



A more reliable method to verify Auditory Evoked Potential repeatability **vivosonic**

Aaron Steinman¹, PhD PEng, Isaac Kurtz¹, MHSc PEng

¹Vivosonic Inc., 120-5525 Eglinton Avenue West, Toronto, ON, Canada, M9C 5K5, 1-416-231-9997, 1-877-255-7685 (Canada & US), www.vivosonic.com

ABSTRACT

One challenge of clinical Auditory Evoked Potential (AEP) measurement is determining whether specific features of the averaged AEP waveform represent true electrophysiological responses or artifactual noise. Recommended practice is to obtain two averages from independent sets of individual stimuli and assess their repeatability. However, repeatability between two averages is not the most reliable indicator of the presence of electrophysiological response. This is demonstrated by storing all unprocessed responses to individual stimuli, randomly changing their order prior to averaging, and recreating the two averages, with half the stimuli in each average. Unprocessed ABRs were recorded from the same subject (2000 responses for both 35 dB nHL clicks and 'no-stimulus'). Random reordering (repeated 1000 times) yielded overlapping distributions of the correlation coefficient ($\mu=0.43$, $\sigma=0.12$ and $\mu=0.05$, $\sigma=0.17$ for 35 dB nHL and 'no-stimulus' cases, respectively) with no statistically reliable difference between the 35 dB nHL and 'no-stimulus' cases.

To overcome this problem, rather than assigning individual responses to one of the two averages based on data collection order, we assigned responses so that the mean-squared difference between the resultant two averages approximates the statistically expected mean-squared error. The correlation coefficient distribution was significantly improved ($\mu=0.41$, $\sigma=0.02$ and $\mu=0.02$, $\sigma=0.02$).

REPEATABILITY IN AEP

"No single factor contributes more to confidence and accuracy in ABR analysis than waveform repeatability" [1, p.215]. As such, it is clinically recommended to repeat each ABR trace [2,3].

Typically, thousands of averaged responses are required in order to obtain two repeatable ABR traces. These multiple averages may be collected sequentially or from interleaved data. Using the interleaved approach, each response to individual stimuli is assigned to one of two buffers — designated Buffer A and Buffer B. Subsequent single-sweep responses are alternately assigned to a buffer, e.g., the odd-numbered responses in Buffer A and the even-numbered responses in Buffer B. The responses are combined with previous data in the assigned buffer using a conventional method such as averaging or weighted averaging. The statistics of the waveform generated by subtracting the buffers (A-B in the typical case) then provides an estimate of the statistics of the noise that contaminates the estimated ER signal. Quantitatively, the signal-to-noise ratio may be calculated as the variance ratio $\frac{\sigma_{A+B}^2}{\sigma_{A-B}^2}$. Furthermore, the correlation coefficient between buffers is a measure of response repeatability.

POSING THE PROBLEM

The variance ratio $\frac{\sigma_{A+B}^2}{\sigma_{A-B}^2}$ and correlation between buffers A and B are not statistically powerful indicators of the presence of a repeatable response in the data because their value depends on the random assignment to two buffers. The variance estimates in both the numerator and denominator are based on a limited number of linearly independent observations - constrained by the window size for data collection and lowest frequency components of the acquired signal. This problem is exemplified in Figure 1 below. The buffer assignment in Figure 1a yields a large difference between A and B while that in Figure 1b yields a correlation close to 1, leading to clinically opposite conclusions. Yet, both figures are taken from the same 4000 sweeps.

