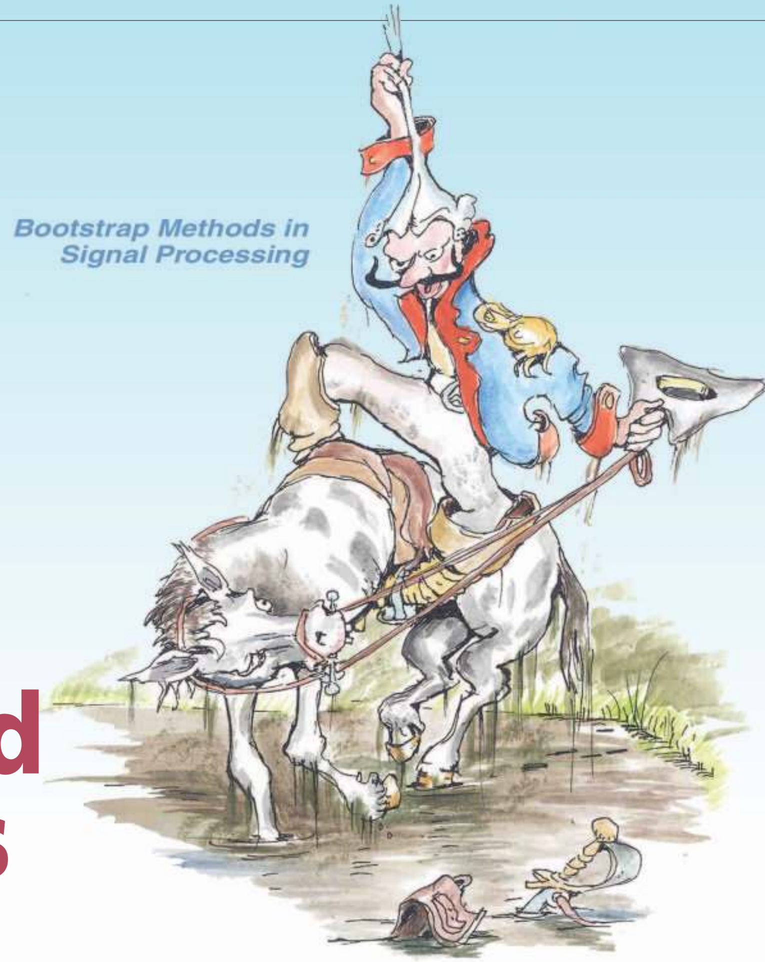
Zoubir, Abdelhak M. and Iskander, D. Robert (2007) Bootstrap Methods and Applications : A Tutorial for the Signal Processing Practitioner. *IEEE - Signal Processing Magazine* 24(4):pp. 10-19.

Abdelhak M. Zoubir and D. Robert Iskander



PROF. DR. KARL HEINRICH HOFMANN

Bootstrap Methods and Applications

A tutorial for the signal processing practitioner

This year marks the pearl anniversary of the bootstrap. It has been 30 years since Bradley Efron's 1977 Reitz lecture, published two years later in [1]. Today, bootstrap techniques are available as standard tools in several statistical software packages and are used to solve problems in a wide range of applications. There have also been several monographs written on the topic, such as [2], and several tutorial papers written for a nonstatistical readership, including two for signal processing practitioners published in this magazine [4], [5].

Given the wealth of literature on the topic supported by solutions to practical problems, we would expect the bootstrap to be an off-the-shelf tool for signal processing problems as are maximum likelihood and least-squares methods. This is not the case, and we wonder why a signal processing practitioner would not resort to the bootstrap for inferential problems.

We may attribute the situation to some confusion when the engineer attempts to discover the bootstrap paradigm in an overwhelming body of statistical literature. To give an example and ignoring the two basic approaches of the bootstrap, i.e., the parametric and the nonparametric bootstrap [2], there is not only one bootstrap. Many variants of it exist, such as the small bootstrap [6], the wild bootstrap [7], the naïve bootstrap (a name often given to the standard bootstrap resampling technique), the block (or moving block) bootstrap (see the chapter by Liu and Singh in [8]) and its extended circular block bootstrap version (see the chapter by Politis and Romano in [8]), and the iterated bootstrap [9]. Then there are derivatives such as the weighted bootstrap or the threshold bootstrap and some more recently introduced methods such as bootstrap bagging and bumping. Clearly, this wide spectrum of bootstrap variants may be a hurdle for newcomers to this area.

 1053-5888/07/$25.00©2007IEEE

The name *bootstrap* is often associated with the tale of Baron von Münchhausen who pulled himself up by the bootstraps from a sticky situation. This analogy may suggest that the bootstrap is able to perform the impossible and has resulted sometimes in unrealistic expectations, especially when dealing with real data. Often, a signal processing practitioner attempting to use the basic concepts of the bootstrap is encouraged by his or her early simulation studies. However, this initial fascination is often followed by fading interests in the bootstrap, especially when the technique did not prove itself with real data. Clearly, the bootstrap is not a magic technique that provides a panacea for all statistical inference problems, but it has the power to substitute tedious and often impossible analytical derivations with computational calculations [3], [5], [10]. The bootstrap indeed has the potential to become a standard tool for the engineer. However, care is required with the use of the bootstrap as there are situations, discussed later, in which the bootstrap fails [11].

The first question a reader unfamiliar with the topic would ask is, "what is the bootstrap used for?" In general terms, the answer would be "the bootstrap is a computational tool for statistical inference." Specifically, we could list the following tasks: estimation of statistical characteristics such as bias, variance, distribution functions and thus confidence intervals, and more specifically, hypothesis tests (for example for signal detection), and model selection. The following question may arise subsequently, "when can I use the bootstrap?" A short answer to this is, "when I know little about the statistics of the data or I have only a small amount of data so that I cannot use asymptotic results."

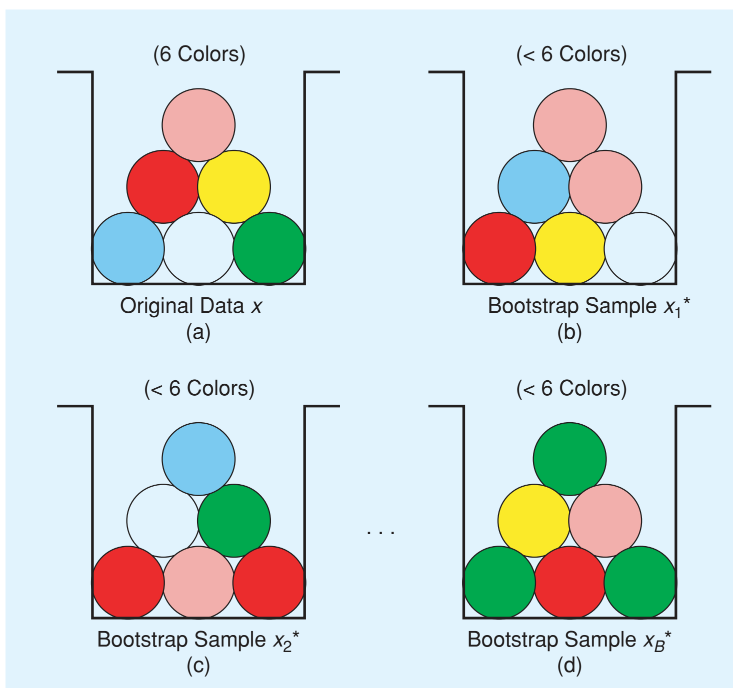
Our aim is to give a short tutorial of bootstrap methods supported by real-life applications so as to substantiate the answers to the questions raised above. This pragmatic approach is to serve as a practical guide rather than a comprehensive treatment, which can be found elsewhere; see for example [2]–[5].

