DETERMINING WHICH SINE WAVE FREQUENCIES CORRESPOND TO SIGNAL AND WHICH CORRESPOND TO NOISE IN EYE-TRACKING TIME-SERIES

A PREPRINT

Mehedi H. Raju*
Texas State University
601 University Drive
San Marcos, Texas, 78640, USA
m.raju@txstate.edu

Lee Friedman
Texas State University
601 University Drive
San Marcos, Texas, 78640, USA
lfriedman10@gmail.com

Troy M. Bouman
Department of Mechanical Engineering-Engineering Mechanics
Michigan Technological University
1400 Townsend Dr.
Houghton, MI 49931, USA
tmbouman@mtu.edu

Oleg V. Komogortsev
Texas State University
601 University Drive
San Marcos, Texas, 78640, USA
ok11@txstate.edu

January 31, 2023

ABSTRACT

The Fourier theorem proposes that any time-series can be decomposed into a set of sinusoidal frequencies, each with its own phase and amplitude. The literature suggests that some of these frequencies are important to reproduce key qualities of eye-movements ("signal") and some of these frequencies are not important ("noise"). We looked at three types of analysis: (1) visual inspection of plots of saccade, microsaccade and smooth pursuit exemplars; (2) an analysis of the percentage of variance accounted for (PVAF) in each of 1,033 unfiltered saccade trajectories by each frequency cutoff; (3) an analysis of saccade peak velocity in the unfiltered and various filtered conditions. Visual inspection suggested that frequencies up to 75 Hz are required to represent microsaccades. Our PVAF analysis indicated that data in the 0-25 Hz band are sufficient to account for nearly 100% of the variance in unfiltered saccade trajectories. Our analysis indicated that frequencies below 100 Hz are sufficient to maintain peak velocities. Therefore, our overall conclusion is that to maintain eye-movement signal and reduce noise, a cutoff frequency of 100 Hz is appropriate. Our results have implications for the proposed sampling rate of eye-tracking recordings. If one is working in the frequency domain and 100 Hz needs to be preserved, the minimum required sampling rate would be 200 Hz. However, in a time domain analysis, a minimum 1000 Hz sampling rate is required.

Keywords Eye Tracking · Signal · Noise · Fourier

1 Introduction

Fourier analysis models a time-series as the sum of a set of sine-waves with variable frequencies, phases, and amplitudes. In many cases (but not all, e.g., nystagmus [Rosengren et al., 2020]), the lower frequencies are required to preserve a time-series feature of interest (e.g., saccade peak velocity), and higher frequencies may not be needed and thus represent noise. In this common case, a low-pass filter can be used to keep the signal part of the time-series and remove the noise

*corresponding author

part. The best practice would be for researchers to evaluate what frequencies are needed, and which are not, for any particular goal, prior to data collection. This preliminary analysis would allow the researcher to design an appropriate data collection scheme. In this part of the study, signals should be collected at the highest possible frequency so that an analysis of which frequencies are needed can be performed. Once this information is known, filter settings and sampling rate could be optimized. For example, if the signal is in analog form, and if it is determined in this preliminary analysis that frequencies above 25 were not needed, then for the actual data collection one could place an anti-aliasing filter (low pass, 25 Hz cutoff) before the A/D conversion with a sampling rate of 250 Hz.

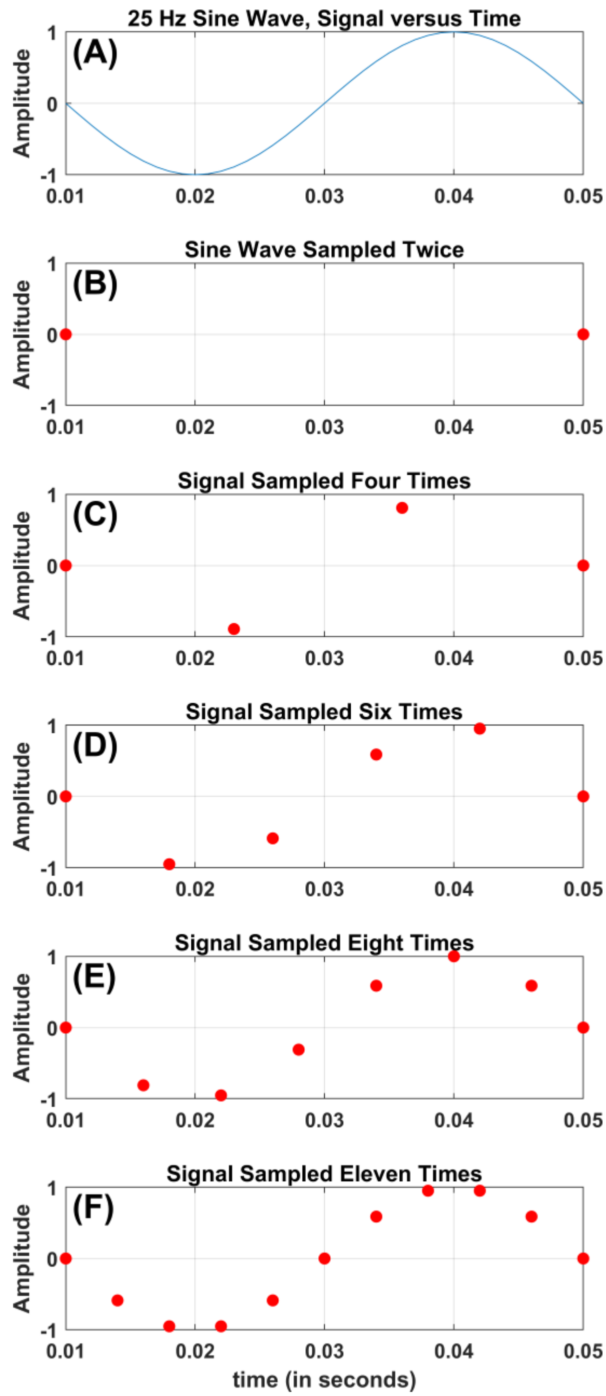

Figure 1: Visual representation of the 10x rule

One's study goals are very important when trying to determine a required sampling rate. If we were interested in the frequency domain, then a minimum of 2 samples per wave is required [Shannon, 1949]. If the fastest frequency we need was $x\ Hz$, then the minimum sampling frequency needs to be $2 \times\ x\ Hz$. However, if we were interested in the time domain, as we believe that most eye-movement researchers are, then the minimum sampling frequency (Fs) needed to be $Fs\ >= x \times\ 10\ Hz$[2]. We refer to this as the "10x rule". The basis for the rule is obvious if one considers this question: How many samples are needed per sine-wave to resolve and see that sine wave? If one only samples a sine wave twice, then the sine-wave will not look like a sine wave. It is a rule of thumb that to accurately visualize a sine-wave, the sine wave needs to be sampled at least 10 times, and preferably 20 times. See Fig. 1 illustrating the need for this rule.

We are not aware of any paper in the eye-movement field that used this 10x rule, including all the papers cited in this study. Here, a leading research group makes this statement:

> For oscillating eye-movements, such as tremors, we can argue based on the Nyquist-Shannon sampling theorem (Shannon, 1949) that the sampling frequency should be at least twice the speed of the particular eye movement (e.g., behavior at 150 Hz requires $>$ 300 Hz sampling frequency) [Andersson et al., 2010].

Of course, this rule is only appropriate if one's focus is in the frequency domain (e.g., Fourier amplitude or power spectra). But eye-movement researchers are interested in the time domain, i.e., the trajectories of saccades, or PSOs, or the length and stability of fixation etc. Therefore, the correct rule of thumb is the 10x rule described above.

