ARM PROCESSES AND MODELING METHODOLOGY

Benjamin Melamed

Rutgers University
Faculty of Management
Department of MSIS
94 Rockafeller Rd.
Piscataway, NJ 08854
melamed@rbs.rutgers.edu

ABSTRACT

ARM (*Auto-Regressive Modular*) processes constitute a broad class of nonlinear autoregressive schemes with modulo-1 reduction and additional transformations. Unlike their TES (*Transform-Expand-Sample*) precursors, which only admit iid innovation sequences, ARM processes admit dependent innovation sequences as well, so long as they are independent of the initial ARM variate. As such, the class of ARM processes constitutes a considerable generalization of the TES class, endowed with enhanced modeling flexibility. For example, a Markovian innovation sequence can model burstiness in traffic processes far better by making use of the Markovian state to capture the structure of bursts. This paper introduces ARM processes and derives their fundamental properties in terms of marginal distributions and autocorrelation functions. It defines several useful subclasses that illustrate the modeling flexibility of ARM classes. Finally, it outlines a modeling methodology of empirical time series that can simultaneously fit the empirical marginal distribution (histogram) and empirical autocorrelation function.

Key Words and Phrases: Autoregressive Modular Processes, ARM Processes, Stochastic Processes, Nonlinear Processes

1 Introduction

This paper introduces a new and versatile class of nonlinear stochastic sequences, called ARM (*Auto-Regressive Modular*) processes. ARM processes are essentially (nonlinear) autoregressive sequences, with modulo-1 reduction (hence the appellation "modular") and additional transformations. Two classes of ARM processes have been previously investigated. The class of TES (*Transform-Expand-Sample*) processes was introduced first [25]. It has been investigated theoretically [10, 11, 12, 13, 14, 26] and as the arrival component of queueing systems [6, 9, 17, 33], and exercised extensively in practical modeling, especially teletraffic [20, 27, 28, 31, 30]. QTES (Quantized TES), a discretized variant of TES, has been developed more recently [29, 18]. Unlike their TES and QTES precursors, which only admit only iid so-called innovation sequences, ARM processes constitute a considerable generalization in that the ARM class admits dependent innovation sequences as well, so long as they are independent of the initial ARM variate.

As with TES and QTES processes, the motivation for defining the class of ARM processes is the need to construct high-fidelity models from empirical records, which capture an extensive statistical signature of the empirical data by simultaneously fitting the empirical marginal distribution and the empirical leading (low-lag) autocorrelations. More specifically, suppose that an empirical time series has been observed, and assume that its underlying probability law has time invariant 2-dimensional joint distributions. Assume further that the corresponding empirical density (histogram) and leading empirical autocorrelations have been computed (estimated); in view of the aforementioned invariance assumption, the former is an estimate of the underlying marginal density, and the latter estimate the autocorrelation function as a function of the lag only. The ARM modeling methodology deems a prospective model a good fit to an empirical record, if it simultaneously satisfies the following three modeling requirements:

Requirement 1. The marginal distribution of the model should exactly fit its empirical counterpart.

Requirement 2. The leading autocorrelations of the model should approximate their empirical counterparts.

Requirement 3. The model should give rise to Monte Carlo simulated paths which are visually "similar" or substantially "resemble" the empirical records.

The first two modeling requirements constitute clear-cut quantitative goodness-of-fit criteria; the exact fit of histograms (Requirement 1) is well-defined, while the approximation of leading autocorrelations (Requirement 2) can be made precise by selecting some metric on finite vectors. In contrast, the notions of sample path "similarity" or "substantial resemblance" (Requirement 3) are heuristic and cannot be defined with mathematical crispness; note that "resemblance" does not mean that the sample paths are identical, point for point, but rather that human observers perceive them as "similar" in appearance. It should be pointed out that these notions are intuitively clear and that the principle of "path similarity" is routinely practiced by modelers as a means for increasing their confidence in a model. Note carefully that Requirement 3 is imposed *in addition* and *not instead* of the quantitative first two. Clearly, the latter are strong requirements, as they call for the simultaneous capture of first-order and second-order statistics of the empirical data, and can therefore be expected to yield better models than those obtained by capturing a weaker statistical signature. For similar modeling approaches, see [21, 23, 32, 3], and especially [3] which is capable of matching precisely the empirical autocorrelation function in addition to the empirical distribution. For a comprehensive survey of this topic, refer to [16].

The rest of the paper is organized as follows. Section 2 contains some notational and technical preliminaries. Section 3 introduces the requisite definitions for ARM processes. Section 4 presents some stationarity properties of ARM processes. Section 5 investigates the transition structure of ARM processes and derives formulas for their transition densities, while Section 6 investigates second-order properties of ARM processes and derives formulas for their autocovariance/autocorrelation functions. Several ARM processes of theoretical and practical interest are exhibited in Sections 7 and 8, respectively, by specializing certain ARM modeling parameters. Section 9 describes the ARM modeling methodology. Finally, Section 10 contains some concluding remarks.

2 Preliminaries

The following notation will be used throughout this paper. For a random variable X with distribution F, the notation $X \sim F$ means that X has distribution F. If $\{X_n\}$ is a real-valued random sequence with finite second moments, then its autocorrelation function is

$$\rho_X(k,\tau) \;=\; \frac{E[X_k\, X_{k+\tau}] - \mu_{X_k}\,\mu_{X_{k+\tau}}}{\sigma_{X_k}\,\sigma_{X_{k+\tau}}}, \quad k \ge 0, \tau \ge 0, \tag{2.1}$$

where μ_Z and σ_Z denote the mean and standard deviation, respectively, of random variable Z. When $\rho_X(k,\tau)$ does not depend on k but only on the lag τ, we denote the autocorrelation function by $\rho_X(\tau)$. The modulo-1 (fractional part) operator, $\langle\cdot\rangle$, is defined by $\langle x\rangle = x - \max\{\text{integer } n :\ n \le x\}$. The Laplace Transform of a function f is $\tilde{f}(s) = \int_{-\infty}^{\infty} e^{-sx} f(x)dx$. The imaginary number $\sqrt{-1}$ is always denoted by i. The indicator function of a set A is written as $1_A(x)$. The density of a random variable Y is denoted by f_Y; similarly, the marginal densities of a random sequence $\{Y_n\}$ are denoted by f_{Y_n}, but for stationary sequences we simplify the notation of the common density to f_Y.

The next result makes use of the Poisson summation formula (see, e.g., [4], p. 52). Let $h(x)$ be a function for which the series $\eta(z) = \sum_{\nu=-\infty}^{\infty} h(z+\nu)$ converges uniformly for $z \in [0,1]$ and $\eta(z)$ is representable by its Fourier series (for sufficient conditions, see [34], p. 79), then $\eta(z)$ has the representation

$$\sum_{\nu=-\infty}^{\infty} h(z+\nu) \;=\; \sum_{\nu=-\infty}^{\infty} \tilde{h}(i2\pi\nu)\, e^{i2\pi\nu z}. \tag{2.2}$$

For other versions of the Poisson summation formula, see [22, 4]; for a more restricted version requiring continuous densities, see [7], Volume II, Chapter XIX, Equation (5.9), with $\zeta = 0$ and $\lambda = \pi$.

The following lemma is a general structural result, relating the marginal density of a random variable Y to that of its modular counterpart $\langle Y \rangle$; it will prove essential to rapid computation of transition densities and autocorrelations of ARM processes.

Lemma 1
For any random variable Y with density f_Y which satisfies the conditions of (2.2),

$$f_{\langle Y \rangle}(y) \;=\; \sum_{\nu=-\infty}^{\infty} \tilde{f}_Y(i2\pi\nu)\, e^{i2\pi\nu y}, \quad y \in [0,1), \tag{2.3}$$

and

$$\tilde{f}_{\langle Y \rangle}(s) \;=\; \sum_{\nu=-\infty}^{\infty} \tilde{f}_Y(i2\pi\nu)\, \frac{1-e^{-s}}{s - i2\pi\nu}. \tag{2.4}$$

Proof
Observe that

$$f_{\langle Y \rangle}(y) \;=\; \sum_{\nu=-\infty}^{\infty} f_Y(y+\nu).$$

Eq. (2.3) now follows as a direct consequence of (2.2).

Next, we use Eq. (2.3) to write

$$\tilde{f}_{\langle Y \rangle}(s) \;=\; \int_0^1 e^{-sy} \sum_{\nu=-\infty}^{\infty} \tilde{f}_Y(i2\pi\nu)\, e^{i2\pi\nu y}\, dy \;=\; \sum_{\nu=-\infty}^{\infty} \tilde{f}_Y(i2\pi\nu) \int_0^1 e^{(-s+i2\pi\nu)y}\, dy,$$

where the interchange of integration and summation is always permissible (see [34], p. 125). Eq. (2.4) now follows readily, noting that $e^{i2\pi\nu} = 1$ for all integer ν. $\bowtie$

Finally, let $U \sim$ Unif[0,1) be a random variable, uniformly distributed on the interval $[0,1)$. Then,

$$\tilde{f}_U(s) \;=\; \int_0^1 e^{-sx}\, dx \;=\; \begin{cases} 1, & s = 0 \\ \dfrac{1-e^{-s}}{s}, & s \neq 0 \end{cases} \tag{2.5}$$

whence, for any integer ν,

$$\tilde{f}_U(i2\pi\nu) = \begin{cases} 1, & \nu = 0 \\ 0, & \nu \neq 0. \end{cases} \tag{2.6}$$

3 Definitions of ARM Processes

Let $U_0 \sim$ Unif[0,1) be a random variable, uniformly distributed on the interval $[0, 1)$. An *innovation process* is any real-valued stochastic sequence, $\{V_n\}_{n=0}^{\infty}$, with the proviso that the sequence $\{V_n\}$ is independent of U_0. All random variables are defined over some common probability space $(\Omega, \mathcal{F}, P)$.

