Trust but Verify – A Case History

John Van Baren, Vibration Research, Jenison, Michigan

During the initial stages of a random vibration test at full level, some controllers present an inaccurate power spectral density (PSD) by combining both low- and high-level measurements in their PSD calculations. By nature, random averaging requires time to present in-tolerance and smooth spectral lines after level changes. This article discusses both the correct and incorrect methods of random averaging and explains why Vibration Research developed and patented a new random PSD averaging technique to protect customer products.

Trust but verify. An age-old Russian adage often employed by President Ronald Reagan in the 1980s. A proverb that, used properly, can serve us well in many areas of life. But what does it have to do with vibration testing?

A few years ago, we had an opportunity to provide a demonstration of our product to the manager of a vibration testing lab in the aerospace/defense industry. He was in need of another vibration controller and it was immediately obvious that he was a sophisticated operator of vibration testing equipment; very knowledgeable and experienced. Ours would not be the only vibration controller in the lab, but he was impressed with our system's capabilities and chose to requisition his newest controller from our company.

Shortly thereafter, the customer contacted us with a pointed question. He was running a short-duration, high-level random test with two different controllers in a comparison test. He wanted to verify the accuracy of our vibration controller compared to his previous controller. He was seeing different results.

The previous controller was displaying all random spectral lines in tolerance immediately after making the transition from low level to the full test level. But the VR controller display showed many random spectral lines out of tolerance, some high, some low, for some time after the transition to full level. Reasonably, the customer was confused: "Why can't your controller control this test as well as my existing controller?"

Investigation

It wasn't a new question. Our controller resets the averaging (it does not use old low-level data) when transitioning from low level to full level, so a few other customers had noticed this same phenomenon when switching controllers.

The customer believed the existing controller's PSD display when it immediately showed all lines in tolerance (although this is a statistical impossibility[1]). So we asked him to "trust but verify." We asked him to connect an independent signal analyzer to the response accelerometer during a repeat test with the existing controller. He was startled by the results. The signal analyzer results, supporting our controller's readings, indicated that indeed there *were* spectral lines outside of tolerances at a certain frequencies during the initial stages at full level. But the existing controller's results did not match; it showed all spectral lines completely smooth and in tolerance during this time.

To explain what happened, we need to describe how random averaging works, as well as the different techniques utilized by controllers to present the PSD.

DOF – FFT to PSD

During a random test, time-domain data are transformed into frequency-domain data using the fast-Fourier transform (FFT). Control systems collect time-domain data and divide them into equally sized time periods (frames). The time-domain data from each frame are transformed using the FFT, and the power density is computed at each frequency to a get PSD estimate for that frame. Multiple frames of data are processed and averaged together to get a better PSD estimate for the signal. As more frames are included in the average, the randomness averages out, and the PSD estimate improves, as illustrated in Figure 1. For linear averaging of independent frames of data, the amount of roughness on the PSD estimate is described by the chi-squared distribution. The statistical degrees of freedom (DOF) of the chi-squared distribution will be two times the number of frames included in the average.[2]

Modern controllers enable the test engineer to set the degrees of freedom (DOF); that is, the amount of averaging to be done when estimating the PSD of the random signal. The higher the DOF setting, the more data included in the average, which translates to a smoother estimate. A typical random test might use 120 DOF, meaning that with linear averaging, there must be 60 independent frames of time-domain data available to average together to achieve a PSD estimate of the desired smoothness. The astute reader will note that windowing, overlap processing, and exponential averaging will add nuances to this example. For the sake of illustrating the characteristics of the averaging, we will limit this discussion to linear averaging with no overlap. Suffice it to say here that the characteristics described for simple linear averaging also hold when including windowing, overlap processing, and exponential averaging.

DOF – Traditional Averaging Process

time our customer performed his controller comparison.

During the random averaging process, the estimated PSD is updated on the user display each time another frame of data has been averaged in to the set. Immediately after the transition to full level there will be only a limited amount of data available at that level (see Figure 2). The controller uses the available data to estimate the PSD, but with only a few frames to average, the estimate will be quite ragged. As more data are measured, more averaging will be done, and the estimate will get progressively less ragged until it reaches a point where a total of 60 frames (120 DOF) have been averaged together.