THE BOOTSTRAP PRINCIPLE

Suppose that we have measurements collected in $x = \{x_1, x_2, \ldots, x_n\}$, which are realizations of the random sample $X = \{X_1, X_2, \ldots, X_n\}$, drawn from some unspecified distribution F_X. Let $\hat{\theta} = \hat{\theta}(X)$ be an estimator of some parameter θ of F_X, which could be, for example, the mean $\theta = \mu_X$ of F_X estimated by the sample mean $\hat{\theta} = \hat{\mu}_X = 1/n \sum_{i=1}^{n} X_i$. The aim is to find characteristics of $\hat{\theta}$ such as the distribution of $\hat{\theta}$. Sometimes, the parameter estimator $\hat{\theta}$ is computed from a collection of n independently and identically distributed (i.i.d.) data $X_1, X_2, \ldots, X_n$. If the distribution function F_X is known or is assumed to be known and given that the function $\hat{\theta}(X)$ is relatively simple, then it is possible to exactly evaluate the distribution of the parameter estimator $\hat{\theta}$. Textbook examples of this situation are the derivations of the distribution functions of the sample mean $\hat{\mu}_X$ and its variance when the data is Gaussian.

In many practical applications, either the distribution F_X is unknown or the parameter estimator $\hat{\theta}(X)$ is too complicated for its distribution to be derived in a closed form. The question is then how to perform statistical inference. Specifically, we wish to answer the following question: how reliable is the parameter estimator $\hat{\theta}$? How could we, for example, test that the parameter θ is significantly different from some nominal value (hypothesis test)? Clearly, we could use asymptotic arguments and approximate the distribution of $\hat{\theta}$. In the case of the sample mean $\hat{\mu}_X$ above, we would apply the central limit theorem and assume that the distribution of $\hat{\mu}_X$ is Gaussian. This would lead to answering inferential questions. But how would we proceed if the central limit theorem does not apply because n is small and we cannot repeat the experiment? The bootstrap is the answer to our question. Its paradigm suggests substitution of the unknown distribution F_X by the empirical distribution of the data, $\hat{F}_X$. Practically, it means that we reuse our original data through resampling to create what we call a bootstrap sample. The bootstrap sample has the same size as the original sample, i.e., $x_b^* = \{x_1^*, x_2^*, \ldots, x_n^*\}$ for $b = 1, 2, \ldots, B$, where x_i^*, $i = 1, 2, \ldots, n$ are obtained, for example, by drawing at random with replacement from x. The simplest form of resampling is pictured in Figure 1. Each of the bootstrap samples in the figure is considered as new data. Based on the bootstrap sample x_b^*, bootstrap parameter estimates $\hat{\theta}_b^* = \hat{\theta}(x_b^*)$ for $b = 1, \ldots, B$ are calculated. Given a large number B of bootstrap parameter estimates, we can then approximate the distribution of $\hat{\theta}$ by the distribution of $\hat{\theta}^*$, which is derived from the bootstrap sample x^*, i.e., we approximate the distribution $F_{\hat{\theta}}$ of $\hat{\theta}$ by $\hat{F}_{\hat{\theta}^*}$, the distribution of $\hat{\theta}^*$.

From a practical point of view, a limitation of the bootstrap may appear to be the i.i.d. data assumption, but we will show



[FIG1] The independent data bootstrap resampling principle.

[FIG2] Typical example of a slit lamp image of an eye with manually selected points (yellow crosses) around the limbus area.

later how this assumption can be relaxed. There are, however, several other technical points that need to be addressed. The sample length n is also of great importance. Bootstrap methods have been promoted as methods for small sample sizes when asymptotic assumptions may not hold. However, as with any statistical problem, the sample size will influence the results in practice. The number of bootstrap samples B necessary to estimate the distribution of a parameter estimator has also been discussed in the statistical literature [12]. One rule of thumb is for the number of bootstrap samples B to take a value between 25 and 50 for variance estimation and to be set to about 1,000 where a 95% confidence interval is sought. However, with the fast increasing computational power, there are no objections to exceeding these numbers.

Note that the bootstrap simulation error, which quantifies the difference between the true distribution and the estimated distribution, comprises two independent errors of different sources, i.e., a bootstrap (statistical) error and a simulation (Monte Carlo) error. The first error is unavoidable and does not depend on the number of bootstrap samples B but on the size n of the original sample. The second one can be minimized by increasing the number of bootstrap samples. The aim is therefore to choose B so that the simulation error is no larger than the bootstrap error. For a large sample size n, we would reduce the number of samples B to reduce computations. However, the larger the size n of the original data, the smaller the bootstrap error. Thus, a larger number of bootstrap samples is required for the simulation error to be smaller than the bootstrap error. We found that the rule of thumb of choosing $B = 40n$, proposed by Davison and Hinkley [13], is appropriate in many applications. If desired, a method called jackknife-after-bootstrap [14] can be used to assess the contribution of each of these errors (i.e., bootstrap error versus Monte Carlo error). In practice, the value of B is application dependent and is left to the experimenter to choose.

The assumption of the original data being a good representation of the unknown population is not well articulated in the statistical literature. However, it is quite intuitive to a signal processing practitioner who is familiar with the jargon *garbage in→garbage out*. The issue essentially concerns the allowed number of outliers contained in the original data sample for the bootstrap to work because when we resample with replacement, it is likely that we produce bootstrap samples with significantly higher numbers of outliers than the original sample. The issue of a good original sample is closely related to that of sample size.

Many success stories have been reported by both statisticians and engineers, while little is shown on bootstrap failures. Cases indeed exist where bootstrap procedures fail no matter how good the original sample is and no matter how large n is. A classical example of bootstrap failure is when we apply the independent data bootstrap to find the distribution of the maximum (or the minimum) of a random sample. Another example is when the mean of a random variable with infinite variance (e.g., from the family of α-stable distributions) is of interest. This implies that standard bootstrap techniques may produce uncontrolled results for heavy-tailed distributions. See the work of Mammen [15] for more details.

A promising method that has been reported to work when the conventional independent data bootstrap fails is subsampling. Note that subsampling has been developed as a method for resampling dependent data under minimal assumptions and is based on drawing at random subsamples of consecutive observations of length less than the original data size n. See the book by Politis et al. [16] for more details.

AN IMAGE PROCESSING EXAMPLE

We consider an example of fitting a circle to a set of two-dimensional (2-D) data. This is common in many image processing and pattern recognition applications. The application we consider is to fit functions to the outlines of the pupil and limbus (iris outline) in eye images [17]. The performance of automatic procedures for extracting these features can be evaluated with synthetic images. However, for real images and particularly for the limbus, these procedures need to be benchmarked against manual operators assisted with computer-based procedures for point selection. An operator is asked to select a small number of points (x_i, y_i), $i = 1, 2, \ldots, n$, where x_i and y_i denote the horizontal and vertical point position (see the eight yellow crosses in Figure 2).

A linear least-squares procedure can be applied to fit a circle, modeled by the equation $x^2 + y^2 + 2xx_0 + 2yy_0 + x_0^2 + y_0^2 - R^2 = 0$, to the data (x_i, y_i), $i = 1, 2, \ldots, n$, so that

$$\begin{bmatrix} x_1^2 + y_1^2 \\ \vdots \\ x_n^2 + y_n^2 \end{bmatrix} = \begin{bmatrix} 2x_1 & 2y_1 & 1 \\ \vdots & \vdots & \vdots \\ 2x_n & 2y_n & 1 \end{bmatrix} \cdot \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix},$$

[TABLE 1] BOOTSTRAP PROCEDURE FOR THE ESTIMATION OF THE DISTRIBUTION OF $\hat{P}$.

