L^1-Norm Convergence Properties of Correlogram Spectral Estimates

G. Casinovi

IEEE Transactions on Signal Processing
vol. 55, no. 9, pp. 4354–4365, September 2007

Abstract

This paper establishes the following results concerning the estimation of the power spectrum of a single, deterministic, infinitely long signal. a) If $\hat{s}_x$ is the signal's power spectral density, correlogram spectral estimates obtained from increasingly longer signal segments tend to $\hat{s}_x * \hat{w}/2\pi$ in the L^1-norm, where $\hat{w}$ is the Fourier transform of the window used to generate the estimates. b) The L^1-norm of $\hat{s}_x - \hat{s}_x * \hat{w}/2\pi$ can be made arbitrarily small by an appropriate choice of window. It is further shown that the accuracy of the spectral estimates generated by a given window is related to a newly introduced function, termed the windowing error kernel and that this function yields bounds on the asymptotic error of the estimates. As an example, correlogram spectral estimates are used to analyze spectral regrowth in an amplifier.

Copyright Notice

L^1-Norm Convergence Properties of Correlogram Spectral Estimates

Giorgio Casinovi, *Senior Member, IEEE*

***Abstract*—This paper establishes the following results concerning the estimation of the power spectrum of a single, deterministic, infinitely long signal. a) If $\hat{s}_x$ is the signal's power spectral density, correlogram spectral estimates obtained from increasingly longer signal segments tend to $\hat{s}_x * \hat{w}/2\pi$ in the L^1-norm, where $\hat{w}$ is the Fourier transform of the window used to generate the estimates. b) The L^1-norm of $\hat{s}_x - \hat{s}_x * \hat{w}/2\pi$ can be made arbitrarily small by an appropriate choice of window. It is further shown that the accuracy of the spectral estimates generated by a given window is related to a newly introduced function, termed the windowing error kernel and that this function yields bounds on the asymptotic error of the estimates. As an example, correlogram spectral estimates are used to analyze spectral regrowth in an amplifier.**

***Index Terms*—Estimation, Fourier transforms, nonlinear distortion, numerical analysis, spectral analysis, windowing.**

I. Introduction

THE estimation of the power spectrum of deterministic or stochastic signals has been the object of intense research for over a century, although the focus has been almost entirely on spectral analysis of stochastic processes. The periodogram was probably the first method to be developed for this purpose [1]: it generates a spectral estimate based on the Fourier transform of segments of the signal being analyzed. The correlogram method was developed later [2]: its spectral estimates are obtained from the Fourier transform of approximations to the signal's autocorrelation function.

In addition to the periodogram and the correlogram, other approaches to spectral estimation exist, which do not rely directly on the Fourier transform. For instance, so-called parametric methods are based on specific models for the signal being analyzed and/or the system that generated it. Typically, the models contain parameters whose values are estimated from the available signal data. A power spectrum estimate is then obtained, based on the model and the computed parameter values. With a correct choice of model, parametric methods can generate more accurate spectral estimates than the periodogram or the correlogram, especially if the signal data is available only over a short time interval [3]. On the other hand, completely erroneous spectral estimates can be obtained from an incorrect choice of the model.

The particular application that prompted the research described in this paper is the estimation of regrowth in signal spectra caused by nonlinearities in power amplifiers and other components of telecommunications systems. Virtually all modulation schemes used in modern digital communications generate signals whose power spectra have a continuous component [4]. Simulating the effects of nonlinearities on signal spectra provides valuable information in the design of communications circuit, and specialized algorithms have been devised for this purpose [5], [6]. The results of the simulations, however, are time-domain signals; their power spectra must then be computed using appropriate techniques. Virtually all parametric spectral estimation methods are based on linear system models, and are therefore unsuited for this particular application. For this reason, this paper focuses on one of the methods based on the Fourier transform: the correlogram.

The specific problem being examined is the estimation of the power spectrum of a single continuous-time or discrete-time signal. No assumption is made about the signal being a realization of a particular stochastic process. It is shown that, if the signal's power spectrum is continuous, correlogram estimates can theoretically approximate it with an accuracy that depends only on the window used, provided that sufficiently long signal segments are available. It is also shown that, under broad conditions, the window can be chosen to make the difference between the correlogram estimates and the signal's power spectral density arbitrarily small. This means that, if sufficiently long signal segments and an appropriate window are used, a continuous power spectrum can theoretically be approximated by correlogram estimates with arbitrary accuracy. Finally, it is shown that the accuracy of a window can be measured in terms of a function associated to it, which is termed the windowing error kernel. This function provides a way to make both qualitative and quantitative comparisons between windows, in terms of the error that they introduce in spectral estimates.

Section II explains the notation used throughout this paper. Section III contains the formal mathematical definition of the power spectrum of a signal, and Section IV proves a number of properties regarding correlogram spectral estimates. Section V examines the effect of windows on the accuracy of spectral estimates, introduces the windowing error kernel associated with a given window, and compares a number of commonly used windows in terms of their accuracy. Finally, Section VI presents a few numerical examples illustrating the theoretical results obtained in the previous sections.

Manuscript received July 11, 2006; revised January 3, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Patrice Abry. This work was supported in part by the National Science Foundation under Grant CCR-0306343.

The author is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: Giorgio.Casinovi@ece.gatech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2007.896257

II. Notation

Most of the results presented in this paper apply both to continuous-time and discrete-time signals, and the notation has been selected to accommodate either case. Thus, the domain of the time variable is denoted by G, with the understanding that $G = \mathbb{R}$ for continuous-time signals, or $G = \mathbb{Z}$ for discrete-time signals. This choice of notation is motivated by the fact that both $\mathbb{R}$ and $\mathbb{Z}$ are locally compact, commutative topological groups [7]. Accordingly, $\int_G f\,dt$ denotes the integral of f with respect to the Haar measure on G. If $G = \mathbb{R}$, this is the usual Lebesgue integral; if $G = \mathbb{Z}$, the Haar measure of G is the counting measure, and the integral is in fact a summation: $\int_{\mathbb{Z}} f\,dt = \sum_{t\in\mathbb{Z}} f(t)$. In either case, $x * y$ denotes convolution

$$(x * y)(t) = \int_G x(\tau)y(t-\tau)\,d\tau$$

while $\hat{x}$ is the Fourier transform of x

$$\hat{x}(\omega) = \int_G x(t)e^{-j\omega t}\,dt.$$

$\hat{G}$ denotes the dual group of G [7, p. 6], that is, the domain of the frequency variable in the Fourier transform operation. Therefore, in the continuous-time case, $\hat{G} = \mathbb{R}$; in the discrete-time case, $\hat{G}$ can be identified with the interval $(-\pi, \pi]$. $\int_{\hat{G}} f\,d\omega$ denotes the integral of f with respect to the Haar measure on $\hat{G}$, which, as before, is the ordinary Lebesgue integral if $\hat{G} = \mathbb{R}$. In the discrete-time case, $\int_{\hat{G}} f\,d\omega$ is the Lebesgue integral of the periodic function of period 2π that coincides with f on $(-\pi, \pi]$. Consequently, in the discrete-time case the convolution operation

$$(f * g)(\omega) = \int_{\hat{G}} f(\theta)g(\omega-\theta)\,d\theta$$

must be understood as the circular convolution of two periodic functions of period 2π. In either case, $\mathcal{F}^{-1}$ denotes the inverse Fourier transform operation:

$$[\mathcal{F}^{-1}(\hat{x})](t) = \frac{1}{2\pi}\int_{\hat{G}} \hat{x}(\omega)e^{j\omega t}\,d\omega.$$

$L^p(H)$, with $1 \le p < \infty$ and $H = G$ or $H = \hat{G}$, denotes the space of all functions f that satisfy the condition $\int_H |f|^p\,dh < \infty$, where the integral is evaluated with respect to the Haar measure on H. $L^p(H)$ is a Banach space with the norm $\|f\|_p = (\int_H |f|^p\,dh)^{1/p}$.

III. Power Spectrum

Let x be a complex-valued, bounded, continuous-time or discrete-time signal with the following properties:

- the limit:

$$s_x(t) = \lim_{T\to+\infty} \frac{1}{T}\int_{-T/2}^{T/2} x(t+\tau)\bar{x}(\tau)\,d\tau$$

  exists for all $t \in G$;
- s_x is a continuous function of t (if $G = \mathbb{Z}$, this condition is satisfied automatically).

It was shown by Wiener [8] that s_x is a positive-definite function. A theorem by Bochner [7, p. 19] states that a continuous, positive-definite function on G can be expressed in a unique way as the inverse Fourier transform of a bounded, positive measure σ_x on $\hat{G}$, as follows:

$$s_x(t) = \frac{1}{2\pi}\int_{\hat{G}} e^{j\omega t}\,d\sigma_x(\omega).$$

σ_x is an element of $M(\hat{G})$, the space of bounded, regular Borel measures on $\hat{G}$: following Wiener's definition, it will be referred to as the *power spectrum* of x. As the inverse Fourier transform of a bounded measure, s_x is an element of $B(G)$, the Fourier–Stieltjes algebra of G (see the Appendix).

As an example, consider the following signal: $x(t) = e^{j\omega_0 t}$. Then, $s_x(t) = e^{j\omega_0 t}$, and $\sigma_x(\omega) = 2\pi\delta(\omega - \omega_0)$, where $\delta(\omega)$ is Dirac's atomic measure located at $\omega = 0$. More generally, it is fairly straightforward to verify that the power spectrum of $x(t) = \sum_k X_k e^{j\omega_k t}$ is $\sigma_x(\omega) = 2\pi\sum_k |X_k|^2\delta(\omega - \omega_k)$.

