THE EMPIRICAL TES METHODOLOGY: MODELING EMPIRICAL TIME SERIES

BENJAMIN MELAMED

Rutgers University
Faculty of Management, Department of MSIS
New Brunswick, NJ 08903

Abstract

TES (Transform-Expand-Sample) is a versatile class of random sequences defined via an autoregressive scheme with modulo-1 reduction and additional transformations. The scope of TES encompasses a wide variety of sample path behaviors, which in turn give rise to autocorrelation functions with diverse functional forms — monotone, oscillatory, alternating and others. TES sequences are readily generated on a computer, and their autocorrelation functions can be numerically computed from accurate analytical formulas at a modest computational cost.

This paper presents the empirical TES modeling methodology which uses TES process theory to model empirical records. The novel feature of the TES methodology is that it expressly aims to simultaneously capture the empirical marginal distribution (histogram) and autocorrelation function. We draw attention to the non-parametric nature of TES modeling in that it always guarantees an exact match to the empirical marginal distribution. However, fitting the corresponding autocorrelation function calls for a heuristic search for a TES model over a large parametric space. Consequently, practical TES modeling of empirical records must currently rely on software assistance. A visual interactive software environment, called TEStool, has been designed and implemented to support TES modeling. The paper describes the empirical TES modeling methodology as implemented in TEStool and provides numerically-computable formulas for TES autocorrelations. Two examples illustrate the efficacy of the TES modeling approach. These examples serve to highlight the ability of TES models to capture first-order and second-order properties of empirical sample paths and to mimic their qualitative appearance.

Keywords and Phrases: Autocorrelation Functions, Histograms; Modeling Time Series; Model Fitting; TES Processes; TES Sequences.

1 INTRODUCTION

Fitting mathematical models to empirical time series often presents modelers with a standard problem: how to capture the greatest range of statistical aspects of the empirical data in a model with the smallest number of parameters. In other words, one attempts to simultaneously achieve both a high degree of model accuracy and a low level of model complexity. These opposing requirements of fidelity versus parsimony present a classical modeling trade-off and must often be settled with an unsatisfactory compromise between the desirable and the practical.

The general problem setup assumes that an empirical sample path (record) is given as a partial history of a stationary time series. This paper proposes a modeling approach that aims for a more satisfactory modeling compromise. We seek to identify candidate models which simultaneously capture first-order and second-order statistics of the empirical data, and in addition, give rise to sample paths that mimic the "appearance" of the empirical record. Thus, we require candidate models to satisfy both quantitative requirements (local distributional aspects and global temporal dependence aspects), as well as qualitative aspects (sample path "resemblance"). The stringency of the requirements ranges from the mathematically precise to the heuristic; a formal precise statement of these requirements is deferred until Section 2.3. Intuitively, model compliance with these requirements can be expected to result in more accurate models, as more statistical aspects are targeted for capture. On the other hand, satisfying such multiple requirements is a particularly difficult problem, as one tries, in effect, to reconcile two conflicting objectives: generality and accuracy, the combination of which we term *versatility.* To illustrate this point by an example, consider the popular class of $AR(n)$ models (autoregressive models of order n); see, e.g., Box and Jenkins (1976). If one chooses, say, an $AR(1)$ scheme to model a prescribed geometric autocorrelation function, it often becomes necessary to assume that the marginal distribution is normal, in order to preserve stationarity. The converse case, in which the marginal distribution is prescribed, often leads to restricted feasible forms of the attendant autocorrelation functions; see Jagerman and Melamed (1992a) for more details. Such methods, therefore, have a limited degree of versatility.

The subject of this paper is a versatile modeling methodology, called TES *(Transform-Expand-Sample)*, which is naturally suited for modeling empirical time series, in the spirit of the modeling requirements above. The computer implementation (Monte Carlo simulation) of TES is computationally efficient and its autocorrelation function can be computed by fast and accurate numerical methods from analytical formulas. The modeling activity consists of a heuristic search for a TES model that gives rise to a suitable autocorrelation function. As will be explained later, TES enjoys a great deal of freedom in this activity, since an exact match with the empirical histogram is automatically guaranteed, regardless of the autocorrelation function. Since the heuristic search for a TES model is conducted over a large parametric space, practical TES modeling of empirical records must currently rely on software assistance. A visual interactive software environment, called TEStool, has been designed and implemented to support TES modeling; see Hill and Melamed (1995), Melamed et al. (1992). TEStool uses extensive visualization and real-time feedback to cast the modeling action into an intuitive activity akin to an arcade game, rendering TES modeling an efficient activity, accessible to experts and non-experts alike. The purpose of this paper is to describe the empirical TES modeling methodology in detail and to provide the computational details used in TEStool to implement the TES methodology.

The rest of the paper is organized as follows. Section 2 defines the problem formally. Section 3 provides a brief review of TES theory germane to the problem at hand. Section 4 outlines the empirical TES methodology for modeling empirical records and Section 5 illustrates it through two examples based on sample data from two application domains: finance and reliability. Section 6 concludes the paper with a summary. Finally, the appendices present explicit formulas for fast and accurate computation of TES model autocorrelations in the context of the empirical TES modeling methodology, as implemented in TEStool.

2 PROBLEM FORMULATION

Let $\{Y_n\}_{n=0}^{N-1}$ be an empirical data sequence (record) of size N, sampled from an unknown real-valued, discrete time stationary time series. Henceforth, we use the standard notational convention of appending a hat symbol to estimators; additionally, objects associated with an empirical sample will be subscripted, to reflect this association. Throughout the paper, 1_A denotes the indicator function of set A.

2.1 Empirical Histograms

The marginal distribution of $\{Y_n\}$ is modeled as an empirical histogram with either continuous or discrete components. This section provides a taxonomy of empirical histograms.

A *continuous histogram* is specified by a finite set of triplets of the form $\hat{\mathcal{H}}_Y^c = \{(l_j, r_j, \hat{p}_j) : j \in \mathcal{J}\}$, where $\mathcal{J}$ is the index set of histogram cells, $[l_j, r_j)$ is the interval of cell j with width $w_j = r_j - l_j > 0$, and $\hat{p}_j$ is the probability estimator of cell j. For a pure continuous histogram, $\mathcal{J} = \{1, \ldots, J\}$, where J is the number of histogram cells, but this is too restrictive for the general case, to be described below. We assume, for simplicity, that the cell intervals are sorted in increasing order and disjoint (i.e., $[l_i, r_i) \cap [l_j, r_j) = \emptyset$, $i \neq j$). Recall that the parameters J and the intervals $[l_j, r_j)$, $1 \leq j \leq J$ are prespecified and used to calculate the $\hat{p}_j$ from $\{Y_n\}$ as relative frequencies. Consequently, rather than assuming, to avoid trivialities, that $\hat{p}_j > 0$ for all $1 \leq j \leq J$, we define the index set $\mathcal{J}^+ = \{j \in \mathcal{J} : \hat{p}_j > 0\}$. The continuous histogram $\hat{\mathcal{H}}_Y^c$ induces the empirical density,

$$\hat{h}_Y^c(y) \; = \; \sum_{j \in \mathcal{J}^+} 1_{[l_j, r_j)}(y) \, \frac{\hat{p}_j}{w_j}, \quad -\infty < y < \infty, \tag{2.1}$$

called *continuous histogram density*. Thus, the empirical density $\hat{h}_Y(y)$ is a probabilistic mixture of J uniform densities over the intervals $[l_j, r_j)$ with mixing probabilities $\hat{p}_j$. These will be referred to as *continuous components*, where component j is characterized by triplet $(l_j, r_j, \hat{p}_j)$.

