Efficient generation of correlated random numbers using Chebyshev-optimal magnitude-only IIR filters

Alejandro Rodriguez[1,*] and Steven G. Johnson[2]

[1]*Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139*
[2]*Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139*

We compare several methods for the efficient generation of correlated random sequences (colored noise) by filtering white noise to achieve a desired correlation spectrum. We argue that a class of IIR filter-design techniques developed in the 1970s, which obtain the global Chebyshev-optimum minimum-phase filter with a desired magnitude and arbitrary phase, are uniquely suited for this problem but have seldom been used. The short filters that result from such techniques are crucial for applications of colored noise in physical simulations involving random processes, for which many long random sequences must be generated and computational time and memory are at a premium.

The generation of correlated random sequences, or "colored noise", is important for many physical simulations involving random processes [1–5], and often the required computational time and memory is a critical concern. Although many filter-based techniques have been applied to this problem [1–3, 6–11], in this Letter we point out that these methods are suboptimal: a global Chebyshev-optimum IIR filter for this problem may be designed based on techniques developed in the 1970s [12, 13].

Colored-noise generation is required for many types of numerical simulations, such as thermodynamics [2–4], laser noise and first-passage time problems [1], and chaotic dynamics [5]. In general, any numerical model involving stochastic differential equations in which there is some background distribution, nonlinearity, or external noise associated with the quantities driving the fluctuations will require the use of colored noise [1, 3, 5, 8, 11]. From a computational standpoint, the key problem is that many of these applications require large numbers of random sequences to be generated in parallel, imposing stringent performance and memory constraints on the noise generation. For example, to investigate thermal radiation, Ref. 4 introduces a stochastic partial differential equation (Maxwell's equations) involving spatially-uncorrelated random sources at every point in space with a temporal correlation spectrum given by the Planck spectrum, shown in Fig. 1. In a three-dimensional simulation, this can require millions of simultaneous colored-noise sequences, each of which could consist of hundreds of thousands of random numbers [23].

The fundamental technique for generating correlated random sequences is to start with white noise (uncorrelated random numbers) and to apply a filter whose power spectrum matches the desired correlation spectrum [2, 3, 6, 10, 11]. That is, suppose that we want colored-noise y_n with some desired correlation function $R_m = \overline{y_n y_{n+m}}$, the (discrete-time) Fourier transform of which is the correlation spectrum $R(\omega)$. We then start with white noise x_n, e.g. uncorrelated Gaussian random numbers, with zero mean and $\overline{x_n^2} = 1$, and apply

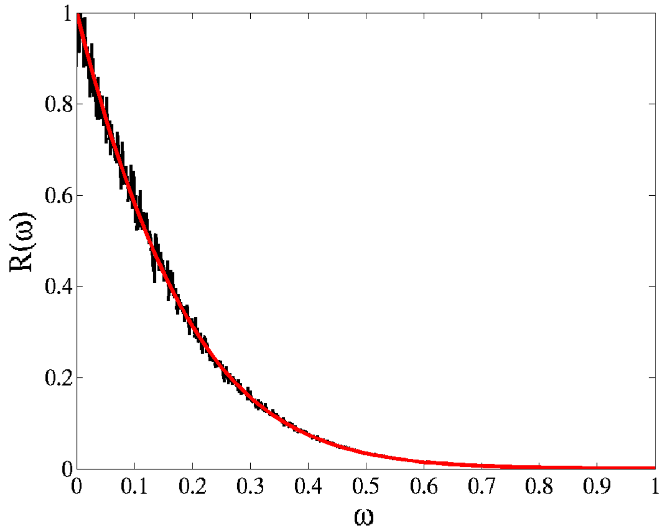

FIG. 1: Plot of Planck distribution $R(\omega) = a\omega/(\exp(a\omega) - 1)$ vs. ω (red). The periodogram (spectrum) of a finite-length correlated random sequence generated by the methods in this paper is also shown (black), computed by Bartlett's method [14].

a filter $H(\omega)$: in frequency domain, $Y(\omega) = H(\omega)\, X(\omega)$. The desired correlation spectrum is achieved if $|H(\omega)|^2 = R(\omega)$.