Based on the properties of σ_x, it is possible to identify several types of power spectra. For example, a measure is *discrete* if it can be expressed as the sum of a finite or countable number of Dirac atomic measures [7, p. 266]. Accordingly, a signal x is said to have a *discrete spectrum* if σ_x is a discrete measure. All quasi-periodic signals have discrete spectra, as shown by the example given above.

A measure σ_x is *absolutely continuous* with respect to the Lebesgue measure on $\hat{G}$ if there exists $\hat{s}_x \in L^1(\hat{G})$ such that $d\sigma_x = \hat{s}_x\,d\omega$. In such case

$$s_x(t) = \frac{1}{2\pi}\int_{\hat{G}} e^{j\omega t}\hat{s}_x(\omega)\,d\omega$$

i.e., s_x is the inverse Fourier transform of $\hat{s}_x$. A signal x is said to have a *continuous spectrum* if σ_x is absolutely continuous with respect to the Lebesgue measure on $\hat{G}$. In this paper, $\hat{s}_x$ will be referred to as the power spectral density (PSD) of x. It follows from this definition that the PSD of a signal is always a real-valued, non-negative function belonging to $L^1(\hat{G})$.[1] Note that x has a continuous spectrum if and only if $s_x \in A(G)$, the Fourier algebra of G (see the Appendix).

In the most general case, the power spectrum of a signal may contain both discrete and continuous components. Signals of this type are said to have a *mixed spectrum*.

IV. Correlogram Spectral Estimation

Broadly speaking, correlogram estimates of a signal's power spectrum are obtained by taking a finite-length segment of the signal and computing the Fourier transform of the segment's autocorrelation function. More precisely, let x be a signal satisfying the assumptions stated at the beginning of the previous

[1] In the literature the terms "power spectrum" and "power spectral density" are often used interchangeably. The distinction drawn in this paper mirrors the difference between "probability measure" and "probability density function" in probability theory.

section. Assuming that the values of x are available for time intervals of arbitrary length, let $\{T_n\}$ be an increasing sequence of positive numbers such that $\lim_{n\to\infty} T_n = +\infty$. Define

$$x_n(t) = \begin{cases} x(t)/\sqrt{T_n}, & |t| \le T_n/2 \\ 0, & |t| > T_n/2 \end{cases}$$
$$s_n = x_n * x_n^*$$

where $x_n^*(t) = \bar{x}_n(-t)$. Clearly, $x_n \in L^2(G)$; hence $\hat{x}_n \in L^2(\hat{G})$ and $\hat{s}_n = |\hat{x}_n|^2 \in L^1(\hat{G})$. Consequently, s_n is a positive-definite function in $A(G)$. This section and the next prove a number of results showing that, under fairly broad conditions, the sequence $\{\hat{s}_n\}$ can be used to generate increasingly accurate estimates of the power spectrum of x. The first step in this direction is the result established in the following theorem.

Theorem 1 [8]: The sequence $\{s_n\}$ converges pointwise to s_x, that is

$$\lim_{n\to\infty} s_n(t) = s_x(t) \quad \forall t \in G.$$

Proof: Since $s_n(-t) = \bar{s}_n(t)$ and $s_x(-t) = \bar{s}_x(t)$, it is sufficient to prove the claim for $t \ge 0$. In such case

$$\begin{aligned} s_n(t) &= \frac{1}{T_n}\int_{-\frac{T_n}{2}}^{\frac{T_n}{2}-t} x(t+\tau)\bar{x}(\tau)\,d\tau \\ &= \frac{1}{T_n}\int_{-\frac{T_n}{2}}^{\frac{T_n}{2}} x(t+\tau)\bar{x}(\tau)\,d\tau \\ &\quad - \frac{1}{T_n}\int_{\frac{T_n}{2}-t}^{\frac{T_n}{2}} x(t+\tau)\bar{x}(\tau)\,d\tau. \end{aligned}$$

Since $|x(t)| \le M$, the following inequality applies:

$$\lim_{n\to\infty}\left|\frac{1}{T_n}\int_{\frac{T_n}{2}-t}^{\frac{T_n}{2}} x(t+\tau)\bar{x}(\tau)\,d\tau\right| \le \lim_{n\to\infty}\frac{tM^2}{T_n} = 0.$$

Therefore

$$\lim_{n\to\infty} s_n(t) = \lim_{n\to\infty}\frac{1}{T_n}\int_{-\frac{T_n}{2}}^{\frac{T_n}{2}} x(t+\tau)\bar{x}(\tau)\,d\tau = s_x(t).$$

■

In light of this theorem, it might be surmised that the sequence $\{\hat{s}_n\}$ converges in some way to the power spectrum of x. This is indeed the case: it can be shown that σ_x is the distributional limit of $\{\hat{s}_n\}$ [9]. This type of convergence, however, is not too useful for computational purposes. To illustrate this point, consider the sequence $\hat{s}_n(\omega) = \cos^2(n\omega)$. Although this sequence converges to the constant function 1/2 in the sense of distributions, it does not provide numerically convenient estimates of a flat spectrum. For computational purposes, a more stringent type of convergence is needed: for example, convergence in the norm of $M(\hat{G})$. Additional conditions, however, are needed to ensure that $\{\hat{s}_n\}$ converges in norm to σ_x.

To understand why, note that $\hat{s}_n \in L^1(\hat{G})$; this holds true regardless of the type of spectrum of x. Since $L^1(\hat{G})$ is a norm-closed subspace of $M(\hat{G})$, $\{\hat{s}_n\}$ cannot converge to σ_x in norm if $\sigma_x \notin L^1(\hat{G})$, i.e., if x does not have a continuous spectrum. Even if $\sigma_x \in L^1(\hat{G})$, however, norm convergence of $\hat{s}_n$ to σ_x cannot be guaranteed: to the author's best knowledge, to date no formal proof of this fact exists, even for particular classes of signals. On the other hand, it will be proven below that the application of a fixed-length window to the sequence $\{s_n\}$ has the effect of generating norm-convergent spectral estimates.

A *window* is a real-valued function w that satisfies the following conditions:

1) $|w(t)| \le 1, w(0) = 1$;
2) $w(t) = w(-t)$;
3) w has finite length, i.e., there exists $T_w > 0$ such that $w(t) = 0$ for $|t| > T_w/2$;
4) $w \in A(G)$ or, equivalently, $\hat{w} \in L^1(\hat{G})$.

The last requirement in the list above merits an explanation. The condition $w \in A(G)$ is needed to satisfy the hypotheses of Theorem 6 in the Appendix, which is used in the proof of Theorem 2 below and of Theorem 3 in Section V. In the discrete-time case ($G = \mathbb{Z}$), the Fourier transform of any finite-length signal belongs to $L^1(\hat{G})$. Therefore, in this case the condition $w \in A(G)$ is automatically satisfied, because of the requirement that w must have finite length. In the continuous-time case, this is no longer true: for example, the rectangular window is not an element of $A(\mathbb{R})$. On the other hand, the Fourier transform of any finite-length signal that is sufficiently smooth belongs to $L^1(\mathbb{R})$; for example, both the triangular window

$$w(t) = 1 - 2|t|/T_w, \quad |t| \le T_w/2$$

and the Hanning window

$$w(t) = [1 + \cos(2\pi t/T_w)]/2, \quad |t| \le T_w/2$$

are elements of $A(\mathbb{R})$. So even in the continuous-time case, it is not overly restrictive to insist that a window must belong to $A(G)$.

The application of a fixed-length window w to s_n yields $s_n w$, whose Fourier transform is $\hat{s}_n * \hat{w}/2\pi$. The significance of Theorem 2 below is that, although $\{\hat{s}_n\}$ may not converge to σ_x in norm, $\{\hat{s}_n * \hat{w}/2\pi\}$ converges in norm to $\sigma_x * \hat{w}/2\pi$. Note that, if x does not have a continuous spectrum, $\sigma_x \notin L^1(\hat{G})$. Nevertheless, $\sigma_x * \hat{w}/2\pi$ is always an element of $L^1(\hat{G})$, because $\hat{w} \in L^1(\hat{G})$, and $L^1(\hat{G})$ is an ideal of $M(\hat{G})$. In nonmathematical terms, this means that $\sigma_x * \hat{w}/2\pi$ is always a continuous spectrum, regardless of whether σ_x itself is continuous or not.[2] Note also that $\sigma_x * \hat{w}/2\pi \in L^1(\hat{G})$ implies $s_x w \in A(G)$, and that $s_n w \in A(G)$, because: $\hat{s}_n * \hat{w}/2\pi \in L^1(\hat{G})$. Theorem 2 can then be stated formally as follows.

Theorem 2: Let $w \in A(G)$. Then, $\{s_n w\}$ converges to $s_x w$ in the norm of $A(G)$. Equivalently, $\{\hat{s}_n * \hat{w}/2\pi\}$ converges to $\sigma_x * \hat{w}/2\pi$ in the norm of $L^1(\hat{G})$.