After the test has reached the required number of frames in the set based on the DOF setting (in our example, 60 frames, which yields 120 DOF), the controller will continue to average data maintaining a fixed amount of averaging. As new data are added into the average, old data will be removed, the average recalculated, and the estimated PSD updated on the display, as shown in Figure 3. From that point onward in the test, there is not a lot of volatility, as most of the data are the same from one average calculation to the next, with only one frame of new information included on each update.

Variance

As we have shown, a significant amount of time is required to gather the measurements that are averaged together to estimate the PSD.[3] As the DOF is increased, the amount of averaging required also increases, and it will take more time for the estimated PSD to reach the desired DOF. Without overlap, for example, a test with a maximum frequency of 2,000 Hz and 1,600 lines of resolution would measure one frame of data every 0.8 seconds. So it will require 48 seconds to measure the 60 frames of data required to achieve 120 DOF. (With overlap processing, this time can be reduced by about half, but the point remains – it still takes significant time to average.)

In the time after the transition to full level before enough data are available to average the full 120 DOF, the displayed PSD estimate will begin with a very ragged appearance, since there are just not enough frames of data averaged in yet to display the desired DOF. A test certainly will appear to be out of a 1.5-dB tolerance for some time after the transition due to the limited number of frames available to average. The variance of the estimated PSD will decrease (curve becomes smoother) gradually as more data are available to average and the DOF increases (see Figure 2b).

Test engineers know that DOF is inversely proportional to the

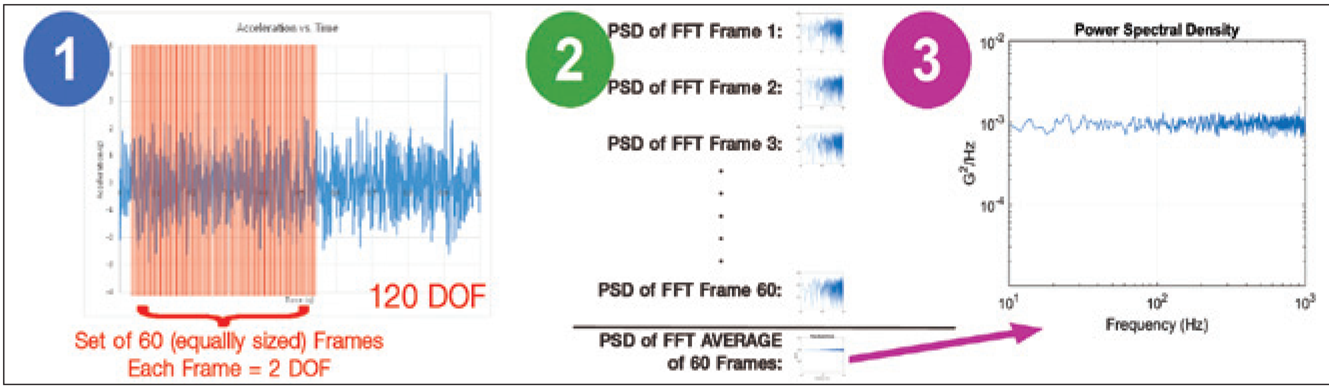

Figure 1. Frames of data are collected (1), transformed, calculated and averaged (2), and then displayed in actual PSD (3).