Although this filtering operation can be performed directly in the frequency domain via a fast Fourier transform (FFT) of the data sequences [2, 3, 11], such an approach is often too inefficient. Filtering in the frequency domain requires the entire data sequence y_n to be computed and stored in advance, and if many long sequences are required the storage becomes prohibitive. The alternative is to perform the filtering in the time domain, using either finite-impulse response (FIR) filters or infinite-impulse response (IIR) filters (also called recursive filters). Since the generation of colored noise is of interest to many researchers without a background in signal processing, let us briefly review the basics of FIR and IIR filtering, which are covered in more detail by numerous textbooks [14]. FIR filters consist of a finite-length sequence b_n that is convolved with the x_n. The Fourier transform of the b_n is the filter $H(\omega)$, but in general a finite-length sequence can only approximate an arbitrary desired spectrum $R(\omega)$. In particular, an FIR filter yields a spectrum $H(\omega)$ which is a polynomial in $z = e^{i\omega}$ with coefficients b_n. A better approximation may be obtained by generalizing to a ratio of two polynomials (i.e., rational functions), which leads to IIR filters. An IIR filter is determined by two finite-length sequences, a_n ($n = 0 \ldots N$, $a_0 = 1$) and b_n ($n = 0 \ldots M$), that determine y_n via the recurrence (which can be written in several equivalent forms):

$$y_n = \sum_{k=0}^{M} b_k x_{n-k} - \sum_{k=1}^{N} a_k y_{n-k} \,. \qquad (1)$$

That is, y_n is a convolution of b_n with x_n and of a_n with the previous values of y_n. The filter design problem is then, given filter orders N and M, to find the a_n and b_n that best approximate the desired spectrum.

Therefore, we must choose what type of filter to apply (IIR or FIR) and a filter-design method. Several choices have been previously proposed in the context of colored-noise generation. The simplest method, as we mentioned above, is to just perform a fast Fourier transform (FFT) of the entire desired spectrum [2, 3, 9, 11]. This is equivalent to an FIR filter of the same length as the data, designed by the "window" method [14]. One can also employ FIR filters of shorter lengths, designed by a variety of standard methods such as Parks-McClellan [14, 15]. For certain noise problems, an FIR filter may also be designed by analytical methods [3, 9]. In order to shorten the length of the required filter, and thus the memory and time requirements, IIR recursive filters have been proposed [3, 6, 10]. In general, the design of IIR filters is a difficult problem [14, 16], and past approaches to colored noise generation by IIR filtering have used local-optimization [10] or Yule-Walker methods [6] that are not guaranteed to yield the global-optimum filter coefficients. One important exception is exponentially correlated noise, for which an exact first-order IIR filter is known analytically and is commonly used (although typically derived from a stochastic differential equation $y' = -ay + x$ and not recognized as an IIR filter *per se*) [1, 5, 7, 8]. However, there is a key property of the IIR filter design problem for colored-noise generation that makes optimal filter design practical: the phase of the filter $H(\omega)$ is irrelevant, since it is multiplied in any case by white noise $X(\omega)$ with a random phase.

In particular, we can exploit results by Dudgeon [12] and Rabiner [13], who demonstrated that the *global* Chebyshev-optimum magnitude-only minimum-phase IIR filter-design problem can be efficiently solved by a sequence of linear-programming problems [17]. Alternative methods with similar properties have also been proposed [18, 19], and the Dudgeon and Rabiner technique was recently generalized to multidimensional IIR filters [20]. However, its applicability to the problem of colored-noise generation does not seem to have been appreciated, and in this Letter we demonstrate that it yields dramatically more efficient filters than previous approaches.

To demonstrate the efficacy of the various filter-design approaches for colored-noise generation, we consider an example drawn from thermodynamic simulations of gray-body thermal radiation [4]. In this problem, thermal effects are modeled as random fluctuating current sources everywhere in space, with a correlation spectrum $R(\omega) = a\omega/(\exp(a\omega) - 1)$, for a constant a determined by the temperature, based on the Planck distribution. This distribution is shown in Fig. 1, along with the periodogram of a finite-length correlated random sequence generated by the methods in this paper.