STEP 1) ESTIMATE THE THREE PARAMETERS OF THE CIRCLE x_0, y_0, AND R, COLLECTED IN $\hat{\boldsymbol{P}}$ AND CONSTRUCT AN ESTIMATE OF THE CIRCLE USING $\hat{\boldsymbol{Y}} = \boldsymbol{X} \cdot \hat{\boldsymbol{P}}$.
STEP 2) CALCULATE THE RESIDUALS $\hat{\boldsymbol{E}} = \boldsymbol{Y} - \hat{\boldsymbol{Y}}$.
STEP 3) SINCE WE HAVE ASSUMED THAT THE RESIDUALS $\hat{\boldsymbol{E}}$ ARE I.I.D., WE CREATE A SET OF BOOTSTRAP RESIDUALS $\hat{\boldsymbol{E}}^*$ BY RESAMPLING WITH REPLACEMENT FROM $\hat{\boldsymbol{E}}$. NOTE THAT THE RESIDUALS NEED TO BE CENTERED (DETRENDED) BEFORE RESAMPLING.
STEP 4) CREATE A NEW ESTIMATE OF THE CIRCLE BY ADDING THE BOOTSTRAPPED RESIDUALS TO THE ESTIMATE, OBTAINED FROM THE ORIGINAL DATA $\boldsymbol{Y}$ IN STEP 1, I.E., $\hat{\boldsymbol{Y}}^* = \boldsymbol{X} \cdot \hat{\boldsymbol{P}} + \hat{\boldsymbol{E}}^*$.
STEP 5) ESTIMATE A NEW SET OF PARAMETERS FROM THE NEWLY CREATED BOOTSTRAP SAMPLE $\hat{\boldsymbol{P}}^* = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\hat{\boldsymbol{Y}}^*$.
STEP 6) REPEAT STEPS 3–5 B TIMES TO OBTAIN A SET OF $\hat{\boldsymbol{P}}_1^*, \hat{\boldsymbol{P}}_2^*, \ldots, \hat{\boldsymbol{P}}_B^*$, FROM WHICH EMPIRICAL DISTRIBUTIONS OF THE CONSIDERED PARAMETER ESTIMATORS CAN BE OBTAINED.

where $p_1 = x_0$, $p_2 = y_0$, $p_3 = x_0^2 + y_0^2 - R^2$ and ε_i, $i = 1, 2, \ldots, n$, is the modeling error.

The above equation can be rewritten in the form $Y = X \cdot P + E$, for which an estimator for P is easily derived, i.e., $\hat{P} = (X^T X)^{-1} X^T Y$.

The question of interest is how well an operator can fit a circle to the limbus. One way of assessing the parameter estimator would be to select eight data point pairs (considered in our example) 1,000 times. Although feasible, this task is laborious, and the results would most definitely be affected by the subsequently decreasing commitment of the operator. The alternative is to use the bootstrap. Clearly, the selected data points (x_i, y_i), $i = 1, 2, \ldots, n$, are not i.i.d, unlike the modeling errors $\varepsilon_i, i = 1, \ldots, n$, collected in the random sample $\varepsilon = \{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$, which can be assumed to be i.i.d. Our bootstrap procedure is described in Table 1.

An example of the distribution (histogram) of the limbus radius obtained with the bootstrap method is shown in Figure 3. Clearly, the bootstrap is capable of providing answers, substituting the tedious manual labor that would have been required to complete this task. In the above example, we used the bootstrap to find the distribution of the limbus radius estimator. This is not the only question of interest in this application. The bootstrap can also be used to estimate the existing bias when fitting the data to ellipses [18], or it can be used for testing whether the limbus or pupil parameters are different from the left to the right eye in anisometropic subjects.

BOOTSTRAP TECHNIQUES FOR DEPENDENT DATA

The assumption that the data is i.i.d. is not always valid. Here we provide some insight as to how to resample dependent data. Note that if the data was i.i.d., standard bootstrap resampling with replacement gives an accurate representation of the underlying distribution. However, if the data shows heteroskedasticity (the random variables in the sequence or vector may have different variances) or serial correlation, randomly resampled data would lead to errors.

One way to extend the basic bootstrap principle to dependent data is the previously mentioned concept of data modeling and the subsequent assumption of i.i.d. residuals that approximate the modeling and measurement errors.

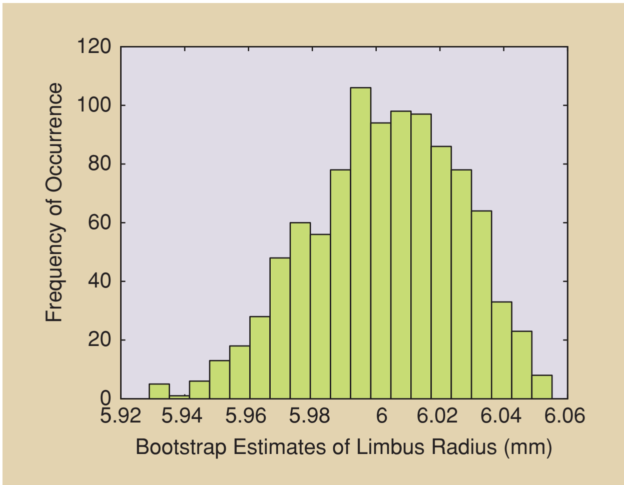
[TABLE 2] RESIDUAL-BASED BOOTSTRAP PROCEDURE FOR DEPENDENT DATA.

STEP 1) FIT A MODEL TO THE DATA.
STEP 2) SUBTRACT THE FITTED MODEL FROM THE ORIGINAL DATA TO OBTAIN RESIDUALS.
STEP 3) CENTER (OR RESCALE) THE RESIDUALS.
STEP 4) RESAMPLE THE RESIDUALS.
STEP 5) CREATE NEW BOOTSTRAP DATA BY ADDING THE RESAMPLED RESIDUALS TO THE FITTED MODEL FROM STEP 1.
STEP 6) FIT THE MODEL TO THE NEW BOOTSTRAP DATA.
STEP 7) REPEAT STEPS 4–6 MANY TIMES TO OBTAIN DISTRIBUTIONS FOR THE MODEL PARAMETER ESTIMATORS.

There have been a variety of bootstrap methods developed for dependent data models such as autoregressive (AR) and moving average models (see [19] and references therein), and Markov chain models (see the chapter by Athreya and Fuh in [8]), in which the concept of i.i.d. residuals has been used. In analogy to the linearization of a nonlinear problem, the idea here is to reformulate the problem so that the i.i.d. component of the data may be used for resampling. In most cases, the procedure follows the structure described in Table 2.

We used the above procedure in many signal processing problems, including those related to higher-order statistics and nonstationary signals with polynomial phase [3]. In some cases, the residuals in Step 2 of the above procedure can be found as a ratio of two parameter estimators. For example, in power spectrum density estimation [20], the ratios between the periodogram and the kernel spectrum density estimator at distinct frequency bins are assumed to be i.i.d. Note, however, that we could not use the same concept for the bispectrum [3]. We also note the approach taken by the authors in some real-life applications where the asymptotic independence of the finite Fourier transform at distinct frequencies was explored so that sampling could be undertaken in the frequency domain [3].

As an example, we describe below the principle of bootstrap resampling for AR models. Given n observations x_t, $t = 1, \ldots, n$, of an AR process of order p and coefficients a_k,



[FIG3] Histogram of $\hat{R}_1^*, \hat{R}_2^*, \ldots, \hat{R}_{1,000}^*$ based on the eight manually selected limbus points.

[TABLE 3] BOOTSTRAP RESAMPLING FOR AR MODELS.

