

Information-theoretic Bounds of Resampling Forensics: New Evidence for Traces Beyond Cyclostationarity

Cecilia Pasquini

Universität Innsbruck, Austria
University of Münster, Germany
pasquini@uni-muenster.de

Rainer Böhme

Universität Innsbruck, Austria
University of Münster, Germany
rainer.boehme@uibk.ac.at

ABSTRACT

Although several methods have been proposed for the detection of resampling operations in multimedia signals and the estimation of the resampling factor, the fundamental limits for this forensic task leave open research questions. In this work, we explore the effects that a downsampling operation introduces in the statistics of a 1D signal as a function of the parameters used. We quantify the statistical distance between an original signal and its downsampled version by means of the Kullback-Leibler Divergence (KLD) in case of a wide-sense stationary 1st-order autoregressive signal model. Values of the KLD are derived for different signal parameters, resampling factors and interpolation kernels, thus predicting the achievable hypothesis distinguishability in each case. Our analysis reveals unexpected detectability in case of strong downsampling due to the local correlation structure of the original signal. Moreover, since existing detection methods generally leverage the cyclostationarity of resampled signals, we also address the case where the autocovariance values are estimated directly by means of the sample autocovariance from the signal under investigation. Under the considered assumptions, the Wishart distribution models the sample covariance matrix of a signal segment and the KLD under different hypotheses is derived.

KEYWORDS

Signal resampling, resampling forensics, Kullback-Leibler divergence, hypothesis distinguishability.

ACM Reference format:

Cecilia Pasquini and Rainer Böhme. 2017. Information-theoretic Bounds of Resampling Forensics: New Evidence for Traces Beyond Cyclostationarity. In *Proceedings of IH&MMSec '17, June 20–22, 2017, Philadelphia, PA, USA*, 12 pages. DOI: <http://dx.doi.org/10.1145/3082031.3083233>

1 INTRODUCTION

The detection of resampling operations in 1D and 2D signals is of great interest in multimedia forensics, as it can indicate that the object under investigation has been resized or subject to other

geometric transformations (e.g., rotation in case of images). Multiple methods have been proposed for this task, relying on different rationales.

When a signal is resampled, its samples are combined by means of an interpolation kernel to obtain new signal values located in a different lattice with respect to the original ones. Commonly used kernels periodically employ the same interpolation coefficients, thus introducing into the signal periodic linear correlations. This effect is exploited by most of existing techniques, which seek for periodicities in the signal itself [22] or in a linear predictor residual [14–16, 18, 23]. In doing so, a frequency analysis is generally employed for the detection, which can also provide an (unavoidably ambiguous [14]) estimate of the resampling factor. Moreover, researchers have recently tackled the problem with different approaches, e.g., by relying on set-membership theory [24], the use of SVD decomposition [21], or the measurement of the normalized energy density [9].

In parallel, the community in multimedia forensics has started to study forensic tasks from a more fundamental perspective, with the goal of assessing the performance limits of detection techniques [1, 4, 5]. In an hypothesis testing framework, this can be interpreted as quantifying the statistical distance between hypotheses and determine under which conditions they are actually distinguishable. In [5], the Kullback-Leibler Divergence (KLD) between the distributions of the signal under investigation (or a feature representation of it) and different processing hypotheses is proposed as a distinguishability measure, and a number of practical scenarios are addressed [6]. In these works, the authors are particularly interested in the distinguishability of operator chains and they consider different combinations of quantization, noise addition, Gamma correction and second-order finite impulse response (FIR) filtering operations.

However, to the best of our knowledge, such perspective has not been adopted yet to the case of resampling detection, for which performance limits are for now empirically established by state-of-the-art techniques. In this work, we start addressing this gap by quantifying how much the statistics of a 1D downsampled signal deviates from a not downsampled one. Under certain assumptions, commonly adopted in the literature, the distribution of the signal under the hypothesis of no downsampling and the hypothesis of downsampling with a certain factor and interpolation kernel is derived. We show that it is possible to compute the KLD in each case as a function of the original signal parameters, revealing the key role of the local correlation among samples. Effects of prefiltering (i.e., the use of a linear predictor) are also studied by means of this measure. Moreover, given the role of the signal's second-order moments (variance and autocovariance) in detection techniques,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

IH&MMSec '17, June 20–22, 2017, Philadelphia, PA, USA

© 2017 ACM. 978-1-4503-5061-7/17/06...\$15.00

DOI: <http://dx.doi.org/10.1145/3082031.3083233>

the KLD analysis is extended to the distribution of standard covariance matrix estimators as a function of the number of independent observations available.

It is to be noted that this work does not propose practical detection algorithms but aims to the assessment of the theoretical difficulties encountered. At the same time, it identifies space for improvement to be filled with not yet invented detectors.

The paper is structured as follows: in Section 2 we illustrate the perspectives adopted in our work and their relationship to existing studies; in Section 3, we formulate the resampling operation and recall the periodicity properties generally exploited by existing detectors; in Section 4, we formalize the assumptions on the original signal (in accordance with previous approaches in the literature), and analyze the effects of resampling in the signal statistics. These findings are then exploited in Section 5 to formulate different hypothesis tests and evaluate hypotheses distinguishability when varying the parameters involved. Finally, Section 6 concludes the paper.

2 PROPOSED APPROACH IN RELATION TO PRIOR WORK

We now describe and motivate the main novelties characterizing our study and how they are related to previous work:

- We directly study the statistics of a segment of the 1D signal under investigation under different hypotheses, although many state-of-the-art methods perform a frequency analysis [14, 18, 23]. This is due to the fact that deterministic processing (like DFT or other transformations) cannot increase the hypothesis distinguishability in terms of KLD [3, Lemma 1]. Thus, from an information-theoretic perspective, the best choice to measure the information contained in a signal segment is to study its very distribution.
- According to assumptions already adopted in the literature [14, 16, 23], we consider the original signal as a wide sense stationary autoregressive model of the 1st order (1-AR) with Gaussian innovations. In this setting, we show how the second-order moments

are transformed by the resampling operation. In particular cases, we analytically explore the role of the correlation coefficient of the 1-AR model in the statistical properties of the resampled signal.

- State-of-the-art techniques are generally evaluated by considering equally spaced values of the resampling factor in a certain range. However, it has been shown that the effects of resampling depend on the representation of the resampling factor as ratio of coprime natural numbers, where the numerator determines the period of the periodic structures introduced. For this reason, we employ sequences of rational numbers with the same numerator in our tests, thus differentiating resampled signals with the same periodicities. Such resampling factors are used, together with the chosen interpolation kernel, to define hypothesis tests and evaluate their distinguishability in terms of KLD.
- In addition to the distribution of the signal segment, we also study the distribution of the related prediction error computed in the same location and the distribution of a standard covariance matrix estimator obtained from a number of independent observations. Hypothesis distinguishability is also evaluated in these cases.
- In our analysis, we consider 1D real-valued signals, thus not accounting for quantization effects. Future work will be devoted to deal with quantization effects in the signal under investigation, as it is explored in [20, 24].

3 PRELIMINARIES

Let $s : \mathbb{R} \rightarrow \mathbb{R}$ represent a 1D real-valued signal. We suppose (without loss of generality) that the signal is originally sampled at integer coordinates, i.e., before resampling the values $s(n)$ with $n \in \mathbb{Z}$ are available. In this case, a resampling operation with factor ξ can be seen as a map from the integer lattice \mathbb{Z} to the lattice $\xi^{-1}\mathbb{Z}$, where $\xi < 1$ corresponds to downsampling and $\xi > 1$ to upsampling. In fact, every value $s(n\xi^{-1})$ is obtained starting from the original samples $s(n)$ as

$$s(n\xi^{-1}) = \sum_{n' \in \mathbb{Z}} h(n\xi^{-1} - n')s(n') \quad (1)$$

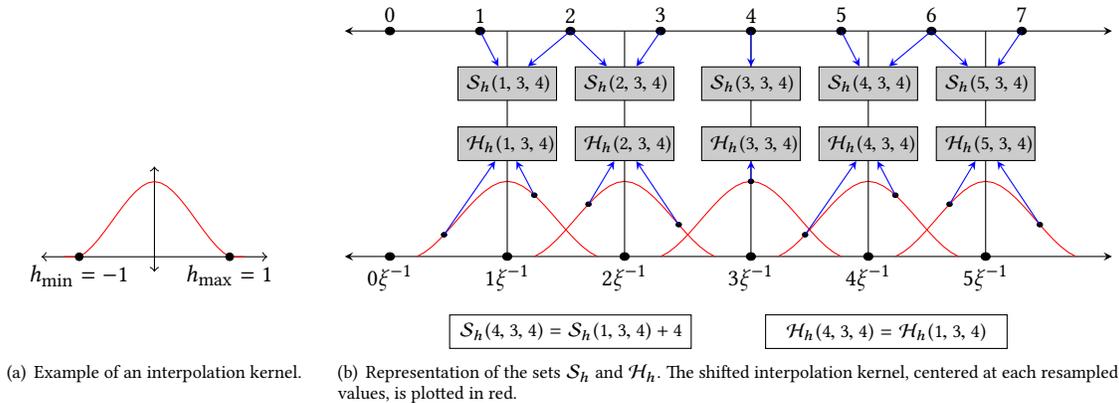


Figure 1: Example of resampling operation with $\xi = \frac{p}{q} = \frac{3}{4}$ and an interpolation kernel with support in $] - 1, 1[$.

where $h : \mathbb{R} \rightarrow \mathbb{R}$ is an interpolation kernel and the values $h(n\xi^{-1} - n')$ are the interpolation coefficients.