An IIR filter is defined by a rational polynomial function in $z = e^{i\omega}$:

$$H(\omega) = \frac{\sum_{k=0}^{M} b_k e^{i\omega k}}{1 + \sum_{k=1}^{N} a_k e^{i\omega k}} \qquad (2)$$

where N and M define the filter order and $a_0 = 1$ for convenience. A stable minimum-phase IIR filter has all of its poles and zeros within the unit circle in z (i.e., for $\operatorname{Im} \omega > 0$) [14]. Given a sequence of uncorrelated random numbers x_n, the output colored-noise sequence y_n is then given by the recurrence (1) above. The required memory, along with the computation time per output, is therefore $O(N + M)$. An FIR filter is the special case $N = 0$. A sequence y_n of length K from an FIR filter can be generated in $O(K \log M)$ time instead of $O(KM)$ by use of fast Fourier transforms, but the memory requirements are not improved.

In Fig. , we plot the L_∞ (Chebyshev) error, $\max_\omega \left| R(\omega) - |H(\omega)|^2 \right|$, as a function of the total filter order $N + M$, for filters $H(\omega)$ designed by several techniques. (Here, we employ the Chebyshev error over the entire frequency bandwidth; for other applications, only a subset of the bandwidth may be of interest.) The best method, i.e. smallest *Chebychev error* for any given-order, is the optimal magnitude-only IIR filter design (implemented using the differential-correction algorithm [12, 17]). The other design methods plotted consist of two FIR and two IIR filter techniques. The two linear-phase FIR filter techniques are (from the Matlab signal-processing toolbox): first, the Parks-McClellan algorithm, which finds the global Chebyshev optimum [14, 15] (black); second, a least-squares FIR optimization (blue), as described in Ref. 16. The two IIR

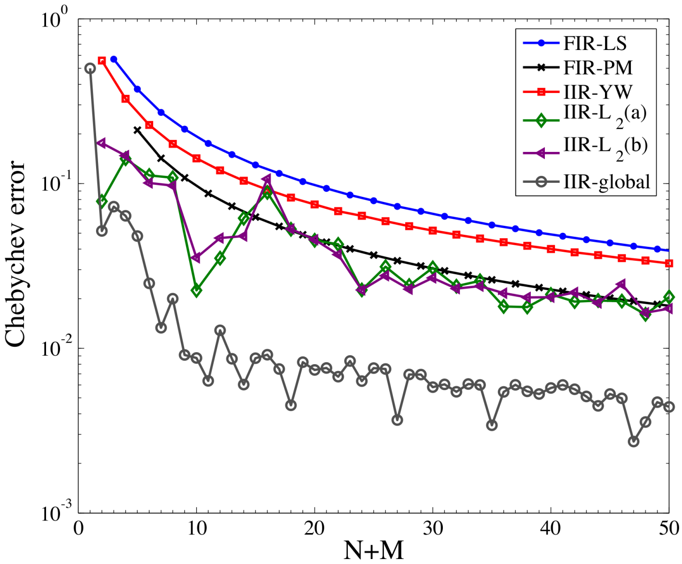

FIG. 2: Plot of Chebychev error, $L_\infty = \max_\omega |R(\omega) - |H(\omega)|^2|$, vs. $N + M$ for the Planck distribution $R(\omega)$ given in Fig. 1. The Chebychev error is plotted for six different filter methods: IIR Yule-Walker (red squares), FIR Parks-McClellan (black crosses), FIR Least-Squares (blue dots), IIR global-optimum filter (gray circles), and two nonlinear conjugate-gradient methods optimizing two different L_2 norms: (a) $L_2(|H| - \sqrt{R})$ (green diamonds) and (b) $L_2(|H|^2 - R)$ (magenta triangles).

filter designs plotted are: first, a nonlinear conjugate-gradient minimization of a least-squares norm proposed by Ref. 21, and suggested by Ref. 10 for use in generation of Rayleigh-correlated noise (gray circles); second, another (non-global) optimization technique based on the modified Yule-Walker algorithm in the Matlab signal-processing toolbox [6, 22]. The conjugate-gradient method only finds a local optimum, and is highly sensitive to the starting point of the optimization for this problem [10]. Here, we use the Yule-Walker IIR filter as the starting point, and conjugate gradient is able to improve upon it significantly (in fact, it happens to nearly match the Parks-McClellan FIR performance). We examined the conjugate-gradient minimization of two different L_2 norms $[L_2(x_{n=1\ldots N}) = \sqrt{\sum_n x_n^2/N}]$ of the error: $L_2(|H| - \sqrt{R})$ [10, 21] and $L_2(|H|^2 - R)$, although it turns out to make little difference in the result.