Observe that any reasonable density function (say, with a finite number of probability atoms and a Riemann-integrable diffuse component) can be approximated arbitrarily closely by mixtures of uniform distributions, and this fact accounts for the widespread practice of modeling marginals of empirical data by continuous histograms. If, however, it is known that $\{Y_n\}$ is discrete-valued, the Y_n are better modeled as a probabilistic mixture of discrete values, and the corresponding *discrete histogram* has the form $\hat{\mathcal{H}}_Y^d = \{(v_i, \hat{p}_i) : \; i \in \mathcal{I}\}$, where $\mathcal{I}$ is the index set of distinct discrete values, and value v_i occurs with probability $\hat{p}_i$. For a pure discrete histogram, $\mathcal{I} = \{1, \ldots, I\}$, where I is the number of histogram atoms, but this is again too restrictive for the general case, to be described below. We assume, for simplicity, that the discrete values are disjoint and sorted in increasing order. Analogously to the reasoning in the continuous case, we define the index set $\mathcal{I}^+ = \{i \in \mathcal{I} : \hat{p}_i > 0\}$, rather than assuming that $\hat{p}_i > 0$, for all $1 \leq i \leq I$. The discrete histogram $\hat{\mathcal{H}}_Y^d$ induces the empirical density,

$$\hat{h}_Y^d(y) \;=\; \sum_{i \in \mathcal{I}^+} 1_{\{v_i\}}(y)\, \hat{p}_i, \quad -\infty < y < \infty, \tag{2.2}$$

called *discrete histogram density.* The probability atoms, characterized by the pairs $(v_i, \hat{p}_i)$, will be referred to as *discrete components.*

A *general histogram* $\hat{\mathcal{H}}_Y = \hat{\mathcal{H}}_Y^c \bigcup \hat{\mathcal{H}}_Y^d$ is any probabilistic mixture of $J \geq 0$ continuous components and $I \geq 0$ discrete components, provided $J + I = M > 0$. Thus, general histograms include continuous and discrete ones as special cases. To simplify matters will assume that all M components have been sorted in increasing order (indexed by $1 \leq m \leq M$), and that all M components are disjoint in the sense that for all $i \in \mathcal{I}$ and $j \in \mathcal{J}$, we have $\{v_i\} \bigcap [l_j, r_j) \subset \{l_j\}$. Although continuous and discrete components may alternate on the real line, the mixing probability of component m (in sorted order) will be denoted $\hat{p}_m$. As before, we let $\mathcal{J}$ and $\mathcal{I}$ be the index sets of continuous and discrete components, respectively, except that the indices are drawn from the set $\{1, \ldots, M\}$. Consequently, we can leave the definitions of $\mathcal{J}^+$ and $\mathcal{I}^+$ unaltered for the general histogram case. Note that all these set definitions are proper generalizations, consistent with the previous ones, the latter merely being special cases. The induced empirical density has the form

$$\hat{h}_Y(y) \;=\; \hat{h}_Y^c(y) + \hat{h}_Y^d(y), \quad -\infty < y < \infty, \tag{2.3}$$

and is simply called a *histogram density.*

2.2 Empirical Autocorrelations

For a stationary time series $\{X_n\}$, with common mean $\mu_X < \infty$ and common variance $\sigma_X^2 < \infty$, the autocorrelation function

$$\rho_X(\tau) \;=\; \frac{E[X_n\, X_{n+\tau}] \;-\; \mu_X^2}{\sigma_X^2}, \quad \tau \geq 1, \tag{2.4}$$

is a measure of the linear dependence of τ-lagged variates. It is usually estimated by the empirical autocorrelation function (see Cox and Lewis

(1968), Chapter 5),

$$\hat{\rho}_Y(\tau) = \frac{\dfrac{1}{N-\tau}\displaystyle\sum_{n=0}^{N-1-\tau} Y_n\, Y_{n+\tau} \; - \; \hat{\mu}_Y(0, N{-}1{-}\tau)\, \hat{\mu}_Y(\tau, N{-}1)}{\hat{\sigma}_Y(0, N{-}1{-}\tau)\, \hat{\sigma}_Y(\tau, N{-}1)}, \quad 1 \le \tau \ll N, \tag{2.5}$$

where $\hat{\mu}_Y(r,s) = \frac{1}{s-r}\sum_{n=r}^{s} Y_n$ and $\hat{\sigma}_Y(r,s)^2 = \frac{1}{s-r}\sum_{n=r}^{s} Y_n^2 - \hat{\mu}_Y^2(r,s)$ are the sample mean and sample variance, respectively, each based on the subsample $\{Y_r, Y_{r+1}, \ldots, Y_s\}$ of $\{Y_n\}$, from index r to index s. Notice that the lag τ in (2.5) is much smaller than the sample size, in order that $\hat{\rho}_Y(\tau)$ be statistically meaningful; for example, $\tau \le \sqrt{N}$ is suggested in Cox and Lewis (1968), Chapter 5.

2.3 Model Fitting Requirements

Recall that, informally, the model we seek is a stationary real-valued random sequence $\{X_n\}_{n=0}^{\infty}$ (with common mean $\mu_X < \infty$ and common variance $\sigma_X^2 < \infty$), such that its marginal distribution matches its empirical counterpart, its autocorrelation function approximates its empirical counterpart, and its sample paths "resemble" the empirical record. More formally, we require $\{X_n\}$ to obey the three requirements listed below in order of decreasing rigor:

Requirement 1: The marginal density $h_X(y)$ of $\{X_n\}$ should equal the empirical histogram density $\hat{h}_Y(y)$ of Eq. (2.3).

Requirement 2: The autocorrelation function $\rho_X(\tau)$ should approximate its empirical counterpart $\hat{\rho}_Y(\tau)$ of Eq. (2.5). The extent of the approximation is left to the analyst.

Requirement 3: The sample paths of $\{X_n$ should bear adequate "resemblance" to the empirical sequence $\{Y_n\}$. Since the notion of "resemblance" is left unquantified, we view this requirement as a highly subjective heuristic, rather than a precise requirement. Nevertheless, some sort of "path resemblance criterion" is often employed in practice to increase one's confidence in a proposed model. We stress that this qualitative property is posited in addition to, not instead of, the previous two more rigorously quantitative ones.

In practice, we may or may not be able to calculate $h_X(y)$ and $\rho_X(\tau)$ neither in closed form nor approximate them numerically. We assume, however, that $\{X_n\}$ sample paths can always be computed in a Monte Carlo simulation. In such cases, $h_X(y)$ and $\rho_X(\tau)$ will have to be estimated statistically from the simulated sample paths. This problem is further exacerbated when high positive autocorrelations are present in $\{X_n\}$, since reliable estimates will then require large sample sizes.

3 TES PROCESSES

This section contains a brief review of general TES processes and their second-order properties. The details may be found in Jagerman and Melamed (1992a) and (1992b); an overview at an introductory level appears in Melamed (1993).

Let $\{V_n\}_{n=1}^{\infty}$ be a sequence of iid (independent identically distributed) random variables, called the *innovation sequence*. Let U_0 be uniform on $[0, 1)$ (denoted $U_0 \sim$ Uniform(0,1)), and independent of the innovations $\{V_n\}$. We define two classes of TES sequences, called TES$^+$ and TES$^-$, each consisting of cid (correlated, identically distributed) Uniform(0,1) sequences, denoted $\{U_n^+\}_{n=0}^{\infty}$ and $\{U_n^-\}_{n=0}^{\infty}$, respectively. These sequences are defined on a common probability space by

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

and

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

The angular brackets in (3.1) denote the modulo-1 (fractional part) operator, defined by $\langle x \rangle = x - \max\{\text{integer } n : \ n \leq x\}$. The superscript notation in (3.1)-(3.2) is motivated by the fact that $\{U_n^+\}$ and $\{U_n^-\}$ can generate lag-1 autocorrelations in the range [0,1], and [-1,0] respectively. In fact, Eq. (3.2) implies $\rho_U^-(\tau) = (-1)^\tau \rho_U^+(\tau)$, where $\rho_U^+(\tau)$ and $\rho_U^-(\tau)$ are the autocorrelation functions corresponding to $\{U_n^+\}$ and $\{U_n^-\}$, respectively. From now on, we shall always append plus or minus superscripts to other mathematical objects associated with $\{U_n^+\}$ and $\{U_n^-\}$ in a natural way. However, the superscript is omitted when the distinction is immaterial.

Intuitively, the modulo-1 arithmetic used in defining TES processes in Eqs. (3.1) - (3.2) has a simple geometric interpretation as a random walk on a circle of circumference 1 (unit circle), with mean step size $E[V_n]$. When $E[V_n] = 0$, the random walk has zero drift around the circle and $\rho_U^+(\tau)$ is monotone decreasing in the lag τ. If $E[V_n] > 0$, the drift is positive, resulting in cyclical sample paths with the attendant $\rho_U^+(\tau)$ oscillating about zero in the lag τ. The case $E[V_n] < 0$ is analogous but with opposite drift.