We begin by defining two stochastic sequences, $\{U_n^+\}_{n=0}^{\infty}$ and $\{U_n^-\}_{n=0}^{\infty}$, to be referred to as *background* ARM sequences due to their auxiliary nature. These are given by

$$U_n^+ = \begin{cases} \langle U_0 + V_0 \rangle, & n = 0 \\ \langle U_{n-1}^+ + V_n \rangle, & n > 0 \end{cases} \tag{3.1}$$

and

$$U_n^- = \begin{cases} U_n^+, & n \text{ even} \\ 1 - U_n^+, & n \text{ odd} \end{cases} \tag{3.2}$$

The plus and minus superscripts are used as mnemonics for the corresponding autocorrelation structure of the underlying sequences: background ARM sequences $\{U_n^+\}$ give rise primarily to a positive lag-1 autocorrelation, whereas background ARM sequences $\{U_n^-\}$ give rise primarily to a negative lag-1 autocorrelation. In fact, since Eq. (3.2) implies $\{U_n^-\} = \{U_0^+, 1 - U_1^+, U_2^+, 1 - U_3^+, \ldots\}$, it readily follows that the associated autocorrelations are related by

$$\rho_U^-(k, \tau) = (-1)^\tau \, \rho_U^+(k, \tau).$$

Henceforth, we shall routinely append a plus or minus superscript (or both) to other associated objects, unless the distinction is immaterial; for example, the random variable U_0 has no superscript, being common to the sequences $\{U_n^+\}$ and $\{U_n^-\}$, while the notation $\{U_n^\pm\}$ pertains to both sequences, in the respective order. We remark that the definitions (3.1) and (3.2) have an obvious geometric interpretation: the modulo-1 arithmetic renders background ARM processes random walks on the unit circle (a circle of circumference 1), where each point on the circle corresponds to a fraction in the range [0,1) (see [10] for more details).

Target ARM processes are transformed background ARM processes, and are consequently referred to as *foreground* ARM sequences. These are of the general form

$$X_n^+ = D(U_n^+) \tag{3.3}$$

and

$$X_n^- = D(U_n^-) \tag{3.4}$$

where D is a measurable transformation from $[0,1)$ to the real line $\mathbb{R}$, called a *distortion*. We assume throughout that $\int_0^1 D^2(x)\,dx < \infty$.

Useful alternative representations of the background ARM processes, (3.1) and (3.2), can be obtained by exploiting properties of modulo-1 arithmetic. First, it is readily seen that for any real numbers, a, b, and c,

$$\langle a \pm b \rangle = \langle \langle a \rangle \pm \langle b \rangle \rangle \tag{3.5}$$

and

$$\langle a + b \rangle = \langle c \rangle \iff \langle a \rangle = \langle c - b \rangle. \tag{3.6}$$

Hence, (3.5) applied to (3.1) implies the alternative representation

$$U_n^+ = \langle U_{n-1}^+ + \langle V_n \rangle \rangle, \quad n \geq 0. \tag{3.7}$$

In view of the representation (3.7), we incur no loss of generality if we replace $\{V_j\}$ by $\{\langle V_j \rangle\}$. More generally, we will henceforth assume that the state space of $\{V_j\}$ is contained in an interval of length 1, such as $[0,1)$ or $[-0.5, 0.5)$. In this case, Eq. (3.6) implies the useful equality of σ-algebras

$$\sigma\{U_k^+, \ldots, U_{k+n}^+\} = \sigma\{U_k^-, \ldots, U_{k+n}^-\} = \sigma\{U_k^+, V_k, \ldots, V_{k+n}\}, \; k, n \geq 0, \tag{3.8}$$

namely, that when the initial ARM variate, U_0, is known, then knowledge of further ARM variates is equivalent to knowledge of the corresponding innovation variates.

Next, define the partial sums

$$S_{m,n} = \begin{cases} 0, & m > n \\ \sum_{j=m}^{n} V_j, & m \leq n \end{cases} \tag{3.9}$$

A simple induction on (3.1), with the aid of (3.5) and (3.9), leads to the second alternative representation

$$U_n^+ = \langle U_0 + S_{0,n} \rangle = \langle U_0 + \langle S_{0,n} \rangle \rangle, \quad n \geq 0. \tag{3.10}$$

We conclude this section with a taxonomy and notational conventions for ARM processes. The foregoing definitions lead naturally to a broad classification of ARM processes into two classes, ARM$^+$ and ARM$^-$; note that this classification pertains both to background and foreground ARM processes, since the former constitute just a special case of foreground ARM processes with the identity distortion. To indicate other subclasses of ARM processes, determined by the probability law of the innovation process (such as AR(p), ARMA(p,q), etc.), we write, e.g., ARM$^+$(AR(p), D), ARM$^-$(ARMA(p,q), D), and so on. Finally, a particular ARM process will be specified by the notation ARM$^+(V,D)$ or ARM$^-(V,D)$, where $V = \{V_n\}$ is some innovation sequence and D is some distortion.

4 Stationarity Properties of ARM Processes

In this section we explore stationarity properties of ARM processes, namely, the invariance of their finite-dimensional distributions under temporal translation. More specifically, we say that a stochastic process $\{Z_n\}_{n=0}^{\infty}$ is *k-stationary*, if the joint distribution of the random vectors of the form $(Z_n, \ldots, Z_{n+k-1})$ and $(Z_{n+m}, \ldots, Z_{n+m+k-1})$ is identical, for all $n \geq 0$ and $m \geq 1$; in particular, the process is said to be *stationary*, if it is 1-stationary, i.e., if it has an invariant marginal distribution. We also deduce some transition properties of background ARM processes pertaining to the presence or absence of a Markovian structure.

A key motivation for studying ARM processes stems from the following lemma on the 1-stationarity of *all* ARM processes, and the form of the marginal distribution of *all* background ARM processes.

Lemma 2
(a) *Every background ARM process of the form (3.1) or (3.2) is 1-stationary with invariant* Unif[0,1) *marginal distribution.*
(b) *Every foreground ARM process of the form (3.3) or (3.4) is 1-stationary.*

Proof
(a) The lemma of General Iterated Uniformity (see Lemma 1 in [10]) states that if $U \sim$ Unif[0,1) and Y is any random variable independent of U, then $\langle U + Y \rangle \sim$ Unif[0,1). The proof in Lemma 1 of [10] used a probabilistic time domain argument. Because this result is central to ARM processes, we also present here a transform-domain proof, for completeness.

First, applying (2.4) to (3.10) yields

$$\begin{aligned}
\tilde{f}_{U_n^+}(s) &= \tilde{f}_{\langle U_0 + S_{0,n} \rangle}(s) = \sum_{\nu=-\infty}^{\infty} \tilde{f}_{U_0 + S_{0,n}}(i2\pi\nu)\, \frac{1 - e^{-s}}{s - i2\pi\nu} \\
&= \frac{1 - e^{-s}}{s} + \sum_{\substack{\nu=-\infty \\ \nu \neq 0}}^{\infty} \tilde{f}_{U_0}(i2\pi\nu)\, \tilde{f}_{S_{0,n}}(i2\pi\nu)\, \frac{1 - e^{-s}}{s - i2\pi\nu}.
\end{aligned}$$

But the sum above vanishes by virtue of (2.6), so that $U_n^+ \sim$ Unif[0,1) by Eq. (2.5). Finally the result follows for $\{U_n^-\}$ from (3.2), since if $U \sim$ Unif[0,1), then its antithetic variate also satisfies $1 - U \sim$ Unif[0,1).
(b) Follows from the fact that the distortion function, D, is common to all random variables in a given ARM process, and is applied to a stochastic sequence with an invariant distribution, as proven in (a) above. $\bowtie$

To appreciate the significance of Lemma 2, recall the desirability of defining a broad class of stochastic processes with versatile and flexible modeling properties. By ensuring that all *background* ARM processes have uniform marginals, Lemma 2 ensures in fact that any marginal distribution can be easily fitted to a *foreground* ARM process, by appeal to the inversion method: If $U \sim$ Unif[0,1) and F is any distribution function, then $F^{-1}(U) \sim F$. See, e.g., [2, 19]. Since this is true *regardless* of the innovation process selected, we can, in principle, attempt the fitting of both the empirical (marginal) distribution *and* empirical autocorrelation function, *simultaneously*. The two fittings are decoupled: the first is guaranteed an exact match via the inversion method, and the second is approximated by a suitable choice of an innovation process.

We next deduce some k-stationarity properties of ARM processes, using the following proposition which is a direct consequence of Lemma 2.

Proposition 1
For every $k \geq 0$, both U_k^+ and U_k^- are each independent of $\{V_n\}$.