STEP 1) WITH THE ESTIMATES $\hat{a}_k$ OF a_k FOR $k = 1, \ldots, p$ (OBTAINED BY SOLVING THE YULE-WALKER EQUATIONS), CALCULATE THE RESIDUALS AS $\hat{z}_t = x_t + \sum_{k=1}^{p} \hat{a}_k x_{t-k}$ FOR $t = p+1, \ldots, n$.
STEP 2) CREATE A BOOTSTRAP SAMPLE $x_1^*, \ldots, x_n^*$ BY DRAWING $\hat{z}_{p+1}^*, \ldots, \hat{z}_n^*$, WITH REPLACEMENT FROM THE RESIDUALS $\hat{z}_{p+1}, \ldots, \hat{z}_n$, THEN LETTING $x_t^* = x_t$ FOR $t = 1, \ldots, p$ AND $x_t^* = -\sum_{k=1}^{p} \hat{a}_k x_{t-k}^* + \hat{z}_t^*$ FOR $t = p+1, \ldots, n$.
STEP 3) OBTAIN BOOTSTRAP ESTIMATES $\hat{a}_1^*, \ldots, \hat{a}_p^*$ FROM $x_1^*, \ldots, x_n^*$.
STEP 4) REPEAT STEPS 2-3 B TIMES TO OBTAIN $\hat{a}_1^{*b}, \ldots, \hat{a}_p^{*b}$ FOR $b = 1, \ldots, B$.

$k = 1, \ldots, p$, we would proceed as summarized in Table 3 to create bootstrap parameter estimates so as to estimate the distribution functions of the parameter estimators, based on the original data [3].

The bootstrap estimates $\hat{a}_1^{*b}, \ldots, \hat{a}_p^{*b}$ for $b = 1, \ldots, B$ are used to estimate the distributions of $\hat{a}_1, \ldots, \hat{a}_p$ or their statistical measures such as means, variances, or confidence intervals.

Practical examples of the above procedure are shown in what follows in the context of hands-free telephony and micro-Doppler radar.

AN EXAMPLE FOR HANDS-FREE TELEPHONY

Hands-free communication in cars can be severely disturbed by car noise. To ensure understandability, noise-reduction algorithms are necessary. These are usually assessed by listening to estimated speech samples. However, a quantitative assessment seems to be a more objective approach. In this example, we propose to assess the confidence intervals of the parameters of an AR model used to represent the recovered speech signal to ultimately compare them with those of the AR parameters corresponding to the original signal. A bootstrap approach is suggested due to the complicated nature of the signals and their statistical properties.

The single-channel recorded signal $x(t)$ is described as a mixture of a clear speech signal $s(t)$ and car noise $n(t)$. The noise reduction approach we use here, proposed in [21], assumes both speech and car noise to be AR processes contaminated by white noise. The algorithm uses subband AR modeling and Kalman filtering to find a noise-reduced estimate $\hat{s}(t)$ of the clear speech signal $s(t)$, as shown in Figure 4.

An overview of the noise-reduction algorithm is as follows (details can be found in [21]):

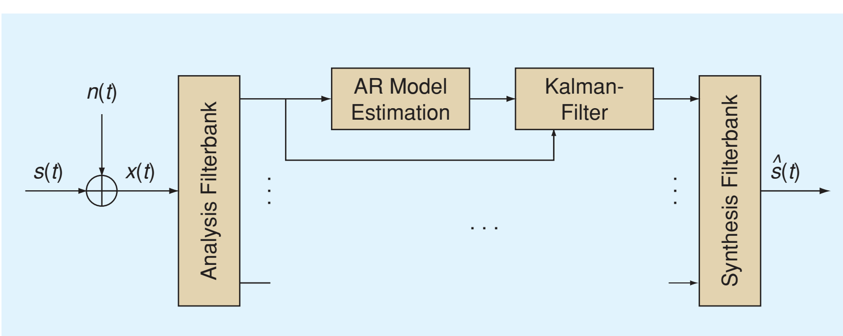
- To obtain small AR model orders, the signal is split into 16 subbands with an undersampling rate of 12. The AR model orders used are 4–6 for clear speech and 2 for car noise.
- Signal segments with 48 ms duration are considered to be quasistationary.
- The block "AR Model Estimation" in Figure 4 can be roughly described as follows.
  - — If only noise is present, the noise spectrum is measured by means of a smoothed periodogram.
  - — If voice activity is detected, the current noise spectrum is held fixed and is subtracted from the disturbed speech spectrum, to obtain an estimate of the speech only spectrum.
  - — AR parameter and input power estimation is performed for both noise and speech separately.

A single-channel speech signal was used. The recording was the German sentence: "Johann Philipp Reis führte es am 26. Oktober 1861, erstmals in Frankfurt am Main, vor und nutzte dazu als einen der ersten Testsätze: 'Pferde fressen keinen Gurkensalat.'" A quasistationary segment of this recording is then chosen corresponding to the vocal part of "Reis." This signal segment, sampled at 8 kHz is shown in Figure 5(a). After estimating the model order to be 11, by means of the minimum description length (MDL) information theoretic criterion, we estimate the AR parameters and use the bootstrap to find their distributions as described in Table 3.

Figure 5(b) shows the residuals, obtained by inverse filtering of the signal with the estimated parameters of the AR model from the recovered speech signal $\hat{s}(t)$. The residuals are close to white, as can be inferred from the covariance function of the residuals shown in Figure 5(c).

A quality assessment of a noise-reduction algorithm should give a measure of how well $\hat{s}(t)$ estimates $s(t)$, or how close $\hat{s}(t)$ is to $s(t)$ in a statistically meaningful way. This could be done based on the bootstrap distributions of the AR parameter estimates of $s(t)$ and $\hat{s}(t)$. We use 90% confidence intervals of the AR parameters based on the estimated values to assess how close the original (clear speech) and the noise-reduced signals are. Figure 5(d) shows the bootstrap 90% confidence intervals for the 11 AR parameters of the noise-reduced signal $\hat{s}(t)$ along with the estimated AR parameters of $\hat{s}(t)$ (black crosses) and the estimated AR parameters of $s(t)$ (red diamonds). The confidence intervals for the fitted AR parameters to the original speech signal $s(t)$ (not shown in the figure) are in close agreement with the confidence intervals shown in Figure 5(d).

From this example, we can deduce that the bootstrap can be used to assess the quality of the noise-reducing algorithm using approximate confidence bounds. As an alternative to the common practice of listening to both the original speech and the noise reduced speech to assess clarity, the confidence bounds of their respective AR parameters are compared.



[FIG4] Noise reduction algorithm using subband AR modeling and Kalman filtering.