Starting from this definition, in the following section we derive the periodicity properties commonly exploited by existing detectors, and then express the resampling operation in matrix form.

3.1 Periodicity properties

Common interpolation kernels have finite support $]h_{\min}, h_{\max}[$, thus the sum in (1) is also finite. Moreover, the resampling factor ξ is in practice expressed in rational form as $\xi = \frac{p}{q}$ where p and q are coprime, and we will equivalently use the notation ξ or p/q in the following. This allows us to derive properties from (1):

- The resampled value $s((n+p)\xi^{-1})$ is obtained by combining samples whose indices are shifted by q with respect to the ones used to calculate $s(n\xi^{-1})$. In fact, we can obtain the set of values n' that are actually employed in the sum in (5) as a function of n, p and q as follows:

$$\mathcal{S}_h(n, p, q) = \left\{ n' \in \mathbb{Z} \left| n \frac{q}{p} - n' \in]h_{\min}, h_{\max}[\right. \right\} \quad (2)$$

$$= \left\{ \left\lfloor n \frac{q}{p} - h_{\max} \right\rfloor, \left\lfloor n \frac{q}{p} - h_{\max} \right\rfloor + 1, \dots, \left\lfloor n \frac{q}{p} - h_{\min} \right\rfloor \right\}.$$

It is easy to show that $\mathcal{S}_h(n+p, p, q)$ contains the very same elements of $\mathcal{S}_h(n, p, q)$ incremented by q .

- The interpolation coefficients used in the resampling operation are periodic with period p . From (1) the interpolation coefficients used for $s(n\xi^{-1})$ are given by

$$\mathcal{H}_h(n, p, q) = \left\{ h \left(n \frac{q}{p} - n' \right) \left| n' \in \mathcal{S}_h(n, p, q) \right. \right\} \quad (3)$$

so that, for the previous point,

$$\begin{aligned} \mathcal{H}_h(n+p, p, q) &= \left\{ h \left((n+p) \frac{q}{p} - n' \right) \left| n' \in \mathcal{S}_h(n+p, p, q) \right. \right\} \quad (4) \\ &= \left\{ h \left(n \frac{q}{p} + q - (n' + q) \right) \left| n' \in \mathcal{S}_h(n, p, q) \right. \right\} \\ &= \left\{ h \left(n \frac{q}{p} - n' \right) \left| n' \in \mathcal{S}_h(n, p, q) \right. \right\} \equiv \mathcal{H}_h(n, p, q). \end{aligned}$$

An example in case of $\xi = n/p = 3/4$ and an interpolation kernel with support in $] -1, 1[$ is reported in Fig. 1.

As it was first observed in [18] for the application in image forensics, this leads to periodic linear correlations in the resampled signal, which motivates the use of a linear predictor as a detection approach [14, 15, 18]. Given a set of prediction weights $\beta_{-T}, \dots, \beta_T, \beta_0 = 1$, the prediction error

$$e(n\xi^{-1}) = \sum_{\substack{|t| \leq T \\ t \in \mathbb{Z}}} \beta_t s(n\xi^{-1} + t\xi^{-1}) \quad (5)$$

is used as a measure of the linear correlation between the interpolated value $s(n\xi^{-1})$ and its $2T$ closest neighbors. A frequency analysis is then generally performed on $e(\cdot)$ to identify periodicities due to resampling operations.

3.2 Matrix formulation

As noted in [14], both (1) and (5) are linear combinations of the original samples $s(n)$ with coefficients that depend on ξ, h and the

β_t . Thus, it is convenient to consider the original signal as a finite vector and express the resampling operation and the prediction error computation in matrix form. Without loss of generality, we can define the vector $\mathbf{s} = [s_1, \dots, s_N]$ where $s_n \doteq s(n)$. Fixing the interpolation kernel with finite support $h(\cdot)$ and the resampling factor $\xi = p/q$, the vector \mathbf{s} can be linearly transformed according to (1) to obtain a vector $\mathbf{r} = [r_1, \dots, r_M]$ containing the resampled signal. The component r_1 is the first interpolated value that can be obtained from the samples $s(n), n \geq 1$, and r_M is the last interpolated value that can be obtained from the samples $s(n), n \leq N$. The correspondance between the indices of \mathbf{r} and the multiples of ξ^{-1} in (1) depends on both the support of h and the value of ξ (for instance, in Fig. 1 we have that $r_1 = s(1\xi^{-1})$, but in case of $\xi > 1$ also $s(0)$ would be necessary to compute $s(1\xi^{-1})$, so $r_1 = s(2\xi^{-1})$), but this does not compromise the generality of our analysis.

The resampling operation can then be written as

$$\mathbf{r} = C\mathbf{s}, \quad (6)$$

where each n -th row of the matrix C contains the values of $\mathcal{H}_h(n, p, q)$ at the locations contained in $\mathcal{S}_h(n, p, q)$. Thus, its entries depend on both ξ and h , so $C = C(\xi, h)$, and its size $M \times N$ is such that $M < N$ in case of downsampling and $M > N$ in case of upsampling, as represented in Fig. 2.

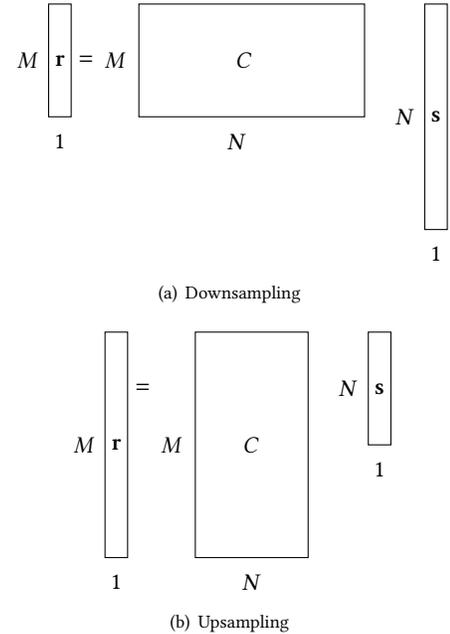


Figure 2: Resampling matrix transformation.

Because of the previous considerations, the nonzero entries in the rows of C will be the same every p rows, but shifted by q positions to the right. For instance, for a linear interpolation kernel h and

$\xi = p/q = 3/4$, we have

$$C\left(\frac{3}{4}, h\right) = \begin{bmatrix} 2/3 & 1/3 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1/3 & 2/3 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 2/3 & 1/3 & \dots \\ 0 & 0 & 0 & 0 & 0 & 1/3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (7)$$

Moreover, the computation of the prediction error in (5) can be seen as a discrete convolution of the resampled signal with vector $[\beta_{-T}, \dots, \beta_T]$. In fact, by removing the condition that $\beta_0 = 1$, the prediction error is essentially a linearly filtered version of the resampled signal and in the literature such operation is also referred to as prefiltering [23]. Then, the elements of \mathbf{r} are transformed as

$$\mathbf{e} = B\mathbf{r}, \quad (8)$$

where B is a $(M - 2T) \times M$ matrix such that

$$B = \begin{bmatrix} \beta_{-T} & \dots & \dots & \dots & \beta_T & 0 & \dots & \dots & \dots & \dots \\ 0 & \beta_{-T} & \dots & \dots & \dots & \beta_T & 0 & \dots & \dots & \dots \\ 0 & 0 & \beta_{-T} & \dots & \dots & \dots & \beta_T & 0 & \dots & \dots \\ \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \beta_{-T} & \dots & \dots & \dots & \beta_T \end{bmatrix} \quad (9)$$

and \mathbf{e} contains $M - 2T$ samples of the prediction error corresponding to the central indices of \mathbf{r} .