Proof
For $n = 0$, U_0 is independent of $\{V_j\}$ by construction. Furthermore, it suffices to prove this assertion for $\{U_n^+\}$ processes only, since by (3.2) this will then extend trivially to $\{U_n^-\}$. For $k > 0$, and $0 \leq x < 1$, one has

$$P\{U_k^+ \leq x \,|\, V_0 = v_0, V_1 = v_1, \ldots\} \;=\; P\{\langle U_0 + \sum_{j=0}^{k} v_j \rangle \leq x\} \;=\; x,$$

where the first equality follows from (3.9) and the second from Lemma 2. Since the righthand side above is not a function of the $V_j = v_j$, the proposition follows. $\bowtie$

At first blush, Proposition 1 appears to be curious or even counter-intuitive. Although U_0 is independent of $\{V_j\}$, the subsequent U_n^+ and U_n^- are both functions of $V_0, \ldots, V_n$, as evidenced by (3.1) and (3.2), and could be expected to depend on them stochastically as well. Nevertheless, Lemma 2, as applied in Proposition 1, shows that this nominal functional dependence does not give rise to stochastic dependence. Note, however, that while Proposition 1 asserts that $\{V_j\}$ is independent of each singleton variate $U_n^{\pm}$, it does not follow that $\{V_j\}$ is independent of subsets $\{U_{n_1}^{\pm}, \ldots, U_{n_m}^{\pm}\}$, for $m > 1$. This can be verified by a straightforward calculation, based on the fact that the random variables U_{n-1}, U_n and V_n are always constrained by the relations in (3.1) and (3.2) (including the case of iid innovations, i.e., TES processes).

One important difference between background ARM processes and background TES processes is that the latter are Markovian, whereas the former are generally non-Markovian, when the innovation sequence is allowed to be dependent. For instance, to see that $\{U_n^+\}$ is not Markovian, one merely notes that (3.1) implies, for $n > 0$,

$$P\{U_n^+ \leq x \,|\, U_j^+, j = 0, \ldots, n-1\} \;=\; P\{\langle U_{n-1}^+ + V_n \rangle \leq x \,|\, U_0, V_j, j = 0, \ldots, n-1\}. \tag{4.1}$$

Now, when $\{V_j\}$ is iid, the rightmost side of (4.1) is readily seen to be a function of U^+_{n-1} alone on events of the form $\{U^+_{n-1} = u\}$, since V_n is independent of both U_0 and $\{V_j\}_{j=0}^{n-1}$ by construction. However, for general ARM processes, V_n is independent only of U_0 (by Proposition 1), but is dependent on $\{V_j\}_{j=0}^{n-1}$, and consequently U^+_n depends on the history $\{U^+_j\}_{j=0}^{n-1}$, rather than on U^+_{n-1} alone. Equation (4.1) suggests how to deduce when background ARM processes are k-stationary.

Proposition 2

If $\{V_j\}$ is a k-stationary sequence, then

(a) $\{U^+_n\}$ *is k-stationary*

(b) $\{U^-_n\}$ *is k-stationary relative to even translations only.*

Proof

(a) Recall that by Proposition 1, U^+_k is independent of the sequence $\{V_j\}$, for all $k \geq 0$; that both U^+_0 and U^+_k are uniform on $[0,1)$; and that $\{V_j\}$ was assumed k-stationary. Combining these facts with Eqs. (3.1) and (3.10) leads to the conclusion that the σ-algebra $\sigma\{U^+_0, V_0, \ldots, V_n\}$ and its k-step translation, $\sigma\{U^+_k, V_k, \ldots, V_{k+n}\}$, have the same probability law, for all $k \geq 0$. Eq. (3.8) then establishes the k-stationarity of $\{U^+_n\}$.

(b) A similar argument, employing (3.2) instead of (3.1), implies that $\{U^-_n\}$ is only k-stationary for even translations. $\bowtie$

5 The Transition Structure of ARM Processes

In this section we study the 2-dimensional structure of background ARM processes. Note that because the marginal distribution of any ARM background process is uniform on $[0,1)$, it follows that the joint distribution of any pair of background variates, $U^{\pm}_k$ and $U^{\pm}_{k+\tau}$, coincides with the corresponding transition functions (conditional distribution) of $U^{\pm}_{k+\tau}$ given $U^{\pm}_k$. Consequently, our goal is to derive transition densities of the form

$$g^{\pm}_{k+\tau|k}(y\,|\,x) \;=\; (\partial/\partial y) P\{U^{\pm}_{k+\tau} \leq y\,|\,U^{\pm}_k = x\}, \quad k \geq 0, \quad \tau > 0, \tag{5.1}$$

of transitioning on the unit circle from point x at time k to point y in τ steps.

For ARM$^+$ processes, the theorem below generalizes Theorem 1 in [10] and simplifies its proof. Here and elsewhere, $\Re$ is the real-part operator on complex numbers.

Theorem 1

For $\tau \geq 1$,

$$g^{+}_{k+\tau|k}(y\,|\,x) \;=\; \begin{cases} \sum\limits_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, e^{i2\pi\nu(y-x)}, & 0 \leq y, x < 1 \\ 0, & \text{otherwise} \end{cases} \tag{5.2}$$

$$= \begin{cases} 1 + 2\sum_{\nu=1}^{\infty} \Re\Big[\tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, e^{i2\pi\nu(y-x)}\Big], & 0 \le y, x < 1 \\ 0, & \text{otherwise} \end{cases} \tag{5.3}$$

Proof
From the representation (3.10) it follows that for all $n > 0$

$$P\{U^+_{k+\tau} \le y \,|\, U^+_k = x\} \;=\; P\{\langle x + S_{k+1,k+\tau}\rangle \le y\} \;=\; P\{\langle S_{k+1,k+\tau}\rangle \le \langle y - x\rangle\},$$

where the first equality above follows from Proposition 1, and the second from (3.6). Hence,

$$g^+_{k+\tau|k}(y\,|\,x) \;=\; f_{\langle S_{k+1,k+\tau}\rangle}(\langle y - x\rangle). \tag{5.4}$$

To obtain (5.2), it only remains to apply (2.3) to the righthand side of (5.4), noting that the periodicity of complex exponentials implies $e^{i2\pi\langle r\rangle} = e^{i2\pi r}$, for real r. The real-form representation (5.3) follows by noting further that the term corresponding to $\nu = 0$ in (5.2) evaluates to 1, while term pairs indexed there by $-\nu$ and ν are complex conjugates, and so their sum evaluates to twice their real part. $\bowtie$

Corollary 1
If $\{V_n\}$ is stationary of every order, then the quantities $\tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)$ in Theorem 1 may be replaced by $\tilde{f}_{S_{1,\tau}}(i2\pi\nu)$, and $\{U^+_n\}$ has stationary transition densities, i.e., $g^+_{k+\tau|k}(y\,|\,x) \;=\; g^+_\tau(y\,|\,x) = (\partial/\partial y)P\{U^+_\tau \le y \,|\, U^+_0 = x\}$ is independent of k.

Proof
Follows from Part (a) of Proposition 2. $\bowtie$

For ARM$^-$ processes, the next theorem generalizes Theorem 2 in [10] and corrects its statement (the latter is only correct for transition densities of $\{U^-_n\}$, conditioned on an even-index random variable).

Theorem 2
Let $\tau \ge 1$.
(a) *If $k \ge 0$ is even, then*

$$g^-_{k+\tau|k}(y\,|\,x) \;=\; \begin{cases} \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, e^{i2\pi\nu(y-x)}, & 0 \le y, x < 1, \quad \tau \text{ even} \\ \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, e^{-i2\pi\nu(y+x)}, & 0 \le y, x < 1, \quad \tau \text{ odd} \\ 0, & \text{otherwise} \end{cases} \tag{5.5}$$

$$= \begin{cases} 1 + 2\sum_{\nu=1}^{\infty} \Re\Big[\tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, e^{i2\pi\nu(y-x)}\Big], & 0 \le y, x < 1, \quad \tau \text{ even} \\ 1 + 2\sum_{\nu=1}^{\infty} \Re\Big[\tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, e^{-i2\pi\nu(y+x)}\Big], & 0 \le y, x < 1, \quad \tau \text{ odd} \\ 0, & \text{otherwise} \end{cases} \tag{5.6}$$

(b) *If $k \ge 0$ is odd, then*

$$g^{-}_{k+\tau|k}(y\,|\,x) = \begin{cases} \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, e^{-i2\pi\nu(y-x)}, & 0 \le y, x < 1, \quad \tau \text{ even} \\ \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, e^{i2\pi\nu(y+x)}, & 0 \le y, x < 1, \quad \tau \text{ odd} \\ 0, & \text{otherwise} \end{cases} \tag{5.7}$$

$$= \begin{cases} 1 + 2\sum_{\nu=1}^{\infty} \Re\Big[\tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, e^{-i2\pi\nu(y-x)}\Big], & 0 \le y, x < 1, \quad \tau \text{ even} \\ 1 + 2\sum_{\nu=1}^{\infty} \Re\Big[\tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, e^{i2\pi\nu(y+x)}\Big], & 0 \le y, x < 1, \quad \tau \text{ odd} \\ 0, & \text{otherwise} \end{cases} \tag{5.8}$$

Proof
(a) For even k, it follows from (3.2) that for any $\tau > 0$,

$$g^{-}_{k+\tau|k}(y\,|\,x) = \begin{cases} g^{+}_{k+\tau|k}(y\,|\,x), & 0 \le y, x < 1, \quad \tau \text{ even} \\ g^{+}_{k+\tau|k}(1-y\,|\,x), & 0 \le y, x < 1, \quad \tau \text{ odd} \\ 0, & \text{otherwise} \end{cases}$$

The case of τ even has already been derived in (5.2) and (5.3). The case of τ odd is readily obtained by substituting $1-y$ for y in (5.2) and (5.3). Eqs. (5.5) and (5.6) then easily follow.