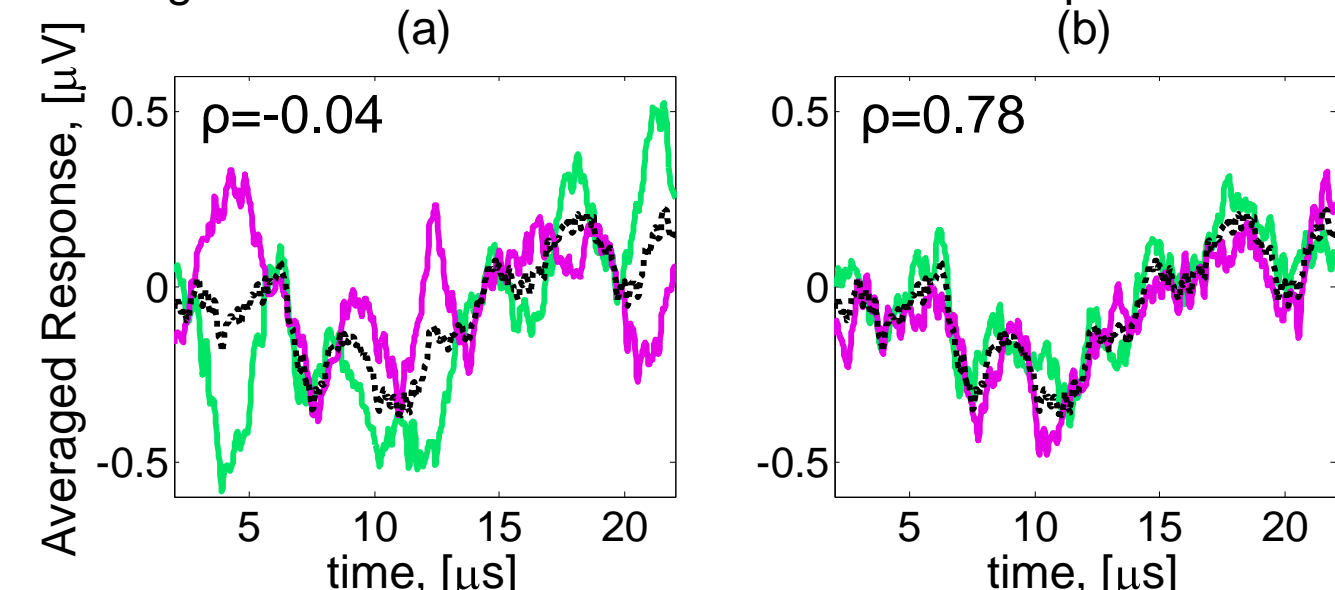


Fig 1. Comparing two sets of buffers created from the same data. The sum of the buffers (black) is identical in both figures.

Histograms of the correlation coefficient are shown in Figure 2. Data was obtained from a moving subject (35 dB nHL clicks and 'no stimulus' conditions). Unprocessed sweeps were randomly reordered 1000 times. For each reordering, the data was assigned to interleaved buffers and averaged, with weighting as in [4]. Separation between the conditions occurs only after a significant number sweeps.