Finally, the whole process can be summarized as

$$\mathbf{e} = A\mathbf{s}, \quad (10)$$

where $A = BC$ and $\mathbf{e} \equiv \mathbf{r}$ if no prefilter is applied.

4 ANALYSIS OF SECOND-ORDER MOMENTS AND ASSUMPTIONS ON THE SIGNAL

Several approaches rely on the assumption that the original signal is a wide sense stationary (WSS) process [14, 16, 23], i.e., the expected value $E\{s(k)\}$ and the autocovariance $E\{(s(k) - E\{s(k)\})(s(k+t) - E\{s(k+t)\})\}$, $\forall t \in \mathbb{Z}$, are constant over $k \in \mathbb{Z}$. This allows us to link the periodicity of the interpolation coefficients to the periodicity of second-order moments in the resulting resampled signal.

In fact, in [19] the authors show that a multirate system performing sampling rate conversion by a factor p/q on a WSS signal produces a cyclostationary signal with period $p/\gcd(p, q)$. With p and q being coprime in our case, this means that for the components of the vector \mathbf{r} we have

$$\text{Cov}\{r_k, r_{k+t}\} = \text{Cov}\{r_{k+jp}, r_{k+jp+t}\} \quad (11)$$

for every integers j and t .

Further results have been obtained by extending the definition of both (1) and (5) to the whole real line instead of the discrete lattice $\xi^{-1}\mathbb{Z}$. In [22, 23], it is proved that the interpolated signal $s(x)$, $x \in \mathbb{R}$, defined as in (1) is cyclostationary with period 1, i.e., the original sampling rate. The author of [14] shows that the variance of the prediction error $e(x)$, $x \in \mathbb{R}$, is also periodic with period 1 regardless of the prediction weights, thus establishing a link between methods based on the prediction error and the ones based on a prefiltering operation like discrete differentiation [7, 11, 16].

However, the vector \mathbf{r} contains only values of $s(\cdot)$ sampled at multiples of ξ^{-1} , so that the periodicity with 1 is not observable. In other words, by considering the indexing of \mathbf{r} , the periodicity would be non-integer with period ξ . In fact, the authors in [22, 23] denote this property as ‘‘almost cyclostationarity’’, referring to the fact that the available vector \mathbf{r} would have non-integer periodicity ξ , in contrast with the ‘‘pure cyclostationarity’’ with integer period p proved in [19].

It is common in the literature to assume statistical models for the distribution of the original signal. While a white Gaussian model would not be accurate, 1D autoregressive models of first-order (1-AR) with Gaussian innovations [12] have been used to capture local correlation in 1D and 2D signals [23]. Moreover, it is also commonly assumed to deal with zero-mean signals, by implicitly supposing that the mean value can be subtracted from the signal under investigation. We can formalize these properties as:

ASSUMPTION 1. *The original signal is a WSS 1-AR model with Gaussian innovations, i.e.:*

$$s_n = \rho s_{n-1} + \varepsilon_n \quad (12)$$

where ρ is a correlation coefficient satisfying $|\rho| < 1$ and ε_n is a zero-mean Gaussian random variable with variance σ_ε^2 such that $\varepsilon_n, \varepsilon_{n'}$ are independent $\forall n, n'$.

Under Assumption 1, the covariance between samples at distance $t \in \mathbb{Z}$ is given by

$$\text{Cov}\{s_n, s_{n+t}\} = \frac{\rho^{|t|}}{1 - \rho^2} \cdot \sigma_\varepsilon^2, \quad \forall n \in \mathbb{Z}, \quad (13)$$

and the covariance matrix of the vector \mathbf{s} has the form

$$\Sigma_s = \begin{bmatrix} \frac{1}{1-\rho^2} \sigma_\varepsilon^2 & \frac{\rho^1}{1-\rho^2} \sigma_\varepsilon^2 & \frac{\rho^2}{1-\rho^2} \sigma_\varepsilon^2 & \dots & \frac{\rho^{N-1}}{1-\rho^2} \sigma_\varepsilon^2 \\ \frac{\rho^1}{1-\rho^2} \sigma_\varepsilon^2 & \frac{1}{1-\rho^2} \sigma_\varepsilon^2 & \frac{\rho^1}{1-\rho^2} \sigma_\varepsilon^2 & \dots & \frac{\rho^{N-2}}{1-\rho^2} \sigma_\varepsilon^2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \frac{\rho^{N-2}}{1-\rho^2} \sigma_\varepsilon^2 & \vdots & \frac{\rho^1}{1-\rho^2} \sigma_\varepsilon^2 & \frac{1}{1-\rho^2} \sigma_\varepsilon^2 & \frac{\rho^1}{1-\rho^2} \sigma_\varepsilon^2 \\ \frac{\rho^{N-1}}{1-\rho^2} \sigma_\varepsilon^2 & \dots & \frac{\rho^2}{1-\rho^2} \sigma_\varepsilon^2 & \frac{\rho^1}{1-\rho^2} \sigma_\varepsilon^2 & \frac{1}{1-\rho^2} \sigma_\varepsilon^2 \end{bmatrix}. \quad (14)$$

Moreover, by iterating (12) we have that each s_n is a linear combination of independent Gaussian realizations, so that every subset of samples is a multivariate normal random variable. Thus, the covariance matrix Σ_s completely determines the distribution of the random vector \mathbf{s} and we can write

$$\mathbf{s} \sim \mathcal{N}_N(\mathbf{0}, \Sigma_s), \quad (15)$$

where $\mathcal{N}_N(\mu, \Sigma)$ indicates an N -dimensional normal distribution with mean vector μ and covariance matrix Σ , and $\mathbf{0}$ is a vector of N zeros.

By the definition of multivariate normal distributions, the transformed vector \mathbf{r} is also multivariate normal, so its distribution is again determined by its covariance matrix:

$$\mathbf{r} \sim \mathcal{N}_M(\mathbf{0}, \Sigma_r). \quad (16)$$

The entries of Σ_r can be obtained by computing

$$\text{Cov}\{r_k, r_{k'}\} = E\{r_k, r_{k'}\} \quad (17)$$

for two arbitrary indices $k, k' \leq M$. Considering that every r_k is the scalar product of the k -th row of C and \mathbf{s} , equation (17) can be expanded to obtain the following equality:

$$\Sigma_r = C\Sigma_s C^T. \quad (18)$$

However, in order for the multivariate normal to be non-degenerate, the covariance matrix must be positive definite. The resulting matrix $\Sigma_r = C\Sigma_s C^T$ preserves the positive definiteness of Σ_s if and only if the matrix C^T has rank equal to M [13, p. 431, Obs 7.1.8]. This can only hold in case of downsampling, while in case of upsampling C^T can have at most rank equal to $N < M$. This is crucial as in the degenerate case the multivariate normal vector does not have a density and, thus, the KLD is not defined. In our analysis we will focus on downsampling, while we leave for future work the problem of dealing with upsampling matrices.

5 HYPOTHESIS TESTS AND DISTINGUISHABILITY

A forensic analysis consists in deciding whether the signal under investigation has been resampled or not. In particular, the decision is usually made on a segment of the signal, that we can represent as a vector \mathbf{x} of M contiguous samples starting at an arbitrary position. Our previous analysis now allows us to obtain information on the binary hypothesis test

$$H_0: \mathbf{x} \text{ has not been downsampled.} \quad (19)$$

$$H_1: \mathbf{x} \text{ has been downsampled}$$

with factor ξ and interpolation kernel h .

As observed in [4], the decision between H_0 and H_1 on \mathbf{x} can be taken according to the distribution of \mathbf{x} itself or a related feature representation $F_{\mathbf{x}}$ defined in a certain space Ω under the two hypotheses. Then, we can define $P_0(\cdot), P_1(\cdot)$ as the distributions of $F_{\mathbf{x}}$ under H_0 and H_1 , respectively. Motivated by known results in information theory, the author of [5] then suggests the use of KLD as a measure hypotheses distinguishability, which in the continuous case is defined as

$$\text{KLD}(P_0, P_1) \doteq \int_{\Omega} P_0(F_{\mathbf{x}}) \log_2 \frac{P_0(F_{\mathbf{x}})}{P_1(F_{\mathbf{x}})} dF_{\mathbf{x}}, \quad (20)$$

where we consider the logarithm to the base 2, without loss of generality.