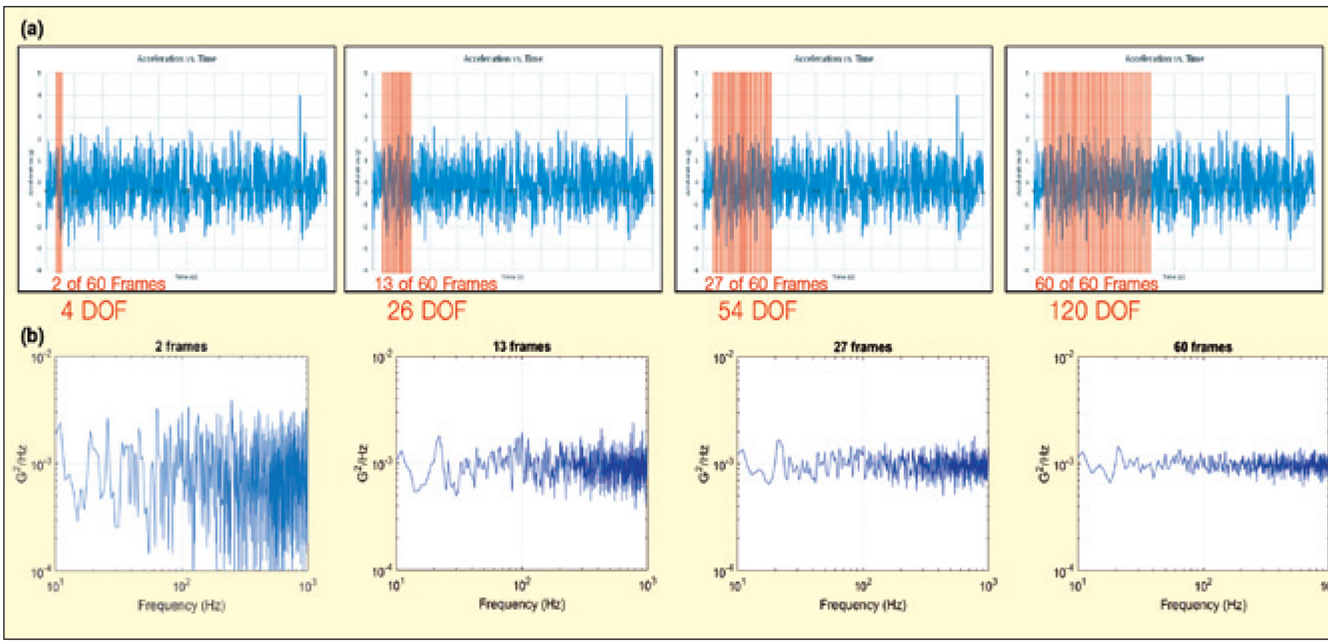

Figure 2. (a) Test builds to total desired DOF; (b) As more frames are averaged, PSD estimate improves.

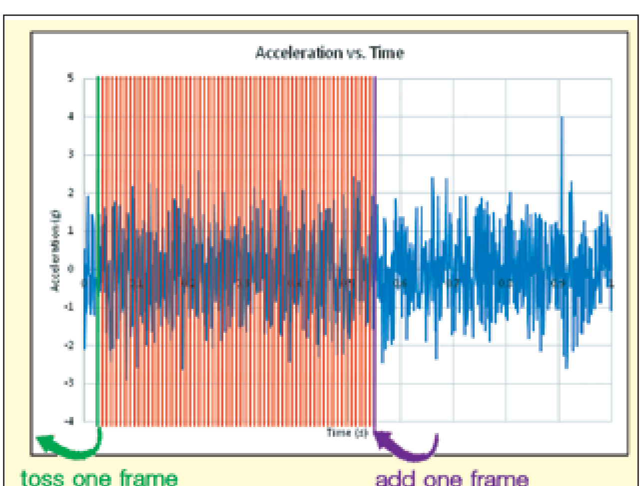

Figure 3. After reaching desired DOF, set no longer expands in size; as new frames are added, old are removed.

variance of the test. A very low DOF means high variance and a high DOF means low variance.

Since the test is random, this behavior can be quantified statistically.[1] This quantification is tabulated in Table 1 as the percentage of lines expected to be high and percentage expected low for a given dB tolerance level and given amount of averaging (expressed as DOF). In our example, with 120 DOF of averaging and a ±1.5 dB tolerance, this table shows that one would expect 0.199% of the 1600 lines, or three lines to be above the +1.5 dB tolerance line and 0.647% of the 1600 lines, or 10 lines to be below the –1.5 dB tolerance line.

It is important to understand that this does not mean the same lines will be outside the tolerance band continuously, but it does mean that at any given time, there will be some set of lines that lie outside of the ±1.5 dB tolerance line. To resolve this, we need to either use a wider tolerance or apply more averaging.[1]

Common Industry Methods

It is typical within the vibration control industry for controllers to retain their averaging while going through a step change in level. In doing so, they continue to present a PSD estimate that has the smoothness of 120 DOF of averaging immediately after a change in level. This allows them to show a PSD trace that is within tolerance immediately after the change in level.

Some of our customers even indicated that this was the desired behavior. They would run a test at 10% of the full level, wait for the averaging to produce a trace within the tolerance lines and then step to full level to execute the test. By doing so they are able to produce plots that appear to be within the tolerance lines from the very beginning through to the end of the test.

As we have already demonstrated in our example, there is not enough full-level data available to produce 120 DOF of averaging until 48 seconds after the change in level. So if the full-level data are not sufficient, where does the additional averaging come from? It may come as a surprise to some engineers, as it did to our customer, that during the initial stages of a test at full level, some controllers display a PSD estimate composed of scaled low-level data combined with the new frames of full-level data. The old, low-level frames of data are multiplied by some factor (determined by the ratio between the full-level and low-level values) and included as if it were full-level data in the full-level average.