Below, we review the prior research on required frequencies for saccades. For our research (and we suspected many others) faithful preservation of saccade trajectories and main-sequence-related saccade metrics would probably be sufficient. We didn't review signal-to-noise determinations for other eye movements with potentially higher-frequency components.

We presented our analysis of the literature in Table 1. One potentially relevant paper was not included in our table [Juhola et al., 1985]. The signals (electrooculography EOG and photoelectric) were analog signals. These analog signals were filtered first with the low-pass analog filter at 30 Hz. Subsequently, the signals were digitally filtered with a low-pass filter with a cutoff of 70 Hz. This creates a very complex situation, and we didn't think that statements about frequencies required to preserve saccade peak velocity were useful given the insertion of this analog filter. Therefore, this paper was not included in Table 1.

We also excluded [Inchingolo and Spanio, 1985]. Their paper was based on EOG signal which was analog-filtered with a cutoff at 100 Hz. Any further statements about the effects of other digital low-pass filtering was confounded by the presence of the analog filter.

From Table 1, despite the difference in recording and other methods, the literature supports the notion that 0-125 Hz frequency components are sufficient to preserve saccade characteristics.

Table 1: Frequency Content of Saccades

Article	Methods	Findings
[Bahill et al., 1981]	Photoelectric techniques	For noisy data, a bandwidth of 0-125 Hz was required to record saccades. Also, a sampling rate of 1000 Hz was suggested.
[Schmitt et al., 2007]	Video-based infrared eye-tracker	Sampling rate should be 250 Hz.
[Wierts et al., 2008]	VOG and Search coil	Saccadic eye movements of $>= 5^o$ amplitude were bandwidth limited up to a frequency of 25 to 30 Hz. A sampling frequency of about 50 Hz was sufficiently high to prevent aliasing.
[Mack et al., 2017]	Synthetic saccades	Signals sampled as low as 240 Hz allow for the good reconstruction of peak velocity. With 240 Hz, the frequencies that can be evaluated were 0-24 Hz (in the time domain).

To summarize, our goal in this study was to determine which frequencies are needed to preserve signal and which frequencies correspond to noise in eye-tracking studies. We evaluate this issue for of saccades, microsaccades and smooth pursuit.

[2]See discussion at `https://community.sw.siemens.com/s/article/digital-signal-processing-sampling-rates-bandwidth-spectral-lines-and-more`

2 Methods

2.1 Subjects

We recorded a total of 23 unique subjects (M=17/ F=6, median age = 28, range = 20 to 69 years). From the total number of unique participants, 14 had normal (not-corrected) vision, and 9 had corrected vision (7 glasses, 2 contact lenses). Nine of the unique participants were left-eye dominant and 14 were right-eye dominant. Subjects were recruited from laboratory personnel, undergraduates taking a class on computer programming, and friends of the experimenters. The Texas State University institutional review board approved the study, and participants provided informed consent.

We report on two datasets, the first dataset ("Fixation"), initially contained data from 15 subjects, but because of blinks and other artifacts, we only analyzed 9 subjects. The second dataset ("RS-SP"), contained data when subjects viewed both a random saccade task and a smooth pursuit task. The RS-SP dataset consisted of 9 subjects.

2.2 Eye Movement Data Collection

Eye movements were collected with a tower-mounted EyeLink 1000 eye tracker (SR Research, Ottawa, Ontario, Canada). The eye tracker operated in monocular mode capturing the participant's dominant eye. Eye dominance was determined using the Miles method [Miles, 1930].

During the collection of eye movements data, each participant's head was positioned at a distance of 550 millimeters from a 19" (48.26 cm) computer screen (474×297 millimeters, resolution 1680×1050 pixels), where the visual stimulus was presented. The sampling rate was 1000 Hz. All data sets were collected with all heuristic filters off, i.e., unfiltered.

For the fixation task, subjects were presented with a single fixation point (white circle, 0.93^o) as the visual stimulus. The point was positioned in the horizontal middle of the screen and at a vertical angle of 3.5^o above the primary position. Participants were instructed to fixate on the stationary point stimulus for a period of 30 seconds.

During the random saccade task, subjects were instructed to follow the same target on a dark screen as the target was displaced at random locations across the display monitor, ranging from ± 15^o and ± 9^o of visual angle in the horizontal and vertical directions respectively. The random saccade task was 30 seconds long. The target positions were randomized for each recording. The minimum amplitude between adjacent target displacements was 2^o of visual angle. The distribution of target locations was chosen to ensure uniform coverage across the display. The delay between target jumps varied between 1 sec and 1.5 sec (chosen randomly from a uniform distribution).

During the smooth pursuit task, subjects were instructed to follow a target on the dark screen as the target moved horizontally from center to right. This ramp was followed by a fixation (length between 1 and 1.5 sec). This was followed by another ramp from the right to the left of the screen, then another fixation, etc. The rest of the task was a series of left-to-right and right-to-left ramps with fixations interposed. The target was moving at velocities of either 5^o/sec, 10^o/sec, or 20^o/sec. For each speed, there were 5 continuous leftward and 5 rightward ramps per set. The order of the velocity sets was random for each participant. There was a 15 sec fixation period at the beginning of the task and between each set. The whole recording was 120 seconds long.

2.3 Signal Processing of Fixation Data

All fixation recordings lasted 30 seconds (30,000 samples). Non-overlapping segments of fixation of 2048 continuous samples were selected. To be included in our analysis, we wanted only segments without saccades and other artifacts. If a segment contained any velocity above 25 deg/sec the segment was rejected[3]. For six of the 15 subjects, we could not find a single segment of 2048 samples that met our criteria.

2.4 Selection of Saccade, Catch-up Saccade and Microsaccade Exemplars

We wanted to have multiple exemplars of saccades, catch-up saccades (CUS), and microsaccades. For the saccade examples, we used the random saccade task. For the CUS, we used the smooth pursuit dataset. Microsaccades were selected from the fixation dataset. Two exemplars were chosen for each eye movement type, a low-noise example and a high-noise example ("clean" and "noisy"). The selection was subjective but incorporated measures of precision to guide this choice. More examples are available as part of our supplementary material.

For those unfamiliar with catch-up saccades, they occur when tracking a smoothly moving target. When gain was less than 1.0, subjects consistently lag behind the smoothly moving signal. In this case, they generate relatively small

[3]Velocity was calculated with a six-point difference approach using t_{+3} and t_{-3} as recommended by Bahill et al. [1982]

saccades to "catch-up" to the target. For a detailed analysis of the relationship between smooth pursuit gain, CUS amplitude, and CUS rate see Friedman et al. [1991].

2.5 Signal Frequency Content Analysis

We wanted to evaluate eye movements after one of seven filtering regimes (unfiltered, low-pass filtered at 25 Hz, band pass filtered from 26-50 Hz, 51-75 Hz, 76-100 Hz, 101-200 Hz, and high-pass filtered at 201 Hz). See Fig. 2 for an illustration of the frequency-response of the various filters. These were created using very sharp high-pass, low-pass, and band-pass Butterworth-style filters (order = 7). To prevent phase effects, all of these filters were zero-phase, which means that after the data were filtered in the forward direction, the signal was flipped and passed through the filter again. This procedure effectively doubled the filters' orders and squares the magnitudes of their transfer functions. The filtering operation was performed in post processing.