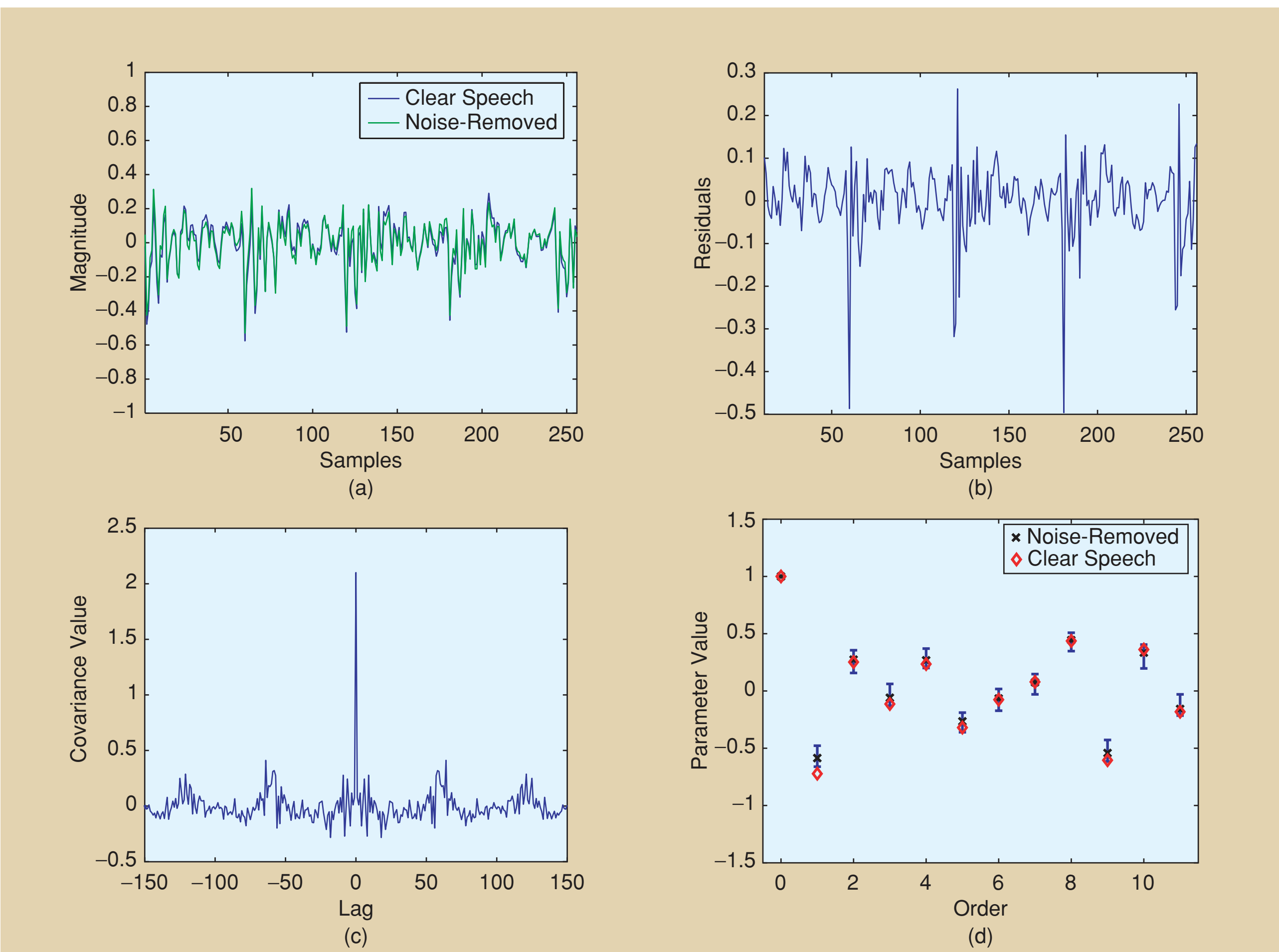
ALTERNATIVE DEPENDENT DATA BOOTSTRAP METHODS

Several questions may be asked at this stage: how can one bootstrap non-i.i.d. data without imposing a parametric model? Can one resample the data nonparametrically? First answers to these questions have been provided by Künsch [22], who introduced the concept of resampling sequences (chunks) of data. The method is referred to as the moving block bootstrap. In essence, rather than resampling with replacement single data points, sets of consecutive points are resampled to maintain, in a nonparametric fashion, the structure between neighboring data points. The segments chosen for bootstrapping can be either nonoverlapping or overlapping. To illustrate this concept, we use a sequence of Iskander's eye aberration data measured by a Hartmann-Shack sensor [23]. The data is sampled at approximately 11 Hz and is composed of 128 data points. We divide the sequence into nonoverlapping blocks of 16 samples each, as illustrated in the top panel of Figure 6. The blocks are then resampled to obtain the bottom panel of Figure 6.

THE BOOTSTRAP HAS THE POWER TO SUBSTITUTE TEDIOUS AND OFTEN IMPOSSIBLE ANALYTICAL DERIVATIONS WITH COMPUTATIONAL CALCULATIONS.

We note that the resampling of blocks of data is based on the assumption that the blocks are i.i.d. This is the case when the data represents a process that is strong mixing. This means, loosely speaking, that the resampling scheme assumes that the data points that are far apart are nearly independent.

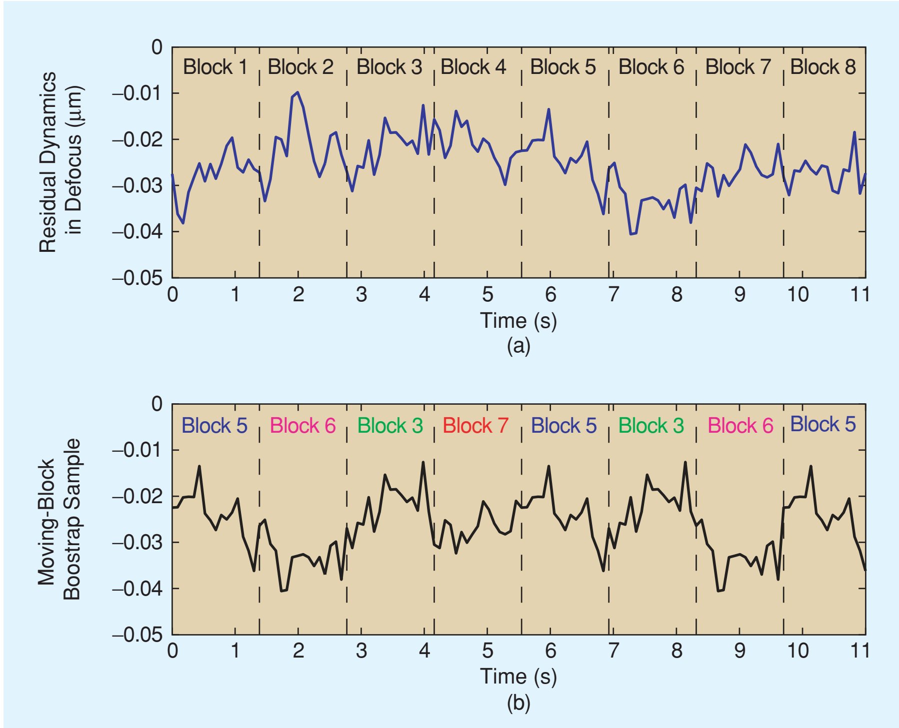
If the data are to be divided into segments, the length of each segment as well as the amount of overlap may become an issue. Although automatic procedures for selecting these parameters have been developed [24], in many practical situations, the dependence structure of the sample may still need to be estimated or at least examined. The problem may become even more complicated if the original data is nonstationary. There are reported cases where moving block bootstrap techniques show a certain degree of robustness for nonstationary data (see the chapter by Lahiri in [8]). On the other hand, it is not guaranteed that the moving block bootstrap estimates from a stationary process would themselves result in stationary processes. This somewhat



[FIG5] Results of the speech signal analysis experiment. (a) The original signal $s(t)$ corresponding to the vocal part of "Reis" and its noise-reduced version $\hat{s}(t)$. (b) The estimated residuals and (c) their covariance structure. (d) The Estimated coefficients of the AR(11) model and 90% confidence intervals.

worrying observation of nonpreservation of stationarity has been reported in [25]. A resampling scheme in which the length of each block is randomly chosen (the so-called stationary bootstrap) provides a solution to this problem [26]. See the paper by Politis [27] for some recent dependent data bootstrap techniques.

The observations made above show that the very appealing simplicity of the standard bootstrap resampling technique is somehow lost in the dependent data bootstrap methods. Also, the amount of evidence supporting the empirical validity of those procedures is still limited. This leads to an unpopular conclusion that the bootstrap novice should attempt a model-based approach when dealing with non-i.i.d. data, especially when only limited knowledge of the data dependence structure is available. Nevertheless, there are practical cases in which a model-based approach combined with dependent data bootstrap is powerful.



