OPEN a textbook on statistical signal processing (SSP) and thumb through the contents. Typically we find stochastic processes, statistical independence, the Fourier transform, the Gaussian, Poisson, Wishart distributions, and Wiener filtering among the topics. There is a distinct mathematical flavor to the subject that is well-earned – radars, sonars, cellphones, and modems are testaments to the successful union of mathematical theory and engineering practice.

Now fast forward a decade or maybe two. What new mathematical content will be added to these texts? Will we find the new processes, counterparts of independence, transforms, distributions and algorithms that are rapidly emerging from random matrix theory (RMT)? What engineering applications will serve as testaments to their importance?

I believe that the body of knowledge we associate with random matrix theory is only a step or two away from being distilled into useful engineering practice. The emergence of compelling paradigms for signal processing is inevitable. Compressive sampling is the first with many more sure to follow. If the past is any indicator of the future, there will be a strong synergy between the development of the theory and applications. In the coming years I look forward to working closely with students and colleagues on bridging this gap between the requisite theory and its implementation, guided by a strong experimental emphasis.

A Fundamental Application-Independent Statistical Limit

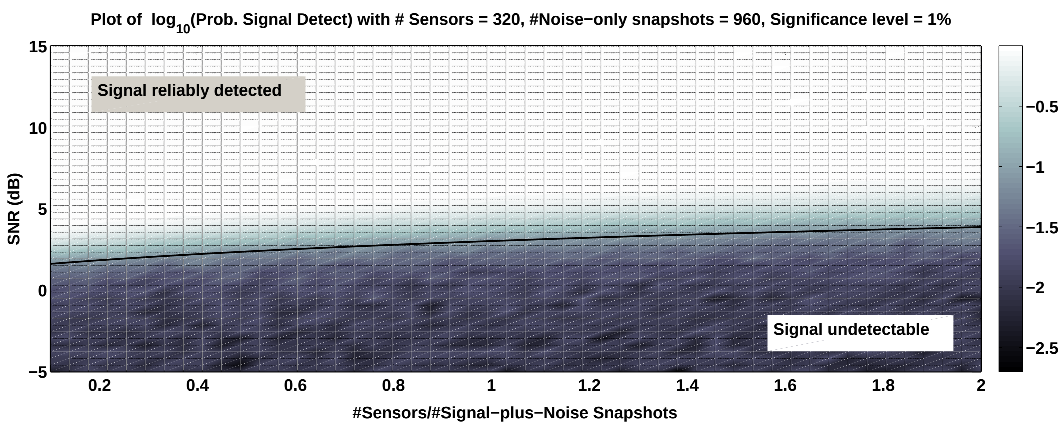

Figure 1: A signal "coverage" map in SNR versus number of signal-plus-noise snapshot phase space. In the light zone, a signal *can* be reliably detected whereas in the dark zone a signal *cannot* be reliably detected. The superimposed solid line demarcates the theoretically predicted boundary as derived in [1, 2].

The development and analysis of SSP algorithms using RMT has been my primary research focus [1, 2, 3, 4]. The detection problem in statistical signal processing can be succinctly formulated: Given m (possibly) signal bearing, $n \times 1$ signal-plus-noise snapshot vectors (samples), $\mathbf{y}_1, \ldots, \mathbf{y}_m$ and p statistically independent $n \times 1$ noise-only snapshot vectors $\mathbf{z}_1, \ldots, \mathbf{z}_p$, can one reliably infer the presence of a signal? This problem arises in the context of applications as diverse as radar, sonar, communications, bioinformatics, and machine learning and is the critical first step before signal parameters are estimated.

In [1, 2] we prove that, in a sweeping application-independent sense, if the signal-to-noise ratio (SNR) is below a critical value, that is a simple function of n, m and p, then reliable signal detection is not possible (in an asymptotic sense). If, however, the SNR is above this critical value then our new, optimal random matrix theory based algorithm will reliably detect the signal (see Figure 1) even at SNR's close to the critical value where widely used algorithms will invariably fail.

For this and other SSP problems, the role of RMT arises because the elements of the n-by-m (or n-by-p) matrix, obtained by stacking the m signal-plus-noise (or p noise-only) snapshot vectors alongside each other, are random variables. It is surprising that RMT tells us something new and unexpected about what can and cannot be done for a problem that current textbooks present as "solved." In recent years, however, RMT has produced other such "surprises" [8, 9] that have fundamentally transformed core electrical engineering theory and practice. This is just the beginning. The trend is accelerating. Larger forces are at play.

Opportunities

Advances in VLSI technology are making sensors cheaper and easier to deploy. Consequently, the sensor count in applications such as radar, sonar, communications, and ad-hoc networks continues to increase. The associated deluge of data provides problems and opportunities. The problematic fact is that widely-used classical multivariate "optimal" SSP algorithms are far from optimal when processing data from a large number of sensors (as in the detection problem earlier). Simply put, the high dimensionality (= number of sensors) of modern systems violates the fundamental assumption of low dimensionality that is hard-wired into the derivation of most current algorithms. This provides us with a tremendous opportunity to use RMT to solve high-dimensional SSP problems that arise in these applications [10].