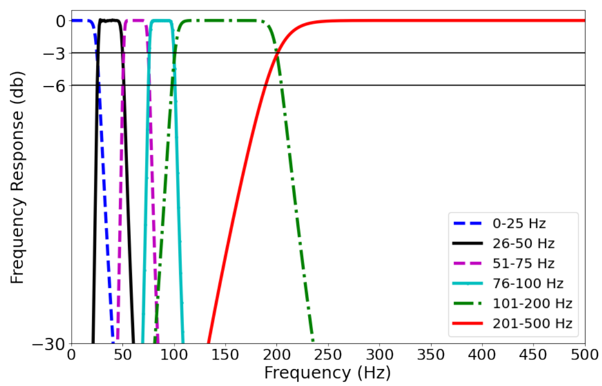

Figure 2: Frequency response of different frequency bandwidths using 7th order Butterworth filters

2.6 Calculation of percentage of variance accounted for (PVAF)

The first step for this analysis was to identify saccades in all of our Random Saccade task data. The identification was initially performed by an updated version of our previously published event detection method [Friedman et al., 2018]. All potential saccades were screened by the authors so that only well-marked saccades were included. There were 1,033 well-marked saccades. A PVAF analysis was performed on each of these saccades. For each of these saccades, data from the unfiltered condition was treated as a dependent variable, and all of the filtered signals were treated as independent variables. We regressed the first filtered signal (0-25 Hz) onto the unfiltered signal and noted the r^2. We then added the data filtered from 26-50 Hz and noted the change in r^2. We kept doing this until all of the filtered bands had been entered into the multiple linear regression model. We multiplied each r^2 by 100 to obtain the percent of variance accounted for (PVAF).

2.7 Study of the Effects of Filtering on Saccade Peak Velocity

We started with the 1,033 saccades discussed above. We created a histogram of saccade duration. We noticed that there were two groups of saccades, short saccades, and long saccades (Fig.3). In this histogram, two groups were obvious: those with durations less than (or equal to) 25 ms (N=415) and those with a duration greater than 25 ms (N=618). The analysis of the effect of filtering on peak velocity was performed on the short saccade group and the long saccade group separately. Snippets of the horizontal position channel were cut from 200 ms prior to each saccade to 200 ms after each saccade. For each snippet, a velocity calculation was performed using the 1st derivative from a Savitzky-Golay filter function with order = 2 and window = 7. The peak (absolute) velocity of the saccade was determined. In the next steps, each snippet was filtered with 7th-order low-pass Butterworth filter with cutoffs of 25, 50, 75, 100, and 200 Hz. Peak velocities were determined for every saccade in the unfiltered state and in each of the filter conditions[4]. A Friedman test was conducted to test median differences between the filter conditions. With N = 415 or 618, and 6 levels, we would have a statistical power of 1.00 to detect a medium effect size[5]. Typically, in planning a study a power of 0.8 is a common goal. In our case, with a power of 1.0, all studies with moderate effect sizes will be statistically significant.

[4]Images of all 1,033 saccades, unfiltered and filtered, are available at https://digital.library.txstate.edu/handle/10877/16437

[5]See https://www.statskingdom.com/34test_power_chi2.html

Therefore, these statistical tests should be considered very powerful (i.e., effectively certain to find a medium effect size if there was one). Post-hoc multiple comparisons were controlled with a Tukey HSD test, at alpha = 0.05.

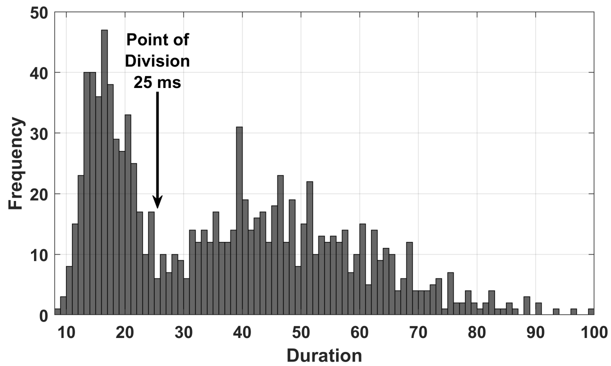

Figure 3: Frequency histogram of saccade duration (N=1,033 saccades)

.

3 Results

3.1 Analysis of Exemplars

3.1.1 Saccades

In Fig. 4, we present the signal frequency content analysis for a "clean" saccade. This saccade has an approximate amplitude of 2.94^o. In plot (A1) we present the unfiltered signal trace for the saccade. In plot (B1 to G1) we present the signal containing frequencies from different bands. All of the plots in the left column were scaled to match the unfiltered saccade in (A1). All of the plots on the right column were scaled individually based on their range of data. The signal in plot (B1) appears very similar to the saccade in plot (A1). However, the post-saccadic activity in (A1) was missing, and there was less noise. The saccade amplitude has not been altered. In plot (C1) we present the signal containing frequencies from 26-50 Hz. There appears to be a minor contribution to signal amplitude from this band. For the remaining plots in the left column (D1 to G1), it appears that no signal remains that was relevant to the trajectory of the unfiltered saccade in (A1).

In the right column, note the range of the data in (D2) to (G2). All of these bands contribute less than 3.0% of the amplitude of the unfiltered saccade. The waveforms of these plots do not appear to be relevant to the unfiltered saccade. So, for this saccade, we would consider that the data below 50 Hz were signal and the data above 50 Hz were noise.

In Fig. 5, we present the signal frequency content analysis for a "noisy" saccade. This saccade has an approximate amplitude of 2.79^o. In plot (A1) we present the unfiltered signal trace for the saccade. In plots (B1 to G1) we present the signal containing frequencies from different bands. The signal in plot (B1) appears very similar to the saccade in plot (A1). However, the post-saccadic activity in (A1) was missing, and there was less noise. The saccade amplitude has not been altered. In plot (C1) we present the signal containing frequencies from 26-50 Hz. There appears to be a minor contribution to signal amplitude from this band. Some of the signals in this band may contribute to the post-saccadic activity in the unfiltered saccade. For the remaining plots in the left column (D1 to G1), it appears that no signal remains that was relevant to the trajectory of the unfiltered saccade in (A1).

In the right column, note the range of the data in (D2) to (G2). All of these bands contribute less than 4.0% of the amplitude of the unfiltered saccade. The waveforms of these plots do not appear to be relevant to the unfiltered saccade. So, for this saccade also, we would consider that the data below 50 Hz were signal and the data above 50 Hz were noise.

3.1.2 Microsaccade

In Fig. 6, we present the signal frequency content analysis for a "clean" microsaccade. This microsaccade has an approximate amplitude of 0.63^o. The saccade detection algorithm determined the end of this saccade later than one would choose manually, but we don't think this difference affects the present analysis. In plot (A1) we present the unfiltered signal trace for the microsaccade. In plots (B1) to (G1) we present the signal containing frequencies from different bands. The signal in plot (B1) appears to be a very smooth version of the waveform in (A1). The microsaccade amplitude may be very slightly less than the amplitude of the unfiltered microsaccade. In plot (C1) we present the