Proof: Note that s_n is a positive-definite function for all values of n; therefore, $\|s_n\|_B = s_n(0)$. It follows that

$$\lim_{n\to\infty}\|s_n\|_B = \lim_{n\to\infty} s_n(0) = s_x(0) = \|s_x\|_B.$$

Being a convergent sequence, $\{\|s_n\|_B\}$ must be bounded, i.e., there exists $M > 0$ such that $\|s_n\|_B \le M$ for all n. By Lemma 1 in the Appendix, $s_x = w^*\text{-}\lim_{n\to\infty} s_n$. Consequently, $\{s_n\}$

[2]Broadly speaking, this explains why the application of a window to a signal has the effect of spreading spectral lines over a continuous range of frequencies.

satisfies all the hypotheses of Theorem 6, and the claim follows. ∎

V. Windowing Error

Theorem 2 states that $\{\hat{s}_n * \hat{w}/2\pi\}$ converges to $\sigma_x * \hat{w}/2\pi$ in the norm of $L^1(\hat{G})$. Consequently, the accuracy of the spectral estimates that can be obtained with a given window is ultimately limited by the difference between σ_x and $\sigma_x * \hat{w}/2\pi$. For this reason, the quantity $\sigma_x - \sigma_x * \hat{w}/2\pi$ will be referred to as the *windowing error*.

The observation above prompts two questions: 1) Can w be chosen to make the norm of the windowing error arbitrarily small? and 2) Is it possible to estimate in some way the error associated with a given window (e.g., to provide an upper bound on the norm of the windowing error)? These two questions are addressed in this section.

As noted earlier, $\sigma_x * \hat{w}/2\pi$ is always a continuous spectrum, even if σ_x itself is not. Since $L^1(\hat{G})$ is a norm-closed subspace of $M(\hat{G})$, it follows that $\|\sigma_x - \sigma_x * \hat{w}/2\pi\|_1$ cannot be made arbitrarily small if $\sigma_x \notin L^1(\hat{G})$, i.e., if σ_x is not a continuous spectrum. On the other hand, if σ_x is a continuous spectrum, the following theorem shows that it is possible to choose w so as to make the L^1-norm of the windowing error arbitrarily small. Recall that, if σ_x is a continuous spectrum, then: $d\sigma_x = \hat{s}_x\, d\omega$ for some $\hat{s}_x \in L^1(\hat{G})$.

Theorem 3: Let $\{w_n\}$ be a sequence of positive definite windows such that: $\lim_{n\to\infty} w_n(t) = 1 \ \forall t \in G$. Then

$$\lim_{n\to\infty} \|\hat{s}_x - \hat{s}_x * \hat{w}_n/2\pi\|_1 = 0$$

for all $\hat{s}_x \in L^1(\hat{G})$.

Proof: Since w_n is positive definite, $\|w_n\|_B = w_n(0) = 1$. Moreover, $1 \in B(G)$ and: $\|1\|_B = 1$. Consequently $\{w_n\}$ satisfies the hypotheses of Lemma 1 and Theorem 6 in the Appendix. It follows that: $\lim_{n\to\infty} \|s_x(w_n - 1)\|_B = 0$ or, equivalently, $\lim_{n\to\infty} \|\hat{s}_x * \hat{w}_n/2\pi - \hat{s}_x\|_1 = 0$. ∎

This theorem ensures that the windowing error can be made arbitrarily small simply by increasing the length of the window. The requirement that the window should be positive definite, however, might prove to be rather restrictive, because many commonly used windows do not satisfy it. It will be shown later, however, that this condition can effectively be dropped, at least in the discrete time case.

Theorem 3 does not provide an answer to the question of how to quantify the windowing error associated with a specific spectral estimate, i.e., the value of $\|\hat{s}_x - \hat{s}_x * \hat{w}/2\pi\|_1$. It will be shown next that an upper bound on this error can be obtained for continuous spectra ($d\sigma_x = \hat{s}_x\, d\omega$), under the additional condition that $\hat{s}_x$ is a function of bounded variation; it can be safely assumed that this condition is always satisfied in practical applications. The bound on the windowing error involves a function, denoted by $\hat{e}_w$, that is associated with the window being used, and which will be referred to as the *windowing error kernel*.

In the continuous-time case, the windowing error kernel is defined as

$$\hat{e}_w(\omega) = \frac{1}{2}\mathrm{sign}(\omega) - \frac{1}{2\pi}\int_0^\omega \hat{w}(\theta)\, d\theta. \tag{1}$$

In the discrete-time case, $\hat{e}_w$ is that periodic function of period 2π whose values are given by (1) for $-\pi < \omega \le \pi$. Recall that, by definition, $w(t)$ must be a real-valued, even function of t; consequently, $\hat{w}(\omega)$ is also a real-valued, even function of ω. It follows that $\hat{e}_w$ is an odd function of ω, that is

$$\hat{e}_w(-\omega) = -\hat{e}_w(\omega).$$

It is also obvious that $\hat{e}_w$ is differentiable everywhere except at $\omega = 0$ (in the continuous-time case) or at $\omega = 2m\pi$ (in the discrete-time case) and that

$$\hat{e}_w(0^+) = -\hat{e}_w(0^-) = \frac{1}{2}.$$

An expression of the windowing error in terms of $\hat{e}_w$ is established by the theorem that follows.

Theorem 4: Let $\hat{s}_x \in L^1(\hat{G})$ be a bounded-variation function of ω. Then the following relationship holds for all values of ω where $\hat{s}_x$ is continuous:

$$\hat{s}_x(\omega) - \frac{1}{2\pi}(\hat{s}_x * \hat{w})(\omega) = (\hat{e}_w * d\hat{s}_x)(\omega) \tag{2}$$

where $(\hat{e}_w * d\hat{s}_x)(\omega) = \int_{\hat{G}} \hat{e}_w(\theta)\, d\hat{s}_x(\omega - \theta)$, and this integral is a Riemann–Stieltjes integral.

Proof: Let $\Omega = +\infty$ in the continuous-time case, or $\Omega = \pi$ in the discrete-time case. In either case, $\omega = 0$ is the only point in $[-\Omega, \Omega]$ where $\hat{e}_w$ is not differentiable. Moreover

$$\frac{1}{2\pi}\int_0^\Omega \hat{w}(\theta)\, d\theta = \frac{1}{4\pi}\int_{-\Omega}^{\Omega} \hat{w}(\theta)\, d\theta = \frac{1}{2}w(0) = \frac{1}{2}$$

Therefore: $\hat{e}_w(\Omega) = \hat{e}_w(-\Omega) = 0$. Then:

$$\begin{aligned}(\hat{e}_w * d\hat{s}_x)(\omega) &= \int_{-\Omega}^{\Omega} \hat{e}_w(\theta)\, d\hat{s}_x(\omega - \theta) \\ &= \int_{-\Omega}^{0^-} \hat{e}_w(\theta)\, d\hat{s}_x(\omega - \theta) \\ &\quad + \int_{0^+}^{\Omega} \hat{e}_w(\theta)\, d\hat{s}_x(\omega - \theta).\end{aligned}$$

Integration by parts yields

$$\begin{aligned}&\int_{-\Omega}^{0^-} \hat{e}_w(\theta)\, d\hat{s}_x(\omega - \theta) \\ &\quad = [-\hat{e}_w(\theta)\hat{s}_x(\omega - \theta)]_{\theta=-\Omega}^{\theta=0^-} \\ &\qquad + \int_{-\Omega}^{0^-} \hat{e}'_w(\theta)\hat{s}_x(\omega - \theta)\, d\theta \\ &\quad = -\hat{e}_w(0^-)\hat{s}_x(\omega) - \frac{1}{2\pi}\int_{-\Omega}^{0^-} \hat{w}(\theta)\hat{s}_x(\omega - \theta)\, d\theta \\ &\quad = \frac{1}{2}\hat{s}_x(\omega) - \frac{1}{2\pi}\int_{-\Omega}^{0^-} \hat{w}(\theta)\hat{s}_x(\omega - \theta)\, d\theta.\end{aligned} \tag{3}$$

Similarly

$$\begin{aligned}&\int_{0^+}^{\Omega} \hat{e}_w(\theta)\, d\hat{s}_x(\omega - \theta) \\ &\qquad = \frac{1}{2}\hat{s}_x(\omega) - \frac{1}{2\pi}\int_{0^+}^{\Omega} \hat{w}(\theta)\hat{s}_x(\omega - \theta)\, d\theta.\end{aligned} \tag{4}$$

Adding (3) and (4) proves the claim. ■

As a corollary to Theorem 4, it should be noted that (2) holds almost everywhere in $\hat{G}$, because the points of discontinuity of a bounded-variation function form a set of measure zero.[3] This means that the following identity is valid in $L^1(\hat{G})$:

$$\hat{s}_x - \hat{s}_x * \hat{w}/2\pi = \hat{e}_w * d\hat{s}_x. \tag{5}$$

Equation (5) can be used to obtain various types of bounds on the windowing error. For instance, the inequality $\|\mu_1 * \mu_2\| \leq \|\mu_1\|\|\mu_2\|$ that holds for the convolution of two measures yields

$$\|\hat{s}_x - \hat{s}_x * \hat{w}/2\pi\|_1 \leq \|\hat{e}_w\|_1 \|d\hat{s}_x\|. \tag{6}$$

In the equation above, $\|d\hat{s}_x\|$ represents the norm of the derivative of $\hat{s}_x$, regarded as a measure, which coincides with the total variation of $\hat{s}_x$. In the type of PSDs that occur in practice, it is reasonable to assume that the total variation of $\hat{s}_x$ is of the same order of magnitude as the largest value of $\hat{s}_x$.

