References

1. J. Lynn Smith, "Breaking the Nyquist Barrier," *IEEE Signal Processing Magazine*, July 1995, pp 41-43.
2. C. Shannon, "Communications in the presence of noise," *Proceedings of the IRE*, January, 1949, pp 10-21.
3. Oppenheim and Wilsky, *Signals and Systems*, Prentice-Hall, 1983, pp 514-526.

Questioning 'Distinction without a Difference' Debate

While I was entertained by the exchange of letters concerning fraction-of-time and probabilistic models, I was startled by the reference (*Signal Processing Magazine*, March 1995, SP Forum) to [1] as a "recent extension and refinement of an important concept..." This contention is too much for me, and I feel compelled to take [1] together with the fraction-of-time versus probablistic models debate as examples to reemphasize a broader point. Engineering is applied science, has a product, and we do not need to carry on "distinction without a difference" debates, either for time series or complex data models. Ideally, authors submitting papers to the IEEE journals have something to contribute to engineering. Contributions include reporting upon techniques that proved effective for achieving an engineering goal, comparing techniques, or explaining in a general context why techniques are effective, and identifying limitations. I see a growing disconnect between scientists and engineers, and the publications of academics, which are too often dubious mathematics.

The reference paper supplies only abstract results, and if better examples were used, it would be apparent to nearly all readers that the introduced terms are unnecessary and overblown. The paper invents jargon to compensate for a lack of a consistent definition for the common concept of "complex random number." A complex random number is nothing more than notation for pairs of random numbers, a notation which simplifies many algebraic manipulations. Measurable physical quantities, including those modeled as random variables, take on values that are real numbers. Imaginary numbers are, well, imaginary. Even in the abstract, there is no useful distinction between a complex random variable, and pairs of real random variables. There is an isomorphism between the two-dimensional Euclidean plane, in which probability measures are naturally defined, and the complex numbers. The probability distribution functions of "complex" random variables are not even functions solely of the complex variables. The probability distribution functions are functions of real combinations of the complex variables and their complex conjugates. In other words, the probability distribution functions of "complex" random variables are real functions of the real and imaginary components of the complex random variables. (And conversely. Every distribution of pairs of random variables, x, y, can be written in terms of $z = x + iy$ and $z^* = x - iy$ using $2x = z + z^*$ and $2iy = z - z^*$). The complex shorthand is convenient for algebraic manipulation of, for example, analytic signals, but this convenience does not make the results of any measurement complex. Complex random numbers are a well accepted, often convenient, notation for pairs of real random numbers. The author suggests no definition for a complex random number that requires an estimation or detection theory distinct from the well established formalisms. Even [1] notes that, "the standard procedure is to use the real and imaginary parts that are the components of a 2-D real random vector. However, doing so results in considering that complex numbers are nothing else but pairs of real numbers, and the complex theory loses most of its interest." Indeed.

The author's "complex LMSE" solutions are constrained to be complex constructs put together for notational convenience. Only by misplacing physical significance to the algebraic construct does there become any need to distinguish to types of linear estimates. A real formulation is the correct formulation for engineering problems, and consequently is the only linear estimate that need be considered. Estimates with independent coefficients for each measured physical quantity are the formulations of interest to engineering, albeit, there are common cases for which the optimal solution for the coefficients has symmetries that permit the optimal linear estimates to be written in a "complex" form. But, many solutions have distinguishable properties, and each distinguishable property of particular solutions does not require a reformulation for the general problem. Taking note of conditions that imply solutions of a particular form does lead to efficiency for solution finding. Nevertheless, alternative algorithms selected for computational efficiency are not reformulations of the problem.

For models exhibiting circularity, the constrained "complex" estimates coincide with the optimal estimates. But, circularity is only a sufficient, and not a necessary condition for the constrained solution to coincide with the optimal solution.

I am not aware of a single system that is suboptimal because a design engineer mistakingly insisted upon solving the overly constrained, "complex" optimization problem. None of the machinations introduced in the cited reference are required if there is no insistence on making too much of the complex notation for pairs of random variables. If there are notable applications of circularity, [1] chooses not to enlighten us. (And I do not count putting already solved problems into new notation.) As a young mathematician once stated, "we cannot continue to forget at the present rate. Total ignorance provides a convenient lower bound."

—Glenn Johnson
TASC, Reston, VA

Reference

1. Bernard Picobono, "On Circularity," IEEE Transactions on Signal Processing, Vol. 42, pp. 3473-3482, Dec. 1994.

Gardner Rests His Case

This is my final letter to *SP Forum* in the debate initiated by Mr. Melvin Hinich's challenge to the resolution made in the book [1], and carried on by Mr. Neil Gerr through his letters to SP *Forum*.