In the next subsections, we explore the behavior of the KLD for different representations $F_{\mathbf{x}}$ for which we derive P_0 and P_1 . In particular, we will first focus on the use of the signal itself (Section 5.1) or a prefiltered version of it (Section 5.2), and then we consider the case where the analyst estimates the second order moments of the signal from the observations available (Section 5.3). While the first two cases result in a multivariate normal distribution of the vector $F_{\mathbf{x}}$, we will see that in the third case $F_{\mathbf{x}}$ is a matrix following a Wishart distribution.

We will suppose that Assumption 1 holds, thus relying on the analysis in Section 4. We fix $M = 30$, i.e., we evaluate the ability of taking a forensic decision on a signal segment of 30 samples. Moreover, we have selected a fixed location of the samples in the signal under investigation in such a way that, in case downsampling occurred, no original samples with indices < 1 would be necessary to obtain the resulting downsampled signal or its pre-filtered version.

Clearly, the test (19) is determined by the chosen interpolation kernel h and the resampling factor $\xi < 1$. Thus, we consider three commonly used interpolation kernels (linear, cubic, Lanczos) represented in Fig. 4. Then, we vary the value of the resampling factor ξ by considering sequences of rational numbers with the same numerator p lying in the interval $[0, 1]$. In fact, we have seen that the resampling factor is in practice expressed as a rational number p/q , so we consider the sequences

$$\xi_{1:10}^a = \left\{ \frac{2^a}{q} \mid q = 2^a + 1, \dots, 2^a + 10 \wedge \text{gcd}(2^a, q) = 1 \right\}, \quad (21)$$

for $a = 0, 1, \dots, 8$, as represented in Figure 3. By doing so, we can separately observe resampled signals with the same pure cyclostationarity. In fact, only combinations of coprime numerator and denominator are included. There is no intersection among sequences for different values of a .

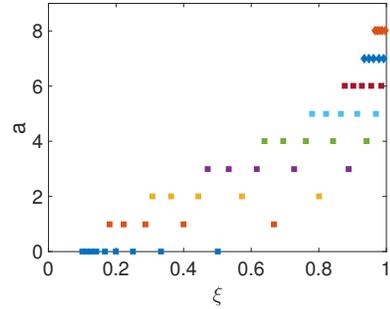


Figure 3: The sequences $\xi_{1:10}^a$ for $a = 0, \dots, 8$.

5.1 Signal distribution

We first analyze the case where the decision on \mathbf{x} is taken according to the multivariate distribution of \mathbf{x} itself, i.e. $F(\mathbf{x}) = \mathbf{x}$. We have already derived P_0 and P_1 in Section 4 and observed that in both cases it is a multivariate normal distribution, but with different covariance matrix. By denoting with Σ_0 and Σ_1 the covariance matrices under H_0 and H_1 , we have that Σ_0 is an $M \times M$ matrix defined as in (14), thus $\Sigma_0 = \Sigma_0(\sigma_\varepsilon, \rho)$. Depending on h and ξ , the hypothesis H_1 indicates that \mathbf{x} is the outcome of a resampling operation from a certain number N of original samples, so

$$\Sigma_1 = C\Sigma_0^{\text{ex}} C^T, \quad (22)$$

where C is the $M \times N$ resampling matrix and Σ_0^{ex} is also defined as in (14) but has size $N \times N$, thus $\Sigma_1 = \Sigma_1(\sigma_\varepsilon, \rho, \xi, h)$.

The test (19) can then be rewritten as

$$H_0: \mathbf{x} \sim \mathcal{N}_M(\mathbf{0}, \Sigma_0(\sigma_\varepsilon, \rho)). \quad (23)$$

$$H_1: \mathbf{x} \sim \mathcal{N}_M(\mathbf{0}, \Sigma_1(\sigma_\varepsilon, \rho, \xi, h)).$$

In case of multivariate Gaussian, the KLD has the following expression [17]:

$$\text{KLD}(\mathcal{N}_M(\mathbf{0}, \Sigma_0), \mathcal{N}_M(\mathbf{0}, \Sigma_1)) = \frac{1}{2} \left[\text{tr}(\Sigma_1^{-1} \Sigma_0) + \log \frac{\det(\Sigma_1)}{\det(\Sigma_0)} - M \right] \quad (24)$$

In order to have a consistent evaluation, we have fixed the location of the $M = 30$ samples contained in \mathbf{x} for every combination of

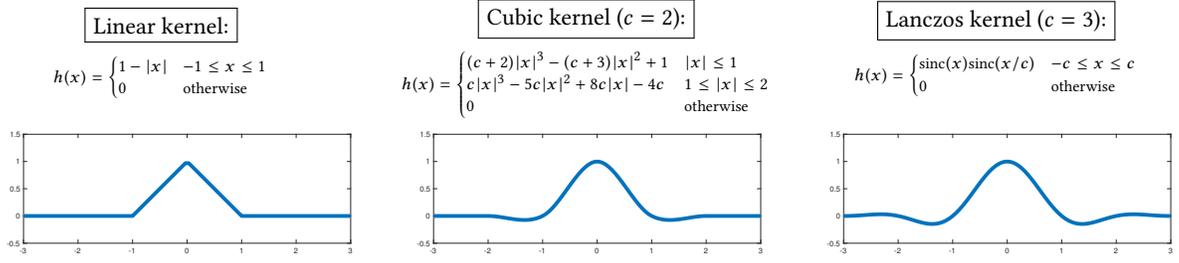


Figure 4: The three kernels used in our tests are reported.

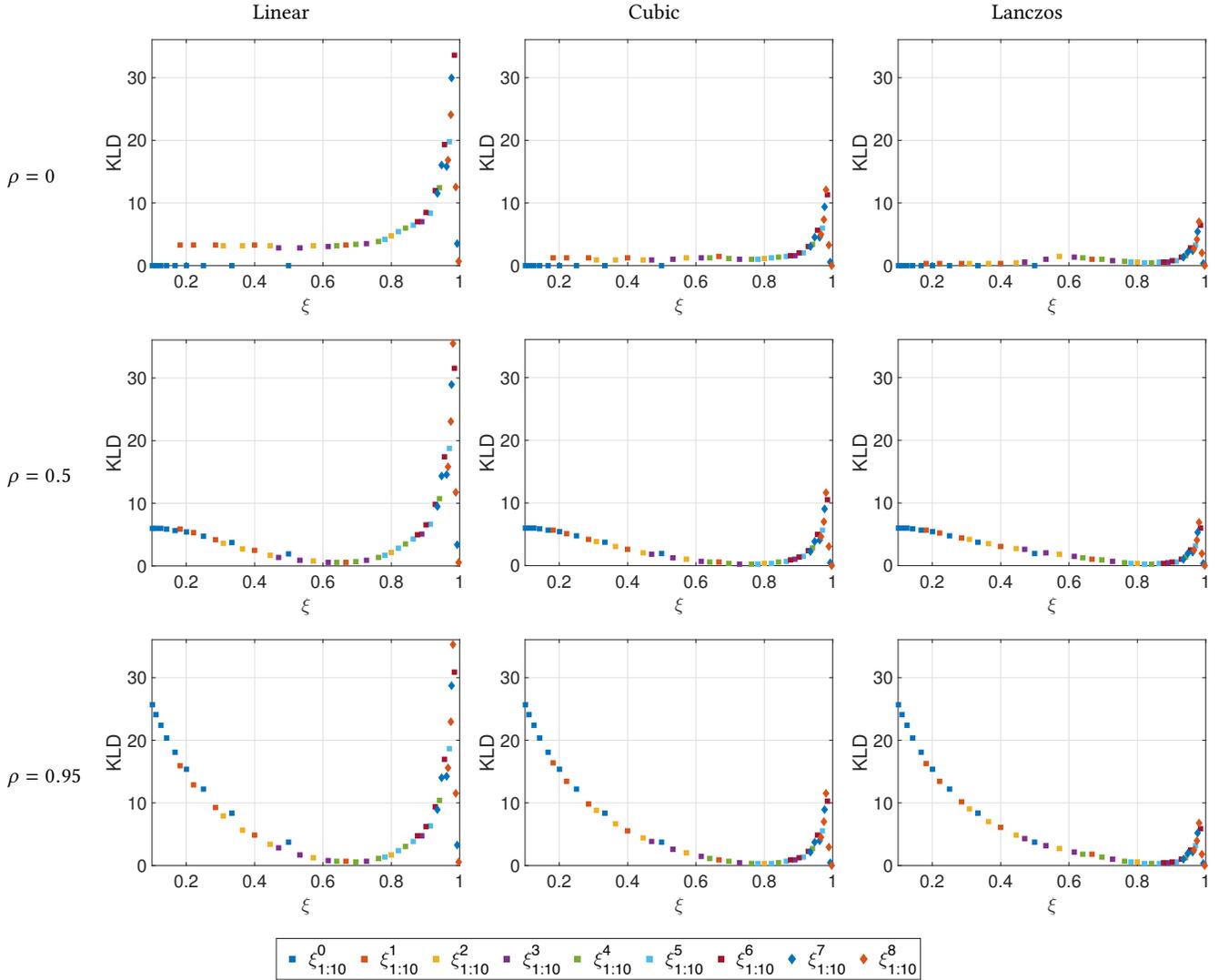


Figure 5: KLD values are reported as a function of ξ for different values of ρ (row-wise) and different interpolation kernels (column-wise). Different marks and colors indicate the different sequences $\xi_{1:10}^a$, according to the legend. The KLD in each case refers to the distributions of a M -variate vector with $M = 30$ under the two hypotheses.

h and ξ , so that N and Σ_0^{ex} are consequently determined to obtain the distribution of \mathbf{x} under H_1 .