If $\hat{s}_x$ is differentiable, i.e., $d\hat{s}_x = \hat{s}'_x\, d\omega$, and $\hat{s}'_x \in L^p(\hat{G})$, the inequality $\|f * g\|_p \leq \|f\|_1 \|g\|_p$ [10, pp. 396–397] is also useful to establish a bound on the windowing error. In particular, setting $p = \infty$ yields

$$\|\hat{s}_x - \hat{s}_x * \hat{w}/2\pi\|_\infty \leq \|\hat{e}_w\|_1 \|\hat{s}'_x\|_\infty. \tag{7}$$

Therefore the value of $\|\hat{e}_w\|_1$ limits both $\|\hat{s}_x - \hat{s}_x * \hat{w}/2\pi\|_\infty$, which is the maximum possible error at any particular frequency, and $\|\hat{s}_x - \hat{s}_x * \hat{w}/2\pi\|_1$, which is the L^1-norm of the total error over the entire range of possible frequencies. For this reason, $\|\hat{e}_w\|_1$ can be used as an approximate measure of the relative error introduced in spectral estimates by a particular window.

The windowing error kernel can also be used to establish theoretical results. As an example of a theoretical application, the inequality in (6) will now be used to prove that, in the discrete-time case, the L^1-norm of the windowing error can be made arbitrarily small under less restrictive conditions than those assumed in Theorem 3.

In preparation for this proof, note that $\hat{w} \in L^1(\hat{G})$ implies that $\hat{e}_w$ is a bounded function of ω. Therefore, in the discrete-time case $\hat{e}_w \in (L^1 \cap L^2)([-\pi, \pi])$, and Hölder's inequality yields

$$\|\hat{e}_w\|_1 \leq \sqrt{2\pi}\|\hat{e}_w\|_2. \tag{8}$$

Moreover, $\hat{e}_w$ can be written explicitly as

$$\begin{aligned}\hat{e}_w(\omega) &= \frac{1}{2}\mathrm{sign}(\omega) - \frac{w(0)}{2\pi}\omega - j\sum_{k\neq 0}\frac{w(k)}{2k\pi}e^{-jk\omega} \\ &= \frac{1}{2}\mathrm{sign}(\omega) - \frac{\omega}{2\pi} - j\sum_{k\neq 0}\frac{w(k)}{2k\pi}e^{-jk\omega}\end{aligned} \tag{9}$$

(note that the summations in the equations above have only a finite number of nonzero terms). Since $\hat{e}_w \in L^2([-\pi, \pi])$, the Plancherel theorem implies that $\hat{e}_w$ is the Fourier transform of some $e_w \in L^2(\mathbb{Z})$. It is fairly straightforward to verify by direct calculation that

$$e_w(k) = \begin{cases} 0 & k = 0, \\ j\dfrac{1 - w(k)}{2k\pi}, & k \neq 0. \end{cases}$$

Then Parseval's equality $\|\hat{e}_w\|_2^2 = 2\pi\|e_w\|_2^2$ can be used to establish the following identity[4]:

$$\sum_{k\neq 0}\left(\frac{1}{2k\pi}\right)^2 = \frac{1}{12}.$$

Theorem 5: Let $\{w_n\}$ be a sequence of discrete-time windows such that $\lim_{n\to\infty} w_n(k) = 1\ \forall k \in \mathbb{Z}$. Then

$$\lim_{n\to\infty}\|\hat{e}_{w_n}\|_1 = 0.$$

Consequently, if $\hat{s}_x$ is a bounded-variation function of ω, then

$$\lim_{n\to\infty}\|\hat{s}_x - \hat{s}_x * \hat{w}_n/2\pi\|_1 = 0$$

Proof: Given an arbitrary $\varepsilon > 0$, there exists a positive integer K_ε such that: $\sum_{|k|>K_\varepsilon}(1/2k\pi)^2 < \varepsilon^2/32\pi^2$. By assumption, there is also a positive index N_ε such that $n > N_\varepsilon$ implies:

$$|1 - w_n(k)| < \sqrt{3/2}\,\varepsilon/\pi$$

for $|k| \leq K_\varepsilon$. Note also that $|w_n(k)| \leq 1$ implies $|1 - w_n(k)| \leq 2$ for all values of n and k. Therefore, the following inequalities hold for $n > N_\varepsilon$:

$$\begin{aligned}\|e_{w_n}\|_2^2 &= \sum_{\substack{|k|\leq K_\varepsilon \\ k\neq 0}}\left[\frac{1 - w_n(k)}{2k\pi}\right]^2 \\ &\quad + \sum_{|k|>K_\varepsilon}\left[\frac{1 - w_n(k)}{2k\pi}\right]^2 \\ &< \frac{3\varepsilon^2}{2\pi^2}\sum_{\substack{|k|\leq K_\varepsilon \\ k\neq 0}}\left(\frac{1}{2k\pi}\right)^2 + \sum_{|k|>K_\varepsilon}\left(\frac{2}{2k\pi}\right)^2 \\ &< \frac{3\varepsilon^2}{2\pi^2}\frac{1}{12} + \frac{4\varepsilon^2}{32\pi^2} = \frac{\varepsilon^2}{4\pi^2}\end{aligned}$$

i.e., $\|e_{w_n}\|_2 < \varepsilon/2\pi$. Then, (8) and Parseval's equality imply

$$\|\hat{e}_{w_n}\|_1 \leq \sqrt{2\pi}\|\hat{e}_{w_n}\|_2 = 2\pi\|e_{w_n}\|_2 < \varepsilon.$$

This shows that $\lim_{n\to\infty}\|\hat{e}_{w_n}\|_1 = 0$. To complete the proof, note that a bounded-variation function is necessarily bounded. Therefore, $\hat{s}_x \in L^1([-\pi, \pi])$, and the second part of the claim follows from Theorem 4 and the inequality in (6). ■

The windowing error kernel can also be used to obtain both qualitative and quantitative information about the error affecting estimates of the PSD of continuous-spectrum signals. As an illustrative example, consider the five windows listed below,

[3]This follows, for instance, from Theorems (8.19) and (17.16) in [10].

[4]Set: $w(0) = 1, w(k) = 0$ for $k \neq 0$.

which are among those most frequently mentioned in the literature [11], [12]:

Rectangular:

$$w(k) = 1$$

Bartlett:

$$w(k) = 1 - |k|/N$$

Hanning:

$$w(k) = [1 + \cos(k\pi/N)]/2$$

Hamming:

$$w(k) = 0.54 + 0.46\cos(k\pi/N)$$

Blackman:

$$w(k) = 0.42 + 0.5\cos(k\pi/N) + 0.08\cos(2k\pi/N)$$

where $k = -N, \ldots, N$. The corresponding windowing error kernels can be computed directly from (9), which can also be written as

$$\hat{e}_w(\omega) = \frac{1}{2}\text{sign}(\omega) - \frac{\omega}{2\pi} - \frac{1}{\pi}\sum_{k=1}^{N}\frac{w(k)}{k}\sin k\omega.$$

Figs. 1 and 2 show the graphs of $|\hat{e}_w(\omega)/\hat{e}_w(0^+)|$, on logarithmic scale, for the five windows listed above, with $N = 32$. A qualitative assessment of the relative accuracy of the various windows can be made by comparing the rate at which $|\hat{e}_w(\omega)|$ decays as $|\omega|$ increases. For example, an inspection of Fig. 1 leads to the conclusion that, despite the absence of oscillations in its error kernel, the Bartlett window is not likely to produce more accurate results than the rectangular window, because the graphs corresponding to those two windows decay at very similar rates. On the other hand, the Hamming window's kernel falls 20 dB below its peak value at $\omega \approx 0.1\pi$, indicating that two features (e.g., two peaks) in $\hat{s}_x(\omega)$ that are separated by a factor of 100 or less in amplitude are likely to be visible in the spectral estimate, provided that their distance is at least 10% of the Nyquist frequency. The graph of the Hamming window, however, never decreases below a floor of approximately 25 dB, which indicates that this window's dynamic range is at most 25 dB, and perhaps less. On the other hand, the amplitudes of the sidelobes of the Hanning and Blackman windows keep decreasing well below those of the three other windows, an indication that they can achieve better dynamic range.

Table I shows the values of $\|\hat{e}_w\|_1$ for the five windows listed above and three different window lengths ($L = 2N + 1$). For each window and window length, three numbers are reported. The first is an approximate value of $\|\hat{e}_w\|_1$, computed by numerical quadrature using the `nag_1d_quad_gen (d01ajc)` routine of the NAG library [13], with `epsabs = 0, epsrel = 1e-4`. The second number is the ratio ω_z/π, expressed as a percentage, where ω_z is the first zero of $\hat{e}_w$. The third is the ratio $R = \int_{-\omega_z}^{\omega_z}|\hat{e}_w(\omega)|\,d\omega/\|\hat{e}_w\|_1$, expressed as a percentage: a value close to 100% indicates that ω_z provides a very good

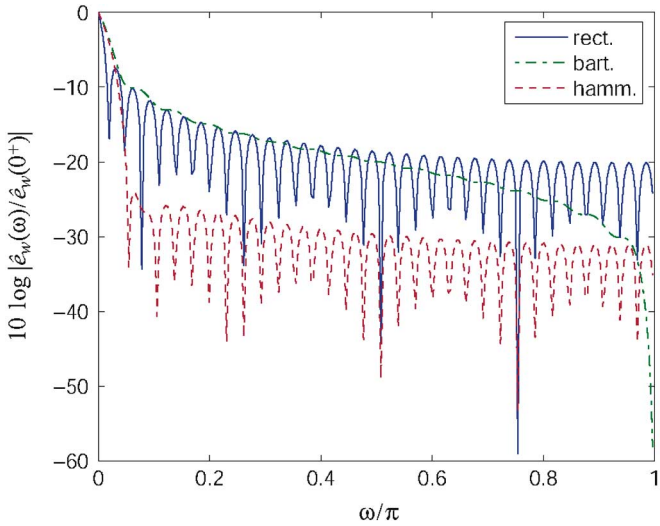


Fig. 1. Normalized windowing error kernel, $N = 32$ ($L = 65$).