This may look like a good thing, since the trace displayed on the screen is smooth and within tolerance, but what it really means is that the displayed PSD is not accurate. The full-level data are diluted by the data manufactured from the low-level data. Immediately after the change in level, most of the information presented in the PSD trace is scaled low-level data, with only a small number of frames of full-level data. As we will see, this can hide potentially damaging responses, and you would not know it.

What's the Difference?

To achieve the goal of accurately displaying what is occurring with the product under test, it is necessary to discard measurements taken at low level and begin averaging anew once the system has reached full level. This is known as "resetting" the averaging.

What is the difference between resetting the averaging, and the method which maintains the averaging by scaling low level data? Resetting the average and starting fresh with new frames of data, as new full-level data are measured, provides the user with a current view of what's happening to the product immediately after the level change. If the world were linear, the product response would be the same at high level as at low level. But as the test engineer is aware, the world is not linear. When the level is changed, product resonances can shift in both frequency and amplitude. If the averaging is not reset at the change in level, then the smooth curve provided by the low-level data will mask these resonance shifts.

Can the test engineer trust multiplied low level data? We have a saying, "You don't know what you don't know." The test engineer is nearly as blind to what is happening to the product during this time as if he had no display. While the multiplied low-level display looks very nice on a report, the test engineer in reality *is not seeing* true product response during this time. In our example of 120 DOF, for up to almost 48 seconds the test engineer wouldn't really know what is happening to the product on the shaker.

- What if there was a shift in a resonance? The test engineer doesn't know.
- Are any lines out of tolerance? The test engineer doesn't know.
- Are any lines outside of abort limits? Should the abort button be pressed right now to protect that high value product? The test engineer doesn't know.
- Worse! The test engineer is led to believe that he *does* know what's happening to the product during this time when, in reality, perhaps the product is being beaten to death.

We find that this is an especially critical issue when running short-duration, high-level screening tests on high-value items – such as those the aerospace industry might use in satellite or rocket testing. If the averaging is not reset at the level transition, then it is completely possible to over- or under-test products during this time, leading to either potential damage during the test or predisposing the item to a field failure without even being aware that this had happened.

It is interesting to note that a recent rocket launch failure post analysis resulted in the report that a structural support member had failed during the launch, causing catastrophic failure. The engineers were "surprised" because the part was designed to support five times the load it should have needed. Could the pre-launch test have broken the part during testing? You would never know.

How long is actual product response at full-level concealed? It depends on the test parameters, including DOF settings, which define the number of frames required to calculate the PSD trace displayed to the operator. And consider that while it may initially be way out of tolerance, the controller will begin to work on the condition, bringing the average back down. So by the time the controller does compute a PSD trace from full-level data, it is likely that the error has been corrected, and the operator is not even aware that it was initially out of tolerance.

Product Response at Different Levels

It is wrong to assume that a product will respond exactly at higher amplitude levels as it would at lower amplitude levels. Since the signal is increasing in power, the product is undergoing a major change, including shifts in resonant frequencies. Test engineers have run sine sweeps with different products and they realize that resonances are not linear with amplitude. Why accept this erroneous assumption in random testing?

Our customer's test proved this out. Immediately after the step change, our controller's PSD estimate showed that at some frequencies the trace was outside tolerance limits, and a separate signal analyzer bore this out. The resonances were controlled appropriately during the initial low-level portion of the test, but after the level change, the product response changed. The shift in resonance frequencies resulted in some control error for a while after the level change until the controller adjusted and brought the response back to the demand level. This control error was readily evident in the display, because the averaging had been reset. When

Table 1. This shows that at 120 DOF, statistically we expect 0.199% of lines to be above +1.5 dB, and 0.647% of lines to be below –1.5 dB.