We believe that the Chebyshev norm is typically the most appropriate one for physical simulations involving random processes. The reason is that a large error in a narrow bandwidth, which might be allowed by a least-square (L_2) norm, could result in a spectral feature that might be mistaken for a spurious physical phenomenon; a large "spike" in error may also interact adversely with nonlinearities in the physics. Nevertheless, in Fig. and Fig. , we show two different L_2 errors, $L_2(|H| - \sqrt{R})$ and

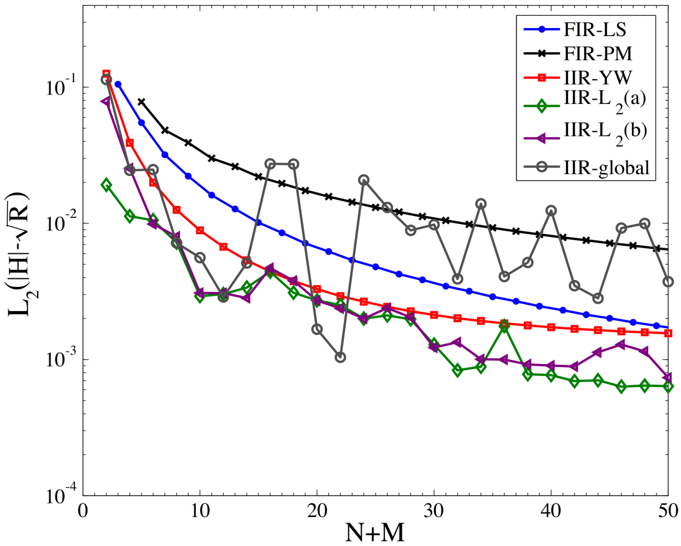

FIG. 3: Plot of L_2 error, $L_2(|H| - \sqrt{R})$, vs. $N + M$ for the Planck distribution $R(\omega)$ given in Fig. 1. The $L_2(|H| - \sqrt{R})$ error is plotted for six different filter methods: IIR Yule-Walker (red squares), FIR Parks-McClellan (black crosses), FIR Least-Squares (blue dots), IIR global-optimum filter (gray circles), and two nonlinear conjugate-gradient methods optimizing two different L_2 norms: (a) $L_2(|H| - \sqrt{R})$ (green diamonds) and (b) $L_2(|H|^2 - R)$ (magenta triangles).

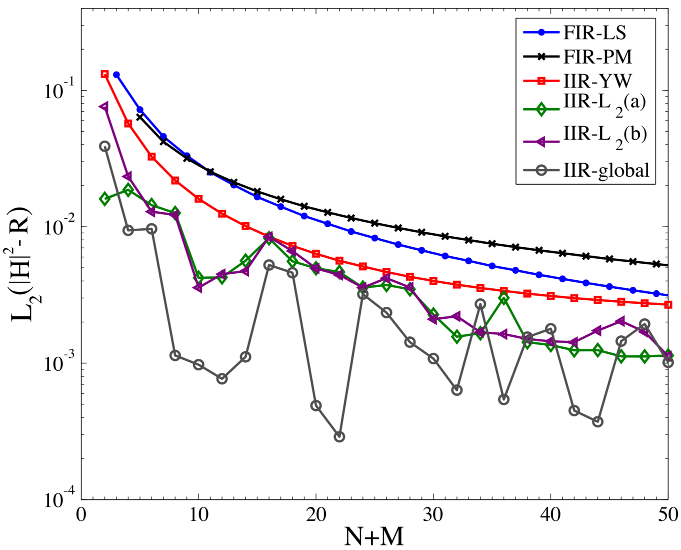

FIG. 4: Plot of L_2 error, $L_2(|H|^2 - R)$, vs. $N + M$ for the Planck distribution $R(\omega)$ given in Fig. 1. The $L_2(|H|^2 - R)$ error is plotted for six different filter methods: IIR Yule-Walker (red squares), FIR Parks-McClellan (black crosses), FIR Least-Square (blue dots), IIR global-optimum filter (gray circles), and two nonlinear conjugate-gradient methods optimizing two different L_2 norms: (a) $L_2(|H| - \sqrt{R})$ (green diamonds) and (b) $L_2(|H|^2 - R)$ (magenta triangles).

