 1994) algorithm is used.

4.3. Avoiding aliasing

There are few compelling reasons to heavily oversample analog signals during data acquisition, as post hoc bandlimited interpolation is a theoretically sound, effective method of achieving the same end. However, before lowering the acquisition rate the potential for signal aliasing must be considered, because the transition bandpass of analog anti-alias filters can be very broad, and the sampling theorem requires zero energy above the Nyquist frequency (Eq. (1)). The risk of aliasing is high if the Nyquist frequency is uncritically equated with the corner or "cut-off" frequency of the amplifier low-pass filters. One way to relate the relevant factors of filter roll-off, the signal (including noise) energy in the transition bandpass, and the level of acceptable distortion from aliasing given the overall dynamic range of the system, is the sampling ratio:

$$\mathrm{SR} = \frac{s}{f_c} = 1 + \frac{f_r}{f_c} \tag{7}$$

where s is the sampling rate, f_r the frequency at the required attenuation, and f_c is the −3 dB corner frequency of the anti-alias filter.

We illustrate this by considering the characteristics of the electrophysiology system used in this study (Fig. 8). To maintain the precision of a 12 bit (72 dB, or ∼0.02% quantization error) acquisition system, with f_c set at 6 kHz and 100 dB/decade anti-alias filters (manufacturer specifications cite five pole RC filters, each of which should provide 20 dB/decade/pole = 100 dB/decade, or 30 dB/octave, of net signal attenuation), the required f_r is 28 kHz. This gives a SR = 5.7 and a sampling rate s = 34 kHz, well above the Nyquist rate calculated on the assumption of a "brick-wall" filter with f_r equal to f_c (Fig. 8A). However, the SR is a very conservative guide, as it assumes a white energy spectrum (i.e. in and above the transition band) and zero tolerance to aliased signals. It also ignores other sources of noise that reduce the overall dynamic range, or effective number of bits (ENOB), of the system as a whole. For our system, a more realistic estimate might put the ENOB of the 12 bit analog ADCs at 11.5bits (69 dB, or ∼0.03% precision) when intrinsic electronic noise, harmonic distortion, and potential for crosstalk in the ADC during high speed signal multiplexing are taken into account. The measured filter decay was closer to 60 dB/decade beyond the corner frequency (Fig. 8B), considerably less than 100 dB/decade. On the other hand, the combined signal and noise power spectra were not white (Fig. 3A). Energy at the corner frequency, of background neural hash recorded in vivo (with all apparatus connected and switched on, but exclud-