[FIG6] An example of the principle of moving block bootstrap. (a) Original data and (b) block bootstrapped data. Note that some blocks from the original data appear more than once and some do not appear at all in the bootstrapped data.

We now close our dependent data bootstrap treatment with an example from radar.

MICRO-DOPPLER ANALYSIS

The Doppler phenomenon often arises in engineering applications where radar, ladar, sonar, and ultra-sound measurements are made. This may be due to the relative motion of an object with respect to the measurement system. If the motion is harmonic, for example due to vibration or rotation, the resulting signal can be well modeled by a frequency modulated (FM) signal [28]. Estimation of the FM parameters may allow us to determine physical properties such as the angular velocity and displacement of the vibrational/rotational motion which can in turn be used for classification. The objectives are to estimate the micro-Doppler parameters along with a measure of accuracy, such as confidence intervals.

Assume the following amplitude modulation (AM)-FM signal model:

$$s(t) = a(t)\exp\{j\varphi(t)\}, \qquad (1)$$

where the AM is described by a polynomial: $a(t;\boldsymbol{\alpha}) = \sum_{k=0}^{q} \alpha_k t^k$ and $\boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_q)$ are the termed AM parameters. The phase modulation for a micro-Doppler signal is described by a sinusoidal function: $\varphi(t) = -D/\omega_m \cos(\omega_m t + \phi)$.

The instantaneous angular frequency (IF) of the signals is defined by

$$\omega(t;\boldsymbol{\beta}) \triangleq \frac{d\varphi(t)}{dt} = D\sin(\omega_m t + \phi), \qquad (2)$$

where $\boldsymbol{\beta} = (D, \omega_m, \phi)$ are termed the FM or micro-Doppler parameters.

The micro-Doppler signal in (1) is buried in additive noise so that the observation process is described by $X(t) = s(t) + V(t)$, where $V(t)$ is assumed to be a colored noise process. Given observations $\{x(k)\}_{k=1}^{n}$ of $X(t)$, the goal is to estimate the micro-Doppler parameters in $\boldsymbol{\beta}$ as well as their confidence intervals.

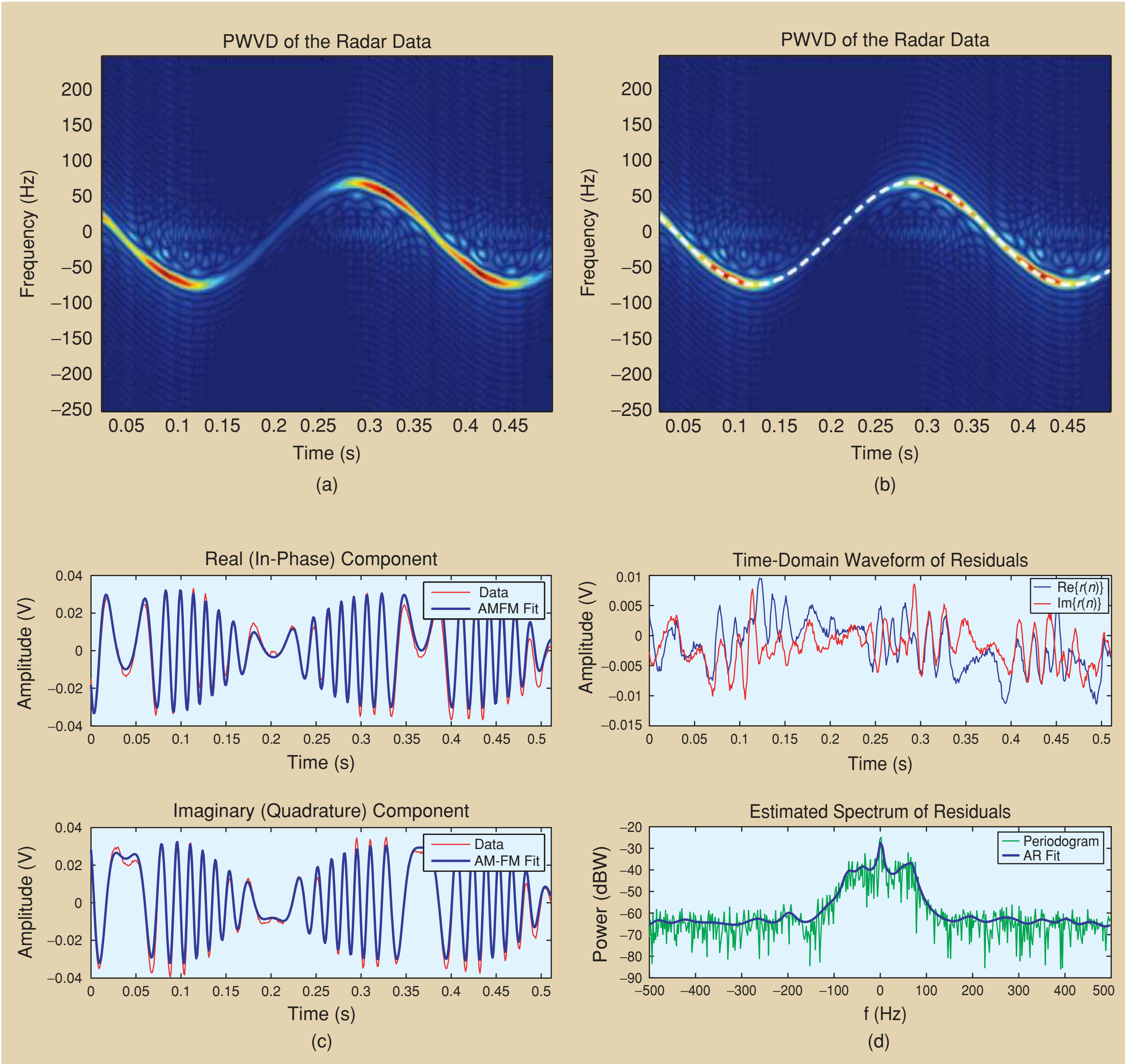
The estimation of the phase parameters is performed using a time-frequency Hough transform (TFHT) [29], [30]. The TFHT we use is given by

$$H(\boldsymbol{\beta}) = \sum_{k=-(L-1)/2}^{n-(L-1)/2+1} P_{xx}[k, \omega_i(n;\boldsymbol{\beta})),$$

where $\omega(t,\boldsymbol{\beta})$ is described in (2), and $P_{xx}[k, \omega_i(n;\boldsymbol{\beta}))$ is the pseudo-Wigner-Ville (PWVD) distribution, defined as

$$P_{xx}[k,\omega) = \sum_{l=-(L-1)/2}^{(L-1)/2} h[k]x[k+l]x^*[k-l]e^{-j2\omega}, \qquad (3)$$

for $k = -(L-1)/2, \ldots, n-(L-1)/2$, where $h[k]$ is a windowing function of duration L. An estimate of $\boldsymbol{\beta}$ is obtained from the location of the largest peak of $H(\boldsymbol{\beta})$, i.e., $\hat{\boldsymbol{\beta}} = \arg\max_{\boldsymbol{\beta}} H(\boldsymbol{\beta})$. Once the phase parameters have been