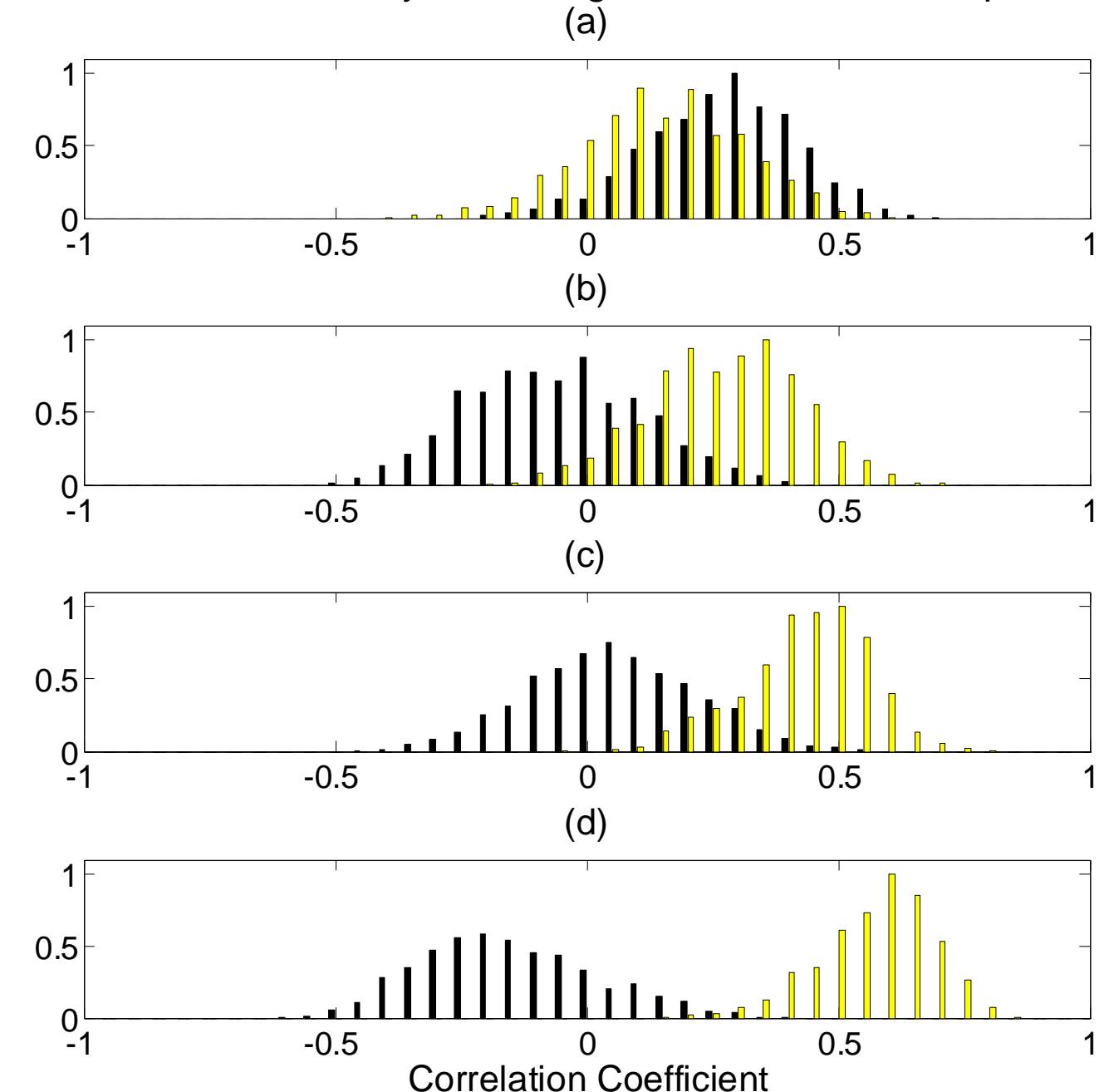


Fig 2. Normalized histograms for stimulus present (yellow) and absent (black) for 1000 reorderings; Number of responses averaged into each waveform: (a) 500, (b) 1000, (c) 2000, (d) 4000.

SEEKING A SOLUTION

To overcome this problem, rather than assigning individual responses to one of the two averages based on data collection order, we assigned responses so that the mean-squared difference between the resultant two averages approximates the statistically expected mean-squared error (calculated as per [4]) (see Figure 3).

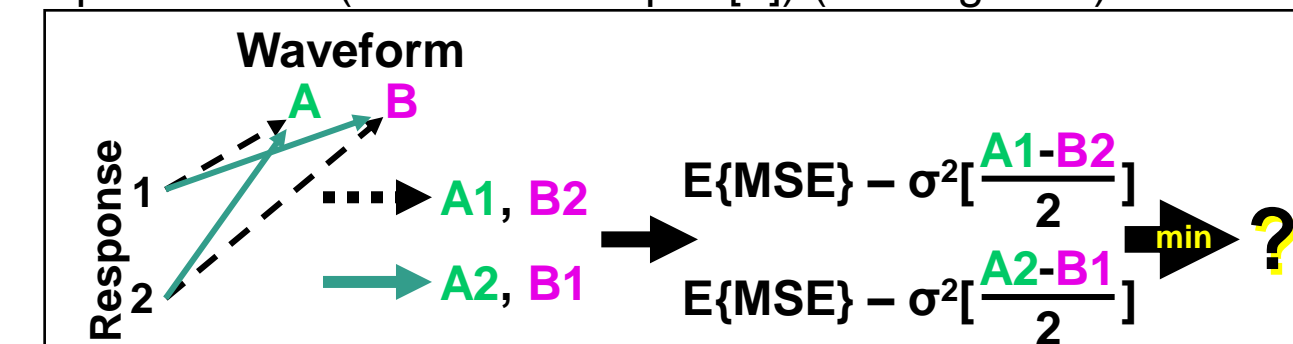


Fig 3. The decision-algorithm.