DOF =	80	100	120	140	160	180	200	220	240	260	280	300
+3 dB	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%
+2 dB	0.068%	0.018%	0.005%	0.001%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%
+1.5 dB	0.892%	0.418%	0.199%	0.095%	0.046%	0.022%	0.011%	0.005%	0.003%	0.001%	0.001%	0.000%
+1 dB	5.867%	4.099%	2.892%	2.056%	1.470%	1.055%	0.761%	0.550%	0.399%	0.290%	0.211%	0.154%
+0.5 dB	21.342%	19.032%	17.058%	15.350%	13.856%	12.541%	11.375%	10.337%	9.409%	8.576%	7.827%	7.152%
-0.5 dB	25.424%	22.628%	20.283%	18.276%	16.534%	15.006%	13.655%	12.452%	11.377%	10.411%	9.541%	8.755%
-1 dB	8.870%	6.426%	4.708%	3.478%	2.586%	1.932%	1.450%	1.091%	0.824%	0.623%	0.473%	0.359%
-1.5 dB	2.214%	1.188%	0.647%	0.356%	0.197%	0.110%	0.062%	0.035%	0.020%	0.011%	0.006%	0.004%
-2 dB	0.402%	0.146%	0.054%	0.020%	0.008%	0.003%	0.001%	0.000%	0.000%	0.000%	0.000%	0.000%
-3 dB	0.006%	0.001%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%

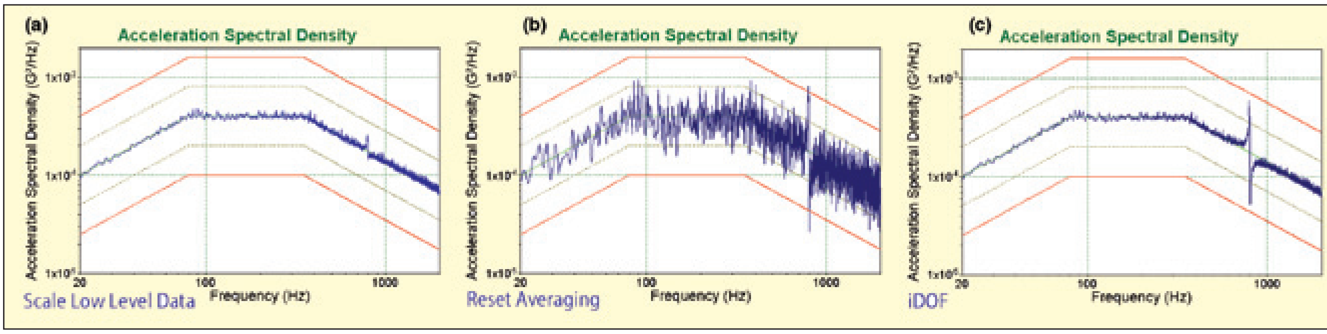

Figure 4. Comparison of PSD estimates after switching from low level to full level: (a) using low level data; (b) averager reset at level change; and (c) iDOF.

run on their other controller, the mix of low and high level data hid the actual change in product response, and they were not even aware of what was happening.

Awareness

While our customer had not been aware of the method being employed on his existing controller, he quickly realized that our traditional averaging method was providing more relevant information, and he understood the math underpinning the PSD. No controller can operate outside the parameters of math, time, or statistics. The desire or expectation that all lines in a PSD plot meet tolerance limits at all times, or that the plot immediately (without any averaging time) be a neat, clean, non-ragged and in-tolerance curve, is unrealistic with the traditional averaging method.[1]

Our customer understood that if the PSD seems to be exhibiting such behavior, the underlying methodology of that PSD display should be re-evaluated; or that if a test specification requires such behavior, that specification ought to be re-evaluated (or interpretation reviewed for clarity).

For example, a test specification may define a 1.5 dB tolerance and a test duration of one minute. From Table 1, we can determine a requirement that a PSD averaged over the full duration of a one-minute-long test be within ±1.5 dB is a reasonable expectation. However, expecting the PSD shown on the controller display to be within that same ±1.5 dB tolerance from the first second of the test to the last second of the test is not reasonable, since it takes a large portion of that minute simply to accumulate the data to average and apply that tolerance.

However, explanations aside, our customer had his own customer to satisfy. His customer would not accept a test report showing any lines out of tolerance, even during the initial averaging period. The specification was interpreted as requiring every line within tolerance for every second of the entire test duration. His customer did not want to hear, "The math doesn't allow it." His customer simply wanted a report showing the test passed with all lines continuously in tolerance. Perhaps this expectation was the impetus for the inaccurate PSD estimation currently employed by other manufacturers. The need to pass a test can sometimes overshadow the original intent of the test.