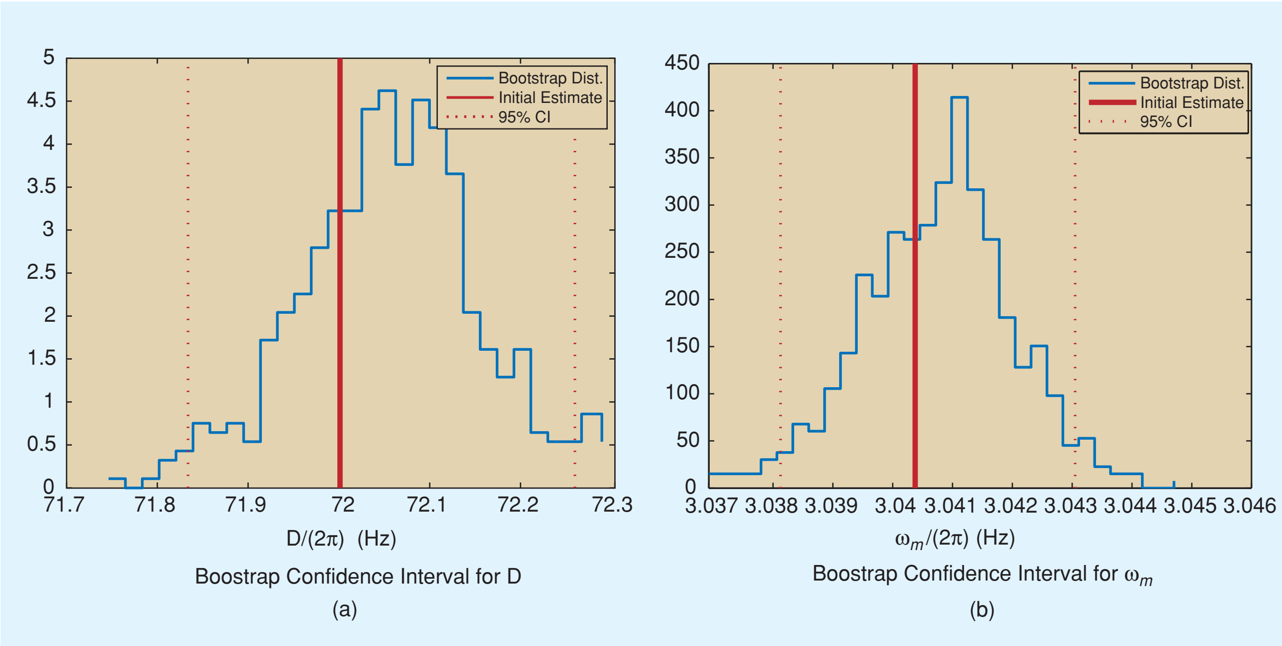


[FIG7] (a) The PWVD of the radar data. (b) The PWVD of the radar data and the micro-Doppler signature estimated using the TFHT. (c) The real and imaginary components of the radar signal with their estimated counterparts. (d) The real and imaginary parts of the residuals and their spectral estimates.

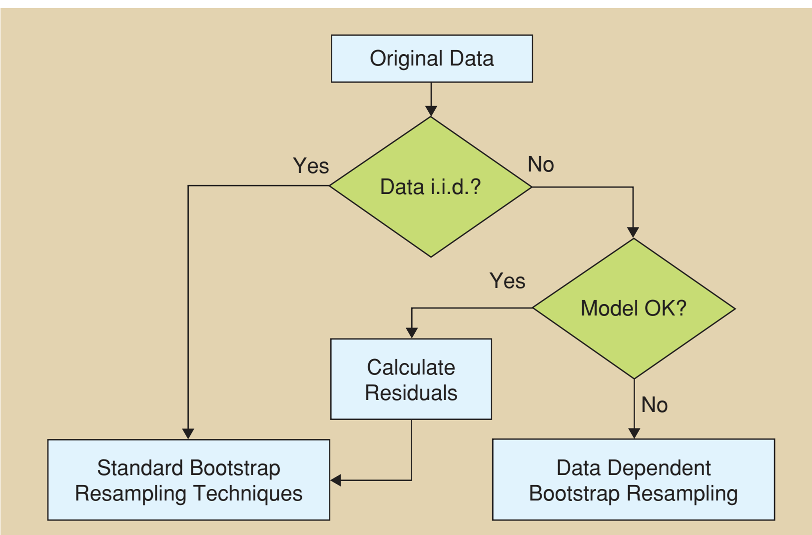
estimated, the phase term is demodulated and the amplitude parameters $\alpha_0, \ldots, \alpha_q$ are estimated via linear least-squares.

We now turn our attention to the estimation of confidence intervals for D and ω_m using the bootstrap. Given estimates for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, the residuals are obtained by subtracting the estimated signal from the observations. The resulting residuals are not i.i.d., and a dependent data bootstrap would seem a natural choice. Due to some difficulties with a dependent data bootstrap approach with real data, we chose to whiten the residuals by estimating parameters of a fitted AR model. The innovations are then resampled, filtered, and added to the estimated signal term to obtain bootstrap versions of the data, as discussed previously. By reestimating the parameters many times from the bootstrap data, we are then able to obtain confidence intervals for the parameters of interest. This is demonstrated using experimental data.

The results shown here are based on an experimental radar system, operating at carrier frequency $f_c = 919.82$ MHz. After demodulation, the in-phase and quadrature baseband channels are sampled at $f_s = 1$ kHz. The radar system is directed towards a spherical object, swinging with a pendulum motion, which results in a typical micro-Doppler signature. The PWVD of the observations is computed according to (3) and shown in Figure 7(a). The sinusoidal frequency modulation is clearly observed.



[FIG8] The bootstrap distributions and 95% confidence intervals for the FM parameters (a) D and (b) ω_m.



[FIG9] A practical strategy for bootstrapping data.

Using the TFHT, we estimate the micro-Doppler signature as discussed above and plot it over the PWVD in Figure 7(b). The AM parameters of the signal are then estimated. The radar data and the estimated AM-FM signal term are shown in Figure 7(c), while the residuals obtained by subtracting the estimated signal from the data are shown in Figure 7(d) together with their periodogram and AR-based spectral estimates. The model appears to fit the data well, and coloration of the noise seems to be well approximated using an AR model.

After applying the bootstrap with $B = 500$, the estimated distribution of the micro-Doppler parameters and the 95% confidence intervals for D and ω_m are obtained and shown in Figure 8.

This example shows that the bootstrap is a solution to finding distribution estimates for $\hat{D}$ and $\hat{\omega}_m$, a task that would be tedious or even impossible to achieve analytically.

GUIDELINES FOR USING THE BOOTSTRAP

Let us summarize the main points from our discussion. Is it really possible to use the bootstrap to extricate oneself from a difficult situation as anecdotally Baron von Münchhausen did? There are many dictionary definitions of the word *bootstrap*. The one we would like to bring to the readers' attention is: "to change the state using existing resources." With this definition, the answer to our question is affirmative. Yes, it is possible to change our state of knowledge (e.g., the knowledge of the distribution of parameter estimators) based on what we have at hand, usually a single observation of the process.

However, for the bootstrap to be successful, we need to identify which resampling scheme is most appropriate. The initial decision must be based on the examination of the data and the problem at hand. If the data can be assumed to be i.i.d. (the unlikely scenario in real world problems, but useful in simulation studies), standard bootstrap resampling techniques such as the independent data bootstrap can be used. Should the data be non-i.i.d., we should consider first a parametric approach in which a specific structure is assumed (see Figure 9). If this can be done, we can reduce the seemingly difficult problem of dependent data bootstrap to standard resampling of the assumed i.i.d. model error estimates (residuals). If a model for the data structure cannot be found, the matter is much more delicate. This is because existing nonparametric bootstrap schemes for dealing with dependent data have not been sufficiently validated as an automatic approach for real-life dependent data problems that a signal processing engineer may encounter.

As we mentioned earlier, a signal processing practitioner always attempts to simplify their work. A nonlinear problem can be either transformed or reduced to a set of linear problems, and a nonstationary signal can be segmented to assume local stationarity. Similarly, bootstrap techniques for real-world data that are often not i.i.d. and nonstationary can be reduced to standard resampling techniques. This is the approach we have taken and the one we advocate. With this approach, the bootstrap may prove itself as an off-the-shelf tool for practical signal processing problems.