In this letter, I supplement my previous remarks aimed at clarifying the precariousness of Hinich's and Gerr's position by explaining the link between my argument in favor of the utility of fraction-of-time (FOT) probability, and the subject of a plenary lecture delivered at ICASSP '94. In the process of discussing this link, I hope to continue the progress made in my previous two letters in discrediting the naysayers and thereby moving toward broader acceptance of the resolution that was made and argued for in [1], and is currently being challenged. My continuing approach is to show that the position taken by the opposition—that the fraction-of-time probability concept and the corresponding time-average framework for statistical signal processing theory and method have nothing to offer in addition to the concept of probability associated with ensembles and the corresponding stochastic process framework—simply cannot be defended if argument is to be based on fact and logic.

Thomson's Waveguide Problem

To illustrate that the stochastic-process conceptual framework is often applied to physical situations where the time-average framework is a more natural choice, I have chosen an example from D. J. Thomson's recent plenary lecture on the project that gave birth to the multiple-window method of spectral analysis [2]. The project that was initiated back in the mid-1960s was to study the feasibility of a transcontinental waveguide for a telecommunications transmission system potentially targeted for introduction in the mid-1980s. It was found that accumulated attenuation of a signal propagating along a circular waveguide was directly dependent on the spectrum of the series, indexed by distance, of the erratic diameters of the waveguide. So, the problem that Thomson tackled was that of estimating the spectrum for the more than 4,000 mile long distance-series using a relatively small segment of this series that was broken into a number of 30-foot long subsegments. (It would take more than 700,000 such 30-foot sections to span 4,000 miles). The spectrum had a dynamic range of more than 100 dB and contained many periodic components, indicating the unusual challenge faced by Thomson.

When a signal travels down a waveguide (at the speed of light), it encounters the distance-series of erratic waveguide-diameters. Because of the constant velocity, the distance-series is equivalent to a time-series. Similarly, the series of diameters that is measured for purposes of analysis is— due to the constant effective velocity of the measurement device—equivalent to a time series. So, here we have a problem where there is one and only one long time-series of interest (which is equivalent to a distance-series)—there is no ensemble of long series over which average characteristics are of interest and, therefore, there is no obvious reason to introduce the concept of a stochastic process. That is, in the physical problem investigated, there was no desire to build an ensemble of transcontinental waveguides. Only one (if any at all) was to be built, and it was the spectral density of distance-averaged (time-averaged) power of the single long distance-series (time-series) that was to be estimated, using a relatively short segment, not the spectral density of ensemble-averaged power.

Similarly, if one wanted to analytically categorize the average behavior of the spectral density estimate (the estimator mean), it was the average of a sliding estimator over distance (time), not the average over some hypothetical ensemble, that was of interest. Likewise, to characterize the variability of the estimator, it was the distance-average squared deviation of the sliding estimator about its distance-average value (the estimator variance) that was of interest, not the variance over an ensemble. The only apparent reason for introducing a stochastic process model with its associated ensemble, instead of a time-series model, is that one might have been trained to think about spectral analysis of erratic data *only* in terms of such a conceptual artifice and might, therefore, have been unaware of the fact that one could think in terms of a more suitable alternative that is based entirely on the concept of time averaging over the single time-series. (Although it is true that the time-series segments obtained from multiple 30 ft. sections of waveguide could be thought of as independent random samples from a population, this still does not motivate the concept of an ensemble of infinitely long time-series—a stationary stochastic process. The fact remains that, physically, the 30-foot sections represent subsegments of one long time-series in the communications system concept that was being studied.)

It is obvious in this example that there is no advantage to introducing the irrelevant abstraction of a stochastic process (the model adopted by Thomson) except to accommodate unfamiliarity with alternatives. Yet Gerr turns this around and says there is no obvious advantage to using the time-average framework. Somehow, he does not recognize the mental gyrations required to force this and other physical problems into the stochastic process framework.

Gerr's Letter

Having explained the link between my argument in favor of the utility of FOT probability and Thomson's work, let us return to Gerr's letter. Mr. Gerr, in discussing what he refers to as "a battle of philosophies," states that I have erred in likening skeptics to religious fanatics. But in the same paragraph, we find him defensively trying to convince his readers that the "statistical/probabilistic paradigm" has not "run out of gas," when no one has even suggested that it has. No one, to my knowledge, is trying to make blanket negative statements about the value of what is obviously a conceptual tool of tremendous importance (probability) and no one is trying to denigrate statistical concepts and methods. It is only being explained that interpreting probability in terms of the fraction-of-time of occurrence of an event is a useful concept in some applications. To argue, as Mr. Gerr does, again in the same paragraph, that in general this concept "has no obvious advantages" and using it is "like building a house without power tools: it can certainly be done, but to what end?" is, as I stated in my previous letter, to behave like a religious fanatic—one who believes there can be only One True Religion. This is a very untenable position in scientific research.