From (23) we see that the test also depends on the signal parameters σ_ε , ρ and their influence on KLD should be studied. However, we can first note that the following Lemma holds:

LEMMA 5.1. *Under Assumption (1), given the correlation coefficient ρ , the resampling factor ξ and the interpolation kernel h , the measure*

$$\text{KLD}(\mathcal{N}_M(\mathbf{0}, \Sigma_0(\sigma_\varepsilon, \rho)), \mathcal{N}_M(\mathbf{0}, \Sigma_1(\sigma_\varepsilon, \rho, \xi, h)))$$

is independent of σ_ε .

PROOF. See Appendix A. \square

It is worth recalling here that we are not considering quantization effects and this result shows that, in this setting, the variance of the original samples turns out to act as a scaling factor (i.e., it does not add nor remove information). However, this may not hold when the signal is discretized before and after the resampling operation, thus leaving as future work the task of assessing the sensitivity of such result when a more or less slight quantization is applied.

In our current study, we can then fix the parameter $\sigma_\varepsilon = 1$ and analyze the variation of ρ , ξ and h , as reported in Fig. 5. We consider $\rho = 0$ (the signal is white Gaussian noise), $\rho = 0.5$ and $\rho = 0.95$, which is adopted in [23] to represent natural images. Different sequences $\xi_{1:10}^q$ (corresponding to locations along the horizontal axis) are identified by different markers and colors.

This shows that in each plot the KLD follows a general trend, although there are oscillations due to the specific interaction between the parameters. In each case, there is a peak in the interval $[0.95, 1]$ followed by a sharp decrease towards 0 when the resampling factor approaches 1. Moreover, we notice that the linear interpolation kernel generally yields a higher KLD, while it decreases when switching to cubic and Lanczos interpolation.

However, the most interesting observation is the different behavior of the KLD when the resampling factor approaches 0 among different rows. In fact, it appears that the correlation coefficient ρ substantially influences the KLD value when downsampling is strong.

In order to examine this phenomenon, we focus on the sequence $\xi_{1:10}^0 = \{1/q, q = 2, \dots, 10\}$ (blue squares in the plots). In fact, we see that for $\rho = 0$ the KLD is always zero, while the left tail significantly increases as ρ increases.

In this case, we can explain analytically the relationship between the entries of Σ_0 and Σ_1 , according to the following Lemma:

LEMMA 5.2. *Under Assumption (1), for every downsampling factor $1/q, q \in \mathbb{N}^0$, and an interpolation kernel such that $h(0) = 1$ and $h(k) = 0$ for every $k \in \mathbb{Z}$, the following expression holds:*

$$\Sigma_1[k, k+t] = \Sigma_0[k, k+t] \cdot \rho^{|t|(q-1)}, \quad t \in \mathbb{Z}. \quad (25)$$

PROOF. See Appendix B. \square

Note that the linear and cubic interpolation kernels analytically fulfill the requirement of the lemma, while the Lanczos kernel yields very small values at integers different than 0.

Given equation (13), we have that (25) does not depend on k but only on the lag t , so that the entries of Σ_1 are constant along each diagonal, just like it happens Σ_0 . Thus, we can evaluate the

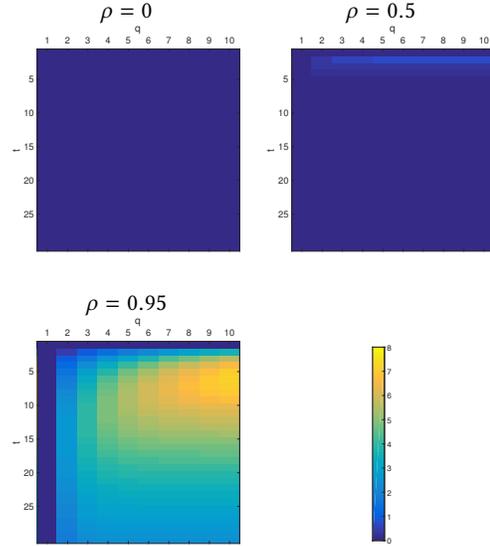


Figure 6: The quantity (6) is plotted as a function of q and t for different values of ρ .

absolute difference between the covariance matrix entries at the t -th diagonal under the two hypotheses as a function of q , which is given by

$$D(q, t) = \frac{\rho^{|t|}}{1 - \rho^2} (1 - \rho^{|t|(q-1)}) \quad (26)$$

In Fig. 6, we report the values of D when varying q and t , so that each vertical bar represents the first row of the covariance matrix for a fixed q , which contains all the values appearing in Σ_1 .

We observe that when $\rho = 0$, the whole covariance matrix is unaltered, thus explaining the zero values of the KLD in the first row of Fig. 5. If $\rho \neq 0$, no changes are clearly introduced when $q = 1$ and the same happens in any case when $t = 0$: the variance of samples is not modified.

It is worth considering that such kind of downsampling factors do not introduce observable periodicities in the covariance matrix, as the downsampled signal has pure cyclostationarity with period 1 (in fact, covariance matrix values are constant along diagonals). This situation has been already pointed out in the literature [18] and represents a major obstacle for existing forensic detectors based on the periodicity analysis, which (therefore) rarely address the detection of strong downsampling ($\xi < 1/2$).

However, our analysis shows that forensically useful information is still present in the signal. In this case, observe the decay of the autocovariance when t increases: expression (25) indicates that after downsampling with factor $1/q$ the autocovariance decays as the power of ρ^q instead of ρ .

5.2 Prediction error distribution

According to the formulation in Section 3, a similar analysis can be performed when the prediction error with coefficients $\beta_{-T}, \dots, \beta_T$ is employed (i.e., when the signal is prefiltered), so to evaluate whether this operation increases the distinguishability of the null and alternative hypothesis. In this case, a transformation by means