(b) For odd k, it follows from (3.2) that for any $\tau > 0$,

$$g^{-}_{k+\tau|k}(y\,|\,x) = \begin{cases} g^{+}_{k+\tau|k}(1-y\,|\,1-x), & 0 \le y, x < 1, \quad \tau \text{ even} \\ g^{+}_{k+\tau|k}(y\,|\,1-x), & 0 \le y, x < 1, \quad \tau \text{ odd} \\ 0, & \text{otherwise} \end{cases}$$

The case of τ even is readily obtained by substituting $1-x$ for x in (5.2) and (5.3). Similarly, the case of τ odd is readily obtained by substituting $1-x$ for x and $1-y$ for y in (5.2) and (5.3). Eqs. (5.7) and (5.8) then easily follow. $\bowtie$

Corollary 2
If $\{V_j\}$ is stationary of every order, then the quantities $\tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)$ in Theorem 2

may be replaced by $\tilde{f}_{S_{1,\tau}}(i2\pi\nu)$, and each set of transition densities, $g^{-}_{2k+\tau|2k}(y\,|\,x)$ and $g^{-}_{2k+1+\tau|2k+1}(y\,|\,x)$, is independent of k.

Proof
Follows from Part (b) of Proposition 2. $\bowtie$

Since background ARM processes are not generally Markovian, neither are their foreground counterparts. A Markovian foreground process will ensue under suitable restrictions. A case in point is an ARM(GI,D), where GI stands for general independent (i.e., iid) innovations and D is 1-1. This case corresponds to TES processes with a 1-1 distortion; see [10, 11, 12, 24, 25]. Based on experience with TES processes, practical ARM models can be expected to be mostly non-Markovian.

6 The Autocorrelation Structure of ARM Processes

In this section we derive autocorrelation functions for ARM processes $\{X^{\pm}_n\}$, of the form (3.3) or (3.4). To this end, we assume the existence of non-zero second moments for all random variables in the ARM processes under consideration. Lemma 2 then ensures the existence of a common mean, μ_X,and common variance, σ^2_X, for the corresponding ARM process $\{X^{\pm}_n\}$, so that the autocorrelation function of Eq. (2.1) becomes

$$\rho^{\pm}_X(k,\tau) \;=\; \frac{E[X^{\pm}_k\, X^{\pm}_{k+\tau}] - \mu^2_X}{\sigma^2_X}, \quad k \ge 0, \quad \tau \ge 0. \tag{6.1}$$

Before proceeding, we make the following observations, which will be needed below.

Proposition 3
Let $\{X^{\pm}_n\}$ be a foreground ARM process.
(a) *The common expectation of random variables in $\{X^{\pm}_n\}$ is*

$$E[X^{\pm}_n] \;=\; \int_0^1 D(x)\,dx \;=\; \tilde{D}(0). \tag{6.2}$$

(b) *The expectation of products of random variables from $\{X^{\pm}_n\}$ is*

$$E[X^{\pm}_k\, X^{\pm}_{k+\tau}] \;=\; \int_0^1\int_0^1 D(x)\,D(y)\,g^{\pm}_{k+\tau|k}(y\,|\,x)\,dx\,dy. \tag{6.3}$$

Proof
(a) Follows from the uniform distribution of any ARM background variate, proven in Lemma 2.
(b) The [0,1) uniformity of any ARM background variate implies that the transition densities $g^{\pm}_{k+\tau|k}$ are also joint densities of pairs, $(U^{\pm}_k, U^{\pm}_{k+\tau})$, of background variates. The result then readily follows. $\bowtie$

The theorem below computes the autocorrelation function of an ARM$^+$ process, generalizing Theorem 3 in [10]. Although the proof is analogous, it is included for completeness.

Theorem 3
Let $\{X_n^+\}$ be an ARM process with finite variance. Then for any $\tau \geq 0$,

$$\begin{aligned} \rho_X^+(k,\tau) &= \frac{1}{\sigma_X^2} \sum_{\substack{\nu=-\infty \\ \nu\neq 0}}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu) \left|\tilde{D}(i2\pi\nu)\right|^2 && (6.4) \\ &= \frac{2}{\sigma_X^2} \sum_{\nu=1}^{\infty} \Re\Big[\tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\Big] \left|\tilde{D}(i2\pi\nu)\right|^2 && (6.5) \end{aligned}$$

Proof
To compute $\rho_X^+(k,\tau)$ from Eq. (6.1), substitute $g^+_{k+\tau|k}$ from (5.2) into (6.3), obtaining

$$\begin{aligned} E[X_k^+ \, X_{k+\tau}^+] &= \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu) \int_0^1 \int_0^1 D(x)\, D(y)\, e^{i2\pi\nu(y-x)}\, dx\, dy \\ &= \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu) \int_0^1 e^{-i2\pi\nu x}\, D(x)\, dx \int_0^1 e^{i2\pi\nu y}\, D(y)\, dy \\ &= \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, \tilde{D}(i2\pi\nu)\, \tilde{D}(-i2\pi\nu) \\ &= \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu) \left|\tilde{D}(i2\pi\nu)\right|^2, \end{aligned}$$

where the last equality follows from the fact that the pairs $\tilde{D}(i2\pi\nu)$ and $\tilde{D}(-i2\pi\nu)$ are complex conjugates. Furthermore, in view of (6.2) and $\tilde{f}_{S_{k+1,k+\tau}}(0) = 1$, it is readily seen that the term corresponding to $\nu = 0$ in the sum above is precisely $E[X_k^+]\, E[X_{k+\tau}^+]$. Thus, (6.4) follows by substituting (6.2) and the sum above into (6.1), while (6.5) follows from (6.4) by noting that terms indexed by $-\nu$ and ν in the sum above are complex conjugates. $\bowtie$

The next theorem computes the autocorrelation function of an ARM$^-$ process. It generalizes the results of Theorem 4 in [10], and corrects an error in its proof (the theorem itself states a correct result).

Theorem 4
Let $\{X_n^-\}$ be an ARM process with finite variance, and let $\tau \geq 0$.

(a) *If $k \geq 0$ is even, then*

$$\rho_X^-(k,\tau) = \begin{cases} \dfrac{1}{\sigma_X^2} \displaystyle\sum_{\substack{\nu=-\infty \\ \nu\neq 0}}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu) \left|\tilde{D}(i2\pi\nu)\right|^2, & \tau \text{ even} \\ \dfrac{1}{\sigma_X^2} \displaystyle\sum_{\substack{\nu=-\infty \\ \nu\neq 0}}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, \tilde{D}^2(i2\pi\nu), & \tau \text{ odd} \end{cases} \qquad (6.6)$$

$$= \begin{cases} \dfrac{2}{\sigma_X^2} \displaystyle\sum_{\nu=1}^{\infty} \Re\Big[\tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\Big] \, \Big|\tilde{D}(i2\pi\nu)\Big|^2, & \tau \text{ even} \\ \dfrac{2}{\sigma_X^2} \displaystyle\sum_{\nu=1}^{\infty} \Re\Big[\tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, \tilde{D}^2(i2\pi\nu)\Big], & \tau \text{ odd} \end{cases} \tag{6.7}$$

(b) *If $k \geq 0$ is odd, then*

$$\rho_X^-(k,\tau) = \begin{cases} \dfrac{1}{\sigma_X^2} \displaystyle\sum_{\substack{\nu=-\infty \\ \nu \neq 0}}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu) \, \Big|\tilde{D}(i2\pi\nu)\Big|^2, & \tau \text{ even} \\ \dfrac{1}{\sigma_X^2} \displaystyle\sum_{\substack{\nu=-\infty \\ \nu \neq 0}}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, \tilde{D}^2(-i2\pi\nu), & \tau \text{ odd} \end{cases} \tag{6.8}$$

$$= \begin{cases} \dfrac{2}{\sigma_X^2} \displaystyle\sum_{\nu=1}^{\infty} \Re\Big[\tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\Big] \, \Big|\tilde{D}(i2\pi\nu)\Big|^2, & \tau \text{ even} \\ \dfrac{2}{\sigma_X^2} \displaystyle\sum_{\nu=1}^{\infty} \Re\Big[\tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, \tilde{D}^2(-i2\pi\nu)\Big], & \tau \text{ odd} \end{cases} \tag{6.9}$$

Proof

(a) To compute $\rho_X^-(k,\tau)$ for k even, from Eq. (6.1), we note that $\rho_X^-(k,\tau) = \rho_X^+(k,\tau)$ has already been computed for even τ in Theorem 3, so it remains to compute $\rho_X^-(k,\tau)$ for odd τ. To this end, substitute $g_{k+\tau|k}^-(y|x)$, τ odd, from (5.5) into (6.3). Arguing similarly to the previous calculation, one obtains

$$\begin{aligned} E[X_k^- \, X_{k+\tau}^-] &= \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu) \int_0^1 \int_0^1 D(x)\, D(y)\, e^{-i2\pi\nu(y+x)}\, dx\, dy \\ &= \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu) \int_0^1 e^{-i2\pi\nu x}\, D(x)\, dx \int_0^1 e^{-i2\pi\nu y}\, D(y)\, dy \\ &= \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, \tilde{D}^2(i2\pi\nu). \end{aligned}$$

Again, the term corresponding to $\nu = 0$ in the sum above is precisely $E[X_k^-]\, E[X_{k+\tau}^-]$, by appeal to (6.2) and the fact that $\tilde{f}_{S_{k+1,k+\tau}}(0) = 1$. Thus, (6.6) follows by substituting (6.2) and the sum above into (6.1), while (6.7) follows from (6.6) by noting that terms indexed by $-\nu$ and ν in the sum above are complex conjugates.