A Lebesgue measurable transformation D from $[0, 1)$ to the reals is called a *distortion*. It is used to transform a sequence $\{U_n\}$ with uniform marginals on [0,1) to a sequence $\{X_n\}$, such that $X_n = D(U_n)$, and the X_n have a common marginal distribution F. When $D = F^{-1}$, then D is called the *inversion method* (Devroye (1986), Bratley et al. (1987), Law and Kelton (1991)). In particular, the stochastic sequences $\{X_n^+\}_{n=0}^{\infty}$ and $\{X_n^-\}_{n=0}^{\infty}$ are obtained as

$$X_n^+ \; = \; D(U_n^+), \quad X_n^- \; = \; D(U_n^-). \tag{3.3}$$

Lemma 1 in Jagerman and Melamed (1992a) provides the theoretical basis for TES. It states that if $U \sim \text{Uniform}(0, 1)$ and V is arbitrarily distributed but independent of U, then

$$\langle U + V \rangle \sim \text{Uniform}(0, 1). \tag{3.4}$$

Consequently, Eq. (3.4) ensures that both $\{U_n^+\}$ and $\{U_n^-\}$ have $\text{Uniform}(0, 1)$ marginals, so $\{X_n^+\}$ and $\{X_n^-\}$ have general marginals. We point out that this fact serves to satisfy Requirement 1 in Section 2.3, for matching the marginals of the empirical data. We only need to construct the distribution function corresponding to the empirical density $\hat{h}_Y(y)$ of Eq. (2.3), and then invert it to obtain the appropriate distortion. The details are deferred, however, until the Section 4.

To satisfy Requirement 2 in Section 2.3, we need a way to calculate the autocorrelation function (2.4) of our model. Such calculations would be used in a search procedure (possibly heuristic) for TES models whose autocorrelations adequately approximate their empirical counterparts. Naturally, a Monte Carlo simulation of Eqs. (3.1)-(3.2) can always provide an estimate of (2.4), if a sufficient sample size is generated. This approach, however, can be costly in terms of computation time, especially when high correlations necessitate large sample sizes for adequate statistical reliability. It is then essential to develop fast and efficient computational techniques for calculating (2.4), over a range of τ. Fortunately, Theorems 3 and 4 in Jagerman and Melamed (1992b) provide numerically-computable formulas for the autocorrelation functions of $\{X_n^+\}$ and $\{X_n^-\}$, respectively,

$$\rho_X^+(\tau) \;=\; \frac{2}{\sigma_X^2} \sum_{\nu=1}^{\infty} Re[\tilde{f}_V^\tau(i2\pi\nu)] \, |\tilde{D}(i2\pi\nu)|^2 \tag{3.5}$$

and

$$\rho_X^-(\tau) \;=\; \begin{cases} \rho_X^+(\tau), & \tau \text{ even} \\ \\ \dfrac{2}{\sigma_X^2} \displaystyle\sum_{\nu=1}^{\infty} Re[\tilde{f}_V^\tau(i2\pi\nu)] \, Re[\tilde{D}^2(i2\pi\nu)], & \tau \text{ odd} \end{cases} \tag{3.6}$$

where $f_V^\tau(x)$ denotes the τ-fold convolution of the innovation density f_V, tilde denotes the Laplace transform, and $i = \sqrt{-1}$.

The sample paths of $\{X_n^+\}$ and $\{X_n^-\}$ often exhibit a visual "discontinuity" whenever the the corresponding TES processes $\{U_n^+\}$ and $\{U_n^-\}$ "cross" point 0 on the unit circle in either clockwise or counter-clockwise direction. Sample path "smoothing" can be accomplished by applying a member from a family of so-called *stitching transformations* S_ξ, $0 \le \xi \le 1$, defined by

$$S_\xi(y) \;=\; \begin{cases} y/\xi, & 0 \le y \le \xi \\ (1-y)/(1-\xi), & \xi \le y < 1 \end{cases} \tag{3.7}$$

where ξ is called a *stitching parameter*. Note that for $0 < \xi < 1$, the S_ξ are continuous on the unit circle. In particular, $S_1(x) = x$ is the identity transformation and $S_0(x) = 1 - x$ is the antithetic transformation. It is quite straightforward to show that S_ξ preserves uniformity for for all $0 \le \xi \le 1$, i.e., if $U \sim$ Uniform$(0,1)$ then $S_\xi(U) \sim$ Uniform$(0,1)$ (see Melamed (1991), Lemma 2). The composite distortion D_ξ, defined by

$$D_\xi(x) \;=\; D(S_\xi(x)), \tag{3.8}$$

is more general than D since $\xi = 1$ is a special case. Naturally, Eqs. (3.5)-(3.6) remain valid when $\tilde{D}_\xi$ is substituted for $\tilde{D}$ throughout. Furthermore, $\tilde{D}_\xi$ is related to $\tilde{D}$ by

$$\tilde{D}_\xi(s) \;=\; \xi \, \tilde{D}(\xi s) \;+\; (1-\xi) \, e^{-s} \tilde{D}(-(1-\xi)s), \quad 0 \le \xi \le 1, \tag{3.9}$$

a fact which is easily verifiable by direct calculation of

$$\tilde{D}_\xi(s) = \int_0^1 e^{-sx} \, D(S_\xi(x)) \, dx,$$

with the aid of Eqs. (3.7) and (3.8).

4 THE EMPIRICAL TES MODELING METHODOLOGY

We are now ready to present a TES-based methodology for modeling empirical data sequences which we call the *empirical TES methodology.* In this section, we merely outline the methodology as a heuristic procedure, omitting most computational details; the latter are provided in the appendices.

A general TES-based modeling methodology would allow the class of innovation densities f_V under consideration to remain unrestricted, since a major advantage of TES modeling is that the choice of an innovation density provides a broad degree of freedom in approximating the empirical autocorrelation function. However, practical implementation considerations will lead us to constrain them to lie in operationally useful classes, without trading off too much generality in return. The main criteria for selecting a suitable class of effective innovations are listed below.

(i) The class should be large enough to approximate any marginal distribution.

(ii) Monte Carlo generation of variates from the class should be computationally efficient.

(iii) The numerical calculation of the associated transform $\tilde{f}_V$ should be fast.

As part of the empirical TES methodology, we select as candidates for innovation variates the class of probabilistic mixtures of uniform variates; the corresponding densities constitute the class of *step-function* densities on the unit circle, taken here as the interval $[-0.5, 0.5)$. We point out that while the concept of a histogram density and a step-function density is probabilistically similar, it is important to maintain a notational and conceptual distinction between histogram entities (associated with distortions) and step-function entities (associated with innovations). To this end, we denote a step-function density by

$$f_V(x) \;=\; \sum_{k=1}^{K} 1_{[L_j,R_j)}(x)\,\frac{P_k}{\alpha_k}, \quad x \in [-0.5, 0.5), \tag{4.1}$$

where $K > 0$ is an integer, L_k and R_k are the left and right endpoints of step k, $\alpha_k = R_k - L_k$ is the width of step k, and P_k is the mixing probability of step k. Notice that the above criteria are satisfied by this selection. Indeed, the class of step-function densities is particularly simple, yet it can approximate arbitrarily closely any reasonable density function. In addition, it enjoys the important advantage of being particularly easy to manipulate graphically in interactive modeling (to be explained later).

In contrast, the class of distortions under consideration has been effectively determined by our decision to model the empirical density as a histogram density $\hat{h}_Y$ of the form (2.3). The generality of this modeling decision is evident, and the fact that histogram densities are step functions as well will simplify the computational details to be presented in the appendices. Denoting the cumulative distribution function attendant to $\hat{h}_Y$ by $\hat{H}_Y$, and the latter's inverse by $\hat{H}_Y^{-1}$, we shall henceforth be solely interested in distortions of the form

$$D_{Y,\xi}(x) = \hat{H}_Y^{-1}(S_\xi(x)), \quad x \in [0,1), \tag{4.2}$$

henceforth referred to as *histogram distortions*. Recall that the inner transformation, S_ξ, is a stitching transformation (3.7), whereas the outer one, $\hat{H}_Y^{-1}$ (called a *histogram inversion*) is the inverse of a histogram distribution function H of the form (2.1), (2.2) or (2.3). In order to distinguish between the continuous and discrete cases (corresponding to a continuous or discrete underlying empirical histogram), we use the notation $D^c_{Y,\xi}$ and $D^d_{Y,\xi}$, respectively, and similarly for other related objects. In all cases, the key fact is that for any background TES sequence $\{U_n\}$, the distorted sequence $\{X_n\}$, with elements

$$X_n = D_{Y,\xi}(U_n) = \hat{H}_Y^{-1}(S_\xi(U_n)), \tag{4.3}$$

will still have marginal distribution $\hat{H}_Y$. To see that, merely recall that every $\{S_\xi(U_n)\}$ is marginally uniform on $[0,1)$, and invoke the inversion method under a histogram inversion to yield a histogram distribution.

For now, we assume that an empirical sample path $\{Y_n\}$ is given, and that an empirical histogram $\hat{\mathcal{H}}_Y$ has been constructed from it. In the procedural outline to follow, we can keep the class of innovations arbitrary, without constraining their densities to step functions.

Outline of the Empirical TES Modeling Methodology

1. Construct the empirical (cumulative) distribution function $\hat{H}_Y$, corresponding to $\hat{\mathcal{H}}_Y$.

2. Construct the associated inverse distribution function $\hat{H}_Y^{-1}$ (this is always possible, since $\hat{H}_Y$ is monotone non-decreasing).