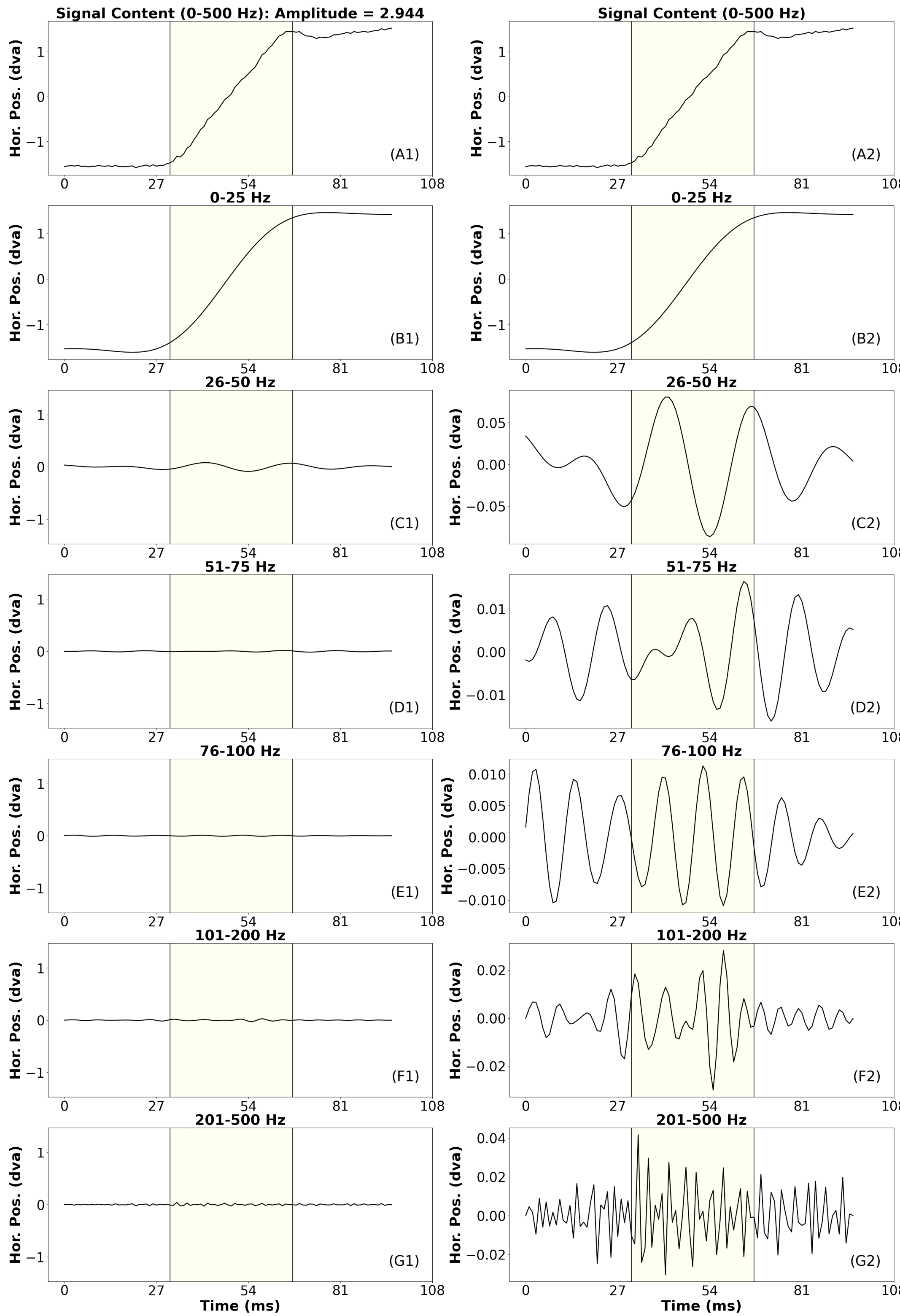

Figure 4: Signal frequency content analysis of a clean saccade. (A1) Exemplar of a clean unfiltered saccade. (B1) The signal in (A1) with only frequencies from 0 to 25 Hz. (C1) The signal in (A1) with only frequencies from 26 to 50 Hz. (D1) The signal in (A1) with only frequencies from 51 to 75 Hz. (E1) The signal in (A1) with only frequencies from 76 to 100 Hz. (F1) The signal in (A1) with only frequencies from 101 to 200 Hz. (G1) The signal in (A1) with only frequencies from 201 to 500 Hz. Note that all plots on the left panel have the same amplitude range as the original saccade. Since we cannot see some of the signals on this scale very well, each plot (A2-G2) on the right panel was y-scaled individually according to the range of the data. Yellow highlighting indicates the saccade.

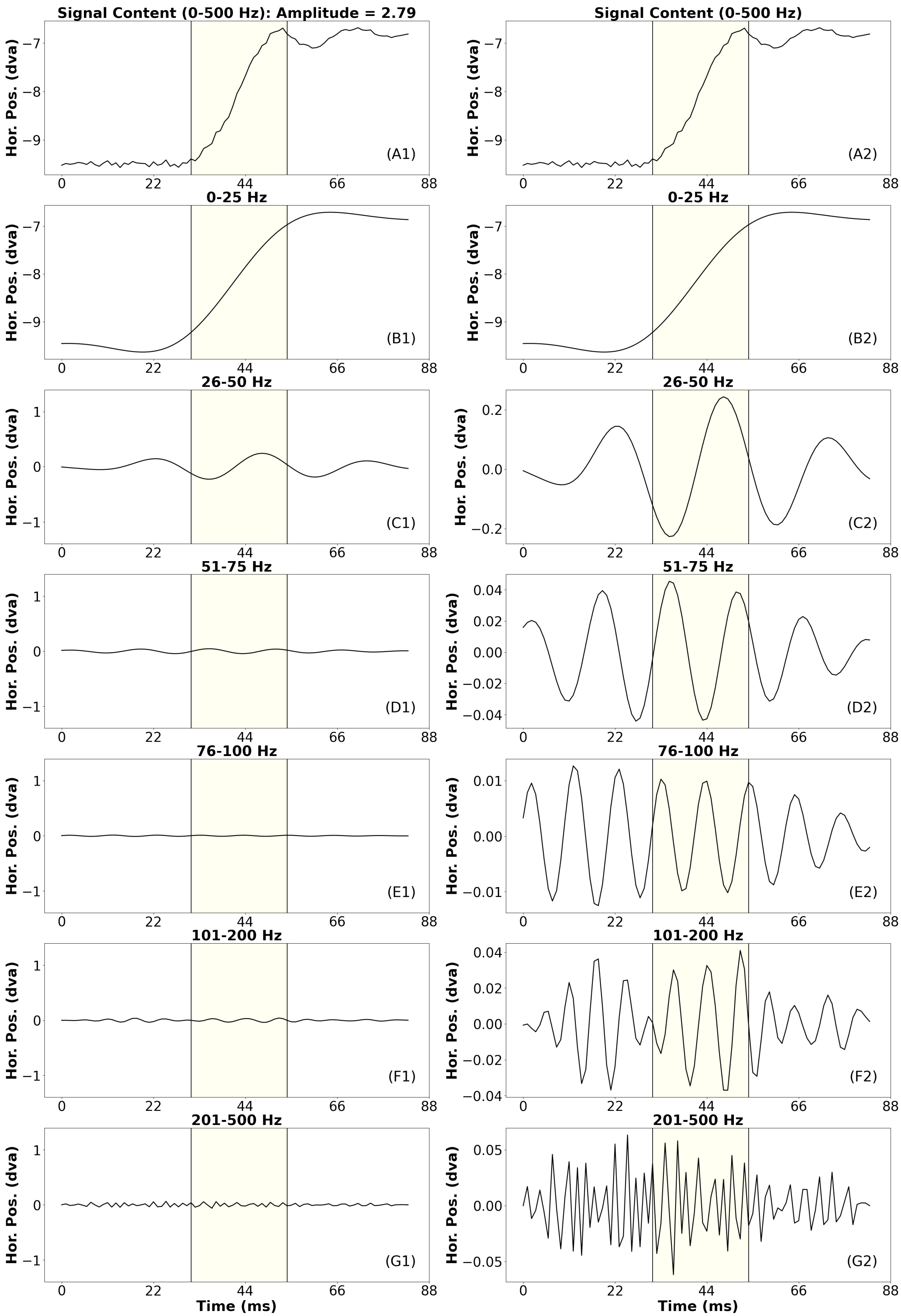

Figure 5: Signal frequency content analysis of a noisy saccade. See caption for Fig. 4 for more details.

signal containing frequencies from 26-50 Hz. The waveform for the data filtered at 51-75 Hz (D1) appears to contain some relevant signal. For the remaining plots in the left column (E1 to G1), it appears that no signal remains that was relevant to the trajectory of the unfiltered microsaccade in (A1).