[THE BOOTSTRAP IS A SOLUTION TO FINDING DISTRIBUTION ESTIMATES FOR THE DOPPLER PARAMETERS ($\hat{D}$ AND $\hat{\omega}_m$), A TASK THAT WOULD BE TEDIOUS OR EVEN IMPOSSIBLE TO ACHIEVE ANALYTICALLY.]

ACKNOWLEDGMENTS

We wish to thank Philipp Heidenreich and Alex Kaps for their assistance with the speech example and Luke Cirillo for his contribution to the micro-Doppler data example. The micro-Doppler data was kindly provided by Prof. M.G. Amin from Villanova University, Pennsylvania, who also has our gratitude. We are grateful for helpful comments from Prof. D.N. Politis. Abdelhak Zoubir thanks his colleague Eberhard Hänsler for the fruitful discussions.

AUTHORS

Abdelhak M. Zoubir (zoubir@ieee.org) received his Dr.-Ing. from Ruhr-Universität Bochum, Germany. He was with Queensland University of Technology, Australia from 1992 until 1999, when he moved to Curtin University of Technology, Australia as Professor of Telecommunications. In 2003, he joined Technische Universität Darmstadt, Germany as professor and head of the Signal Processing Group. His research interest lies in statistical methods for signal processing with emphasis on bootstrap techniques, robust detection and estimation applied to telecommunications, car engine monitoring, and biomedicine. He authored or coauthored more than 200 papers and the book *Bootstrap Techniques for Signal Processing*, Cambridge University Press, 2004.

D. Robert Iskander (d.iskander@qut.edu.au) received his Ph.D. degree in signal processing from Queensland University of Technology (QUT), Australia, in 1997. In 1996–2000 he was a research fellow at Signal Processing Research Centre, Cooperative Research Centre for Satellite Systems, and Centre for Eye Research, QUT, Australia. In 2001, he joined the School of Engineering, Griffith University as a senior lecturer. From 2003, he returned to QUT as a Principal Research Fellow. His current research interests include statistical signal processing, biomedical engineering, and visual optics. He has authored or coauthored more than 100 papers and is a coinventor on seven international patents.

REFERENCES

[1] B. Efron, "Bootstrap methods: Another look at the jackknife," *Ann. Statist.*, vol. 7, no. 1, pp. 1–26, 1979.

[2] B. Efron and R.J. Tibshirani, *An Introduction to the Bootstrap*. London, U.K.: Chapman & Hall, 1993.

[3] A.M. Zoubir and D.R. Iskander, *Bootstrap Techniques for Signal Processing*. Cambridge, U.K: Cambridge Univ. Press, 2004.

[4] D.N. Politis, "Computer-intensive methods in statistical analysis," *IEEE Signal Processing Mag.*, vol. 15, no. 1, pp. 39–55, 1998.

[5] A.M. Zoubir and B. Boashash, "The bootstrap and its application in signal processing," *IEEE Signal Processing Mag.*, vol. 15, no. 1, pp. 56–76, 1998.

[6] L. Breiman, "The little bootstrap and other methods for dimensionality selection in regression," *J. Amer. Stat. Assoc.*, vol. 87, no. 419, pp. 738–754, 1992.

[7] R.Y. Liu, "Bootstrap procedure under some non-i.i.d. models," *Ann. Statist.*, vol. 16, no. 4, pp. 1696–1708, 1988.

[8] R. LePage and L. Billard, Eds. *Exploring the Limits of Bootstrap*. New York: Wiley, 1992.

[9] J.G. Booth and P. Hall, "Monte Carlo approximation and the iterated bootstrap," *Biometrika*, vol. 81, no. 2, pp. 331–340, 1994.

[10] A.M. Zoubir, *The bootstrap: A powerful tool for statistical signal processing with small sample sets, ICASSP-99 Tutorial* [Online]. Available: http://www.spg.tu-darmstadt.de/downloads/bootstrap tutorial.html

[11] G.A. Young, "Bootstrap: More than a stab in the dark?," *Stat. Sci.*, vol. 9, no. 3, pp. 382–415, 1994.

[12] D.W.K. Andrews and M. Buchinsky, "A three-step method for choosing the number of bootstrap repetitions," *Econometrica*, vol. 68, no. 1, pp. 23–51, 2000.

[13] A.C. Davison and D.V. Hinkley, *Bootstrap Methods and their Applications*. Cambridge, U.K.: Cambridge Univ. Press, , 1997.

[14] B. Efron, "Jackknife-after-bootstrap standard errors and in uence functions," *J. Royal Statistical Soc., B*, vol. 54, no. 1, pp. 83–127, 1992.

[15] E. Mammen, *When Does Bootstrap Work? Asymptotic Results and Simulations* (Lecture Notes in Statistics, vol. 77). New York: Springer-Verlag, 1992.

[16] D.N. Politis, J.P. Romano, and M. Wolf, *Subsampling*. New York: Springer-Verlag, 1999.

[17] D.R. Iskander, M.J. Collins, S. Mioschek, and M. Trunk, "Automatic pupillometry from digital images," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 9, pp. 1619–1627, 2004.

[18] J. Cabrera and P. Meer, "Unbiased estimation of ellipses by bootstrapping," *IEEE Trans Pattern Anal. Machine Intell.*, vol. 18, no. 7, pp. 752–756, 1996.

[19] J. Shao and D. Tu, *The Jackknife and Bootstrap*. New York: Springer-Verlag, 1995.

[20] J. Franke and W. Härdle, "On bootstrapping kernel spectral estimates," *Ann. Statist.*, vol. 20, no. 1, pp. 121–145, 1992.

[21] H. Puder, "Noise reduction with Kalman-filters for hands-free car phones based on parametric spectral speech and noise estimates," in *Topics in Acoustic Echo and Noise Control,* E. Hänsler and G. Schmidt, Eds. New York: Springer-Verlag, 2006.

[22] H.R. Künsch, "The jackknife and the bootstrap for general stationary observations," *Ann. Statist*., vol. 17, no. 3, pp. 1217–1241, 1989.

[23] D.R. Iskander, M.J. Collins, M.R. Morelande, and M. Zhu, "Analyzing the dynamic wavefront aberrations in human eye," *IEEE Trans. Biomed. Eng*., vol. 51, no. 11, pp. 1969–1980, 2004.

[24] D.N. Politis and H. White, "Automatic block-length selection for the dependent bootstrap," *Econometric Rev.*, vol. 23, no, 1, pp. 53–70, 2004.

[25] C. Léger, D.N. Politis, and J.P. Romano, "Bootstrap technology and applications," *Technometrics*, vol. 34, no. 4, pp. 378–398, 1992.

[26] D.N. Politis and J.P. Romano, "The stationary bootstrap," *J. Amer. Stat. Assoc.*, vol. 89, no. 428, pp. 1303–1313, 1994.

[27] D.N. Politis, "The impact of bootstrap methods on time series analysis," *Statistical Sci.*, vol. 18, no. 2, pp. 219–230, 2003.

[28] S.-R. Huang, R. Lerner, and K. Parker, "On estimating the amplitude of harmonic vibration from the Doppler spectrum of reflected signals," *J. Acoust. Soc. Amer*., vol. 88, no. 6, pp. 2702–2712, 1990.

[29] S. Barbarossa and O. Lemoine, "Analysis of nonlinear FM signals by pattern recog-nition of their time-frequency representation," *IEEE Signal Processing Lett*., vol. 3, no. 4, pp. 112–115, 1996.

[30] L. Cirillo, A. Zoubir, and M. Amin, "Estimation of FM parameters using a time-frequency Hough transform," in *Proc. ICASSP,* Toulouse, France, May 2006, vol. 3, pp. 169–172.

SP