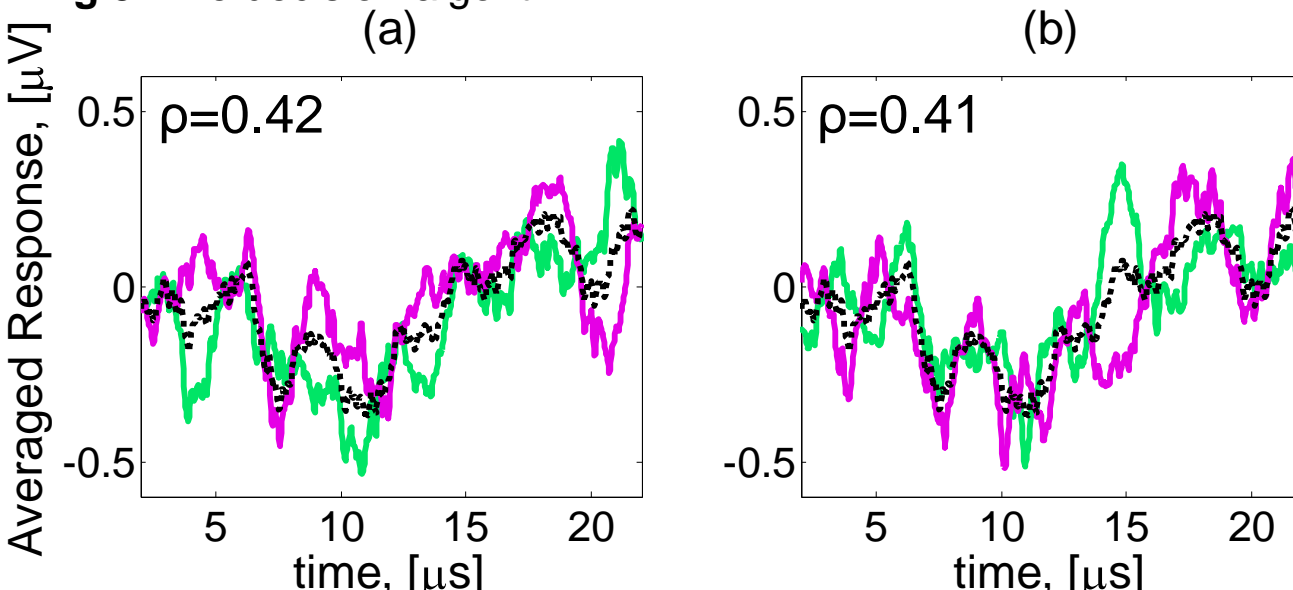


Fig 4. The same buffer order as in Figure 1, with the new algorithm. Note that the waveform repeatability is consistent.

The improvement can clearly be seen in the traces shown in Figure 4, where this decision-algorithm was applied to the same random re-orderings of ABRs used to create Figure 1. The same weighted-average was used to reduce the effect of myogenic artifact. The effectiveness of the algorithm is more pronounced when examining the histograms of Figure 5, produced by the same reorderings of Figure 2.