Note the range of the data in (C2 to G2). Their amplitude range was a much higher ratio to the unfiltered microsaccade than similar bands were to the saccade exemplars. For this exemplar, we consider that the data below 75 Hz were signal and the data above 75 Hz were noise.

Similarly, in Fig. 7, we present the signal frequency content analysis for a “noisy” microsaccade. This microsaccade has an approximate amplitude of 0.651^o. The signal in plot (B1) looks like a very smooth version of the unfiltered saccade. The amplitude of this very smooth waveform is, at most, very slightly less than the unfiltered saccade. In plot (C1), we present the signal containing frequencies from 26-50 Hz. These higher frequencies contribute to the sharpness of unfiltered signal. In plot (D1), we present the signal containing frequencies from 51-75 Hz. As was the case for the waveform in (C1), these higher frequencies contribute to the sharpness of the unfiltered signal.

For the remaining plots in the left column (E1 to G1), it appears that no signal remains that was relevant to the trajectory of the unfiltered microsaccade in (A1). This part of the signal was what makes this a “noisy” saccade. For this microsaccade also, we would consider that the data below 75 Hz were signal and the data above 75 Hz were noise.

3.1.3 Smooth Pursuit and Catch-up saccades (CUS)

In Fig. 8 we present a “clean” segment of smooth pursuit and in Fig. 9 we present a “noisy” segment. Both segments have five or more CUS. The analysis of these figures was identical. In the (A1) plots we present the unfiltered smooth pursuit signal, including catch-up saccades. The (B1) plots appear very similar to that of the unfiltered segment. In the band from 26-50 Hz, There are some very small high-frequency bursts coincident with each saccade. Plot (C2) makes this point more clearly. For the remaining plots (D1 to G1), it appears that no signal remains that was relevant to the pattern of smooth pursuit of the unfiltered catch-up saccade in (A1). In both plots (D2) and (E2), there were bursts of high-frequency noise signal coincident with each CUS. However, the amplitude of these bursts in (E2) was so small that we think data from this frequency band can be ignored. For these smooth pursuit segments, we would consider that the data below 50 or 75 Hz was signal and the data above 75 Hz was noise.

3.2 Percentage of variance accounted for (PVAF)

Our results for the PVAF analysis of all saccade trajectories are presented in Fig. 10. It was clear from this figure that nearly all of the variance in the trajectory of the unfiltered saccade was accounted for by data in the range of 0-25 Hz. None of the high-frequency data contributed to the variance in the original unfiltered saccade in any substantial way. See Table 2 for exact numbers.

Table 2: PVAF

Measure	0-25 Hz	16-50 Hz	51-75 Hz	76-100 Hz	101-200 Hz	201-500 Hz
Median	99.1795	0.5210	0.0302	0.0034	0.0076	0.0247
MAD	1.5289	1.1703	0.2250	0.0499	0.0632	0.2094

3.3 Effects of Filters on Saccade Peak Velocity

The saccade duration histogram clearly indicated two distinct saccade groups, Group 1 had a duration of 25 ms or less, and group 2 had a duration greater than 25 ms (Fig 3). Error bar plots of the median peak velocities for both groups are presented in Fig. 11. The medians (and mean absolute difference) values are also presented in Table 3. The Chi-square for the Friedman test applied to the short saccades was 1,655.64, ($df = 5, p < 0.0001$). Post-hoc tests indicated that the 100 Hz condition and the 200 Hz condition were not statistically significantly different, but all other possible comparisons were statistically significant ($p < 0.05$, Tukey-HSD for multiple comparisons). The Chi-square for the Friedman test applied to the long saccades was 1,176.91, ($df = 5, p < 0.0001$). Post-hoc tests indicated that the 75 Hz condition was not statistically significantly different from the 100 Hz condition, but that all other potential comparisons were statistically significantly different from all others, $p < 0.05$. On the basis of these statistical tests, peak velocity for both short and long saccades never return to the unfiltered level even in data filtered at 200 Hz. It is, however, important to recall the very great power of these tests. With these sample sizes, almost any small difference would be considered statistically significant. On the basis of the median values, it appears that by 100 Hz, the differences in median peak velocity were trivial (Short saccades: 109.9 $^o/s$ vs 107.6 $^o/s$; Long saccades: 328.5 $^o/s$ vs 323.14 $^o/s$).

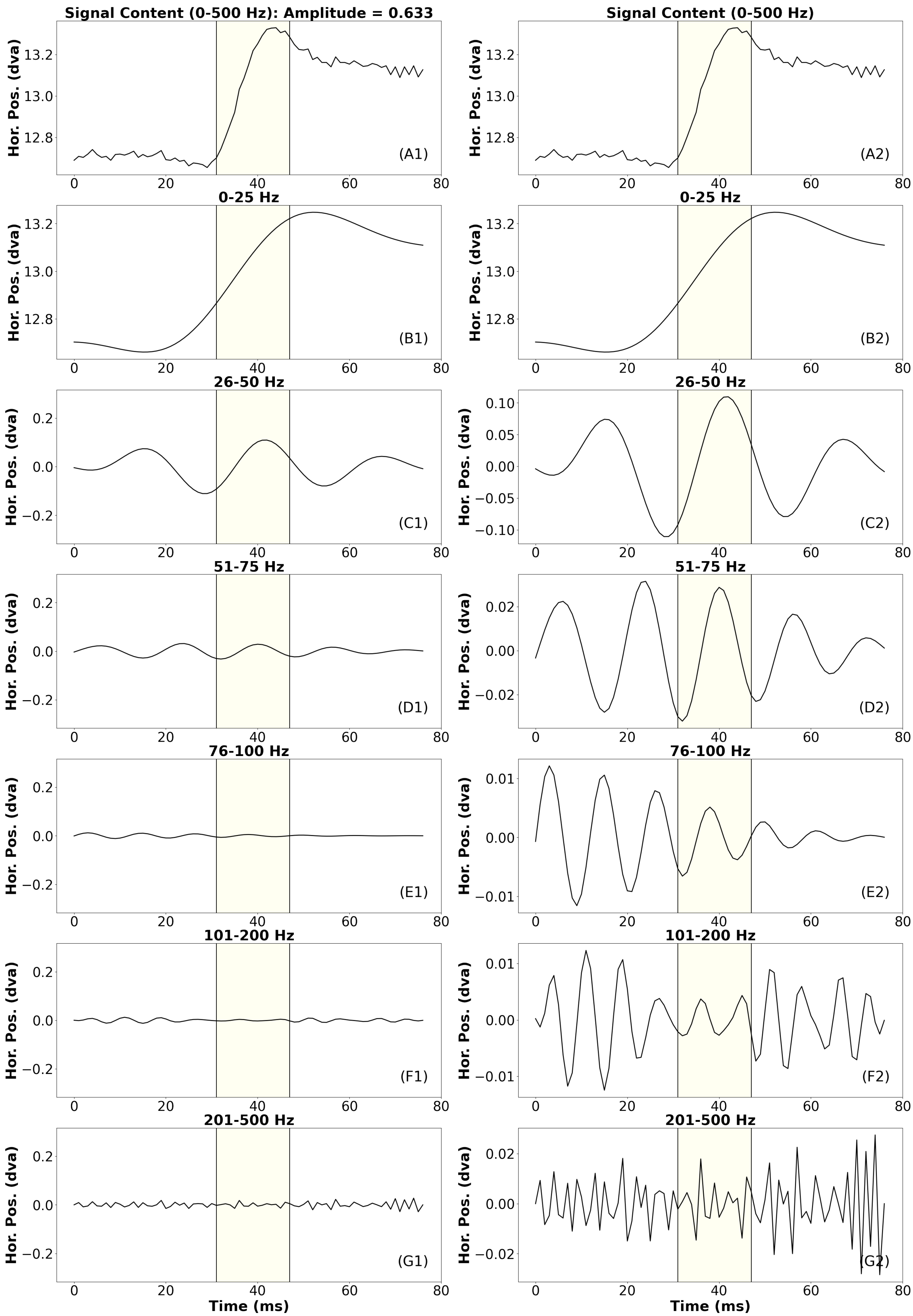

Figure 6: Signal frequency content analysis of a clean microsaccade. See caption for Fig. 4 for more details.