(b) To compute $\rho_X^-(k,\tau)$ for k odd, from Eq. (6.1), we perform the computation for even τ and odd τ separately. Only the expectation of the products will be computed, as the rest of the argument in Part (a) applies to the two remaining cases.

For k odd and τ even, substitute the corresponding $g_{k+\tau|k}^-(y|x)$, from (5.7) into (6.3), yielding

$$E[X_k^- \, X_{k+\tau}^-] = \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu) \int_0^1 \int_0^1 D(x)\, D(y)\, e^{-i2\pi\nu(y-x)}\, dx\, dy$$

$$= \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu) \int_0^1 e^{i2\pi\nu x}\, D(x)\, dx \int_0^1 e^{-i2\pi\nu y}\, D(y)\, dy$$
$$= \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, \tilde{D}(-i2\pi\nu)\, \tilde{D}(i2\pi\nu).$$

from which the requisite result follows, using the arguments of Part (a).

For k odd and τ odd, substitute the corresponding $g^-_{k+\tau|k}(y|x)$, from (5.7) into (6.3), yielding

$$\begin{aligned} E[X_k^-\, X_{k+\tau}^-] &= \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu) \int_0^1 \int_0^1 D(x)\, D(y)\, e^{i2\pi\nu(y+x)}\, dx\, dy \\ &= \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu) \int_0^1 e^{i2\pi\nu x}\, D(x)\, dx \int_0^1 e^{i2\pi\nu y}\, D(y)\, dy \\ &= \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)\, \tilde{D}^2(-i2\pi\nu). \end{aligned}$$

from which the requisite result follows, using the arguments of Part (a). $\bowtie$

From a modeling standpoint, fitting the autocorrelation function of Eq. (6.1) to empirical data involves a highly non-parsimonious set of parameters, as the corresponding $\rho_X^{\pm}(k,\tau)$ is a function of two variables. Thus, practical fitting of ARM processes to empirical data motivates us to seek definitions and conditions that remove the dependence of the requisite autocorrelation functions on k. We therefore seek "simplified" autocorrelation functions $\rho_X^+(k,\tau) = \rho_X^+(\tau)$ for ARM$^+$ processes $\{X_n^+\}$, and $\rho_X^-(k,\tau) = \rho_X^-(\tau)$ for ARM$^-$ processes $\{X_n^-\}$, which depend on the lag τ alone. Furthermore, such $\rho_X^+(\tau)$ and $\rho_X^-(\tau)$ should mirror the corresponding empirical estimators $\hat{\rho}_s(\tau)$, formed from a sample $y = \{y_1, \ldots, y_L\}$ as

$$\hat{\rho}_y(\tau) = \frac{\frac{1}{L-\tau}\sum_{k=1}^{L-\tau} y_k y_{k+\tau} - \hat{\mu}_y^2}{\hat{\sigma}_y^2}, \quad 1 \le \tau < L, \tag{6.10}$$

where $\hat{\mu}_y$ and $\hat{\sigma}_y^2$ are the sample mean and sample variance of y.

To this end the stationarity of the underlying innovation process is necessary. For ARM$^+$ processes, this is also sufficient as demonstrated in the proposition below.

Proposition 4
Let $\{X_n^+\}$ be an ARM process with finite common variance, whose innovation process is stationary for every order k. Then, for any $\tau \ge 0$, Eq. (6.1) is given by

$$\rho_X^+(\tau) = \frac{1}{\sigma_X^2} \sum_{\substack{\nu=-\infty \\ \nu\neq 0}}^{\infty} \tilde{f}_{S_{1,\tau}}(i2\pi\nu) \left|\tilde{D}(i2\pi\nu)\right|^2 \tag{6.11}$$

$$= \frac{2}{\sigma_X^2} \sum_{\nu=1}^{\infty} \Re\left[\tilde{f}_{S_{1,\tau}}(i2\pi\nu)\right] \left|\tilde{D}(i2\pi\nu)\right|^2 \tag{6.12}$$

Proof

Follows immediately by applying Corollary 1 to Theorem 3. $\bowtie$

An analogous definition for $\{X_n^-\}$, however, does not work, since the dependence of $g^-_{k+\tau|k}(y\,|\,x)$ on k cannot be removed completely by dint of the stationarity of $\{V_n\}$ alone; specifically, $g^-_{k+\tau|k}(y\,|\,x)$ would still depend on whether k is even or odd and consequently, so would $E[X_k^-\, X_{k+\tau}^-]$. Looking back at the estimator (6.10), it is clear that even and odd k occur in the sum-of-products estimator of $E[X_k^-\, X_{k+\tau}^-]$ in equal proportions, and this observation motivates the definition

$$\rho_X^-(\tau) \;=\; \frac{\frac{1}{2}E[X_k^-\, X_{k+\tau}^-|k \;\text{ even}] + \frac{1}{2}E[X_k^-\, X_{k+\tau}^-|k \;\text{ odd}] - \mu_X^2}{\sigma_X^2}. \tag{6.13}$$

Eq. (6.13) can be interpreted as an unconditioned (averaged) conditional autocorrelation function, where the index of the starting variate, k, is random with probability 1/2 of being even or odd. This interpretation coincides with the operational meaning of (6.10). We now derive the formula for (6.13) in the following proposition.

Proposition 5

Let $\{X_n^-\}$ be an ARM process with finite common variance, whose innovation process is stationary for every order k. Then, for any $\tau \geq 0$, Eq. (6.13) is given by

$$\rho_X^-(k,\tau) \;=\; \begin{cases} \dfrac{1}{\sigma_X^2} \displaystyle\sum_{\substack{\nu=-\infty \\ \nu\neq 0}}^{\infty} \tilde{f}_{S_{1,\tau}}(i2\pi\nu)\, \left|\tilde{D}(i2\pi\nu)\right|^2, & \tau \text{ even} \\[2ex] \dfrac{1}{\sigma_X^2} \displaystyle\sum_{\substack{\nu=-\infty \\ \nu\neq 0}}^{\infty} \tilde{f}_{S_{1,\tau}}(i2\pi\nu)\, \dfrac{\tilde{D}^2(i2\pi\nu) + \tilde{D}^2(-i2\pi\nu)}{2}, & \tau \text{ odd} \end{cases} \tag{6.14}$$

$$\;=\; \begin{cases} \dfrac{2}{\sigma_X^2} \displaystyle\sum_{\nu=1}^{\infty} \Re\Big[\tilde{f}_{S_{1,\tau}}(i2\pi\nu)\Big]\, \left|\tilde{D}(i2\pi\nu)\right|^2, & \tau \text{ even} \\[2ex] \dfrac{2}{\sigma_X^2} \displaystyle\sum_{\nu=1}^{\infty} \Re\Big[\tilde{f}_{S_{1,\tau}}(i2\pi\nu)\Big]\, \Re\Big[\tilde{D}^2(i2\pi\nu)\Big], & \tau \text{ odd} \end{cases} \tag{6.15}$$

Proof

First, apply Corollary 2 to Theorem 4, to replace the quantities $\tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)$ by the quantities $\tilde{f}_{S_{1,\tau}}(i2\pi\nu)$. Eq. (6.14) now follows easily from Eq.(6.13) by averaging the even-τ cases and odd-τ cases, separately, over k in Eq. (6.6) (k even) and Eq. (6.8) (k odd). To show that Eq. (6.15) follows from Eq. (6.14), notice that the even-τ case is trivial, while for the odd-τ case, a simple computation yields

$$\frac{\tilde{D}^2(i2\pi\nu) + \tilde{D}^2(-i2\pi\nu)}{2} = \Re\Big[\tilde{D}^2(i2\pi\nu)\Big]\,,$$

by noting that the averaged quantities are complex conjugates. The rest follows from arguments similar to those in Theorem 4. $\bowtie$

7 Specialized Distortions for ARM Processes

The theory developed in this paper so far is rather general, the foregoing formulas being valid for any distortion which is square-integrable on the interval $[0, 1)$. Practical ARM modeling, however, focuses on distortions based on inverse distribution functions via an application of the inversion method [2, 19]. An empirical density is often estimated by a histogram $\hat{h}_Y$, which gives the observed relative frequencies of various ranges for a realization of an empirical sequence $\{Y_n\}_{n=0}^N$; here and elsewhere, hats indicate estimated quantities, as opposed to modeler-supplied parameters. More specifically, a histogram $\hat{h}_Y$ is a step function of the form

$$\hat{h}_Y(y) \;=\; \sum_{j=1}^{J} 1_{[l_j,r_j)}(y) \frac{\hat{p}_j}{w_j}, \tag{7.1}$$

where J is the number of histogram cells, each of the form $[l_j, r_j)$; $w_j = r_j - l_j$ is the width of cell j; and $\hat{p}_j$ is the probability estimate (relative frequency) of cell j. The corresponding distribution function, $\hat{H}_Y$, is a piecewise linear function, obtained by integrating (7.1) with respect to y, and its inverse, $\hat{H}_Y^{-1}$, is the piecewise linear function

$$\hat{H}_Y^{-1}(x) \;=\; \sum_{j=1}^{J} 1_{[\hat{C}_{j-1},\hat{C}_j)}(x) \left[l_j + (x - \hat{C}_{j-1}) \frac{w_j}{\hat{p}_j} \right], \tag{7.2}$$

where $\{\hat{C}_j\}$ is the distribution function of the discrete probability density $\{\hat{p}_j\}_{j=1}^J$, namely, $\hat{C}_j = \sum_{i=1}^{j} \hat{p}_i$, with $\hat{C}_0 = 0$ and $\hat{C}_J = 1$.