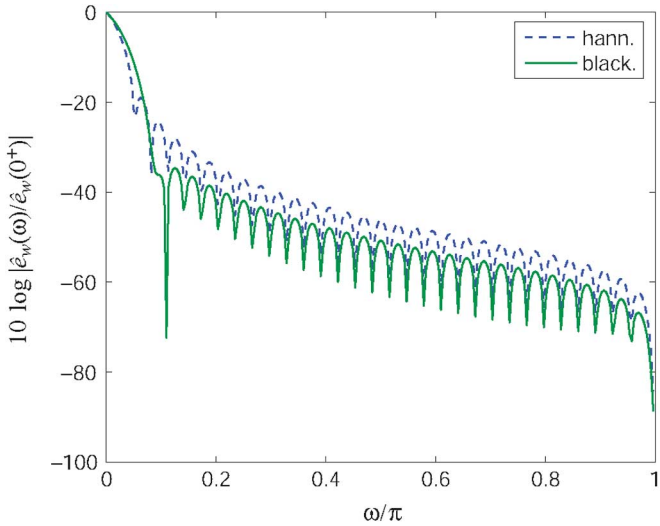


Fig. 2. Normalized windowing error kernel, $N = 32$ ($L = 65$).

approximation to the resolution of the window. Table I provides the basis for a quantitative comparison of the global accuracy of the various windows. For example, the values of $\|\hat{e}_w\|_1$ indicate that, at least in some cases, spectral estimates obtained using the Bartlett window could be less accurate than estimates obtained with the rectangular window. On the other hand, the accuracy of the Hanning and Hamming windows, as measured by $\|\hat{e}_w\|_1$, is effectively the same. In fact, among the five windows examined here the Hanning window appears to offer the best overall performance, when global error bound, resolution, and dynamic range are all taken into account.

VI. Numerical Examples

As mentioned in the Introduction, one of the main advantages of correlogram-based spectral estimates is that they do not rely on any specific model for the signal or the system that generated it. This makes the correlogram particularly suited for the spectral analysis of signals in applications where such models are either not available, or too inaccurate or too complex to be

TABLE I
SELECTED WINDOWING ERROR KERNEL PERFORMANCE INDICATORS

	$L = 33$			$L = 65$			$L = 129$		
	$\|\hat{e}_w\|_1$	ω_z/π (%)	R	$\|\hat{e}_w\|_1$	ω_z/π (%)	R	$\|\hat{e}_w\|_1$	ω_z/π (%)	R
Rectangular	$1.36 \cdot 10^{-1}$	3.715	38.38	$7.73 \cdot 10^{-2}$	1.887	34.18	$4.32 \cdot 10^{-2}$	0.951	30.78
Bartlett	$2.01 \cdot 10^{-1}$	100	100	$1.14 \cdot 10^{-1}$	100	100	$6.40 \cdot 10^{-2}$	100	100
Hanning	$1.19 \cdot 10^{-1}$	10.46	98.05	$5.96 \cdot 10^{-2}$	5.230	98.03	$2.98 \cdot 10^{-2}$	2.615	98.02
Hamming	$1.14 \cdot 10^{-1}$	11.09	96.52	$5.77 \cdot 10^{-2}$	5.501	95.47	$2.92 \cdot 10^{-2}$	2.740	94.36
Blackman	$1.44 \cdot 10^{-1}$	21.88	99.91	$7.21 \cdot 10^{-2}$	10.94	99.91	$3.61 \cdot 10^{-2}$	5.469	99.91

practically useful: this is the case, for instance, in digital communications systems. As an example, in this section, correlogram spectral analysis is used to analyze the effect of system nonlinearities on the power spectrum of a typical digital communications signal.

In a communications system, nonlinear effects are usually most pronounced in power amplifiers, where signal levels reach their highest values. The amplifier's nonlinear behavior can broaden the power spectrum of the output signal, a phenomenon known as spectral regrowth, which, in turn, can cause interference between signals occupying neighboring frequency bands (adjacent channel interference). For this reason, the ability to obtain accurate *a priori* estimates of the amount of spectral regrowth can be extremely valuable in the design phase of a power amplifier [14].

Methods to estimate spectral regrowth have been the object of intense research. In many cases, it is assumed that the amplifier can be adequately described by a memoryless or quasi-memoryless polynomial model [15], [16]. Other approaches [17] rely on Volterra series expansions, which are still polynomial in nature but can be used to represent systems with memory. Because of their limitations, however, none of these models can be used in practice to quantify spectral regrowth in an actual power amplifier. Simple (i.e., low-order) polynomial models are almost never sufficiently accurate, and higher order models are too unwieldy to be practically useful. Moreover, the assumption that the amplifier can be described as a quasi-memoryless system is often unrealistic, especially at the high frequencies used for wireless digital communications. Currently, simulation is the best and most reliable way to predict the behavior of amplifiers and other electronic circuits with reasonable accuracy [18]. For this reason, circuit simulators are heavily relied upon as verification tools in electronic circuit design. In the example described below, simulation and correlogram-based spectral analysis are used to analyze spectral regrowth in the output signal of a MOS amplifier.

Fig. 3 shows the amplifier's schematic diagram; the graph showing the amplifier's input-output gain curve, which was obtained through simulation, has the origin located at the DC operating point. The input signals applied to the amplifier were trains of finite-length, nonoverlapping pulses given by the following expression:

$$x(t) = \sum_{n=-\infty}^{+\infty} b_n p_x(t - nT) \tag{10}$$

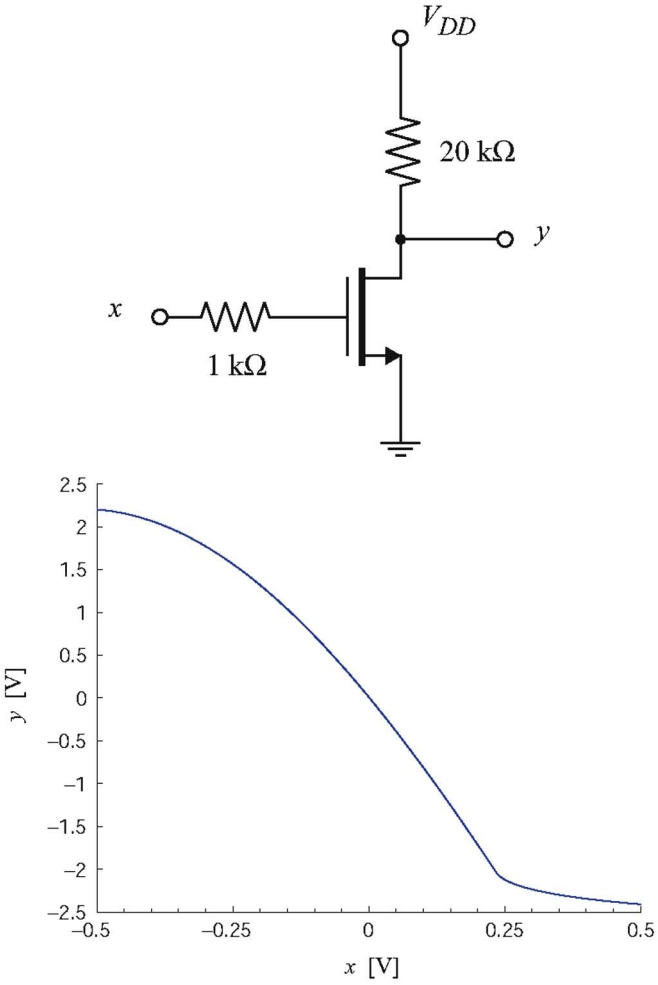


Fig. 3. MOS amplifier and its input–output gain curve.

where $\{b_n\}$ is a sequence of coefficients satisfying the following relationships:

$$b_n = \pm 1$$

$$\lim_{N \to +\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} b_n b_{n+m} = \begin{cases} 1, & m = 0 \\ 0, & m \neq 0. \end{cases}$$

These signals encode pseudorandom bit streams and are common in digital communications systems. It can be shown that their power spectrum is continuous, and their theoretical

PSD can be expressed in terms of $\hat{p}_x$ [4]. For this example, raised cosine pulses were used, as follows:

$$p_x(t) = \frac{A}{2}[1 - \cos(2\pi t/T)], \quad 0 \le t \le T$$

with $T = 5\,\mu$s, corresponding to a symbol frequency of 200 kHz.

A pseudorandom number generator was used to obtain a sequence $\{b_n\}$ containing approximately 6 500 values, and standard time-domain circuit simulation techniques were used to compute numerically the amplifier's output signal. In order to limit aliasing effects, the value of the integration timestep was set to 62.5 ns, which generated slightly over 524 000 signal samples. The CPU time necessary to run each simulation ranged between 350 and 500 s on a Sun Ultra 60 workstation running Solaris 8.