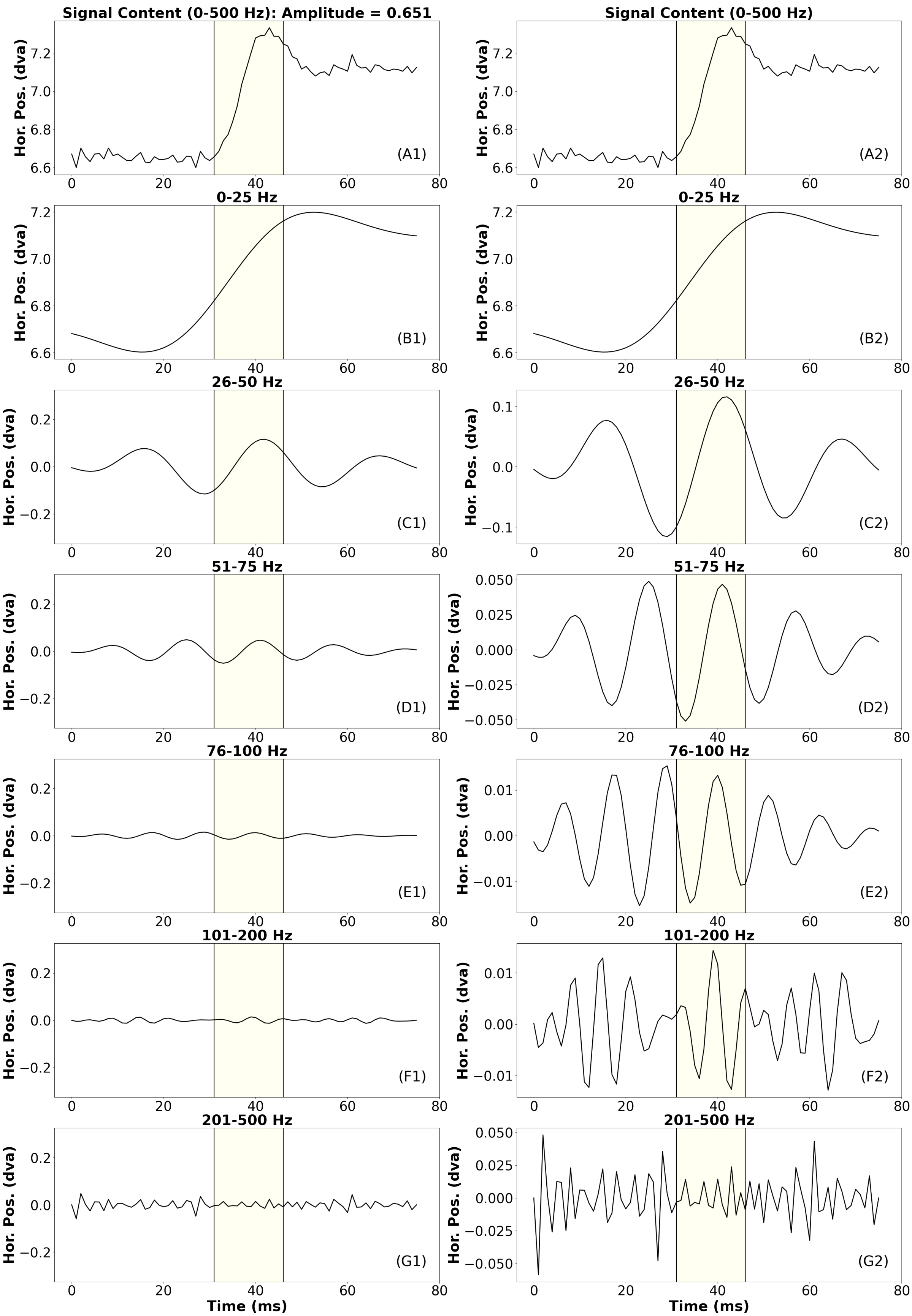

Figure 7: Signal frequency content analysis of a noisy microsaccade. See caption for Fig. 4 for more details.