In particular, I envision a prominent role for *free probability* theory [11] in designing new algorithms that are optimal in high-dimensional settings. Free probability provides the natural probabilistic framework for understanding the average behavior of high-dimensional random matrices. What makes it especially relevant to our situation is its introduction of a new kind of non-linear convolution operators and corresponding transforms that describe the behavior of random matrices in a manner analogous to how filtering of scalar random variables is described by the familiar convolution operator and the Fourier transform. I have developed a computational theory and software package that, for the first time, allows users to make complicated computations using free probability theory [5, 6]. This paves the way for designing optimal high-dimensional algorithms. Our work [7] represents initial progress on that front. More will surely follow.

Testing these algorithms using real-world data sets and experiments will be critical and will ensure that the algorithms we develop are relevant for engineering practice. Sharing these data sets with the broader academic community will stimulate collaboration and provide concrete benchmarks. We can even take it further with a Netflix Prize-like challenge in SSP. This would engage and energize the broader academic community. Striving for a more perfect union of engineering theory and practice remains a worthy goal.

Random Matrix Theory Inspired Design Paradigms

The connection between random matrix theory and the statistical physics of interacting particle systems intrigues me and suggests a possible engineering design paradigm. First, a brief detour.

The theory of stochastic processes provides us with a language for describing how random events occur. In statistical physics terms, the occurrence times may be interpreted as locations of particles placed along a line. The familiar Poisson process is characterized by the observation that the particles do not interact; consequently, the probability of clustering (in time or space) is high. In some systems, however, such clustering might be undesirable so that the particles will want to interact and *repel* each other. For example, cars driving along a smoothly flowing freeway during rush hour will tend to interact and space themselves out more evenly. Similarly, individually owned and operated commuter buses will tend to repel each other to maximize ridership.

At the level of a single car (or bus), the nearest neighbor spacing depends on the idiosyncrasies of the driver (or traffic) and is inherently random. Remarkably, real-world data for the examples mentioned suggests that the average behavior is described by random matrix theory. Specifically, the resulting spacing statistic matches the Gaudin distribution arising from random matrix theory that models this repulsion in interacting particle systems. This and other distributions from random matrix theory, such as the Tracy-Widom distribution for extreme values, are becoming increasingly important as they are starting to repeatedly show up in myriad places in mathematics and physics [12].

This raises the obvious question: Is there some "natural" optimality criterion associated with these distributions and processes? Research suggests that there might be a maximum entropy-like criterion except that this mysterious entropy will necessarily be different from Shannon's entropy. Indeed it must because it will have to account for the interactions between the individual particles. There is much research to be done on this front. I believe that this direction of research could prove fruitful and yield some interesting answers to applications such as multi-agent stochastic optimization and robotic swarm coordination where such a repulsive interaction might be beneficial.

I strongly believe that further development of RMT for SSP applications will not only provide interesting solutions within SSP itself but will also provide new paradigms for core electrical engineering applications. I look forward to embarking on this journey.

REFERENCES

[1] R. Rao Nadakuditi and J. W. Silverstein, "Fundamental limit of sample eigenvalue based detection of signals in colored noise using relatively few samples," *Signals, Systems and Computers, 2007. Conference Record of the Forty-First Asilomar Conference on*, pp. 686–690, Nov. 2007.

[2] R. Rao Nadakuditi and J. W. Silverstein, "Fundamental limit of sample generalized eigenvalue based detection of signals in noise using relatively few signal-bearing and noise-only samples," submitted to IEEE Trans. on Signal Processing.

[3] R. Rao Nadakuditi, "Applied Stochastic Eigen-Analysis," Ph.D. dissertation, Massachusetts Institute of Technology, February 2007, Department of Electrical Engineering and Computer Science.

[4] R. Rao Nadakuditi and A. Edelman, "Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples," *Signal Processing, IEEE Transactions on*, vol. 56, no. 7, pp. 2625–2638, July 2008.

[5] N. Raj Rao and A. Edelman, "The polynomial method for random matrices," 2008, to appear in the *Foundations of Computational Mathematics*.

[6] N. Raj Rao, "RMTool: A random matrix and free probability calculator in MATLAB," `http://www.mit.edu/~raj/rmtool/`.

[7] N. Raj Rao, J. Mingo, R. Speicher, and A. Edelman, "Statistical eigen-inference from large Wishart matrices," 2008, to appear in the *Annals of Statistics*.

[8] A. Tulino and S. Verdu, *Random Matrix Theory And Wireless Communications*. Now Publishers Inc, 2004.

[9] E. J. Candès, "Compressive sampling," in *International Congress of Mathematicians. Vol. III*. Eur. Math. Soc., Zürich, 2006, pp. 1433–1452.

[10] I. M. Johnstone, "High dimensional statistical inference and random matrices," in *International Congress of Mathematicians. Vol. I*. Eur. Math. Soc., Zürich, 2007, pp. 307–333.

[11] D. Voiculescu, "Free probability theory: random matrices and von Neumann algebras," in *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)*. Basel: Birkhäuser, 1995, pp. 227–241.

[12] P. Deift, "Universality for mathematical and physical systems," in *International Congress of Mathematicians. Vol. I*. Eur. Math. Soc., Zürich, 2007, pp. 125–152.