3. Select the sign of the TES class (TES$^+$ or TES$^-$).

4. Select a stitching parameter $0 \leq \xi \leq 1$. This determines a histogram inversion $D_{Y,\xi}(x) = \hat{H}_Y^{-1}(S_\xi(x))$, where ξ is a heuristic search parameter.

5. Select an innovation density f_V, where the innovation density constitutes a set of heuristic search parameters. This determines a uniform TES process $\{U_n\}$ which can be either $\{U_n^+\}$ or $\{U_n^-\}$.

6. At this point, a TES process $\{X_n\}$ has been determined by Eq. (4.3) and its autocorrelation function can be computed from Eq. (3.5) or (3.6) with the aid of the appendices. In addition, generate a simulated sample path of $\{X_n\}$ (the initial variate of the simulated path is typically set to the corresponding empirical observation). If $\rho_X(\tau)$ approximates well its empirical counterpart $\hat{\rho}_Y(\tau)$, and the simulated sample path "resembles" qualitatively its empirical counterpart, then the search is terminated; otherwise, the search is iterated from any previous step.

The procedure outlined above is highly heuristic and should only be viewed as a general guideline. The main heuristic component is a search for a stitching parameter ξ and an innovation density function f_V that together give rise to a TES process whose autocorrelation function approximates its empirical counterpart well, and whose Monte Carlo simulation runs produce sample paths with "acceptable resemblance" to their empirical counterpart.

Our procedural guidelines do not specify how to structure this search. Naturally, a "blind" search in such an enormous parameter space is bound to be inefficient, if not fruitless, barring a lucky guess. It is therefore essential to impose some structure on the search, if we hope to apply TES as a practical modeling methodology. A simple way to structure the search for step-function innovation densities is to set K (the number of steps) to 1, and search among candidate single-step densities. If this does not yield a satisfactory TES model, K may be incremented successively and the search continued. Naturally, by the principle of parsimony, we aim to find the smallest K for which a good model can be found.

Searching the vast parameter space of step function densities over $[-0.5, 0.5)$ is a major problem. To implement a rule-based approach, one needs to have at least qualitative understanding of how the autocorrelation function behaves as a function of the defining triplets $\{(L_k, R_k, P_k)\}_{k=1}^{K}$ and the stitching parameter ξ. For a given step (L_k, R_k, P_k), this behavior can be deduced from the case $K = 1$, for which we have a considerable qualitative understanding (see Jagerman and Melamed (1992a), (1992b) and (1994)). It is worth summarizing this knowledge, but a more useful understanding would be gleaned if we adopt the equivalent parameterization $(\alpha_k, \phi_k, p_k),\ 1 \leq k \leq K$, where

$$\alpha_k \ = \ R_k - L_k, \quad \phi_k \ = \ \frac{R_k + L_k}{R_k - L_k}. \tag{4.4}$$

The interpretation of α_k and ϕ_k is highly germane to an understanding of TES process behavior. Clearly, α_k is just the length of step k. The interpretation of ϕ_k is more complicated; it can be viewed as the angle of rotation of the innovation step relative to symmetric straddle. Symmetric straddle of the current iterate U_n corresponds to $L_k = -R_k$, i.e., the point U_{n+1} is equally likely to lie to the right or to the left of U_n on the unit circle, given that step k of the innovation density was sampled. The qualitative effect of the α, ϕ and ξ parameters on the autocorrelation function is fairly well-understood for single-step innovation densities (in this case $K = 1$ and $p_1 = 1$, so the innovation variates reduce to ordinary uniform variates over a portion of the unit circle). These qualitative effects are summarized below.

Effect of α: α controls the magnitude of the autocorrelation function. The autocorrelation function envelope decay in the lag increases with α.

Effect of ϕ: ϕ controls the frequency of oscillation of the autocorrelation function. The larger the value ϕ, the higher is the frequency of oscillation. When $\phi = 0$, the autocorrelation function is monotone decreasing and a spectral analysis reveals no periodicities. When $\phi \neq 0$, the autocorrelation function is oscillatory, and the sample paths have a cyclical appearance. the presence of periodicities can be confirmed by spectral analysis.

Effect of ξ: The effect of $0 < \xi < 1$ is to “smooth” sample paths. The magnitude of the autocorrelation function increases as ξ approaches 0.5 from the left or from the right. Symmetry about 0.5 manifests itself in another way. While $\{S_\xi(U_n)\}$ and $\{S_{1-\xi}(U_n)\}$ have different sample paths, their autocorrelation functions are identical, for any background TES sequence $\{U_n\}$. In cyclical TES processes (those with$E[V_n] \neq 0$), ξ can be used to skew sample paths in accordance with the corresponding stitching transformation. Here, $\{S_\xi(U_n)\}$ is also cyclical, but its cycle peaks are shifted by a phase proportional to ξ. In particular, $S_0(y) = 1-y$ and $S_1(y) = y$ give rise to "descending" and "ascending" sawtooth cycles, respectively, while $S_{0.5}$ gives rise to stitched sequences with symmetrical cycles.

We remark that TES^+ models are the most common choice, in our experience; TES^- sequences should be considered, however, when empirical records or autocorrelation functions have an alternating (zigzag) appearance.

The heuristic nature of the empirical TES modeling methodology described above requires computer assistance for effective implementation. To this end, a software environment, called TEStool, has been built to support visual interactive TES modeling of empirical records; see Hill and Melamed (1995) for a detailed description, or Melamed et al. (1992) for a brief overview. A salient feature of TEStool is that it casts the heuristic search process in the mold of an arcade game through extensive visualization and real-time interaction. A workstation mouse is used to visually create, delete, resize and relocate innovation density steps. The geometric simplicity of step-functions is exploited graphically, as steps are simply represented as rectangles on the display screen. Thus, modifying a step-function density is simple and intuitive: changing a step size is accomplished by stretching" a side or a corner of the corresponding rectangle, while translating a step on the horizontal axis is just the familiar operation of "dragging" an icon. In a similar vein, the stitching parameter is selected by positioning a slider in the $[0, 1]$ interval, and a TES sign is chosen by pressing a button. In the interactive mode, any change in model specification (e.g., innovation density, stitching parameter etc.) triggers a numerical recomputation of all TES statistics (sample paths, histogram, autocorrelation function and spectral density) with the corresponding graphs redisplayed, superimposed on their empirical counterparts for comparison. This can be executed in real-time, since the autocorrelation function can be computed rapidly and accurately from Eqs. (3.5) and (3.6) with the aid of the appendices. Thus, the heuristic search process can proceed rapidly and efficiently, guided by visual feedback — the goal being to bring the TES autocorrelation graph as close as possible to the corresponding empirical target. At the same time the user may also judge the "qualitative similarity" of the model-generated sample path to the empirical data used in the modeling heuristic. The visual interactive facilities are not only instrumental in facilitating an effective modeling process, but they also serve to hold the modeler's interest by reducing the tedium of repetitive search iterations. An additional advantage of TEStool is that its visual style tends to speed up the learning process, thereby increasing the efficiency of subsequent searches. The graphical user interface encourages tinkering and experimentation, allowing the user to readily study the behavior of TES sample paths and autocorrelations as functions of TES parameters. It is expected that such studies will lead to the identification of qualitative rules which could be used to automate, or at least guide, the search procedure for a TES model.

Recently, an algorithmic modeling approach has been devised and implemented for TES modeling (Jelenkovic et al. (1994)). The algorithm first carries out a brute-force search over a subspace of step-function innovation densities and various stitching parameters; recall that the distortion is completely determined by the empirical record and user-supplied histogram parameters. Of those, the algorithm selects the best n combinations of pairs, (f_V, ξ), in the sense that the associated TES model autocorrelation functions have the smallest error (sum of squared deviations) relative to the empirical autocorrelation function. The algorithm then performs non-linear optimization on each model candidate to further reduce the aforementioned error. Finally, the analyst peruses the results and selects from the n optimized candidate models the one whose Monte Carlo sample paths bear the "most resemblance" to the empirical record, in addition to having a small error. Experience shows that the TES modeling algorithm produces better and faster results than its heuristic counterpart.

5 EMPIRICAL TES MODELING EXAMPLES

This section presents two illustrative examples of empirical TES modeling. The modeling results are graphically summarized in Figures 1 and 2, both of which are actual TEStool displays, copied off a workstation's screen. Figure 1 illustrates a TES$^+$ model of an empirical financial time series, while Figure 2 exhibits a TES$^-$ model of an empirical sequence of machine fault interarrival times (up times). For each model, the corresponding heuristic searches required less than an hour of visual interaction time with TEStool. For more details on these case studies, refer to Melamed and Hill (1993).