After downsampling by a factor of four, correlogram estimates of the PSD's of the input and output signals were computed using a Blackman window with $L = 1025$. The CPU time needed to compute the estimates was a few percent of the simulation time. The data in Table I indicates that, for this window, the value of $\|\hat{e}_w\|_1$ decreases in a manner that is approximately inversely proportional to the window length, and extrapolating the data to $L = 1025$ yields $\|\hat{e}_w\|_1 \approx 4.5 \cdot 10^{-3}$. Using this number as an approximate measure of the windowing error, the spectral estimates generated by this window can be expected to have a relative accuracy on the order of one percent or less, which is more than adequate for this particular application.

For comparison purposes, a different estimate of the output signal's PSD was obtained in the following way. If the amplifier can be modeled as a memoryless system, $y = f(x)$, and if f is an odd function, i.e., $f(-x) = -f(x)$, then it is easy to verify that the input signal defined in (10) generates an output of the same type, as follows:

$$y(t) = \sum_{n=-\infty}^{+\infty} b_n p_y(t - nT)$$

where $p_y(t) = f[p_x(t)]$. Then, the theoretical expression for $\hat{s}_y$ is [4, p. 194]

$$\hat{s}_y(\omega) = \frac{1}{T}|\hat{p}_y(\omega)|^2. \tag{11}$$

For a given p_x, the corresponding p_y was obtained through simulation, and $\hat{p}_y$ was computed from it using a numerical approximation of the Fourier integral.

Fig. 4 shows the PSDs of the input and output signals, obtained from correlogram estimates as explained above, with the amplitude of the input pulses (i.e., the value of A) set to 10 mV. In this case, because the input signal is comparatively small, the amplifier is operating essentially linearly, and the output PSD is just a scaled replica of the input PSD. A comparison with the PSD estimate obtained from (11) is shown in Fig. 5: because of the amplifier's good linearity at these signal levels, there is a substantial agreement between the two estimates.

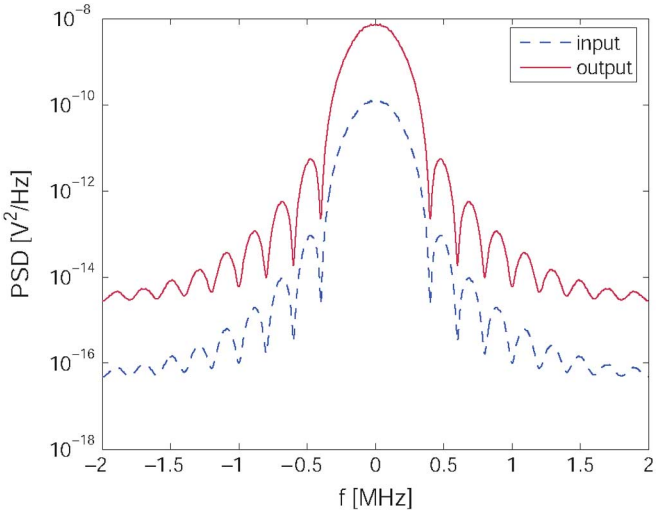


Fig. 4. Correlogram estimates of the PSD of the amplifier's input and output signals, $A = 10$ mV.

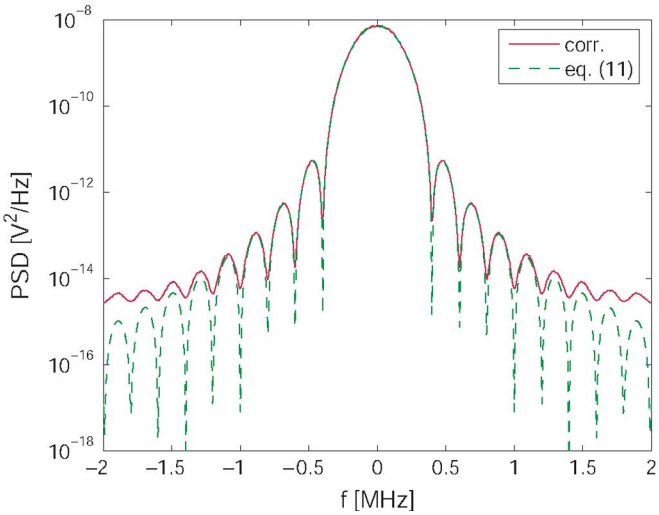


Fig. 5. Estimates of the amplifier's output PSD obtained from the correlogram and (11), $A = 10$ mV.

A similar comparison, with the amplitude of the input pulses increased to 200 mV is shown in Fig. 6. No obvious warping of the output PSD due to nonlinearities in the amplifier can be observed yet, but spikes can be detected at 200 kHz (barely) and more clearly at 400 kHz. These spikes indicate the presence of periodic components in the output signal at multiples of the symbol frequency and are due to the fact that, at these signal levels, the amplifier's gain curve can no longer be regarded as symmetric with respect to the DC operating point (i.e., the origin in Fig. 3). This lack of symmetry means that the Taylor expansions of the system's input–output relationship contains even powers of x, as follows:

$$y = f(x) \approx a_1 x + a_2 x^2 + a_3 x^3 \ldots.$$

It can be readily verified that, for the type of input signals given by (10), even-order powers of x in the equation above create periodic components in the output. The comparison with the spectral estimate obtained from (11) is shown in Fig. 7. It can be seen

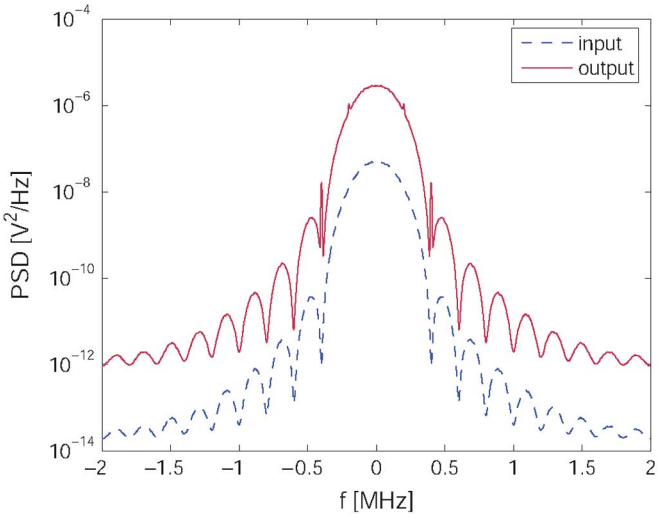


Fig. 6. Correlogram estimates of the PSD of the amplifier's input and output signals, $A = 200$ mV.

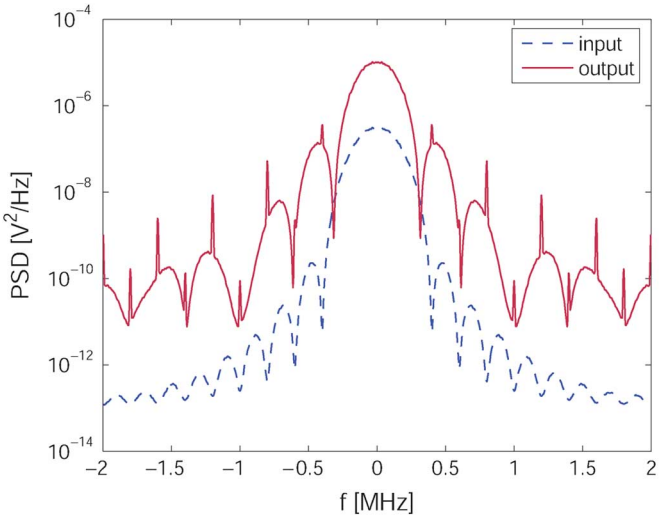


Fig. 8. Correlogram estimates of the PSD of the amplifier's input and output signals, $A = 500$ mV.

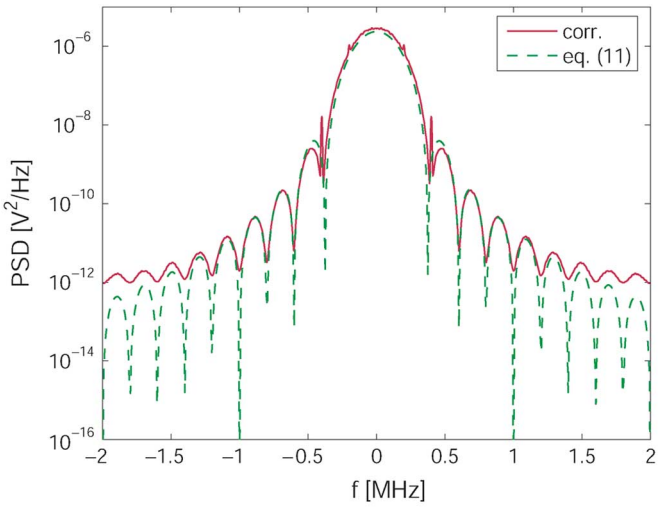


Fig. 7. Estimates of the amplifier's output PSD obtained from the correlogram and (11), $A = 200$ mV.