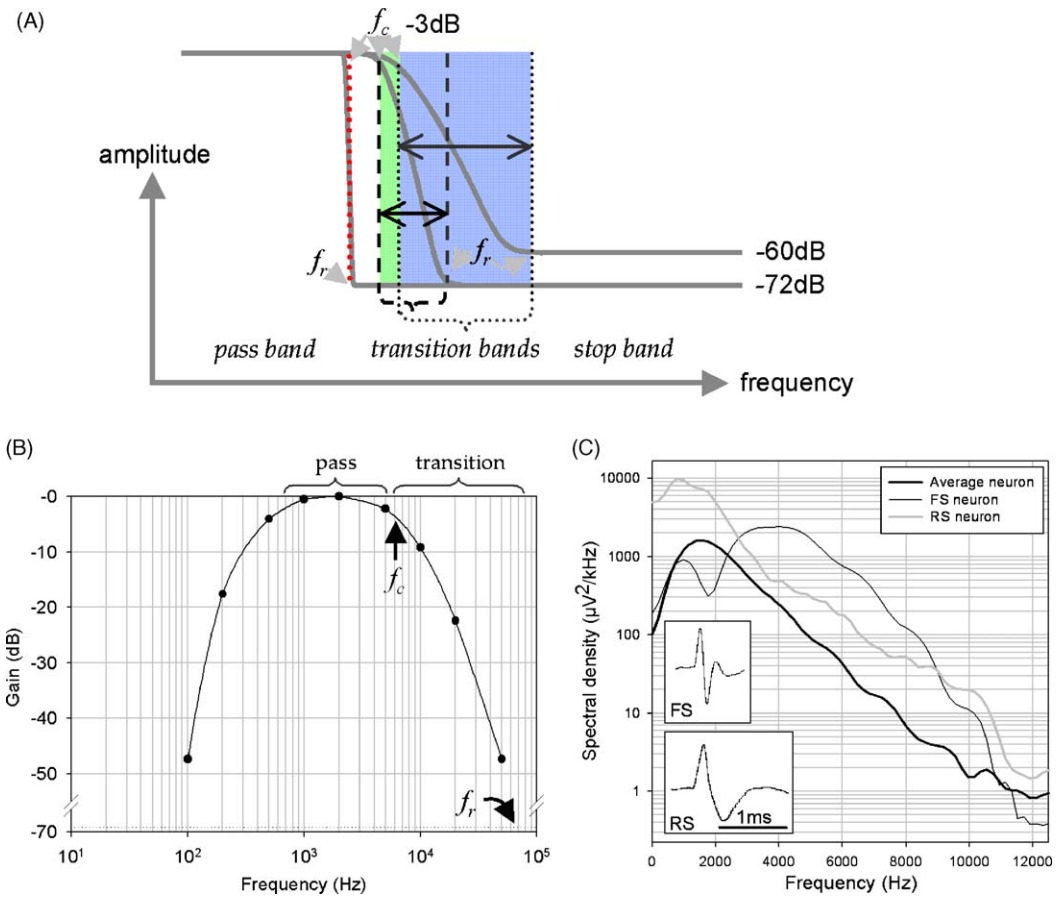

Fig. 8. Avoiding aliasing at lower sampling frequencies. (A) Determining an appropriate sampling rate depends on the corner frequency f_c, the frequency at the required level of attenuation, f_r, and the slope of the anti-alias low-pass filter. The pass band is defined as the range of frequencies below f_c that pass signals without change in amplitude. The frequency range between f_c and f_r is called the transition band. Signals in this range are subject to aliasing if the most common interpretation of the Nyquist sample rule is applied (i.e. $2 \times f_c$). Frequencies above f_r are referred to as the stop band. For an ideal filter [white dotted line (print version); red dotted line (web version)] $f_r = f_c$ so the SR = 1+1/1 = 2, in accord with the Nyquist criterion. For narrow transition bands [dashed lines (print version); green shaded region (web version)], the SR will be slightly higher, for broad transition bands [black dotted lines (print version); blue shaded region (web version)]; the SR will be higher still. (B) Filter characteristics of the amplifier system used in this study. (C) RS and FS neurons with similar amplitudes (insets) have different frequency spectra due to differences in waveform shape. The average spike spectrum ($n = 148$) is dominated by the RS spectrum since FS neurons were much less prevalent and typically of lower amplitude. (For interpretation of the reference to colour in the figure legend, the reader is referred to the web version of the article.)

ing spikes), was ~0.02% full-scale, close to the f_r. Regular spikes (RS) on average had an even lower power at 6 kHz, however, FS neurons had significant energy beyond 6 kHz (Fig. 8C). Taken together, a more realistic SR for this acquisition system would be 3.4, which implies that significant aliasing of noise and spikes from FS neurons would be avoided by sampling at 20 kHz. Alternatively, if anti-alias filters with $f_c = 5$ kHz and a steeper roll-off were substituted into the amplifiers, or if up to 10 μV_{pp} aliasing was deemed acceptable, then the SR would ≈2.4, yielding a sampling rate of about 12 kHz. However, the ability to distinguish FS neurons from RS neurons would be significantly compromised by filtering and sampling at this reduced rate.

From a practical standpoint, there are other reasons for acquiring and storing data at slightly more than twice the Nyquist frequency (even if defined by f_c). Nyquist-sampled spikes monitored online (i.e. without sinc interpolation) have poor definition, so moderate oversampling avoids the need to interpolate online. Oversampling of raw signals allows for considerably shorter sinc FIR filter kernels for an equivalent level of interpolation precision, because the number of zero crossings required is determined by the width of the transition bandpass. A small percentage increase in the original sampling rate affords a larger percentage saving in computation time because the added bandwidth is a larger percentage of the filter transition bandwidth than it is of the original sampling rate (Smith, 2004). For example, given an f_c of 6 kHz, the transition band for a sampling rate of 15 kHz is about 1.5 kHz, while at 18 kHz the transition band is 3 kHz. A 20% increase in sampling rate thus halves the computational load during upsampling, and high precision interpolation is possible with filter kernels containing as few as 13 points (Fig. 3C).

5. Conclusions

The rationale for upsampling was to lessen the detrimental effects of digital sampling related variability in order to obtain more accurate spike waveform measurements and, in turn, improve spike detection and sorting. Sampling at 10–20 kHz

(depending on the specifications of the system's amplifiers, filter characteristics, and biological signals of interest) followed by bandlimited (Nyquist) interpolation to 50 kHz is strongly recommended, and preferable to oversampling the original signal.

Acknowledgments

We thank Martin Spacek for technical assistance during data collection, Kris Gillespie and staff for animal care, and Robert Douglas for valuable discussions and feedback on the manuscript. This research was funded by grants from the Canadian Institutes of Health Research and the Natural Sciences and Engineering Research Council of Canada to NVS.

References

Abeles M, Goldstein M. Multispike train analysis. Proc IEEE 1977;65: 762–73.

Atiya AF. Recognition of multiunit neural signals. IEEE Trans Biomed Eng 1992;39:723–9.

Bai Q, Wise KD, Anderson DJ. A high-yield microassembly structure for three-dimensional microelectrode arrays. IEEE Trans Biomed Eng 2000;47:281–9.

Bartho P, Hirase H, Monconduit L, Zugaro M, Harris KD, Buzsaki G. Characterization of neocortical principal cells and interneurons by network interactions and extracellular features. J Neurophysiol 2004;92:600–8.