The screen displays in both figures have the same format. Each display consists of four tiled canvases (subwindows). The buttons at the top of the screen and at the bottom of each canvas control various modeling services; these include reading and writing datasets, subdividing the screen real estate, opening a TES specification window or menu, performing various computations and terminating the session. The lower-right canvas displays a graphical TES model specification, while the remaining three canvases each display a pair of statistics, such as sample paths, histograms and autocorrelation functions. In each of the statistical display canvases, the TES model statistics graphs are superimposed on their empirical counterparts for comparison; the empirical statistics are always represented by a solid-line curve, and their TES counterparts by a dashed-line curve.

Each of the two upper-left canvases displays an empirical record superimposed on a TES model sample path generated by Monte Carlo simulation; the TES model was created by applying the empirical TES modeling methodology to the corresponding empirical record. Each upper-right canvas displays the histograms calculated from the sample paths in the corresponding upper-left canvas. Similarly, each lower-left canvas displays the corresponding empirical autocorrelation function (computed from the upper-left canvas) superimposed on its numerically-computed TES model counterpart, according to the formulas provided in the appendices. Finally, each lower-right canvas consists of a joint visual specification of TES model parameters. The upper part of this canvas displays a visual specification of a step-function innovation density f_V; the control panel below it displays a joint specification of a TES sign (plus or minus), a stitching parame-

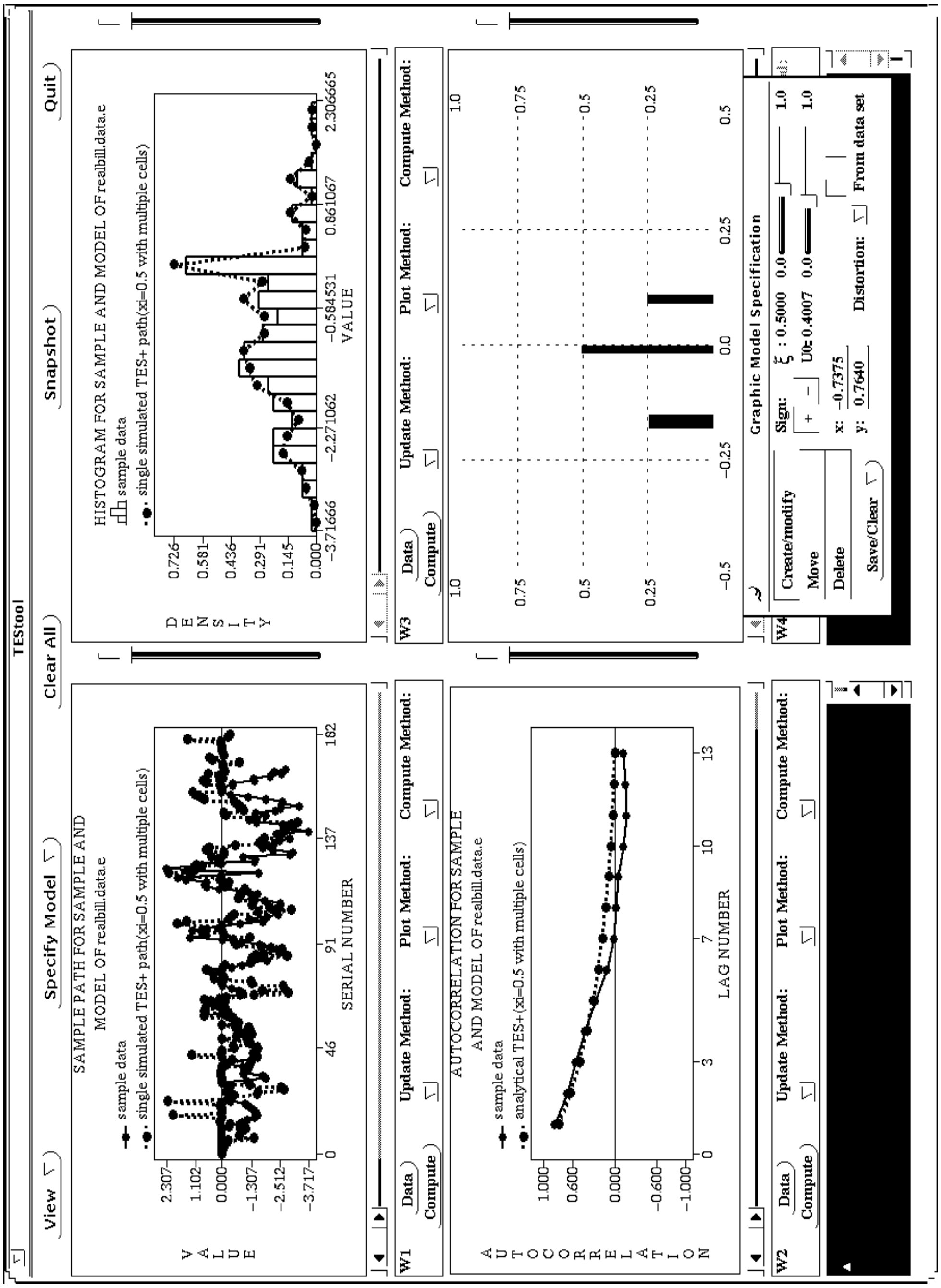

Figure 1: A TEStool screen display of a TES^{+} model for a financial time series.

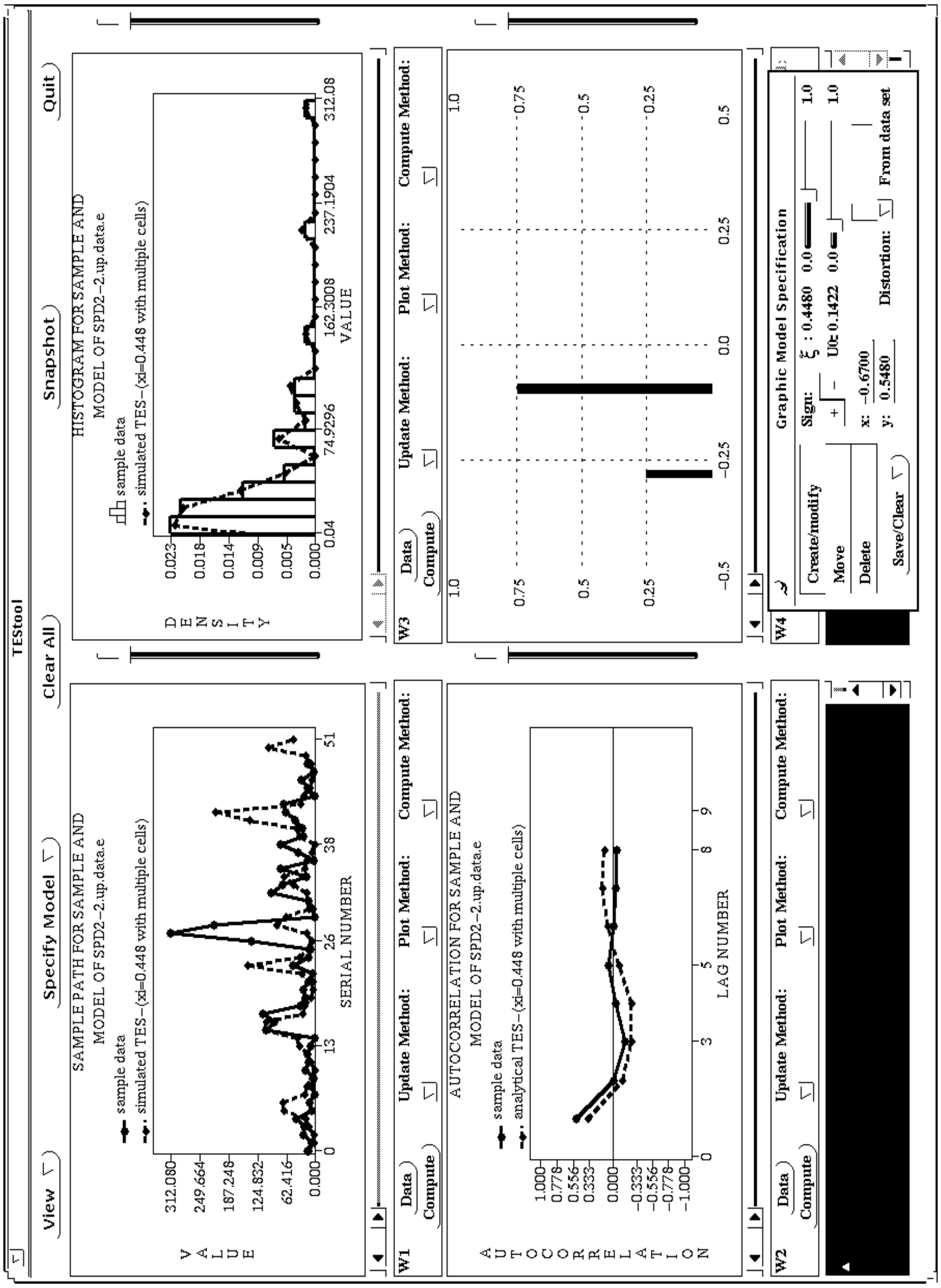

Figure 2: A TEStool screen display of a TES^- model for a machine up times record.