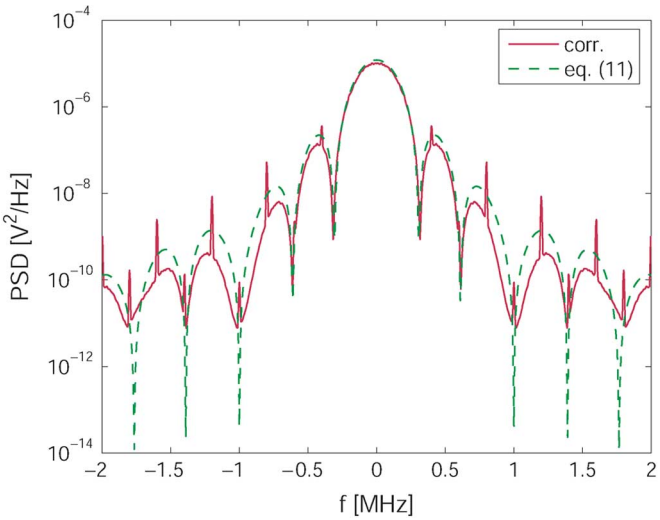


Fig. 9. Estimates of the amplifier's output PSD obtained from the correlogram and (11), $A = 500$ mV.

that there is still a generally good agreement between the two estimates, although the latter lacks the spikes at 200 and 400 kHz. The reason is that (11) assumes that the system's input–output relationship is an odd function and, consequently, it cannot predict phenomena caused by deviations from this assumption.

Spectral regrowth in the output PSD caused by the amplifier's nonlinear behavior becomes manifest when the amplitude of the input pulses is further increased to 500 mV, as can be seen in Fig. 8. Narrowing of the main lobe and widening of the secondary lobes can readily be observed, together with the presence of periodic components at multiples of the symbol frequency, which now reach the upper limit of the frequency range shown in the graph. A further loss of accuracy in the spectral estimate obtained from (11), which is compared with the correlogram estimate in Fig. 9, can also be observed. While there are no noticeable differences between the two graphs on the main lobe, the amplitudes of the secondary lobes are clearly overestimated, and, as in the previous case, the periodic components at multiples of the symbol frequency are completely absent.

These results illustrate the advantages of using the correlogram to obtain spectral estimates in applications where nonlinearities are present, the numerical accuracy of the estimates is an important requirement, and arbitrarily long signal segments can be obtained without too much difficulty (e.g., by simulation, as in this example). On the one hand, the need for accuracy usually requires the use of complex models to describe the system that generates the signals being analyzed. On the other hand, the very complexity of the system equations effectively rules out any hope of obtaining analytical models for the signal spectra that are equally accurate, yet not so unwieldy as to be unusable for actual numerical calculations. In contrast, numerical estimates of a signal's PSD based on the correlogram can be computed in a straightforward manner, and the theoretical

results presented in the previous sections ensure that, in principle, arbitrarily accurate estimates can be obtained with an appropriate choice of the window and of the length of the signal segment.

VII. Conclusion

The results established in this paper can be summarized as follows. Windowed correlogram estimates generated by signal segments of increasing length converge in the L^1-norm to the signal's windowed PSD (i.e., $\hat{s}_x * \hat{w}/2\pi$). The L^1-norm of the windowing error, i.e., the difference between the exact and windowed PSDs, can be made arbitrarily small by a suitable choice of window (Theorems 3 and 5). It follows that, in principle, the correlogram can estimate the PSD of a continuous-spectrum signal with arbitrary accuracy, provided that an appropriate window and sufficiently long signal segments are used.

It was also shown that bounds on the asymptotic error of the estimates generated by a given window (i.e., the quantity $\hat{s}_x - \hat{s}_x * \hat{w}/2\pi$) can be obtained from the L^1-norm of the windowing error kernel. Since this kernel can always be computed explicitly from the window coefficients, it provides a way to compare the accuracy of different windows in both qualitative and quantitative terms. As an illustration, selected performance parameters related to the resolution and overall accuracy of five commonly used windows have been computed from the corresponding error kernels. These calculations have led to the conclusion that, for example, the Hanning window should generate more accurate estimates of the PSD of a continuous-spectrum signal than the Bartlett window.

Thus, if a correlogram-based estimate of a signal PSD is needed, the windowing error kernel can be used as a guide to select the type and length of the window, based on considerations such as relative accuracy, frequency resolution, and so on. After a window has been selected, signal segments of increasing length can be used to evaluate the correlogram and generate corresponding spectral estimates, until the difference between successive estimates becomes negligible.

Although only deterministic signals were considered in this paper, some of the results presented here are directly applicable to the spectral analysis of stochastic processes. For example, given a discrete-time stationary process with autocorrelation function s_x, it is shown in [11, p. 745] that the expected value of a windowed correlogram spectral estimator is

$$\mathcal{E}\{\hat{s}_n * \hat{w}/2\pi\} \approx \hat{s}_x * \hat{w}/2\pi.$$

Therefore, the bias of this estimator is $\hat{s}_x - \hat{s}_x * \hat{w}/2\pi$, and the results presented in Section V can be used directly to place an upper bound on it.

In conclusion, among the many spectral analysis techniques that are now available, methods based on the Fourier transform remain useful in many applications, particularly those in which signals can be observed over arbitrarily long time intervals, and little or no *a priori* knowledge is available about the type of signals or the systems that generated them. In those cases where signal segments of sufficient length are not available, additional information about the signal is necessary to establish theoretical bounds on the accuracy of spectral estimates obtained from the correlogram. Specific expressions for such bounds, if they can be derived, will almost certainly be highly dependent on what is known about the types of signals under analysis or the systems that generated them. On the other hand, other techniques (e.g., parametric methods) might be better suited for spectral analysis when such information is available.

Appendix

A. Fourier and Fourier–Stieltjes Algebras

Let $M(\hat{G})$ denote the set of all complex-valued, bounded, regular Borel measures on $\hat{G}$. It is possible to define a convolution operation between two measures $\mu, \nu \in M(\hat{G})$ [7, p. 13]: the result, which is still an element of $M(\hat{G})$, will be denoted by $\mu * \nu$. Equipped with convolution and with the usual vector space operations, $M(\hat{G})$ is a complex Banach algebra under the norm defined by $\|\mu\| = |\mu|(\hat{G})$, where $|\mu|$ is the total variation of μ [7], [10].

A measure $\mu \in M(\hat{G})$ is *absolutely continuous* with respect to the Lebesgue measure on $\hat{G}$ if there exists $\hat{m} \in L^1(\hat{G})$ such that

$$\mu(S) = \int_S \hat{m}(\omega)\, d\omega$$

for all measurable sets $S \subseteq \hat{G}$. In such case, the following equality holds: $\|\mu\| = \|\hat{m}\|_1$. The notation $d\mu = \hat{m}\, d\omega$ will be used as a shorthand to denote the relationship between an absolutely continuous measure $\mu \in M(\hat{G})$ and the corresponding function $\hat{m} \in L^1(\hat{G})$.

It can be shown that the convolution of two absolutely continuous measures corresponds to the usual convolution operation between functions. In other words, if $d\mu_1 = \hat{m}_1\, d\omega$ and $d\mu_2 = \hat{m}_2\, d\omega$, then $d(\mu_1 * \mu_2) = (\hat{m}_1 * \hat{m}_2)\, d\omega$. Therefore, $L^1(\hat{G})$ can be identified with the set of absolutely continuous measures on $\hat{G}$; with this identification, $L^1(\hat{G})$ becomes a subalgebra of $M(\hat{G})$. In fact, $L^1(\hat{G})$ is a norm-closed ideal of $M(\hat{G})$ [7, p. 16]. This means that the result of the convolution between an element of $L^1(\hat{G})$ and any other measure in $M(\hat{G})$ is always an element of $L^1(\hat{G})$, that is

$$\mu_1 \in L^1(\hat{G}),\ \mu_2 \in M(\hat{G}) \Longrightarrow \mu_1 * \mu_2 \in L^1(\hat{G}).$$

In particular, since the convolution of absolutely continuous measures is identical to function convolution, this implies that $L^1(\hat{G})$ is closed under convolution

$$\hat{m}_1, \hat{m}_2 \in L^1(\hat{G}) \Longrightarrow \hat{m}_1 * \hat{m}_2 \in L^1(\hat{G}).$$

Let $B(G)$ denote the set of all complex-valued functions that can be expressed as the inverse Fourier transform of a measure in $M(\hat{G})$. In other words, $x \in B(G)$ if

$$x(t) = \left[\mathcal{F}^{-1}(\mu)\right](t) = \frac{1}{2\pi} \int_{\hat{G}} e^{j\omega t}\, d\mu(\omega)$$

for some $\mu \in M(\hat{G})$. It is straightforward to verify that all the elements of $B(G)$ are bounded, continuous (in fact, uniformly continuous) functions. Furthermore, convolution of measures in $M(\hat{G})$ is mapped by the inverse Fourier transform

into a product of functions in $B(G)$. In other words, if $x_1 = \mathcal{F}^{-1}(\mu_1), x_2 = \mathcal{F}^{-1}(\mu_2)$, then $x_1 x_2 = \mathcal{F}^{-1}(\mu_1 * \mu_2)/2\pi$. It follows that $x_1, x_2 \in B(G) \Longrightarrow x_1 x_2 \in B(G)$, i.e., $B(G)$ is a function algebra.

A norm can be defined on $B(G)$ in the following way:

$$\|x\|_B = \frac{1}{2\pi}\|\mu\|, \quad x = \mathcal{F}^{-1}(\mu).$$

Note that $\|x\|_\infty \leq \|x\|_B$. Under this norm, $B(G)$ is a Banach algebra: it is referred to as the *Fourier–Stieltjes algebra* of G [19].