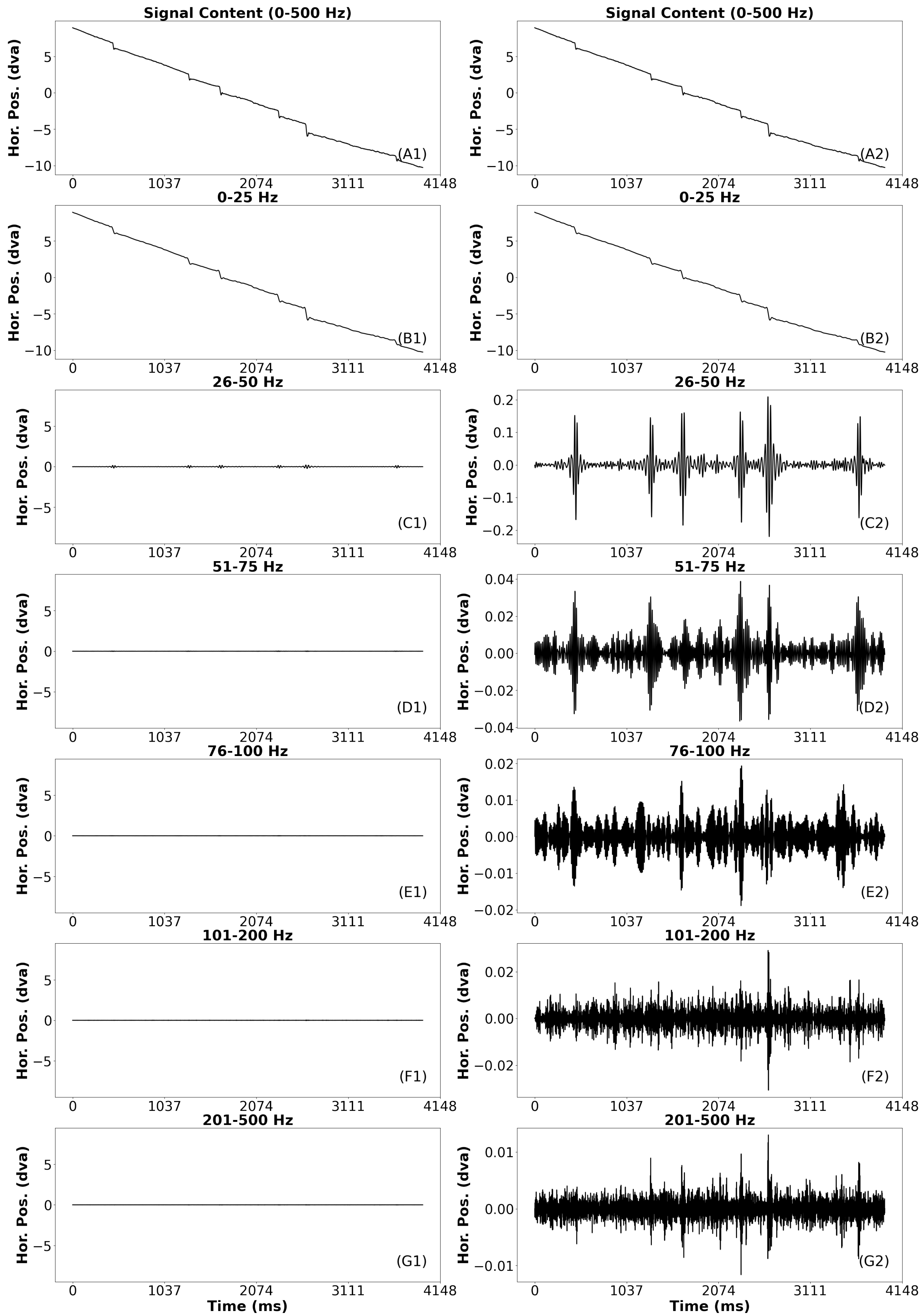

Figure 8: Signal frequency content analysis of a relatively clean smooth pursuit segment with catch-up saccades. See caption for Fig. 4 for more details.

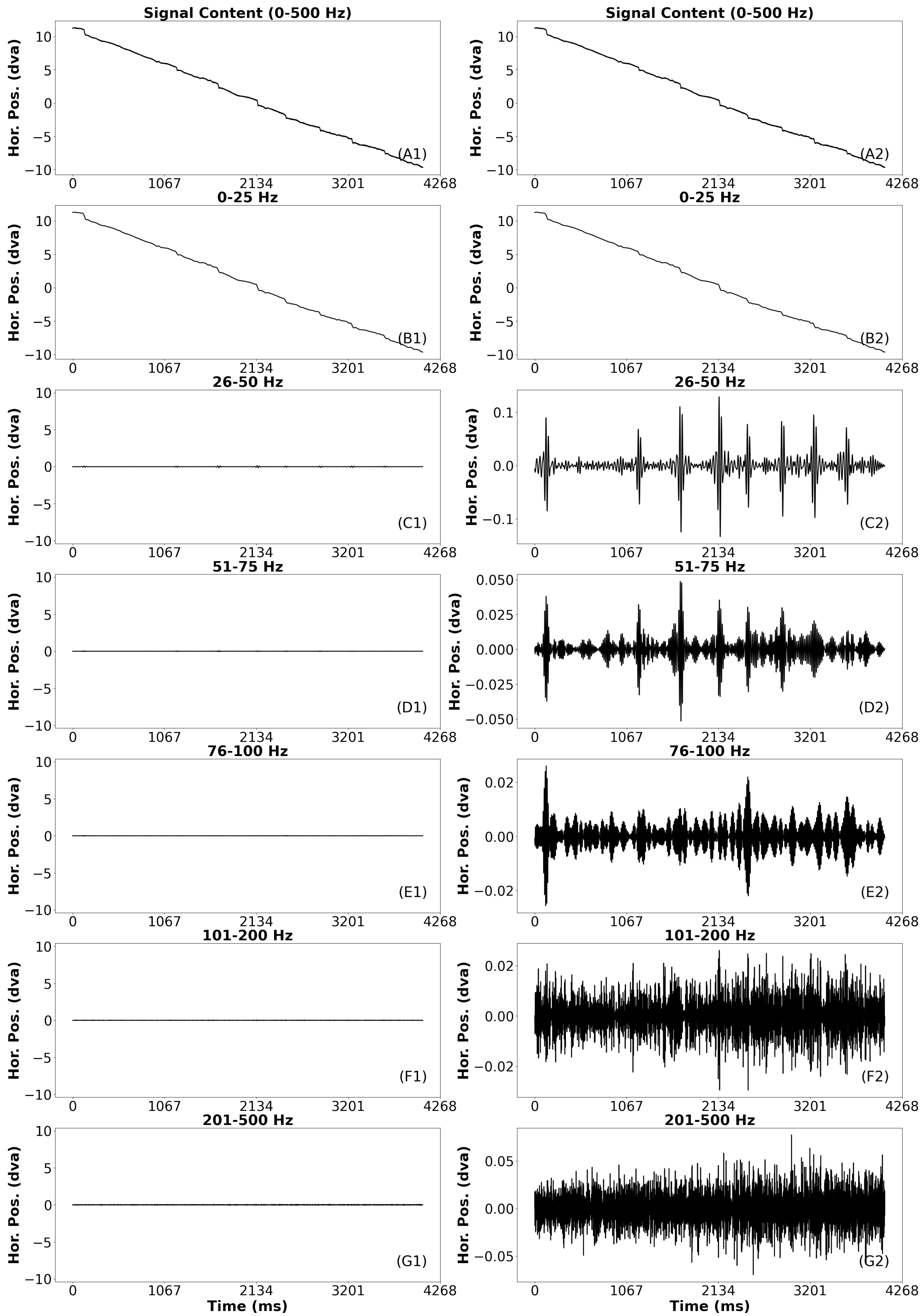

Figure 9: Signal frequency content analysis of a relatively noisy smooth pursuit segment with catch-up saccades. See caption for Fig. 4 for more details.

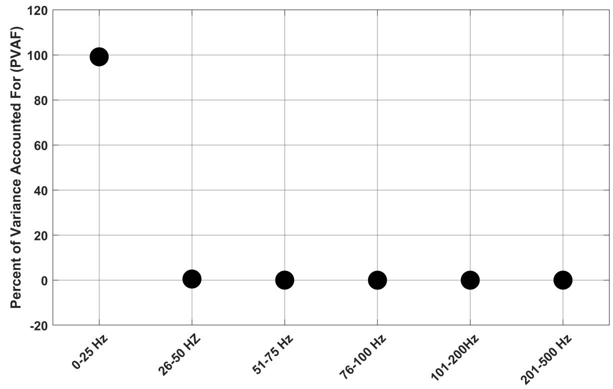

Figure 10: PVAF Error Bar Chart. Each circle was the median PVAF for a particular filter level. There are error bars in this plot based on the median absolute deviation, but they are so small as to be invisible in this range.

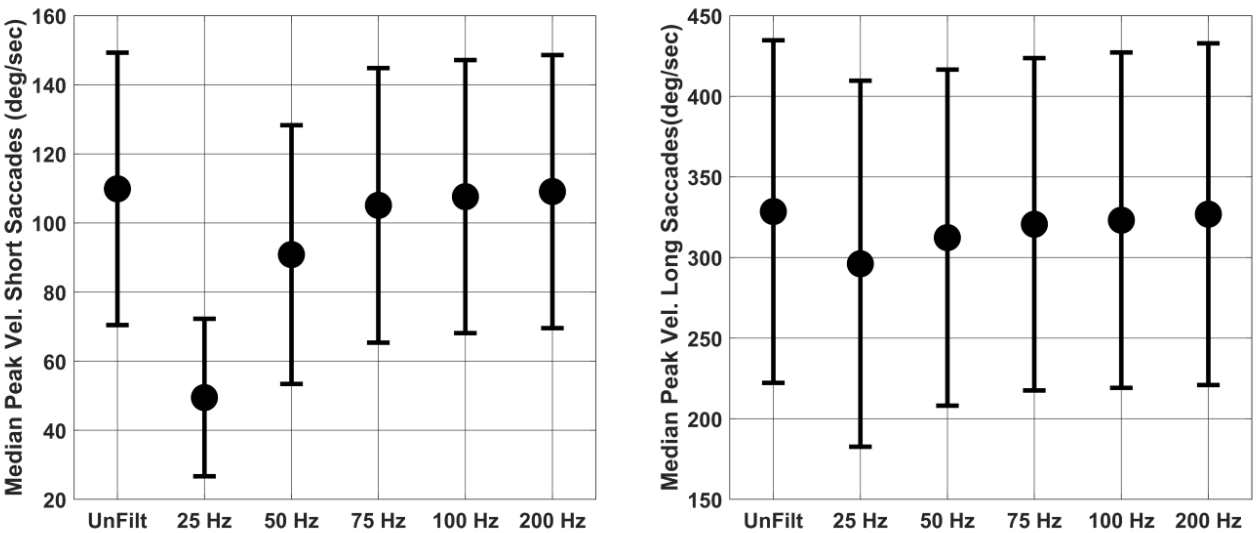

Figure 11: Peak-velocity Error Bar plots. The plot on left is for saccades $\leq$ 25ms. The plot on the right is for saccades $>$ 25 ms. Each circle is the median for a particular filter level. The error bars are ± the mean absolute deviation.