Blanche TJ, Hetherington PA, Rennie CJ, Spacek MA, Swindale NV. Model-based 3D cortical neuron localization and classification with silicon electrode arrays. In: Proceedings of the 32nd annual meeting of the society for neuroscience; 2003.

Blanche TJ, Spacek MA, Hetke JF, Swindale NV. Polytrodes: high density silicon electrode arrays for large scale multiunit recording. J Neurophysiol 2005;93:2987–3000.

Brand A, Behrend O, Marquardt T, McAlpine D, Grothe B. Precise inhibition is essential for microsecond interaural time difference coding. Nature 2002;417:543–7.

Campbell PK, Jones KE, Huber RJ, Horch KW, Normann RA. A silicon-based, three-dimensional neural interface: manufacturing processes for an intracortical electrode array. IEEE Trans Biomed Eng 1991;38:758–68.

Chandra R, Optican LM. Detection, classification, and superposition resolution of action potentials in multiunit single-channel recordings by an online real-time neural network. IEEE Trans Biomed Eng 1997;44:403–12.

Csicsvari J, Henze DA, Jamieson B, Harris KD, Sirota A, Bartho P, et al. Massively parallel recording of unit and local field potentials with silicon-based electrodes. J Neurophysiol 2003;90:1314–23.

DeWeese MR, Wehr M, Zador AM. Binary spiking in auditory cortex. J Neurosci 2003;23:7940–9.

Gadicke R, Albus K. Real-time separation of multineuron recordings with a DSP32C signal processor. J Neurosci Methods 1995;57:187–93.

Gray CM, Maldonado PE, Wilson M, McNaughton B. Tetrodes markedly improve the reliability and yield of multiple single-unit isolation from multi-unit recordings in cat striate cortex. J Neurosci Methods 1995;63:43–54.

Henze DA, Borhegyi Z, Csicsvari J, Mamiya A, Harris KD, Buzsaki G. Intracellular features predicted by extracellular recordings in the hippocampus in vivo. J Neurophysiol 2000;84:390–400.

Horowitz P, Hill W. The Art of Electronics. Cambridge: Cambridge University Press; 1989.

Kawasaki M, Rose G, Heiligenberg W. Temporal hyperacuity in single neurons of electric fish. Nature 1988;336:173–6.

Kayser C, Salazar RF, Konig P. Responses to natural scenes in cat V1. J Neurophysiol 2003;90:1910–20.

Lewicki MS. Bayesian modeling and classification of neural signals. Neural Comput 1994;6:1005–30.

Lewicki MS. A review of methods for spike sorting: the detection and classification of neural action potentials. Network 1998;9:R53–78.

Litke AM, Bezayiff N, Chichilnisky EJ, Cunningham W, Dabrowski W, Grillo AA, et al. What does the eye tell the brain? Development of a system for the large-scale recording of retinal output activity. IEEE Trans Nucl Sci 2004;51:1434–40.

Mahalanobis PC. On the generalized distance in statistics. Proc Natl Inst Sci (India) 1936;2:49–55.

Mainen ZF, Sejnowski TJ. Reliability of spike timing in neocortical neurons. Science 1995;268:1503–6.

McGill KC, Dorfman LJ. High-resolution alignment of sampled waveforms. IEEE Trans Biomed Eng 1984;31:462–8.

Meijering EH. A chronology of interpolation: From ancient astronomy to modern signal and image processing. Proc IEEE 2002;90:319–42.

Nicolelis MA, Dimitrov D, Carmena JM, Crist R, Lehew G, Kralik JD, et al. Chronic, multisite, multielectrode recordings in macaque monkeys. Proc Natl Acad Sci USA 2003;100:11041–6.

Nyquist H. Certain topics in telegraph transmission theory. Trans AIEE 1928;47:617–44.

Pouzat C, Mazor O, Laurent G. Using noise signature to optimize spike-sorting and to assess neuronal classification quality. J Neurosci Methods 2002;122:43–57.

Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical Recipes in Fortran 77: The Art of Scientific Computing. Cambridge, Mass: Cambridge University Press; 1994.

Schafer RW, Rabiner LR. A digital signal processing approach to interpolation. Proc IEEE 1973;61:692–702.

Shannon CE. Communication in the presence of noise. Proc Inst Radio Eng 1949;37:10–21.

Smith JO, 2004. Digital Audio Resampling Home Page. http://www-ccrma.stanford.edu/~os/resample/.

Smith SW. The Scientist and Engineer's Guide to Digital Signal Processing. San Diego: California Technical Publishing; 1997.

Wagner H, Brill S, Kempter R, Carr CE. Microsecond precision of phase delay in the auditory system of the barn owl. J Neurophysiol 2005;94:1655–8.

Wheeler BC, Smith SR. High-resolution alignment of action potential waveforms using cubic spline interpolation. J Biomed Eng 1987;10:47–53.

Whittaker JM. Interpolatory Function Theory in Cambridge Tracts in Mathematics and Mathematical Physics. Cambridge: Cambridge University Press; 1935.