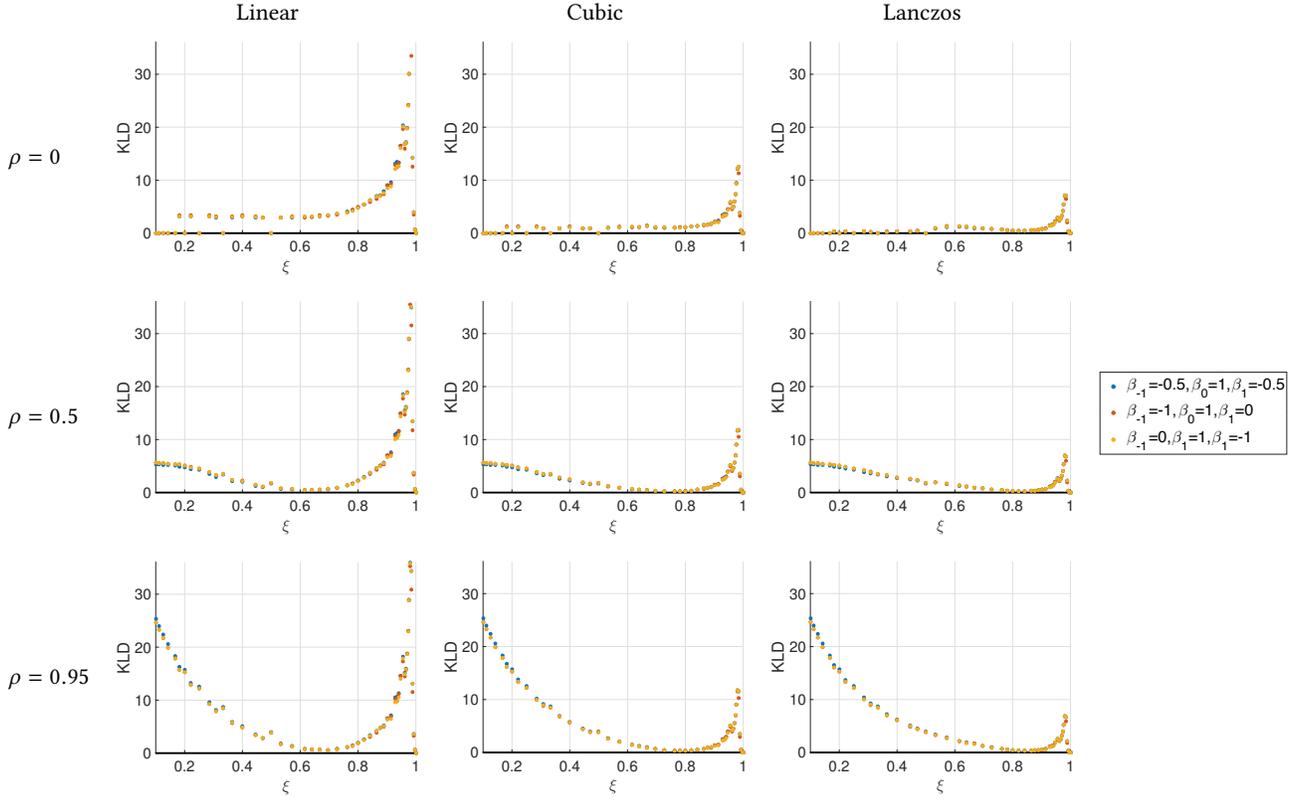


Figure 7: KLD values are reported as a function of ξ for different values of ρ (row-wise) and different interpolation kernels (column-wise). Different colors indicate different prefilters, according to the legend. The KLD in each case refers to the distributions of a M -variate vector with $M = 30$ under the two hypotheses.

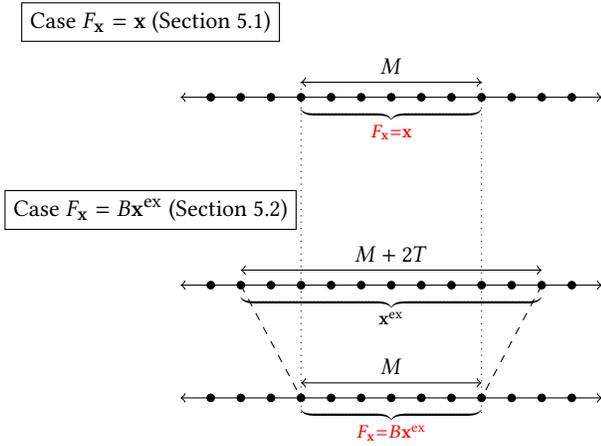


Figure 8: The figure shows the signal samples considered in the hypothesis tests of Sections 5.1 and 5.2, respectively. It can be noted that F_x in both cases refers to the same M locations in the signal, although $2T$ additional samples are used in the second case.

of the matrix B defined as in (9) is applied. However, it is to be noted that applying B to \mathbf{x} would not increase the hypothesis distinguishability with respect to the test considered in 5.1. This is due to the result recalled in [3, Lemma 1], stating that deterministic processing of data cannot increase the KLD of between two distributions. Thus, we define \mathbf{x}^{ex} as an extended version of \mathbf{x} containing in the first (last) positions the T samples of the signal under investigation lying before (after) \mathbf{x} . Then, we consider $F_x = B\mathbf{x}^{\text{ex}}$, so that the obtained vector F_x has length M and contains the values of the prediction error at the same locations of the values of \mathbf{x} , although more information is used with respect to the test in Section 5.1 (i.e., the additional samples). An illustration is given in Fig. 8.

By replicating the reasoning of Section 4, F_x is also a multivariate normal random vector and the covariance matrices under the two hypotheses are computed as in (22) with an additional pre-multiplication by B and post-multiplication by B^T . Thus, the KLD of $P_0(F_x)$ and $P_1(F_x)$ can be computed in the same way as in (5).

In Fig. 7 we report the KLD values when varying ξ and h as in Section 5.1. We considered three different kind of prefilters represented with different colors in the plots: one related to the discrete second order derivative ($\beta_{-1} = -0.5, \beta_0 = 1, \beta_1 = -0.5$) and two corresponding to discrete first order derivative, namely

the backward finite difference ($\beta_{-1} = -1, \beta_0 = 1, \beta_1 = 0$) and the forward finite difference ($\beta_{-1} = 0, \beta_1 = 1, \beta_1 = -1$).

We can observe that the values resemble the ones of Fig. 5, although a direct comparison would not be fair due to the different amount of information used (cf. Fig. 8). Moreover, the different prefilters do not lead to significantly different results, except for the case of the second order derivative filter with strong down-sampling. This is certainly related to the specific interaction of the prediction coefficients in β and the interpolation coefficients in C . An analytical characterization of the results reported in Fig. 7 as a function of ρ would explain the general trend observed in every plot and will be subject of future investigation. With this respect, our current conjecture is that, apart from information introduced by the additional samples, pre-filtering cannot substantially increase hypothesis distinguishability from an information-theoretic perspective, while its effect strongly depends on the decision method adopted (e.g., in [23] it proves to be crucial when performing a frequency analysis).

5.3 Autocovariance estimator distribution

As emphasized before, one of the properties exploited by many existing techniques is the periodicity of the elements of the covariance matrix Σ of \mathbf{x} . Moreover, we have seen in Section 5.1 that the covariance matrix can play a key role also when no periodicity is present. This calls for studying the situation where the entries of Σ are estimated starting from S realizations of \mathbf{x} by means of standard estimators, like the sample covariance matrix

$$\frac{\mathbf{X}^T \mathbf{X}}{S}, \quad \mathbf{X} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \dots \\ \mathbf{x}_S^T \end{bmatrix}. \quad (27)$$

It is to be noted that the estimation of second-order moments from the signal samples more or less implicitly happens in the majority of existing detectors: in [14] it is observed that the p -map is a simplified model for the error variance, so that every value of the prediction error is considered as an estimate of its variance at the specific location; in [23], the correlation of an image block with lag 0 is estimated before performing the frequency analysis; in [16], the autocovariance function after the Radon transform is estimated; in [11], different image rows are used to estimate the autocovariance in the horizontal direction.

A simple example is given by the case where the autocovariance function of the signal is considered as dependent only on the lag t and $\text{Cov}\{t\}$ is estimated as

$$\frac{1}{S} \sum_{i=1}^S s_i s_{i+t}, \quad (28)$$

where the s_i are the samples of the zero-mean signal under investigation. Note that this is the approach adopted in [16] to estimate the autocovariance function of the Radon transform vectors. By referring to the notation in formula (27), this is equivalent to considering S consecutive M -dimensional vectors as observations and use their values to estimate the entries of Σ , as represented in Fig. 9. However, this approach is only suited for WSS processes, including not resampled signals under Assumption 1 or downsampled signals with factor $1/q$, as shown in Section 5.1.

Other practical examples are given by sequential images taken by a fixed camera (where signal segments at fixed locations and different frames can be considered as equidistributed) or a single image (where horizontal signal segments at fixed location and different rows can be considered as equidistributed over homogeneous regions, as done in [11]).