Table 3: Saccade Peak Velocities (degrees per second)

Measure	UnFiltered	0-25 Hz	0-50 Hz	0-75 Hz	0-100 Hz	0-200 Hz
Short Medians	109.9	49.5	90.9	105.1	107.6	109.1
Short MAD	39.4	22.8	37.4	39.7	39.5	39.5
Long Median	328.5	296.1	312.3	320.6	323.1	326.8
Long MAD	106.3	113.5	104.1	103.0	104.0	105.9

3.4 Summary of Results

In Table 4 we summarize our conclusions about which frequencies correspond to signal and which correspond to noise.

4 Discussion

We provided a set of analyses that each provide estimates of which sine-wave frequencies are signal and which are noise in eye-tracking data. The different analyses provide different answers but can be summarized in a final single rule. The visual analysis of our microsaccade and smooth pursuit exemplars suggested that frequencies up to 75 Hz were required to retain signal whereas waves above 75 Hz represent noise. Our analysis of the percent of variance accounted for in unfiltered saccade trajectories by different filter bands indicated that essentially all of the variance in saccade

Table 4: Summary of Results

Evidence	Method	What is Signal	What is Noise
Exemplars	Visual Inspection	0-75 Hz	76-500 Hz
Variance Explained	Compute percent of variance accounted for in unfiltered saccades	0-25 Hz	26-500 Hz
Saccade Peak Velocity [1]	Short Saccades	0-100 Hz	101-500 Hz
Saccade Peak Velocity [1]	Long Saccades	0-100 Hz	101-500 Hz

[1] Formal statistical testing found a difference in median peak velocity between the unfiltered data and all filter bands. However, these tests were extremely statistically powerful. We chose 100 Hz as the cutoff because the medians at 100 Hz were very close to the medians in the unfiltered condition.

shape was accounted for with data in the 0-25 Hz band. Saccade peak velocity was reduced when data were low-pass filtered at 25, 50, or 75 Hz. Data filtered at 100 Hz had peak velocities only trivially lower than the peak velocities of unfiltered saccades. Taken together, we conclude, that, if the goal is to preserve saccade (including microsaccade) and smooth pursuit characteristics, that frequencies up to 100 Hz are required but that frequencies above this frequency are noise. Juhola [Juhola, 1986] discuss the importance of correctly designing digital low pass files to preserve saccade peak velocity, However, they do not offer specific recommendations of cutoff frequencies and filter orders. Mack et al [Mack et al., 2017] concluded that a sampling rate of 240Hz was the minimum required to preserve estimates of saccade peak velocity.

In our proposed follow-up to this article, we plan to compare various filter schemes in terms of their frequency response. We will make a recommendation of the best filter to retain signal and remove noise in eye-movement time-series. Since such filters can introduce or increase temporal autocorrelation, we will also evaluate and compare filters in terms of their autocorrelation effects.

These results have implications for proposed sampling frequencies for future data collection. If our studies only involved the frequency domain, we would only need two samples at 100 Hz, so a sampling rate of 200 Hz would suffice. However, because we were interested in evaluating eye movements in the time domain, the 10x rule discussed in the introduction applies. Therefore, the minimum acceptable sampling rate for eye-tracking studies is 1000 Hz.

Acknowledgments

This work was funded by grant from the NSF (1714623) (PI: Oleg Komogortsev).

Open Practices Statement

The data and code for the present study is available at `https://digital.library.txstate.edu/handle/10877/16437`. Download Supplementary_Materials.zip. This study was not pre-registered.

Conflict of interest

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

William Rosengren, Marcus Nyström, Björn Hammar, Markus Rahne, Linnea Sjödahl, and Martin Stridh. Modeling and quality assessment of nystagmus eye movements recorded using an eye-tracker. *Behavior Research Methods*, 52 (4):1729–1743, 2020.

C.E. Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37(1):10–21, jan 1949. doi:10.1109/jrproc.1949.232969. URL `https://doi.org/10.1109/jrproc.1949.232969`.

Richard Andersson, Marcus Nyström, and Kenneth Holmqvist. Sampling frequency and eye-tracking measures: how speed affects durations, latencies, and more. *Journal of Eye Movement Research*, 3(3), 2010.

Martti Juhola, Ville Jäntti, and Ilmari Pyykkö. Effect of sampling frequencies on computation of the maximum velocity of saccadic eye movements. *Biological cybernetics*, 53(2):67–72, 1985.

Paolo Inchingolo and Massimo Spanio. On the identification and analysis of saccadic eye movements-a quantitative study of the processing procedures. *IEEE Transactions on Biomedical Engineering*, (9):683–695, 1985.

AT Bahill, A Brockenbrough, and BT Troost. Variability and development of a normative data base for saccadic eye movements. *Investigative ophthalmology & visual science*, 21(1):116–125, 1981.

Kai-Uwe Schmitt, Markus H Muser, Christian Lanz, Felix Walz, and Urs Schwarz. Comparing eye movements recorded by search coil and infrared eye tracking. *Journal of clinical monitoring and computing*, 21(1):49–53, 2007.

Roel Wierts, Maurice JA Janssen, and Herman Kingma. Measuring saccade peak velocity using a low-frequency sampling rate of 50 hz. *IEEE Transactions on Biomedical Engineering*, 55(12):2840–2842, 2008.

David J Mack, Sandro Belfanti, and Urs Schwarz. The effect of sampling rate and lowpass filters on saccades–a modeling approach. *Behavior Research Methods*, 49(6):2146–2162, 2017.

Walter R. Miles. Ocular dominance in human adults. *The Journal of General Psychology*, 3(3):412–430, 1930. doi:10.1080/00221309.1930.9918218. URL `https://doi.org/10.1080/00221309.1930.9918218`.

A Terry Bahill, Jeffrey S Kallman, and Jon E Lieberman. Frequency limitations of the two-point central difference differentiation algorithm. *Biological cybernetics*, 45(1):1–4, 1982.

Lee Friedman, John A Jesberger, and Herbert Y Meltzer. A model of smooth pursuit performance illustrates the relationship between gain, catch-up saccade rate, and catch-up saccade amplitude in normal controls and patients with schizophrenia. *Biological psychiatry*, 30(6):537–556, 1991.

Lee Friedman, Ioannis Rigas, Evgeny Abdulin, and Oleg V Komogortsev. A novel evaluation of two related and two independent algorithms for eye movement classification during reading. *Behavior Research Methods*, 50(4): 1374–1397, 2018.

Martti Juhola. The effect of digital lowpass filters on the maximum velocity of saccadic eye movements. *Computers in biology and medicine*, 16(5):361–370, 1986.