$L_2(|H|^2 - R)$, for the same filter designs as in Fig. , and the results demonstrate that the Chebychev-optimal IIR filter is at least comparable, and often superior to, the other methods, even in norms that it does not strictly optimize. In particular, the $L_2(|H|^2 - R)$ norm of the Chebyshev-optimum IIR filter does better than all other local optimization techniques for most $N + M$.

ACKNOWLEDGEMENTS

We are grateful to Alan V. Oppenheim, Stephen Boyd, and Almir Mutapcic for helpful discussions. This work was supported in part by the Materials Research Science and Engineering Center program of the National Science Foundation under award DMR 02-13282, by a Department of Energy (DOE) Computational Science Fellowship under grant DE-FG02-97ER25308, and also by the Paul E. Gray Undergraduate Research Opportunities Program Fund at MIT.

---

* Electronic address: `alexrod7@mit.edu`

[1] R. F. Fox, I. R. Gatland, R. Roy, and G. Vemuri, Phys. Rev. A **38**, 5938 (1988).
[2] K. Y. Billah and M. Shinozuka, Phys. Rev. A **42**, 7492 (1990).
[3] J. N. Kasdin, Proc. IEEE **83**, 802 (1995).
[4] C. Luo, A. Narayanaswamy, G. Ghen, and J. D. Joannopoulos, Phys. Rev. Lett. **93**, 213905 (2004).
[5] A. Traulsen, K. Lippert, and U. Behn, Phys. Rev. E **69**, 026116 (2004).
[6] Y. K. Chan and R. W. Edsinger, IEEE Trans. Automatic Control **26**, 676 (1981).
[7] R. Manella, Phys. Rev. A **40**, 3381 (1989).
[8] J. Garcia-Ojalvo, Phys. Rev. A **46**, 4670 (1992).
[9] D. J. Young and N. C. Beaulieu, IEEE Trans. Comm. **48**, 1114 (2000).
[10] C. Komninakis and J. F. Kirshman, Satellite Communications pp. 24–34 (2004).
[11] K. Lu and J.-D. Bao, Phys. Rev. E **72**, 067701 (2005).
[12] D. E. Dudgeon, IEEE Trans. Acoust, Speech and Sig. Proc. **22**, 443 (1974).
[13] L. R. Rabiner, N. Y. Graham, and H. D. Helms, IEEE Trans. Acoust, Speech and Sig. Proc. **22**, 117 (1974).
[14] A. V. Oppenheim and R. A. Schafer, *Discrete-Time Signal Processing: Second Edition* (Prentice-Hall, Englewood Cliffs, NJ, 1999).
[15] L. R. Rabiner, J. H. McClellan, and T. W. Parks, Proc. of the IEEE **63**, 595 (1975).
[16] T. W. Parks and C. S. Burrus, *Digital Filter Design* (John Wiley and Sons, New York, 1987).
[17] S. Crosara and G. A. Mian, IEEE Trans. Circ. and Sys. **30**, 898 (1983).
[18] A. Alkhairy, IEEE Trans. Circ. and Sys. II **42**, 618 (1995).
[19] X. Zhang and H. Iwakura, IEEE Trans. Sig. Proc. **44**, 1325 (1996).
[20] D. Gorinevsky and S. Boyd, IEEE Trans. Circ. Syst. I **53**, 372 (2006).
[21] K. Steiglitz, IEEE Trans. Audio and Electroacoustics **18**, 123 (1970).
[22] B. Friedlander and B. Porat, IEEE Trans. Aero. and Elec. Sys. **20**, 158 (1984).
[23] In the specific problem considered in Ref. 4, the correlation function could be renormalized to an uncorrelated sequence due to the linearity of the system, but this trick would not work in nonlinear or nonequilibrium systems.