ter ξ, an initial TES variate U_0 and a selection of a histogram inversion computed from the empirical record (the histogram itself is displayed in the corresponding upper-right canvas). The steps of f_V are created and modified visually with the aid of the mouse by "sweeping out" steps and "dragging" them around. The TES sign is chosen via a selection button and the ξ parameter by a slider in the range $[0, 1]$. An inversion transformation to a requite marginal distribution is selected from a menu; in addition to empirical histograms, available marginal distributions include uniform and exponential as well as discrete distributions. A TES model can also be specified in TEStool in standard text mode by populating text fields in a popup subwindow, but the visual specification is more efficient, particularly when modifying the TES process specification in the context of an interactive heuristic search.

Consider now the visual fit depicted in Figures 1 and 2 vis-a-vis the three modeling requirements of Section 2. Note, that Requirements 1 and 2 are apparently satisfied, as evidenced by the excellent agreement of the curves in the lower-left and upper-right canvases. It is also interesting to note that the upper-left canvases exhibit considerable "similarity" between the empirical and simulated sample paths, in apparent compliance with Requirement 3. We conclude that both Figures 1 and 2 represent successful TES modeling efforts according to the three modeling requirements of Section 2.

6 CONCLUSION

The empirical TES modeling methodology is a novel input analysis approach, which strives to *simultaneously* fit both the marginal distribution and the leading autocorrelations of empirical records, and additionally to mimic their qualitative "appearance". This paper has presented the empirical TES modeling methodology; it has summarized the TES process theory underlying TES modeling, and briefly overviewed the TEStool visual interactive softare environment, designed to support TES modeling via a heuristic search for an appropriate TES model. Two examples have illustrated the versatility and efficacy of the empirical TES modeling methodology as implemented in TEStool.

Acknowledgments

I would like to thank Jon Hill and David Jagerman for reading and commenting on the manuscript.

References

[1] Bendat, J.S. and Piersol, A.G. (1986) *Random Data*. Wiley, 1986.

[2] Box, G.E.P. and Jenkins, G.M. (1976) *Time Series Analysis: Forecasting and Control*, Prentice Hall.

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

[4] Cox, D.R. and Lewis P.A.W. (1968) *The Statistical Analysis of Series of Events*, Methuen.

[5] Devroye, L. (1986) *Non-Uniform Random Variate Generation*, Springer-Verlag.

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

[7] Jagerman, D.L. and Melamed, B. (1992a) "The Transition and Autocorrelation structure of TES Processes Part I: General Theory", *Stochastic Models*, Vol. 8, No. 2, 193–219.

[8] Jagerman, D. L. and Melamed, B. (1992b) "The Transition and Autocorrelation structure of TES Processes Part II: Special Cases", *Stochastic Models*,Vol. 8, No. 3, 499–527.

[9] Jagerman, D.L. and Melamed, B. (1994) "The Spectral Structure of TES Processes", *Stochastic Models*, Vol. 10, No. 3, 599–618.

[10] Jelenkovic, P. and Melamed, B. (1995) "Algorithmic Modeling of TES Processes", *IEEE Trans. on Automatic Control*, Vol. 40, No. 7, 1305–1312.

[11] Law, A.M. and Kelton, W.D. (1991) *Simulation Modeling & Analysis*, (second edition), McGraw-Hill.

[12] Melamed, B. (1991) "TES: A Class of Methods for Generating Autocorrelated Uniform Variates", *ORSA J. on Computing* Vol. 3, No. 4, 317–329.

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

[14] Melamed, B., Goldsman D. and Hill J.R. (1992) "The TES Methodology: Nonparametric Modeling of Stationary Time Series" *Proceedings of WSC '92*, 135–144, Arlington, Virginia, 1992.

[15] Melamed, B. and Hill, J.R. (1993) "Applications of the TES Modeling Methodology", *Proceedings of WSC '93*, 1330–1338, Los Angeles, California, 1993.

APPENDICES: NUMERICAL COMPUTATION OF TES FORMULAS

The following appendices provide the fine computational detail for autocorrelations of TES processes. Specifically, we derive explicit numerically computable formulas for the Laplace Transforms $\tilde{f}_V$ of step-function innovation densities, as well as $\tilde{D}^c_{Y,\xi}$ and $\tilde{D}^d_{Y,\xi}$, both evaluated at $i2\pi\nu$. These are needed in Eqs. (3.5) and (3.6) within a larger computation of the autocorrelation functions of $\{X_n^+\}$ and $\{X_n^-\}$. Recall that the empirical TES modeling methodology relies on fast calculation of the corresponding autocorrelation functions, in the context of an interactive visual software support environment. The formulas in the following appendices may look messy, but they are easy to implement in software and lead to fast and accurate results by truncating the infinite sums in Eqs. (3.5) and (3.6) (TEStool uses only the first 100 terms in each formula).

APPENDIX A: COMPUTATION OF $\tilde{f}_V(i2\pi\nu)$

The notation in this appendix was defined in Section 4.

Proposition 1 *For all $\nu \geq 1$,*

$$\tilde{f}_V(i2\pi\nu) = \sum_{k=1}^{K} \frac{P_k}{\alpha_k} \left[\frac{\sin(2\pi\nu R_k) - \sin(2\pi\nu L_k)}{2\pi\nu} + i\,\frac{\cos(2\pi\nu R_k) - \cos(2\pi\nu L_k)}{2\pi\nu} \right] \tag{A.1}$$

Proof From Eq. (4.1),

$$\tilde{f}_V(s) = \sum_{k=1}^{K} \int_{L_k}^{R_k} e^{-sx} \frac{P_k}{\alpha_k}\, dx,$$

whence,

$$\tilde{f}_V(s) = \sum_{k=1}^{K} \frac{e^{-sL_k} - e^{-sR_k}}{s} \times \frac{P_k}{\alpha_k}. \tag{A.2}$$

Eq. (A.1) follows by setting $s = i2\pi\nu$ in Eq. (A.2) and rewriting in terms of the trigonometric definition of complex exponentials. $\square$

APPENDIX B: COMPUTATION OF $\tilde{D}^c_{Y,\xi}(i2\pi\nu)$

The notation in this appendix was defined in Section 2, in the context of continuous histograms and their associated objects.

The cumulative distribution function $\hat{H}^c_Y$ corresponding to the continuous histogram density $\hat{h}^c_Y$ of Eq. (2.1) is the piecewise linear function

$$\hat{H}^c_Y(y) \;=\; \begin{cases} 0, & y < l_1 \\ \hat{C}_{j-1} + (y - l_j)\dfrac{\hat{p}_j}{w_j}, & y \in [l_j, r_j),\; 1 \le j \le J \\ \hat{C}_j, & y \in [r_j, l_{j+1}),\; 1 \le j \le J-1 \\ 1, & y \ge r_J \end{cases} \tag{B.1}$$

where $\{\hat{C}_j\}_{j=0}^J$ is the cumulative distribution function corresponding to $\{\hat{p}_j\}_{j=1}^J$, given by

$$\hat{C}_j \;=\; \sum_{i=1}^{j} \hat{p}_i, \quad 0 \le j \le J. \tag{B.2}$$

Note that Eq. (B.2) implies that $\hat{C}_0 = 0$ and $\hat{C}_J = 1$. From Eqs. (B.1) and (D.2) it follows that the inverse $(\hat{H}^c_Y)^{-1}$ can be written as

$$(\hat{H}^c_Y)^{-1}(x) = \sum_{j \in \mathcal{J}^+} 1_{[\hat{C}_{j-1}, \hat{C}_j)}(x) \left[l_j \;+\; (x - \hat{C}_{j-1})\,\frac{w_j}{\hat{p}_j} \right], \quad 0 \le x \le 1. \tag{B.3}$$