We consider the case where variances and covariances are jointly estimated from a number of observations, as summarized in formula (27). Different observations correspond here to different available multivariate vectors $\mathbf{x}_1, \dots, \mathbf{x}_S$ that are assumed drawn from the same distribution and, additionally, *independent*. Under Assumption (1) and according to the analysis in previous sections, the observations are zero-mean multivariate normal vectors with covariance matrix Σ_0 and Σ_1 in case of null and alternative hypothesis, respectively. It is known that in this case, the matrix $\mathbf{X}^T \mathbf{X}$ (also called scatter matrix) follows a M -variate Wishart distribution with S degrees of freedom [8] and scale matrix Σ_0 or Σ_1 , according to the verified hypothesis. So, we can analyze the test

$$H_0: \mathbf{X}^T \mathbf{X} \sim \mathcal{W}_M(\Sigma_0(\sigma_\varepsilon, \rho), S), \quad (29)$$

$$H_1: \mathbf{X}^T \mathbf{X} \sim \mathcal{W}_M(\Sigma_1(\sigma_\varepsilon, \rho, \xi, h), S),$$

where $\mathcal{W}_M(\Sigma, S)$ indicates a M -dimensional Wishart distribution with S degrees of freedom and scale matrix Σ , with the constraint that $S \geq M$ in order for the Wishart distribution to have a density in the space of $M \times M$ matrices [8]. The KLD between two Wishart distributions with the same number of degrees of freedom is given by [10]

$$\text{KLD}(\mathcal{W}_M(\Sigma_0, S), \mathcal{W}_M(\Sigma_1, S)) = S \left[\frac{\text{tr}(\Sigma_0^{-1} \Sigma_1) + \text{tr}(\Sigma_1^{-1} \Sigma_0)}{2} - M \right]. \quad (30)$$

It is worth noticing that the KLD increases *linearly* with the number of observations. Similarly as in Lemma 5.1, we can prove that expression (30) does not depend on σ_ε , so we perform in Fig. 10 the same analysis as in Section 5.1. The trend observed in Fig. 10 further confirms the considerations made in Section 5.1. It is to be noted that a direct comparison of the KLD values would again be unfair, as in this case the whole matrix \mathbf{X} is used, which contains a considerably higher amount of information with respect to \mathbf{x} . However, we can notice that the effect observed when the resampling factor approaches 0 seems to be here even amplified (see the case $\rho = 0.95$ and linear interpolation).

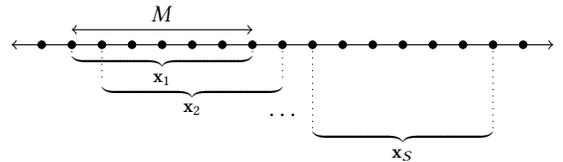


Figure 9: Example of different observations.

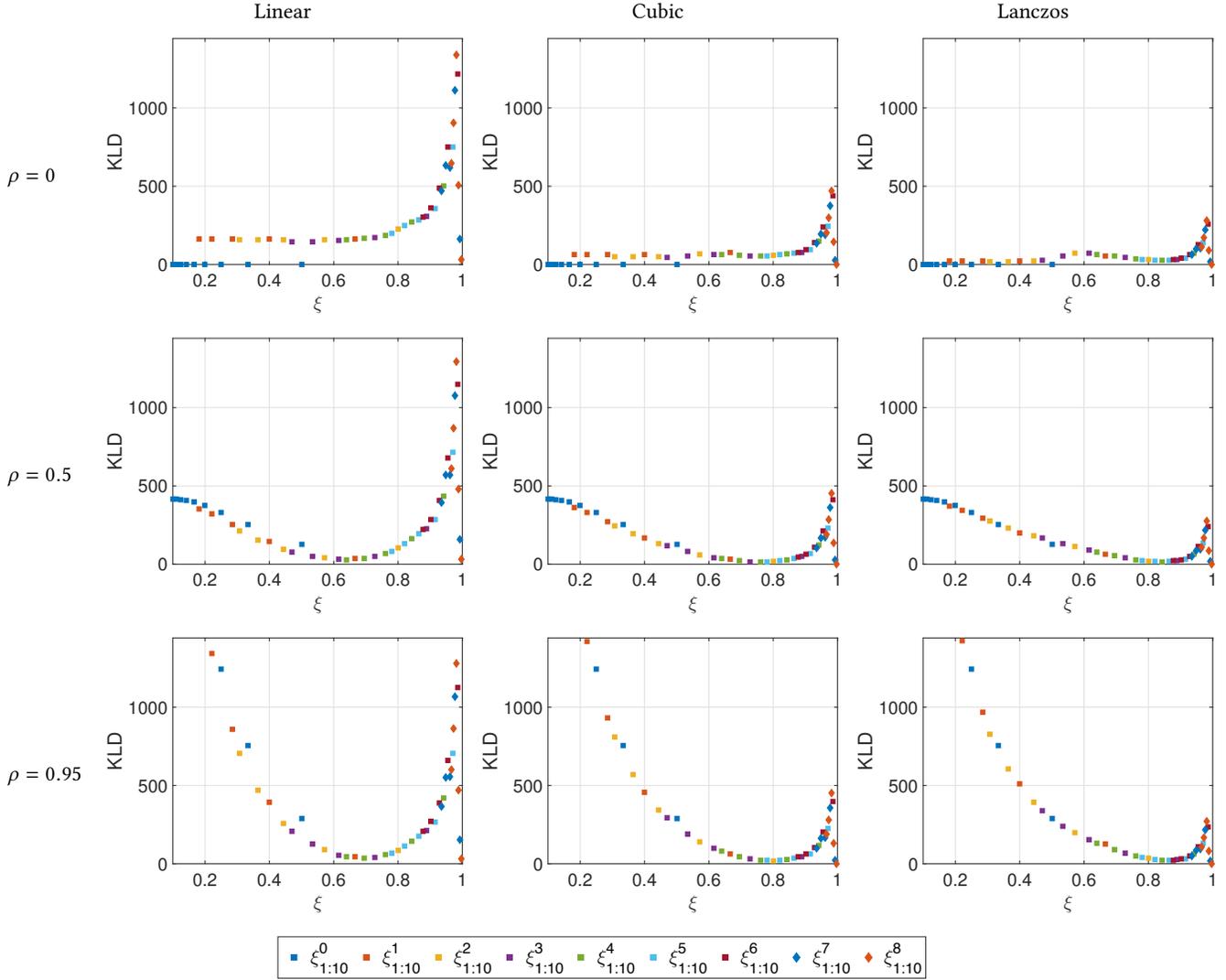


Figure 10: KLD values are reported as a function of ξ for different values of ρ (row-wise) and different interpolation kernels (column-wise). Different marks and colors indicate the different sequences $\xi_{1:10}^a$, according to the legend. The KLD in each case refers to the distributions of a $M \times M$ matrix with $M = 30$ under the two hypotheses.

6 CONCLUDING REMARKS

We have studied the statistical distance of downsampled 1D signals with respect to their non-downsampled counterparts. Under common assumptions on the original signal statistics, we have observed that all the information about a signal segment statistics is contained in its covariance matrix. This allowed us to assess the influence of both signal and resampling parameters on the distinguishability between the hypothesis of no downsampling and the hypothesis of downsampling in terms of KLD. In doing so, we have studied the distribution of the signal segment itself, of different prefiltered versions of it and also of a standard estimator of its covariance matrix.

We can first observe that the use of different interpolation kernels (namely linear, cubic, and Lanczos) does not seem to have a substantial effect on the KLD with respect to different resampling factors. However, we notice that the KLD values when the Lanczos kernel is used are upper bounded by the ones obtained with the cubic interpolation, and the same happens for cubic and linear interpolation, respectively.

When varying the resampling factor $\xi \in [0, 1]$, we can identify in each case a common trend when ξ approaches 1: the KLD increases approximately until 0.95–0.98, and then sharply drops to 0, as expected.

However, the most interesting finding is observed when the correlation coefficient ρ among consecutive original samples is modified. In fact, while the KLD behavior remains stable for ξ

approaching 1, the hypothesis distinguishability in case of strong downsampling significantly increases as ρ increases. With this respect, the choice of considering sequences of rational numbers with the same numerator (i.e., introducing the same periodicity in the resampled signal) proves to be crucial since it allows us to obtain theoretical results for specific sequences, as we did for the factors $\xi = 1/q, q \in \mathbb{N}$. In this case, we have analytically characterized the transformation of single covariance matrix entries, showing that the underlying correlation structure in the original signal plays a key role and introduces statistical distance between a non-downsampled and a downsampled signal even when *no periodicity* is introduced. This obviously calls for research on new classes of detectors which exploit these traces in real signals. For the design of practical approaches, we cannot assume the knowledge of ρ in the original signal. If and how well it can be approximated is subject to future investigations. Other issues arise for signals where ρ is spatially varying.

Also the theoretical analysis leaves room for future extensions in several directions. As mentioned, a theoretical description of different prefiltering effects on the hypothesis distinguishability as well as upsampling in general still represent open questions and would complement the present analysis. Moreover, future work will be devoted to the extension of the model to 2D signals, thus accounting for interpolation kernels and correlation structures involving both the row and column dimension. In addition, the current approach deals with continuous signals, while discretized signals should be considered in order to account for quantization effects, which cannot be avoided in practice and already proved to be a valuable help in in resampling forensics [20, 24] as well as many other forensic decision problems [2].

ACKNOWLEDGMENTS

This research was funded by Deutsche Forschungsgemeinschaft (DFG) under grant ‘‘Informationstheoretische Schranken digitaler Bildforensik’’ and by Archimedes Privatstiftung, Innsbruck, Austria.