However, a naive application of a distortion of the form $D = F^{-1}$ is often inadequate (F is the fitted distribution function, such as a histogram distortion of the form (7.2)). The reason is that background ARM processes, being random walks on the unit circle, produce visual "discontinuities" when the background process crosses the circle origin. This phenomenon stems from the fact that the effect of crossing the origin in the clockwise direction is to produce relatively large values (close to 1), followed by relatively small ones (close to 0); a counter-clockwise crossing is analogous (but the results are in reverse order.) Furthermore, a monotone transformation (including F^{-1}) preserves this effect. For a more detailed discussion refer to [10] or [25].

To circumvent this problem, we use, in practice, 2-stage composite distortions of the form

$$D_{Y,\xi}(x) \;=\; \hat{H}_Y^{-1}(S_\xi(x)). \tag{7.3}$$

Here, the role of the first stage, S_ξ, is to "smooth" the sample paths of ARM processes, with the uniform marginals preserved intact, while the second stage, $\hat{H}_Y^{-1}$, carries out a standard inversion (in this case a histogram inversion). The first stage is called a *stitching transformation* and is given by

$$S_\xi(x) \;=\; \begin{cases} x/\xi, & 0 \le x < \xi \\ (1-x)/(1-\xi), & \xi \le x < 1. \end{cases} \tag{7.4}$$

The parameters ξ, $0 \le \xi \le 1$, are called *stitching parameters*. We point out that the smoothing action of stitching transformations is effected for $0 < \xi < 1$, due to their continuity on the unit circle; in particular, the motivation for the term "stitching" stems from the fact that $S_\xi(0) = S_\xi(1-)$ for all $0 < \xi < 1$. Note that for the extremal cases, $\xi = 1$ and $\xi = 0$, the corresponding stitching transformations are $S_1(x) = x$ (the identity), and $S_0(x) = 1 - x$ (the anti-thetic). These two are "non-smoothing", as they are not continuous on the unit circle. Practical modeling most often employs $0 < \xi < 1$, and primarily the mid-point $\xi = 0.5$. For a discussion of the smoothing properties of stitching transformations and their uniformity-preserving property, refer to [10, 24, 25].

The introduction of stitching as a smoothing operation does not introduce undue difficulty in calculating the Laplace transform of the resulting composite distortion. More specifically, for any general composite distortion of the form

$$D_\xi(x) \;=\; D(S_\xi(x)), \tag{7.5}$$

where D is any distortion, the corresponding Laplace transform is (see [10] for details)

$$\tilde{D}_\xi(i2\pi\nu) \;=\; \xi\,\tilde{D}(i2\pi\nu\xi) \;+\; (1-\xi)\,\tilde{D}(i2\pi\nu(1-\xi)), \quad \nu \ge 0.$$

Thus, if the original $\tilde{D}(s)$ is known, then any composite counterpart, $\tilde{D}_\xi(i2\pi\nu)$, is readily computable as well; the latter may then be used to compute the autocorrelation functions (6.4) – (6.9), as well as the computational formulas (6.11) – (6.12) and (6.14) – (6.15).

Finally, we mention that the quantities $\tilde{D}(i2\pi\nu)$ and $\tilde{D}_\xi(i2\pi\nu)$ have been computed for various distributions, including the histogram distortions $D = \hat{H}_Y^{-1}$ of the form (7.2). For details, refer to [11].

8 Specialized Innovations for ARM Processes

Sections 5 and 6 have shown that the Laplace transform quantities $\tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)$ are essential to the calculation of various ARM process statistics, including transition densities and autocorrelation functions. In this section we discuss several important innovation processes and the calculation of the associated $\tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)$.

8.1 Independent Identically Distributed Innovations

The simplest innovation processes of practical use consist of iid sequences, resulting in ARM(GI,D) processes which correspond to the special case of TES processes. Here, the calculation of the Laplace transform quantities is simplified by the fact that

$$\tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu) = \tilde{f}_V^{\tau}(i2\pi\nu),$$

so only knowledge of $\tilde{f}_V(i2\pi\nu)$ is required, as their powers are readily computed. The $\tilde{f}_V(i2\pi\nu)$ have been computed in [11] for various densities, including step-function densities, which have been widely used in practical TES modeling [25].

8.2 Moving Average Innovations

Let $\{Z_j\}_{j=-q}^{\infty}$ be an iid sequence of random variables for some $q \geq 0$. Define a moving average innovation sequence, $\{V_n\}$, by

$$V_n \;=\; \sum_{j=0}^{q} a_j \, Z_{n-j}, \quad n \geq 0, \tag{8.1}$$

for some real constants $a_0, \ldots, a_q$. The resulting model is an ARM(MA(q), D) process. This model was studied in [5] in the context of compressed video traffic modeling.

8.3 Markovian and Markov-Modulated Innovations

When the innovation sequence is Markovian, the resulting ARM(M, D) processes enjoy a marked modeling flexibility. For example, TES processes, being ARM(GI,D) processes, are entirely subsumed as a rather restricted special case.

For general ARM(M, D) processes, the computation of the Laplace transform quantities, $\tilde{f}_{S_{k+1,k+\tau}}(i2\pi\nu)$, can be difficult or even impractical. In this section, we focus on a useful class of randomized (finite-state) Markov-modulated innovations, which can be thought of as Markovian processes whose 1-step transition densities are step functions, obtained as a discretized approximation of a general transition density on the unit circle. It is important to realize though, that the step-function structure presents no practical loss of generality, since step functions can approximate diffuse densities arbitrarily closely.

Let $\{K_n\}_{n=0}^{\infty}$ be a Markov chain, with state space $\mathcal{K} = \{0, 1, \ldots, M-1\}$, for some positive integer M, and an $M \times M$ transition matrix $\mathbf{Q} = [q_{i,j}]$. Let the vector $\mathbf{p} = (p_0, p_1, \ldots, p_{M-1})$ be some (initial) distribution vector for K_0. Define an innovation sequence, $\{V_n\}_{n=0}^{\infty}$, modulated by $\{K_n\}$, as

$$V_n \;=\; \sum_{k=0}^{M-1} 1_{I_k}(K_n) \, W_n^{(k)}, \tag{8.2}$$

where $I_k = [kh, (k+1)h)$, $h = 1/M$, and the $W_n^{(k)} \sim \mathrm{Unif}[kh, (k+1)h)$ are mutually independent random variables in n and k. Note that for $\{V_n\}$ to be stationary, one must select $\mathbf{p}$ to be an invariant distribution of $\mathbf{Q}$, that is,

$$\mathbf{p}\,\mathbf{Q} = \mathbf{p}. \tag{8.3}$$

For simplicity, assume that $\mathbf{p}$ in (8.3) is the unique invariant distribution of $\mathbf{Q}$.

The intuition behind this construction is straightforward. First, the unit interval (or unit circle) is partitioned into M subintervals I_k of equal length h, each corresponding to a state $k \in \mathcal{K}$. The process $\{K_n\}$ can be thought of as visiting the intervals I_k, governed by the Markovian transition matrix $\mathbf{Q}$. Next, given that $K_n = k$, the current innovation variate, V_n, is obtained by sampling a value uniformly in the interval I_k, independently of the past. Thus, the innovation sequence $\{V_n\}$ is Markovian with step-function transition density. Note that the probability of assuming the next value,V_{n+1}, depends only on the (random) index, K_n, of the previous interval visited, I_{K_n}, rather than on V_n. Moreover, once the next subinterval, $I_{K_{n+1}}$, is selected, the next innovation variate, V_{n+1}, is selected uniformly over it, independently of all past selections. The foregoing construction represents an approximation of a general continuous-state Markov chain by an essentially discretized, though continuous-state, version which is analogous to the standard approximation of measurable functions by simple functions (step functions). Clearly, the transition matrix $\mathbf{Q}$ completely determines the probability law of $\{V_n\}$ in (8.2), up to an initial distribution.