Recall that for $0 \le \xi \le 1$, we have

$$D^c_{Y,\xi}(x) = (\hat{H}^c_Y)^{-1}(S_\xi(x)),\; 0 \le x \le 1. \tag{B.4}$$

Proposition 2 *For all $\nu \ge 1$, and $0 \le \xi \le 1$,*

$$\begin{aligned}
\tilde{D}^c_{Y,\xi}(i2\pi\nu) \;=\; \sum_{j \in \mathcal{J}^+} \Bigg[& \frac{r_j \sin(2\pi\nu\xi\hat{C}_j) \;-\; l_j \sin(2\pi\nu\xi\hat{C}_{j-1})}{2\pi\nu} \\
& + \frac{\cos(2\pi\nu\xi\hat{C}_j) \;-\; \cos(2\pi\nu\xi\hat{C}_{j-1})}{\xi(2\pi\nu)^2} \times \frac{w_j}{\hat{p}_j} \\
& + \frac{r_j \sin(2\pi\nu(1-\xi)\hat{C}_j) \;-\; l_j \sin(2\pi\nu(1-\xi)\hat{C}_{j-1})}{2\pi\nu} \\
& + \frac{\cos(2\pi\nu(1-\xi)\hat{C}_j) \;-\; \cos(2\pi\nu(1-\xi)\hat{C}_{j-1})}{(1-\xi)(2\pi\nu)^2} \times \frac{w_j}{\hat{p}_j} \Bigg]
\end{aligned}$$

$$
\begin{aligned}
&+\ i \sum_{j\in\mathcal{J}^+} \Big[\frac{r_j \cos(2\pi\nu\xi\hat{C}_j) \ -\ l_j \cos(2\pi\nu\xi\hat{C}_{j-1})}{2\pi\nu} \\
&\qquad - \frac{\sin(2\pi\nu\xi\hat{C}_j) \ -\ \sin(2\pi\nu\xi\hat{C}_{j-1})}{\xi(2\pi\nu)^2} \times \frac{w_j}{\hat{p}_j} \\
&\qquad - \frac{r_j \cos(2\pi\nu(1-\xi)\hat{C}_j) \ -\ l_j \cos(2\pi\nu(1-\xi)\hat{C}_{j-1})}{2\pi\nu} \\
&\qquad - \frac{\sin(2\pi\nu(1-\xi)\hat{C}_j) \ -\ \sin(2\pi\nu(1-\xi)\hat{C}_{j-1})}{(1-\xi)(2\pi\nu)^2} \times \frac{w_j}{\hat{p}_j}\Big]
\end{aligned}
\tag{B.5}
$$

and for the limiting cases $\xi = 1$ and $\xi = 0$, this simplifies into

$$
\begin{aligned}
\tilde{D}^c_{Y,1}(i2\pi\nu) \ &= \ \tilde{D}^c_{Y,0}(i2\pi\nu) \ = \\
&\sum_{j\in\mathcal{J}^+} \Big[\frac{r_j \ \sin(2\pi\nu\hat{C}_j) \ -\ l_j \ \sin(2\pi\nu\hat{C}_{j-1})}{2\pi\nu} \\
&\qquad + \frac{\cos(2\pi\nu\hat{C}_j) \ -\ \cos(2\pi\nu\hat{C}_{j-1})}{(2\pi\nu)^2} \times \frac{w_j}{\hat{p}_j}\Big] \\
&+\ i \sum_{j\in\mathcal{J}^+} \Big[\frac{r_j \ \cos(2\pi\nu\hat{C}_j) \ -\ l_j \ \cos(2\pi\nu\hat{C}_{j-1})}{2\pi\nu} \\
&\qquad - \frac{\sin(2\pi\nu\hat{C}_j) \ -\ \sin(2\pi\nu\hat{C}_{j-1})}{(2\pi\nu)^2} \times \frac{w_j}{\hat{p}_j}\Big].
\end{aligned}
\tag{B.6}
$$

Proof The quantities $D^c_{Y,\xi}(i2\pi\nu)$ will be derived in two steps. First, we derive the simple case $D^c_{Y,1}(i2\pi\nu)$ (for $\xi = 1$), where the stitching transformation S_1 is just the identity, and so it effectively drops out of the calculation. The general case $D^c_{Y,\xi}(i2\pi\nu)$ is then obtained via the relation (3.9).

We begin by substituting Eq. (B.3) into the definition of $\tilde{D}^c_{Y,1}$ in Eq. (B.4), resulting in

$$
\begin{aligned}
\tilde{D}^c_{Y,1}(s) \ &= \ \int_0^1 e^{-sx} \, D^c_{Y,1}(x)\, dx \ = \ \int_0^1 e^{-sx} \, (\hat{H}^c_Y)^{-1}(x)\, dx \\
&= \ \sum_{j\in\mathcal{J}^+} \int_{\hat{C}_{j-1}}^{\hat{C}_j} e^{-sx} \left[l_j + (x - \hat{C}_{j-1})\frac{w_j}{\hat{p}_j}\right] dx,
\end{aligned}
$$

whence,

$$\tilde{D}^c_{Y,1}(s) \;=\; \sum_{j\in\mathcal{J}^+}\left[\frac{l_j\,e^{-s\hat{C}_{j-1}} \;-\; r_j\,e^{-s\hat{C}_j}}{s} \;+\; \frac{e^{-s\hat{C}_{j-1}} \;-\; e^{-s\hat{C}_j}}{s^2}\times\frac{w_j}{\hat{p}_j}\right] \tag{B.7}$$

Thus, substituting $s = i2\pi\nu$ into Eq. (3.9) for $\tilde{D}^c_{Y,\xi}$ yields

$$\tilde{D}^c_{Y,\xi}(i2\pi\nu) \;=\; \xi\,\tilde{D}^c_{Y,1}(\xi i2\pi\nu) \;+\; (1-\xi)\,\tilde{D}^c_{Y,1}(-(1{-}\xi)i2\pi\nu), \tag{B.8}$$

because $e^{-i2\pi\nu} = 1$. Hence, from Eq. (B.7),

$$\begin{aligned}\xi\,\tilde{D}^c_{Y,1}(\xi s) \;&=\; \sum_{j\in\mathcal{J}^+}\Bigg[\frac{l_j\,e^{-\xi s\hat{C}_{j-1}} \;-\; r_j\,e^{-\xi s\hat{C}_j}}{s} \;+\\ &\qquad\qquad \frac{e^{-\xi s\hat{C}_{j-1}} \;-\; e^{-\xi s\hat{C}_j}}{\xi\,s^2}\times\frac{w_j}{\hat{p}_j}\Bigg],\end{aligned} \tag{B.9}$$

and

$$\begin{aligned}(1-\xi)\,\tilde{D}^c_{Y,1}(-(1{-}\xi)s) \;&=\; \sum_{j\in\mathcal{J}^+}\Bigg[-\;\frac{l_j\,e^{(1-\xi)s\hat{C}_{j-1}} - r_j\,e^{(1-\xi)s\hat{C}_j}}{s} \;+\\ &\qquad\qquad \frac{e^{(1-\xi)s\hat{C}_{j-1}} \;-\; e^{(1-\xi)s\hat{C}_j}}{(1-\xi)s^2}\times\frac{w_j}{\hat{p}_j}\Bigg].\end{aligned} \tag{B.10}$$

Next, set $s = i2\pi\nu$ in Eqs. (B.9) and (B.10), and substitute in Eq. (B.8). The general result (B.5) now follows, by expanding the complex exponentials in terms of their trigonometric representation in moderately tedious algebra.

Finally, to obtain Eq. (B.6), take limits in Eq. (B.5) as $\xi \downarrow 0$ or as $\xi \uparrow 1$, and use L'Hôpital's rule in each case. In both cases, we obtain vanishing terms, and Eq. (B.5) simplifies into (B.6). $\square$

APPENDIX C: COMPUTATION OF $\tilde{D}^d_{Y,\xi}(i2\pi\nu)$

The notation in this appendix was defined in Section 2, in the context of discrete histograms and their associated objects.