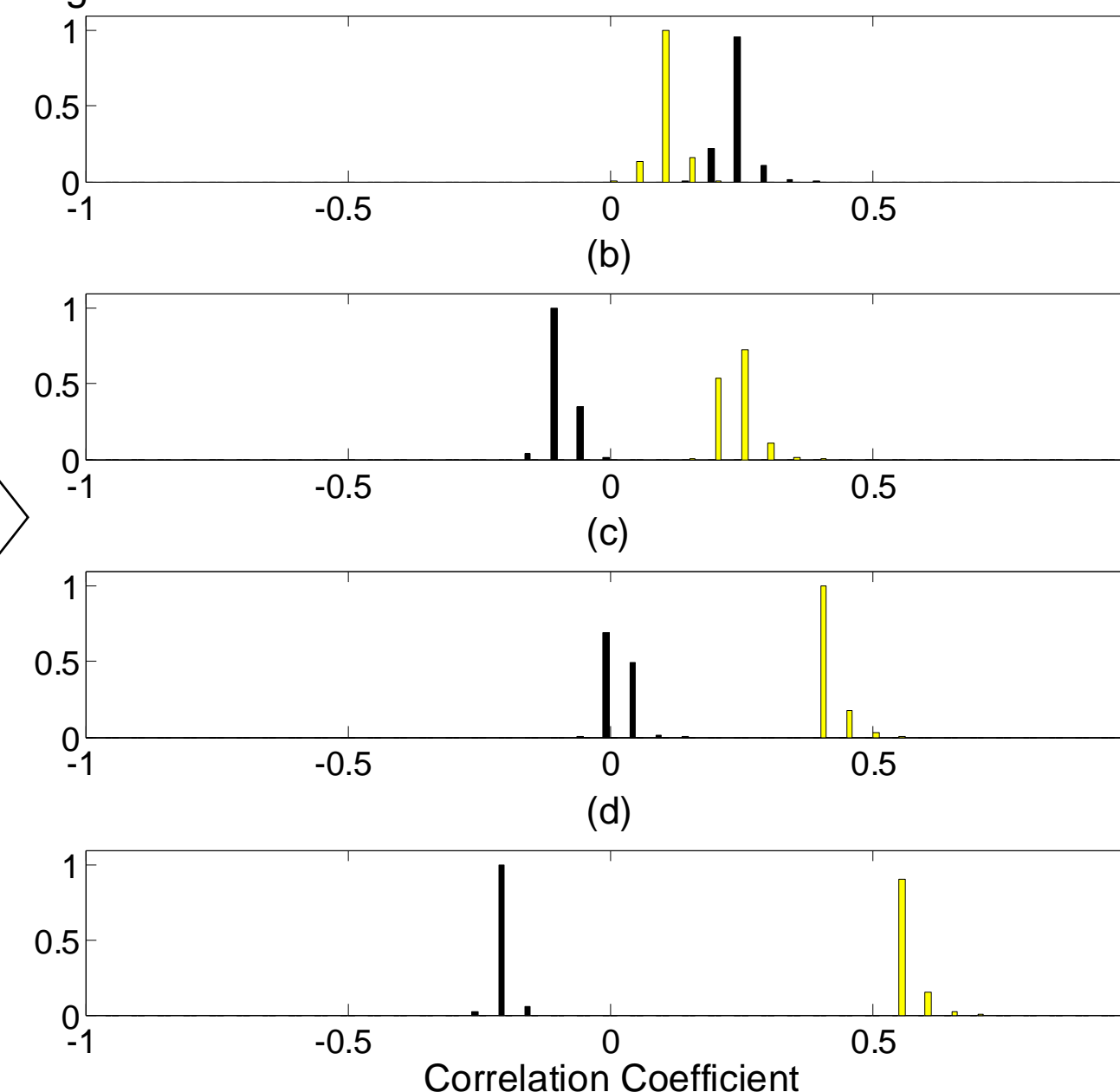


Fig 5. Normalized histograms of correlation for stimulus present (yellow) and absent (black) using the new algorithm for the same 1000 reorderings as in Figure 2. Number of responses averaged into each waveform: (a) 500, (b) 1000, (c) 2000, (d) 4000.

DISCUSSION / CONCLUSION

Table 1 summarizes the details of the distributions presented in Figures 2 and 5. Of primary note is that the significant difference between random buffering and the new method is the tightness of the distribution. The mean values, however, are independent of buffering method. It is important to recognize that these conclusions are based on data obtained for a moving subject with a low level of stimulation, where we would expect poorer performance, which is not apparent in the new method. As well, tighter distributions require fewer averages to separate the signal present and signal absent conditions, and hence determine repeatability.

Table 1: Comparing methods by examining the mean correlation coefficient ± 1 standard deviation for the distributions shown in Figures 2 and 5.

Number of Responses	Stimulus Absent		Stimulus Present	
	Random Buffering	New Method	Random Buffering	New Method
(a) 2 x 500	0.27 \pm 0.15	0.24 \pm 0.03	0.14 \pm 0.16	0.10 \pm 0.03
(b) 2 x 1000	-0.06 \pm 0.17	-0.09 \pm 0.02	0.27 \pm 0.15	0.24 \pm 0.03
(c) 2 x 2000	0.05 \pm 0.17	0.02 \pm 0.02	0.43 \pm 0.12	0.41 \pm 0.02
(d) 2 x 4000	-0.15 \pm 0.17	-0.19 \pm 0.01	0.57 \pm 0.11	0.56 \pm 0.02

It is apparent from the difference in the correlation coefficient distributions for the 'stimulus present' and 'stimulus absent' conditions of Figure 2 that there is sufficient information in the raw data to determine whether an electrophysiological response is present. A single correlation coefficient, however, cannot be relied upon to distinguish between the conditions because its value depends on the order in which the data is received. However, confidence in repeatability can be obtained either by increasing the number of responses in an average or by randomly reordering the responses and finding the mean of the resulting distribution[§].

Conversely, it can be extrapolated that the new method alleviates the need to perform random reorderings of the responses in order to get a reliable measure of repeatability and confidence in the result.

For clinical ABR, and by extension other AEP waveform measurements, we recommend to first use this new method to generate confidence in the repeatability of the waveforms. Once confidence is established, we recommend to then use the sum of the two buffers for diagnostic measurements (such as latency).

§ - Patent Pending

References

- [1] Hall JW (2006). New Handbook for Auditory Evoked Responses. Boston, Pearson Education, Inc.
- [2] Hood L (1998). Clinical applications of the auditory brainstem response. California, Singular Publishing Group, Inc.
- [3] Stapells DR (2000). Frequency-specific evoked potential audiometry in infants. In R. C. Seewald (Ed.) A Sound Foundation Through Early Amplification. Basel: Phonak AG. pp. 13-31.
- [4] Leski JM (1991). New concept of signal averaging in time domain. Annual International Conference of IEEE Eng Med Biol 13(1), p.367-368.

Contact Information

Aaron Steinman – steinman@vivosonic.com; Isaac Kurtz – kurtz@vivosonic.com