Solution – iDOF

Presented with this dilemma, our engineers asked the question, "Is there any way to satisfy such a requirement without using the low-level pretest data in the full-level averaging calculation?" And as engineers tend to do, they came up with a solution – iDOF®, or Instant Degrees of Freedom®.

With this innovative, patent-pending algorithm, the controller can provide a precise e*stimated* PSD with as few as five frames of data. The controller is able to display smooth curves early in a test and do so using only measurements made at full level. It does this by recognizing the difference between estimation error and Control error. The estimation error in the PSD trace is an undesired characteristic of the estimation algorithm, typically understood as the variance, or roughness, of the estimated PSD; the quantity that is reduced through averaging to achieve a smoother curve. It is an error in the perception of the PSD. The PSD is "out of focus." By predicting and correcting for the estimation error, the iDOF algorithm is, in effect, a pair of glasses that corrects our vision and allows the test engineer to clearly and accurately see the PSD of the signal.

On the other hand, we have Control error, which is the true difference between the PSD of the signal and the Demand PSD required by the test specification. It is the Control error that the test engineer is most concerned with, because it shows actual over- or under-testing of the product.

By removing estimation error and providing a clear view of the Control error, iDOF enables the test engineer to make appropriate and informed decisions on how the test is affecting the product. Is it indeed within tolerance? Is it significantly over-testing or under-testing the product? The details that previously were obscured either through averaging with low-level data or by the estimation error due to low averaging, now become abundantly clear.[3]

Figure 4 compares results using the three averaging methods, demonstrating how iDOF reveals details in the response that are obscured in the other methods. In addition, since the iDOF algorithm allows one to achieve smoother PSD estimates in a shorter time, it can be used to provide a clear view of the Control error using many fewer frames of data than traditional averaging tech-

Table 2. iDOF provides smoothest spectral lines while maintaining accuracy at level.

		Averaging Completed:			
		BEFORE	BEFORE	AFTER	
AVERAGING TECHNIQUE	LEVEL	Variance	DATA	Variance	Why?
COMMON IN INDUSTRY	LOW (Eg: -20 dB)	High	ACCURATE ✔	Low	Averaging based on user defined DOF takes time.
COMMON IN INDUSTRY	FULL (0 dB)	Low	INACCURATE! ⚠	Low	Averaging NOT RESET. Initial PSD is multiplication of low level measurements (inaccuracte assumption of product response at different levels), then converges into Actual PSD.
VR TRADITIONAL	FULL (0 dB)	High (Supported by Statistics)	ACCURATE ✔	Low	Averaging IS RESET. Averaging based on user defined DOF takes time. Actual PSD of level is always displayed.
iDOF™	FULL (0 dB)	Lowest (Smoothest Lines!)	ACCURATE 👍	Lowest (Smoothest Lines!)	Averaging IS RESET. Estimated PSD based on full level data is displayed, converges into Actual PSD. IDOF™ algorithm projects the defined DOF instantly and accurately.

niques. For example, 10 frames of data can be processed with the iDOF algorithm to produce a PSD estimate for which traditional averaging would require 500 frames. This can be valuable in detecting changing responses such as shifting resonances (due to a product beginning to fatigue) quickly without the long averaging times associated with high DOF (see Table 2).

Conclusion

Is your controller giving you a clear picture of what's happening to your product(s) on a shaker during the initial stages of a test at full level? You could take someone's word for it. Or you could trust but verify.

Supplements

See how common industry methods of random averaging completely overlook resonances at: http://www.vibrationresearch.com/university/lesson/signal-averaging-dangers

What is the probability of your random test satisfying tolerances based on DOF, lines of resolution, and max frequency? Check out our DOF calculator at http://go.vibrationresearch.com/download-dof-calculator

References:

1. Van Baren, Philip, "Statistical Properties of the Random PSD," Vibration Research Technical Paper, http://www.vibrationresearch.com/university/knowledge-base/statistical-properties-of-the-random-psd/.
2. Lang, George, and Van Baren, Philip, "Does Your Random Controller Square With Chi?" Vibration Research Corporation Technical Paper, http://www.vibrationresearch.com/articles/Square_with_Chi.html.
3. Van Baren, Minderhoud, and Maatman, "An Exploration of Power Spectral Density (PSD) Estimation," Vibration Research Corporation Technical Paper, http://www.vibrationresearch.com/university/an-exploration-of-power-spectral-density-estimation.

S/V

The author can be contacted at: jvb@vibrationresearch.com.