The cumulative distribution function $\hat{H}_Y^d$ corresponding to the discrete histogram density $\hat{h}_Y^d$ of Eq. (2.2) is the step function

$$\hat{H}_Y^d(y) \;=\; \begin{cases} 0, & y < v_1 \\ \hat{C}_i, & y \in [v_i, v_{i+1}),\ 1 \le i \le I-1 \\ 1, & y > v_I \end{cases} \tag{C.1}$$

where $\{\hat{C}_i\}_{i=0}^I$ is the cumulative distribution function corresponding to $\{\hat{p}_i\}_{i=1}^I$, given by

$$\hat{C}_i \;=\; \sum_{j=1}^{i} \hat{p}_j, \quad 0 \le j \le I. \tag{C.2}$$

Note that Eq. (C.2) implies that $\hat{C}_0 = 0$ and $\hat{C}_I = 1$. From Eqs. (C.1) and (C.2) it follows that the inverse $(\hat{H}_Y^d)^{-1}$ can be written as

$$(\hat{H}_Y^d)^{-1}(x) \;=\; \sum_{i \in \mathcal{I}^+} 1_{[\hat{C}_{i-1}, \hat{C}_i)}(x)\, v_i, \tag{C.3}$$

Recall that for $0 \le \xi \le 1$, we have

$$D_{Y,\xi}^d(x) = (\hat{H}_Y^d)^{-1}(S_\xi(x)),\ 0 \le x \le 1. \tag{C.4}$$

Proposition 3 *For all $\nu \ge 1$, and $0 \le \xi \le 1$,*

$$\begin{aligned} \tilde{D}_{Y,\xi}^d(i2\pi\nu) \;=\; & \sum_{i \in \mathcal{I}^+} \frac{v_i}{2\pi\nu}[\sin(2\pi\nu\xi\hat{C}_i) \;-\; \sin(2\pi\nu\xi\hat{C}_{i-1}) \\ & \qquad + \sin(2\pi\nu(1-\xi)\hat{C}_i) \;-\; \sin(2\pi\nu(1-\xi)\hat{C}_{i-1})] \\ & + \; i \sum_{i \in \mathcal{I}^+} \frac{v_i}{2\pi\nu}[\cos(2\pi\nu\xi\hat{C}_i) \;-\; \cos(2\pi\nu\xi\hat{C}_{i-1}) \\ & \qquad - \cos(2\pi\nu(1-\xi)\hat{C}_i) \;-\; \cos(2\pi\nu(1-\xi)\hat{C}_{i-1})] \end{aligned} \tag{C.5}$$

and for the cases $\xi = 1$ and $\xi = 0$, this simplifies into

$$\begin{aligned} \tilde{D}_{Y,1}^d(i2\pi\nu) \;=\; \tilde{D}_{Y,0}^d(i2\pi\nu) \;=\; & \\ & \sum_{i \in \mathcal{I}^+} \frac{v_i}{2\pi\nu}\,[\sin(2\pi\nu\hat{C}_i) \;-\; \sin(2\pi\nu\hat{C}_{i-1})] \\ + \; i & \sum_{i \in \mathcal{I}^+} \frac{v_i}{2\pi\nu}\,[\cos(2\pi\nu\hat{C}_i) \;-\; \cos(2\pi\nu\hat{C}_{i-1})]. \end{aligned} \tag{C.6}$$

Proof We proceed analogously to the previous appendix, except that the computations are somewhat simpler.

Substituting Eq. (C.3) into the definition of $\tilde{D}^d_{Y,1}$ in Eq. (C.4) yields

$$\begin{aligned}\tilde{D}^d_{Y,1}(s) &= \int_0^1 e^{-sx}\, D^d_{Y,1}(x)\, dx \;=\; \int_0^1 e^{-sx}\, (\hat{H}^d_Y)^{-1}(x)\, dx \\ &= \sum_{i\in\mathcal{I}^+} \int_{\hat{C}_{i-1}}^{\hat{C}_i} e^{-sx}\, v_i\, dx,\end{aligned}$$

whence

$$\tilde{D}^d_{Y,1}(s) \;=\; \sum_{i\in\mathcal{I}^+} \frac{e^{-s\hat{C}_{i-1}} - e^{-s\hat{C}_i}}{s} \times v_i. \tag{C.7}$$

Thus, substituting $s = i2\pi\nu$ into Eq. (3.9) yields, analogously to (B.8),

$$\tilde{D}^d_{Y,\xi}(i2\pi\nu) \;=\; \xi\, \tilde{D}^d_{Y,1}(\xi i2\pi\nu) \;+\; (1-\xi)\, \tilde{D}^d_{Y,1}(-(1-\xi) i2\pi\nu). \tag{C.8}$$

Hence, from Eq. (C.7),

$$\xi\, \tilde{D}^d_{Y,1}(\xi s) \;=\; \sum_{i\in\mathcal{I}^+} \left[\frac{e^{-\xi s\hat{C}_{i-1}} - e^{-\xi s\hat{C}_i}}{s} \times v_i\right], \tag{C.9}$$

and

$$(1-\xi)\, \tilde{D}^d_{Y,1}(-(1-\xi)s) \;=\; \sum_{i\in\mathcal{I}^+} \left[-\,\frac{e^{(1-\xi)s\hat{C}_{i-1}} - e^{(1-\xi)s\hat{C}_i}}{s} \times v_i\right]. \tag{C.10}$$

Next, set $s = i2\pi\nu$ in Eqs. (C.9) and (C.10), and substitute in Eq. (C.8). The general result (C.5) now follows, by expanding the complex exponentials in terms of their trigonometric representation in moderately tedious algebra.

Finally, to obtain the special cases (C.6), just set $\xi = 1$ or $\xi = 0$ into the general case (C.5). □

APPENDIX D: COMPUTATION OF $\tilde{D}_{Y,\xi}(i2\pi\nu)$

The notation in this appendix was defined in Section 2, in the context of general histograms and their associated objects.

In this section we find it notationally useful to consider a discrete component of a general histogram as a special case of a continuous (uniform) component with $v_m = l_m = r_m$. With this notational convention, the cumulative distribution function $\hat{H}_Y^c$ corresponding to the continuous histogram density $\hat{h}_Y^c$ of Eq. (2.3) is the piecewise linear function

$$\hat{H}_Y(y) = \begin{cases} 0, & y < l_1 \\ \hat{C}_{m-1} + (y - l_m)\frac{\hat{p}_m}{w_m}, & y \in [l_m, r_m),\ m \in \mathcal{J} \\ \hat{C}_m, & y \in [r_m, l_{m+1}),\ 1 \le m \le M-1 \\ 1, & y \ge r_M \end{cases} \tag{D.1}$$

where $\{\hat{C}_m\}_{m=0}^M$ is the cumulative distribution function corresponding to mixing probabilities $\{\hat{p}_m\}_{m=1}^M$ (of the sorted components), given by

$$\hat{C}_m = \sum_{n=1}^{m} \hat{p}_n, \quad 0 \le m \le M. \tag{D.2}$$

Note that Eq. (D.2) implies that $\hat{C}_0 = 0$ and $\hat{C}_M = 1$. From Eqs. (D.1) and (D.2) it follows that the inverse $\hat{H}_Y^{-1}$ can be written as

$$\begin{aligned} \hat{H}_Y^{-1}(x) &= \sum_{j \in \mathcal{J}^+} 1_{[\hat{C}_{j-1}, \hat{C}_j)}(x) \left[l_j + (x - \hat{C}_{j-1}) \frac{w_j}{\hat{p}_j} \right] \\ &+ \sum_{i \in \mathcal{I}^+} 1_{[\hat{C}_{i-1}, \hat{C}_i)}(x)\, v_i, \quad 0 \le x \le 1. \end{aligned} \tag{D.3}$$

Recall that for $0 \le \xi \le 1$, we have

$$D_{Y,\xi}(x) = \hat{H}_Y^{-1}(S_\xi(x)),\ 0 \le x \le 1. \tag{D.4}$$

Proposition 4 *For all $\nu \ge 1$, and $0 \le \xi \le 1$,*

$$\tilde{D}_{Y,\xi}(i2\pi\nu) = (B.5) + (C.5), \tag{D.5}$$

and for $\xi = 1$ or $\xi = 0$,

$$\tilde{D}_{Y,1}(i2\pi\nu) = \tilde{D}_{Y,0}(i2\pi\nu) = (B.6) + (C.6). \tag{D.6}$$

Proof We show that the results can be formed as additive composites of the corresponding continuous and discrete components.

We begin by substituting Eq. (D.3) into the definition of $\tilde{D}_{Y,1}$ in Eq. (D.4), resulting in

$$\begin{aligned}\tilde{D}_{Y,1}(s) \;&=\; \int_0^1 e^{-sx}\, D_{Y,1}(x)\, dx \;=\; \int_0^1 e^{-sx}\, \hat{H}_Y^{-1}(x)\, dx \\ &=\; \sum_{j\in\mathcal{J}^+} \int_{\hat{C}_{j-1}}^{\hat{C}_j} e^{-sx} \left[l_j + (x-\hat{C}_{j-1})\frac{w_j}{\hat{p}_j} \right] dx \;+\; \sum_{i\in\mathcal{I}^+} \int_{\hat{C}_{i-1}}^{\hat{C}_i} e^{-sx}\, v_i\, dx,\end{aligned}$$

whence, from Propositions 2 and 3,

$$\tilde{D}_{Y,1}(s) \;=\; (B.7) \;+\; (C.7). \tag{D.7}$$

The rest of the computations proceed exactly as in Proposition 2 for the continuous components and as in Proposition 3 for the discrete case. □