The positive-definite functions in $B(G)$ are the inverse Fourier transforms of positive measures in $M(\hat{G})$. Hence, if $x = \mathcal{F}^{-1}(\mu) \in B(G)$ is positive definite, then

$$\|x\|_B = \frac{1}{2\pi}\mu(\hat{G}) = \frac{1}{2\pi}\int_{\hat{G}} d\mu = x(0) \leq \|x\|_\infty.$$

Since $\|x\|_\infty \leq \|x\|_B$, it follows that $\|x\|_B = \|x\|_\infty = x(0)$. This equality, however, does not generally hold for elements of $B(G)$ that are not positive definite.

Let $A(G)$ denote the set of all complex-valued functions that are the inverse Fourier transform of a function in $L^1(\hat{G})$, as follows:

$$x(t) = \frac{1}{2\pi}\int_{\hat{G}} e^{j\omega t}\hat{x}(\omega)\, d\omega, \quad \hat{x} \in L^1(\hat{G}).$$

Since $L^1(\hat{G})$ is a subspace of $M(\hat{G}), A(G)$ is a subspace of $B(G)$. Moreover, as noted earlier, if $d\mu = \hat{m}\, d\omega$, then $\|\mu\| = \|\hat{m}\|_1$. Therefore, if $x \in A(G)$, then $\|x\|_B = \|\hat{x}\|_1$.

$A(G)$ is closely related to $L^2(G)$. If $x_1, x_2 \in L^2(G)$, then $\hat{x}_1, \hat{x}_2 \in L^2(\hat{G})$. Therefore, $\hat{x}_1\hat{x}_2 \in L^1(\hat{G})$, and this implies $x_1 * x_2 \in A(G)$. Conversely, every function in $L^1(\hat{G})$ can be written (in many different ways) as the product of two functions in $L^2(\hat{G})$. This implies that every function in $A(G)$ can be written (again, in many different ways) as the convolution of two functions in $L^2(G)$. Therefore, $A(G)$ can also be defined as the set of all functions that can be written as the convolution of two functions in $L^2(G)$.

Finally, the product of two functions in $A(G)$ is still an element of $A(G)$. This follows from the fact, mentioned earlier, that $L^1(\hat{G})$ is closed under convolution. Therefore, $A(G)$ is a Banach subalgebra (in fact, an ideal) of $B(G)$: it is referred to as the *Fourier algebra* of G [19].

Norm convergence in $B(G)$ is equivalent to norm convergence in $M(\hat{G})$. More precisely, let $x = \mathcal{F}^{-1}(\mu) \in B(G)$, and let $\{x_n = \mathcal{F}^{-1}(\mu_n)\}$ be a sequence in $B(G)$. Then

$$\lim_{n\to\infty}\|x_n - x\|_B = 0 \Longleftrightarrow \lim_{n\to\infty}\|\mu_n - \mu\| = 0.$$

This follows immediately from the definition of the norm in $B(G)$. Similarly, if $x \in A(G)$ and $\{x_n\} \subseteq A(G)$, then

$$\lim_{n\to\infty}\|x_n - x\|_B = 0 \Longleftrightarrow \lim_{n\to\infty}\|\hat{x}_n - \hat{x}\|_1 = 0.$$

Since every function in $B(G)$ is bounded, $B(G)$ is a subspace of $L^\infty(G)$, which is the dual space of $L^1(G)$ [10]. Therefore, a weak-* topology exists on $B(G)$, induced by duality with $L^1(G)$. In this topology, a sequence of functions $\{x_n\}$ converges to a function x if and only if

$$\lim_{n\to\infty}\int_G x_n(t)\phi(t)dt = \int_G x(t)\phi(t)\, dt$$

for every $\phi \in L^1(G)$. The notation

$$x = \underset{n\to+\infty}{w^*\text{-}\lim}\, x_n$$

will be used to denote convergence in the weak-* topology. Weak-* convergence in $B(G)$ is closely related to pointwise convergence, as the following lemma shows.

Lemma 1: Let $\{x_n\}$ be a sequence in $B(G)$ that convergences pointwise to a function $x \in B(G)$, that is

$$\lim_{n\to\infty} x_n(t) = x(t) \quad \forall t \in G$$

Assume that $\{x_n\}$ is bounded in the norm of $B(G)$, i.e., there exists $M > 0$ such that $\|x_n\|_B \leq M$ for all n. Then

$$x = \underset{n\to+\infty}{w^*\text{-}\lim}\, x_n.$$

Proof: Let $\phi \in L^1(G)$. The inequality $\|x_n\|_\infty \leq \|x_n\|_B$ implies $\|x_n\|_\infty \leq M$. Hence

$$|x_n(t)\phi(t)| \leq M|\phi(t)| \quad \forall t \in G.$$

Since $M|\phi| \in L^1(G)$, Lebesgue's dominated convergence theorem [10], applied to the sequence $\{x_n\phi\}$, yields the following equality:

$$\lim_{n\to\infty}\int_G x_n(t)\phi(t)\, dt = \int_G x(t)\phi(t)\, dt.$$

Because this is true for every $\phi \in L^1(G)$, it follows that $\{x_n\}$ converges to x in the weak-* topology of $B(G)$. ■

The next theorem establishes a result that is, in some respect, unexpected, because it is specific to $B(G)$ and cannot generally be extended to other function algebras (e.g., $L^\infty(G)$).

Theorem 6 [20, Theorem A]: Let $x \in B(G)$ be the weak-* limit of a sequence $\{x_n\} \subseteq B(G)$

$$\underset{n\to+\infty}{w^*\text{-}\lim}\, x_n = x$$

and assume that: $\lim_{n\to+\infty}\|x_n\|_B = \|x\|_B$. Then

$$\lim_{n\to+\infty}\|a(x_n - x)\|_B = 0 \quad \forall a \in A(G).$$

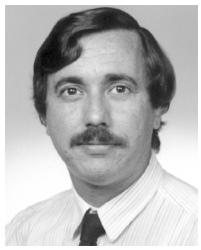
Theorems 2 and 3 in the main body of the paper depend in an essential way on this theorem, whose proof can be found in [20].

REFERENCES

[1] A. Schuster, "On the investigation of hidden periodicities with application to a supposed 26 day period of metereological phenomena," *Terrestrial Magnetism*, vol. 3, pp. 13–41, 1898.

[2] R. B. Blackman and J. W. Tukey, *The Measurement of Power Spectra From the Point of View of Communications Engineering*. New York: Dover, 1959.

[3] S. M. Kay and S. L. Marple, Jr., "Spectrum analysis—A modern perspective," *Proc. IEEE*, vol. 69, no. 11, pp. 1380–1419, Nov. 1981.
[4] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.
[5] G. Casinovi, "An algorithm for frequency-domain noise analysis in nonlinear systems," in *Proc. 39th Design Automation Conf.*, New Orleans, LA, Jun. 2002, pp. 514–517.
[6] G. Casinovi, "Numerical computation of signal power spectral density in nonlinear systems," *IEEE Trans. Circuits Syst. I*, submitted for publication.
[7] W. Rudin, *Fourier Analysis on Groups*. New York: Wiley, 1990.
[8] N. Wiener, "Generalized harmonic analysis," *Acta Mathematica*, vol. 55, pp. 117–258, 1930.
[9] A. Papoulis, *The Fourier Integral and its Applications*. New York: McGraw-Hill, 1962.
[10] E. Hewitt and K. Stromberg, *Real and Abstract Analysis*. New York: Springer-Verlag, 1965.
[11] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 1999.
[12] F. J. Harris, "On the use of windows for harmonic analysis with discrete Fourier transform," *Proc. IEEE*, vol. 66, no. 1, pp. 51–83, Jan. 1978.
[13] *"NAG C Library Manual,"* The Numerical Algorithms Group Ltd. (NAG), Oxford, U.K., Dec. 2005 [Online]. Available: http://www.nag.co.uk
[14] P. B. Kenington, *High-Linearity RF Amplifier Design*. Boston, MA: Artech House, 2000.
[15] G. T. Zhou, "Analysis of spectral regrowth of weakly nonlinear power amplifiers," *IEEE Commun. Lett.*, vol. 4, no. 11, pp. 357–359, Nov. 2000.
[16] G. T. Zhou and J. S. Kenney, "Predicting spectral regrowth of nonlinear power amplifiers," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 718–722, May 2002.
[17] S. A. Maas, "Volterra analysis of spectral regrowth," *IEEE Trans. Commun.*, vol. 7, no. 7, pp. 192–193, Jul. 1997.
[18] K. Mayaram, D. C. Lee, S. Moinian, D. A. Rich, and J. Roychowdhury, "Computer-aided circuit analysis tools for RFIC simulation: Algorithms, features, and limitations," *IEEE Trans. Circuits Syst. II*, vol. 47, no. 4, pp. 274–286, Apr. 2000.
[19] P. Eymard, "L'algèbre de Fourier d'un groupe localement compact," *Bull. Soc. Math. France*, vol. 92, pp. 181–236, 1964.
[20] E. E. Granirer and M. Leinert, "On some topologies which coincide on the unit sphere of the Fourier–Stieltjes algebra $B(G)$ and of the measure algebra $M(G)$," *Rocky Mountain J. Math.*, vol. 11, no. 3, pp. 459–472, 1981.



Giorgio Casinovi (M'89–SM'93) received the B.S. degrees in electrical engineering and in mathematics from the University of Rome, Italy, in 1980 and 1982, respectively, and the M.S. degree and the Ph.D. degree in electrical engineering, both from the University of California, Berkeley, in 1984 and 1988, respectively.

In 1989, he joined the School of Electrical and Computer Engineering of the Georgia Institute of Technology, Atlanta, His research interests include computer-aided design and simulation of electronic devices and circuits and mixed-technology systems.