We now proceed to compute the associated Laplace transforms $\tilde{f}_{S_{k+1,k+\tau}}$. Since we are interested primarily in stationary innovations, we restrict the discussion to the computation of $\tilde{f}_{S_{1,\tau}}$. First, observe that every conditional Laplace transform $\tilde{f}_{V_n}(s \,|\, K_n = k)$ coincides with the Laplace transform $\tilde{f}_{W_n^{(k)}}(s)$, by construction. Thus, for all $k \in \mathcal{K}$,

$$\tilde{f}_{V_n}(s \,|\, K_n = k) \;=\; \tilde{f}_{W_n^{(k)}}(s) \;=\; \int_{kh}^{(k+1)h} e^{-sx}\,\frac{1}{h}\,dx \;=\; \frac{e^{-skh} - e^{-s(k+1)h}}{s\,h}. \tag{8.4}$$

But by construction, the innovations $\{V_0, V_1, \ldots, V_\tau\}$ are conditionally independent, given the jump-chain event, $\{K_0 = k_0, K_1 = k_1, \ldots, K_\tau = k_\tau\}$. Hence, by (8.4), we can write

$$\tilde{f}_{S_{1,\tau}}(s \,|\, K_0 = k_0, K_1 = k_1, \ldots, K_\tau = k_\tau) \;=\; \prod_{j=0}^{\tau} \frac{e^{-sk_jh} - e^{-s(k_j+1)h}}{s\,h} \;=\; \left(\frac{1-e^{-sh}}{sh}\right)^{\tau} e^{-sh\sum_{j=0}^{\tau} k_j}. \tag{8.5}$$

where the products in (8.5) represent convolutions. It remains to uncondition (8.5). Since the joint distribution of jump chains of $\{K_n\}$ are governed by powers of $\mathbf{Q}$, we have

$$\tilde{f}_{S_{1,\tau}}(s) \;=\; \mathbf{p}\,\mathbf{R}^{\tau}(s)\,\mathbf{1}, \quad n \geq 1, \tag{8.6}$$

where $\mathbf{1}$ is a column vector of 1's, and the $M \times M$ matrix $\mathbf{R}(s)$ is given by

$$\mathbf{R}(s) = \begin{bmatrix} q_{0,0}\,\tilde{f}_{W^{(0)}}(s) & q_{0,1}\,\tilde{f}_{W^{(1)}}(s) & \dots & q_{0,M-1}\,\tilde{f}_{W^{(M-1)}}(s) \\ q_{1,0}\,\tilde{f}_{W^{(0)}}(s) & q_{1,1}\,\tilde{f}_{W^{(1)}}(s) & \dots & q_{1,M-1}\tilde{f}_{W^{(M-1)}}(s) \\ \vdots & \vdots & & \vdots \\ q_{M-1,0}\,\tilde{f}_{W^{(0)}}(s) & q_{M-1,1}\,\tilde{f}_{W^{(1)}}(s) & \dots & q_{M-1,M-1}\,\tilde{f}_{W^{(M-1)}}(s) \end{bmatrix} \tag{8.7}$$

To see that, note that the k-th component in the row vector $\mathbf{p}\,\mathbf{R}^{\tau}(s)$ (with the usual convention that $\mathbf{R}^0(s) = I$ is the identity matrix) corresponds to all possibilities of landing in a state $k \in \mathcal{K}$, starting with initial distribution $\mathbf{p}$, and accumulating the appropriate partial sum $S_{1,\tau}$. Indeed, the matrix $\mathbf{R}(s)$ is constructed in such a way that every transition to state $k \in \mathcal{K}$ contributes a factor $\tilde{f}_{W^{(k)}}(s)$ to the corresponding product.

Next, in view of Eq. (8.4), the elements of $\mathbf{R}(s)$, $r_{j,k}(s)$, in (8.7) can be written explicitly as

$$r_{j,k}(s) = q_{j,k}\tilde{f}_{W^{(k)}}(s) = q_{j,k}\,e^{-skh}\,\frac{1-e^{-sh}}{s\,h}, \quad 0 \le j,k \le M-1. \tag{8.8}$$

An efficient recursive computation of $\tilde{f}_{S_{1,\tau}}(s)$ is given by the following algorithm (the explicit quantities from Eqs. (8.4) and (8.8) should be substituted in actual computations):

1. Set $\mathbf{w}^{(1)}(s) \leftarrow (p_0\tilde{f}_{W^{(0)}}(s), p_1\tilde{f}_{W^{(1)}}(s), \dots, p_{M-1}\tilde{f}_{W^{(M-1)}}(s))$.
2. Set $\tilde{f}_{S_{1,1}}(s) \leftarrow \tilde{f}_{V_1}(s) = \sum_{k=0}^{M-1} p_j\,\tilde{f}_{W^{(k)}}(s) = \mathbf{w}^{(1)}(s)\,\mathbf{1}$.
3. Continuing by induction, assume that $\mathbf{w}^{(\tau-1)}(s)$ has already been computed, and set $\mathbf{w}^{(\tau)}(s) \leftarrow \mathbf{w}^{(\tau-1)}(s)\,\mathbf{R}(s)$.
4. Set $\tilde{f}_{S_{1,\tau}}(s) \leftarrow \mathbf{w}^{(\tau)}(s)\,\mathbf{1}$.

When the process $\{K_n\}$ is started in equilibrium, an extra initial step is needed, i.e., the computation of the initial distribution vector $\mathbf{p}$ as the solution of (8.3).

9 The ARM Modeling Methodology

Suppose that we wish to fit an ARM model to some empirical record $\{Y_n\}$, from which the empirical distribution (histogram) and empirical autocorrelation function are calculated as estimates. The ARM modeling methodology searches for an innovation sequence, $\{V_n\}$, and a stitching parameter, ξ, such that the resulting ARM$(V, \xi, \hat{H}_Y^{-1})$ process gives rise to an autocorrelation function which approximates well its empirical counterpart; recall that ARM theory ensures that the associated marginal distribution is automatically matched to its empirical counterpart.

To fix the ideas, we shall focus throughout this section on the ARM$(\mathbf{Q}, \xi, \hat{H}_Y^{-1})$ of Section 8.3 with a Markov-modulated randomized innovation sequence, specified by the an $M \times M$ transition matrix $\mathbf{Q}$, for some $M > 0$. As was demonstrated in the TEStool modeling environment [8], we proceed to sketch how ARM modeling may be implemented as a heuristic interactive search with extensive modeler interaction, or as an algorithmic search with minimal modeler intervention. Both approaches would require software support.

The first step is to select the ARM process sign (ARM$^+$ or ARM$^-$). Such a selection would be typically guided by the modeler's experience; however, trying both classes would simply double the modeling time.

The second step is to construct the empirical inverse distribution function, $\hat{H}_Y^{-1}$, from the empirical data. Once the modeler specifies the number of histogram cells and their width, this function is easily computed from (7.3).

The third step, and the hardest one, is to search for a pair, $(\mathbf{Q}, \xi)$, consisting of an innovation process (in this case, a finite-state Markov transition matrix, $\mathbf{Q}$, and a stitching parameter, ξ), which jointly give rise to an ARM process that approximates well the leading empirical autocorrelations.

In the heuristic approach, the modeler embarks on a heuristic search of pairs $(\mathbf{Q}, \xi)$, guided by visual feedback [25]. The modeler would repeatedly perturb the pair $(\mathbf{Q}, \xi)$, and would observe the resultant autocorrelation fit to guide the next perturbation, if any. Though "hands-on" and therefore instructive, this approach may be time-consuming and consequently can engender considerable tedium. The modeler may have to acquire considerable experience for the heuristic search to be sufficiently efficient (a blind trial-and-error search is bound to be inefficient, since the search space is quite large).

Alternatively, an algorithmic search approach can be devised, analogously to [15], which describes in detail an algorithmic search approach for TES processes, namely, the case of iid innovations. In this approach, a brute-force search is combined with non-linear optimization to produce a set of ARM candidate models. First, the parameter space, $\{(\mathbf{Q}, \xi)\}$, is quantized into a moderately large, but finite, set $\mathcal{A}$ of ARM model specifications, which "sample" the underlying infinite space of ARM$(\mathbf{Q}, \xi, \hat{H}_Y^{-1})$ models. The autocorrelation function is computed for each model in $\mathcal{A}$ via the computationally-fast formulas (6.11) – (6.12) and (6.14) – (6.15). A "best set" of m models is selected from $\mathcal{A}$ (the value of m would be typically small, on the order of 10); the associated "goodness-of-fit" criterion would select ARM models whose autocorrelation functions have the smallest distance to the their empirical counterpart in some prescribed metric. A standard choice of such a metric is the weighted sum of the squared autocorrelation differences, the sum ranging over some span of lags. Next, a Steepest-descent non-linear optimization is performed for models in the "best set" only, to further reduce the aforementioned autocorrelation distance. The modeler then selects the "best of the best", either as the model with the overall smallest distance, or an optimized model in the "best set" that gives rise to Monte Carlo sample paths which appear most "similar" to the empirical data.

Formally, the algorithmic search approach seeks to solve the following nonlinear optimization problem:

Problem 1 *For a fixed histogram inverse distribution (7.2), find an optimal transition matrix and stitching parameter,* $(\mathbf{Q}^*, \xi^*)$*, such that*

$$(\mathbf{Q}^*, \xi^*) \;=\; \operatorname*{argmin}_{(\mathbf{Q},\xi)} \left\{ \sum_{\tau=1}^{L} a_\tau \left[\rho_{\mathbf{Q},\xi}(\tau) - \hat{\rho}_Y(\tau)\right]^2 \right\}, \tag{9.1}$$

where L *is the maximal autocorrelation lag to be approximated, the* $0 \le a_\tau \le 1$ *are fixed weighting coefficients, and* $\rho_{\mathbf{Q},\xi}(\tau)$ *is the autocorrelation function of the ARM model specified by the parameters* $\mathbf{Q}$ *and* ξ*.*

Note that the sum on the righthand side of (9.1) is a metric in terms of a weighted sum of squared differences (errors) between the empirical and model autocorrelation functions. The weighting factors are typically chosen as a geometric series, $a_\tau = a^\tau$, for some positive fraction, a, so as to give preference to fitting the head, rather than the tail, of the empirical autocorrelation function. Problem 1 is a finite-dimensional nonlinear optimization problem with two nice properties. First, it is subject to simple linear constraints; and second, one can derive analytical formulas for the objective function and its partial derivatives with respect to every optimization variable (see [15] for a detailed derivation for TES processes.) Consequently, Problem 1 is amenable to a variety of standard nonlinear programming techniques [1].