As I have also pointed out in my

previous letter, Mr. Gerr is not at all careful in his thinking. To illustrate his lack of care, I point out that Gerr's statement "Professor Gardner has chosen to work within the context of an alternative paradigm (fraction-of-time probability)," and the implications of this statement in Gerr's following remarks completely ignore the facts that I have written entire books and many papers within the stochastic process framework, that I teach this subject to my students, and that I have always extolled its benefits where appropriate. If Mr. Gerr believes in set theory and logic, then he would see that I cannot be "within" paradigm A and also within paradigm B, unless A and B are not mutually exclusive. But he insists on making them mutually exclusive, as illustrated in the statement "From my perspective, developing signal processing results using the fraction-of-time approach (and not probability/statistics)..." (The parenthetical remark in this quotation is part of Mr. Gerr's statement.) Why does Gerr continue to deny that the fraction-of-time approach involves both probability and statistics?

Another example of the lack of care in Mr. Gerr's thinking is the convoluted logic that leads him to conclude "Thus, spectral smoothing of the biperiodogram is to be preferred when little is known of the signal *a priori*." As I stated in my previous letter, it is mathematically proven in [1] that the frequency smoothing and time averaging methods yield approximately the same result (a more detailed and tutorial proof of this fundamental equivalence is given in the article "The history and the equivalence of two methods of spectral analysis," in review for this publication). Gerr has given us no basis for arguing that one is superior to the other and yet he continues to try to make such an argument. And what does this have to do with the utility of the fraction-of-time concept anyway? These are data processing methods; they do not belong to one or another conceptual framework.

To further demonstrate the indefensibility of Gerr's claim that the fraction-of-time probability concept has "no obvious advantages," I cite two more examples to supplement the advantage of avoiding "unnecessary mental gyrations" that was illustrated using Thomson's waveguide problem. The first example stems from the fact that the fundamental equivalence between time averaging and frequency smoothing referred to above was first derived by using the fraction-of-time conceptual framework [1]. If there is no conceptual advantage to this framework, why wasn't such a fundamental result derived during the half century of research based on stochastic processes that preceded [1]?

The second example is taken from the first attempt to develop a theory of higher-order cyclostationarity for the conceptualization and solution of problems in communication system design. In [3], it is shown that a fundamental inquiry into the nature of communication signals subjected to nonlinear transformations led naturally to the fraction-of-time probability concept, and to a derivation of the cumulant as the solution to a practically motivated problem. This is, to my knowledge, the first *derivation* of the cumulant. All other work, which is based on stochastic processes (or non-fraction-of-time probability) and which dates back to the turn of the century, cumulants are defined, by analogy with moments, to be coefficients in an infinite series expansion of a transformation of the probability density function (the characteristic function), which has some useful properties. If there is no conceptual advantage to the fraction-of-time framework, why wasn't the cumulant derived as the solution to the above-mentioned practical problem or some other practical problem using the orthodox stochastic-probability framework?

Conclusion

Since no one in the preceding year has entered the debate to indicate that they have new arguments for or against the philosophy and corresponding theory and methodology presented in [1], it seems fair to proclaim the debate closed. The readers may decide for themselves whether the resolution put forth in [1] was defeated or was upheld. But regarding the skeptics, I sign off with a humorous anecdote:

When Mr. Fulton first showed off his new invention, the steamboat, skeptics were crowded on the bank, yelling "It'll never start, it'll never start."

It did. It got going with a lot of clanking and groaning and, as it made its way down the river, the skeptics were quiet.

For one minute.

Then they started shouting, "It'll never stop, it'll never stop."

—William A. Gardner
University of California, Davis

References

1. W. A. Gardner, *Statistical Spectral Analysis: A Nonprobabilistic Theory*. Prentice-Hall, Englewood Cliffs, NJ, 1987.

2. D. J. Thomson. "An overview of multiple-window and quadratic-inverse spectrum estimation methods," Plenary Lecture. In *Proceedings of the 1994 International Conference on Acoustics, Speech, and Signal Processing*, pp. VI-185 - VI-194.

3. W. A. Gardner and C. M. Spooner, "The cumulant theory of cyclostationary time-series, Part I: Foundation," *IEEE Transactions on Signal Processing*, Vol. 42, December 1994, pp. 3387-3408.