A PROOF OF LEMMA 5.1

From expression (14) we have that

$$\Sigma_0(\sigma_\varepsilon, \rho) = \sigma_\varepsilon^2 \Phi_0, \quad \Phi_0 = \begin{bmatrix} \frac{1}{1-\rho^2} & \frac{\rho^1}{1-\rho^2} & \frac{\rho^2}{1-\rho^2} & \cdots & \frac{\rho^{N-1}}{1-\rho^2} \\ \frac{\rho^1}{1-\rho^2} & \frac{1}{1-\rho^2} & \frac{\rho^1}{1-\rho^2} & \cdots & \frac{\rho^{N-2}}{1-\rho^2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \frac{\rho^{N-2}}{1-\rho^2} & \vdots & \frac{\rho^1}{1-\rho^2} & \frac{1}{1-\rho^2} & \frac{\rho^1}{1-\rho^2} \\ \frac{\rho^{N-1}}{1-\rho^2} & \cdots & \frac{\rho^2}{1-\rho^2} & \frac{\rho^1}{1-\rho^2} & \frac{1}{1-\rho^2} \end{bmatrix}, \quad (31)$$

so that $\Sigma_1 = C \Sigma_0^{\text{ex}} C^T = \sigma_\varepsilon^2 C \Phi_0^{\text{ex}} C^T = \sigma_\varepsilon^2 \Phi_1$, where Φ_0^{ex} is defined in the same way as Σ_0^{ex} and $\Phi_1 \doteq C \Phi_0^{\text{ex}} C^T$ does not depend on σ_ε .

We can now consider the parts of (24) that depends on σ_ε :

$$\text{tr}(\Sigma_1^{-1} \Sigma_0) = \text{tr}((\sigma_\varepsilon^2 \Phi_1)^{-1} \sigma_\varepsilon^2 \Phi_0) = \text{tr}\left(\frac{1}{\sigma_\varepsilon^2} \Phi_1^{-1} \sigma_\varepsilon^2 \Phi_0\right) = \text{tr}(\Phi_1^{-1} \Phi_0), \quad (32)$$

and

$$\frac{\det(\Sigma_1)}{\det(\Sigma_0)} = \frac{\det(\sigma_\varepsilon^2 \Phi_1)}{\det(\sigma_\varepsilon^2 \Phi_0)} = \frac{\sigma_\varepsilon^{2M} \det(\Phi_1)}{\sigma_\varepsilon^{2M} \det(\Phi_0)} = \frac{\det(\Phi_1)}{\det(\Phi_0)}. \quad (33)$$

Neither (32) nor (33) depend on σ_ε , but only on ρ , and so does (24). \square

B PROOF OF LEMMA 5.2

We can observe that, under the premises of the lemma, the matrix C has the form

$$C = \begin{bmatrix} 1 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & \underbrace{\cdots}_q & 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \underbrace{\cdots}_q & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \underbrace{\cdots}_q & 1 & \cdots \\ \vdots & \ddots \end{bmatrix}. \quad (34)$$

Thus, every k -th row has only one nonzero element equal to 1 at the $(q(k-1)+1)$ -th column.

Coupled with (22) and (13), it means that every entry of Σ_1 is given by

$$\begin{aligned} \Sigma_1[k, k'] &= \sum_{n=1}^N C[k, n] \sum_{n'=1}^N C[k', n'] \Sigma_0^{\text{ex}}[n, n'] \\ &= C[k, q(k-1)+1] C[k', q(k'-1)+1] \frac{\rho^{q|k-k'|}}{1-\rho^2} \\ &= \frac{1}{1-\rho^2} \rho^{|k-k'|q}. \end{aligned}$$

By fixing $k' = k + t$, we have expression (25). \square

REFERENCES

- [1] M. Barni and B. Tondi. 2016. Source distinguishability under distortion-limited attack: an optimal transport perspective. *IEEE Transactions on Information Forensics and Security* 11, 10 (2016), 2145–2159.
- [2] R. Böhme and M. Kirchner. 2016. Media Forensics. *S. Katzenbeisser and f. Petitcolas, eds., Information Hiding* (2016), 231–259.
- [3] C. Cachin. 1998. An information-theoretic model for steganography. In *International Workshop on Information Hiding*, 306–318.
- [4] X. Chu, Y. Chen, M. Stamm, and K.J. Ray Liu. 2016. Information theoretical limit of media forensics: the Forensicability. *IEEE Transactions on Information Forensics and Security* 11, 4 (2016), 774–788.
- [5] P. Comesaña-Alfaro. 2012. Detection and information theoretic measures for quantifying the distinguishability between multimedia operator chains. In *IEEE Workshop on Information Forensics and Security (WIFS)*, 211–216.
- [6] P. Comesaña-Alfaro and F. Pérez-González. 2012. Multimedia operator chain topology and ordering estimation based on detection and information theoretic tools. In *International Workshop on Digital Watermarking*, 213–227.
- [7] N. Dalgaard, C. Mosquera, and F. Pérez-González. 2010. On the role of differentiation for resampling detection. In *IEEE International Conference on Image Processing (ICIP)*, 1753–1756.
- [8] M. L. Eaton. 2007. The Wishart distribution. *IMS Lecture Notes Monograph Series. Multivariate statistics: a vector space approach* (2007).
- [9] X. Feng, I. J. Cox, and G. Doërr. 2012. Normalized energy density-based forensic detection of resampled images. *IEEE Transactions on Multimedia* 14, 3 (2012), 536–545.
- [10] A. C. Frery, A. D. C. Nascimento, and R. J. Cintra. 2014. Analytic expressions for stochastic distances between relaxed complex Wishart distributions. *IEEE Transactions on Geoscience and Remote Sensing* 52, 2 (2014), 1213–1226.
- [11] A.C. Gallagher. 2005. Detection of linear and cubic interpolation in JPEG compressed images. In *Proceedings of the 2nd Canadian Conference on Computer and Robot Vision*, 65–72.
- [12] G. K. Grunwald, R. J. Hyndman, and L. Tedesco. 1995. *A unified view of linear AR(1) models*. Technical Report. Department of Statistics, University of Melbourne.
- [13] R. A. Horn and C. R. Johnson. 2013. *Matrix Analysis*. Vol. Second Edition. Cambridge University Press.
- [14] M. Kirchner. 2008. Fast and Reliable Resampling Detection by Spectral Analysis of Fixed Linear Predictor Residue. In *ACM Multimedia and Security Workshop (MM&Sec)*, 11–20.
- [15] M. Kirchner. 2010. Linear row and column predictors for the analysis of resized images. In *ACM Multimedia and Security Workshop (MM&Sec)*, 13–18.
- [16] B. Mahdian and S. Saic. 2008. Blind authentication using periodic properties of interpolation. *IEEE Transactions on Information Forensics and Security* 3, 3 (2008), 529–538.
- [17] F. Nielsen and R. Nock. 2009. Clustering multivariate normal distributions. *Emerging trends in visual computing (LNCS)* (2009).
- [18] A.C. Popescu and H. Farid. 2005. Exposing digital forgeries by detecting traces of resampling. *IEEE Transactions on Signal Processing* 53, 2 (2005), 758–767.
- [19] V. Sathe and P. Vaidyanathan. 1993. Effects of multirate systems on the statistical properties of random signals. *IEEE Transactions on Signal Processing* 41, 1 (1993), 131.
- [20] D. Vázquez-Padín and P. Comesaña-Alfaro. 2012. ML estimation of the resampling factor. In *IEEE International Workshop on Information Forensics and Security (WIFS)*, 205–210.
- [21] D. Vázquez-Padín, P. Comesaña-Alfaro, and F. Pérez-González. 2015. An SVD approach to forensic image resampling detection. In *European Signal Processing Conference (EUSIPCO)*, 2112–2116.
- [22] D. Vázquez-Padín, C. Mosquera, and F. Pérez-González. 2010. Two-dimensional statistical test for the presence of almost cyclostationarity on images. In *IEEE International Conference on Image Processing (ICIP)*, 1745–1748.
- [23] D. Vázquez-Padín and F. Pérez-González. 2011. Prefilter design for forensic resampling estimation. In *IEEE International Workshop on Information Forensics and Security (WIFS)*, 1–6.
- [24] D. Vázquez-Padín and F. Pérez-González. 2013. Set-membership identification of resampled signals. In *IEEE International Workshop on Information Forensics and Security (WIFS)*, 150–155.