A practical implementation of an algorithmic solution for Problem 1 is a two-stage algorithm, dubbed the GSLO (Global Search Local Optimization) algorithm in [15]. The algorithm input consists of four (integer) parameters:

1. The dimension, M, of the matrix $\mathbf{Q}$.

2. The number, B, of "best" solutions to be returned.

3. The discretization level, N_Q, for elements $q_{i,j}$ of the matrix $\mathbf{Q}$, namely, the number of values assumed by each element.

4. The discretization level, N_ξ, for the stitching parameter, ξ.

The GSLO algorithm is sketched below.

<u>GSLO Algorithm Outline</u>

GS Discretize the parameter space into a finite set, $\mathcal{G}_M$, of points of the form $(\mathbf{Q}, \xi)$. Then, for each such point, evaluate the objective function in Eq. (9.1), and keep the best B points (those points, $\mathbf{x} \in \mathcal{G}_M$, which give rise to the B of the objective function.)

LO Using each of these B points as a starting point, find a local minimum of the objective function in Eq. (9.1) via a nonlinear programming algorithm. $\bowtie$

Note that the global search first selects the B most promising initial points for the local search, so as to increase the chance that the best local minimum found is relatively close to the global minimum. However, the analyst is free to select a less optimal model, if its simulated realizations are judged to bear "closer resemblance" to the empirical record.

A computer implementation of the **GS** step is straightforward. The objective function is calculated inside $M^2 + 1$ nested loops (one loop for the discretization of each $q_{i,j}$, and one for that of ξ); the computed value is then compared with the current best set, namely, the running set of the (at most) B best values in some sorting order. If the newly computed value improves on the worst value in the current best set, then the worst value is discarded and the new value is added in the sorting order. We continue this way until the whole discretized parameter space is searched.

The local optimization in the **LO** step above may be implemented, using the nonlinear programming method, called *Zoutendijk Feasible Direction Method*; see [1], pp. 409. This method is an iterative procedure where at each iteration one determines (i) the optimal feasible direction for the choice of the next point, and (ii) the optimal step size in that direction. The details of this method appear in [1] and are omitted here. The interested reader is referred to [15] for the details of its application to TES processes.

We conclude by mentioning that both approaches (heuristic and algorithmic) have proved efficacious in TEStool, yielding excellent TES models for a variety of application domains [27]. Similarly good results can be expected for ARM processes, and especially the Markovian innovations of Section 8.3.

10 Conclusion

This paper generalized the class of TES processes to the broader class of ARM processes. Formulas were derived for the transition structure and second-order statistics of the ARM class, and specializations of ARM processes were exhibited. Finally, an ARM modeling methodology was outlined.

It is hoped that ARM processes would provide better models of empirical time series, in view of their enhanced modeling flexibility. Experience with ARM modeling will be reported elsewhere.

Acknowledgments

I would like to thank David Jagerman and Andrzej Ruszczynski for carefully reading and commenting on this paper.

This work was supported in part by NSF grant DMI-9812858, and by DARPA grant AO-F316, contract #N6601-97-1-8913.

References

[1] M.S. Bazaraa, H.D. Sherali and C.M. Shetty, *Nonlinear Programming*, John Wiley & Sons, 1993.

[2] Bratley, P., Fox, B.L. and Schrage, L.E., *A Guide to Simulation*, Springer-Verlag, 1987.

[3] Cario, M.C. and Nelson, B.L., "Autoregressive to Anything: Time-Series Input Processes for Simulation", *OR Letters* **19**, 51–58, 1996.

[4] De Bruijn, N.G., "Asymptotic Methods in Analysis", *North-Holland*, 1958.

[5] Fang, Y. and Lambadaris, I., "Modeling Compressed Video Traffic Using ARM(MA(k), D) Processes", Manuscript, Carleton University, 1995.

[6] Fang, Y., Devetsikiotis, M., Lambadaris, I. and Kaye, A.R., "Exponential Bounding Techniques for the Waiting Time Distribution of TES/GI/1 Queues", *ACM SIGMETRICS and Performance '95*, 108–115, Ottawa, May 1995.

[7] Feller W., *An Introduction to Probability Theory and its Applications*, Wiley, 1971.

[8] Hill, J.R. and Melamed, B., "TEStool: A Visual Interactive Environment for Modeling Autocorrelated Time Series", *Performance Evaluation* **24(1&2)**, 3–22, 1995.

[9] Ismail, M.R., Lambadaris, I., Devetsikiotis, M. and Kaye, A. R., "Simulation and modeling of Variable Bit Rate MPEG Video Transmission over ATM Networks", *International Journal of Communication Systems* **9**, 283-297, 1996.

[10] Jagerman, D.L. and Melamed, B., "The Transition and Autocorrelation Structure of TES Processes Part I: General Theory", *Stochastic Models* **8(2)**, 193–219, 1992.

[11] Jagerman, D.L. and Melamed, B., "The Transition and Autocorrelation Structure of TES Processes Part II: Special Cases", *Stochastic Models* **8(3)**, 499–527, 1992.

[12] Jagerman, D.L. and Melamed, B., "The Spectral Structure of TES Processes", *Stochastic Models* **10(3)**, 599–618, 1994.

[13] Jagerman, D.L. and Melamed, B., "The Run Probabilities of TES Processes", *Stochastic Models*, Vol. 10, No. 4, 831–851, 1994.

[14] Jagerman, D.L. and Melamed, B., "Bidirectional Estimation and Confidence Regions for TES Processes", *Proceedings of MASCOTS '95*, 94–98, Durham, January 1995.

[15] Jelenkovic, P. and Melamed, B., "Algorithmic Modeling of TES Processes", IEEE Trans. on Automatic Control **40(7)**, 1305–1312 , 1995.

[16] Johnson, G.E., "Construction of Particular Random Processes", *Proc. of the IEEE* **82(2)**, 270–285, 1994.

[17] Karademir, S., Lambadaris, I., Devetsikiotis, M. and Kaye, A. R., "Dynamic Rate Control of VBR MPEG Video Transmission over ATM Networks," *IEEE Global telecommunications Conference, GLOBECOM '96*, 1509-1515, London, 1996.

[18] Klaoudatos, G., Devetsikiotis, M. and Lambadaris, I., "Automated Modeling of Broadband Network Data Using the QTES Methodology", *IEEE ICC '99*, Vancouver, June 1999 (to appear).

[19] Law, A. and Kelton, D.W., *Simulation Modeling and Analysis*, McGraw-Hill, 1991.

[20] Lee, D.S., Melamed, B., Reibman, A. and Sengupta, B., "TES Modeling for Analysis of a Video Multiplexer", *Performance Evaluation* **16**, 21–34, 1992.

[21] Lewis, P.A.W. and McKenzie, E., "Minification Processes and Their Transformations", *J. Appl. Prob.* **28**, 45–57, 1991.

[22] Lighthill, M.J., *Introduction to Fourier Analysis and Generalised Functions*, Cambridge University Press, 1960.

[23] Liu, B. and Munson, D.C., "Generation of a Random Sequence Having a Jointly Specified Marginal Distribution and Autocovariance", *IEEE Trans. on Acoustics, Speech and Signal Processing* **30(6)**, 973–983, 1982.

[24] Melamed, B., "TES: A Class of Methods for Generating Autocorrelated Uniform Variates", *ORSA J. on Computing* **3**, 317–329, 1991.

[25] Melamed, B., "An Overview of TES Processes and Modeling Methodology", in *Performance Evaluation of Computer and Communications Systems* (L. Donatiello and R. Nelson, Editors), 359–393, *Lecture Notes in Computer Science*, Springer-Verlag, 1993.

[26] Melamed, B., "The Empirical TES Methodology: Modeling Empirical Time Series", *J. of Applied Mathematics and Stochastic Analysis* **10(4)**, 333–353, 1997.

[27] Melamed, B. and Hill, J.R., "A Survey of TES Modeling Applications", *SIMULATION* **64(6)**, 353–370, 1995.

[28] Melamed, B., Raychaudhuri, D., Sengupta B. and Zdepski, J., "TES-Based Video Source Modeling For Performance Evaluation of Integrated Networks", *IEEE Transactions on Communications*, **42(10)**, 2773–2777, 1994.

[29] Melamed, B., Ren, Q. and Sengupta, B. "The QTES/PH/1 Queue", *Performance Evaluation* **26**, 1–20, 1996.

[30] Melamed, B. and Pendarakis, D., "Modeling Full-Length VBR Video Using Markov-Renewal-Modulated TES Models", *IEEE JSAC*, **16(5)**, 600–611, 1998.

[31] Reininger, D., Melamed, B. and Raychaudhuri, D., "Variable Bit Rate MPEG Video: Characterization, Modeling and Multiplexing", *Proceedings of the 14-th International Teletraffic Congress*, Antibes Juan-les-Pins, France, Vol. 1a, 295–306, June 1994.

[32] Schmeiser, B.W., "Simulation Experiments", in *Handbook of Operations Research and Management Science*, Vol. 2, (D. Heyman and M. Sobel, editors), North Holland, 1990.

[33] Sharon, C.M., Lambadaris, I., Devetsikiotis M. and Kaye A.R., "Modeling and Control of VBR H.261 Video Transmission over Frame Relay Networks", *IEEE Transactions on Circuits and Systems for Video Technology*, **7(3)**, 527–538, 1997.

[34] Tolstov, G.P., *Fourier Series*, Prentice-Hall, 1962.