2009 American Control Conference
Hyatt Regency Riverfront, St. Louis, MO, USA
June 10-12, 2009

Symbolic Analysis of Time Series Signals Using Generalized Hilbert Transform★

Soumik Sarkar szs200@psu.edu Kushal Mukherjee kum162@psu.edu Asok Ray axr2@psu.edu
Department of Mechanical Engineering
The Pennsylvania State University
University Park, PA 16802, USA

Keywords: *Hilbert Transform; Symbolic Time Series Analysis; D-Markov Machines*

***Abstract*— A recent publication has shown a Hilbert-transform-based partitioning method, called analytic signal space partitioning (*ASSP*). When used in conjunction with *D*-Markov machines, also reported in recent literature, *ASSP* provides a fast tool for pattern recognition. However, Hilbert transform does not specifically address the issue of noise reduction and the usage of *D*-Markov machines with a small depth *D* could potentially lead to information loss for noisy signals. On the other hand, a large *D* tends to make execution of pattern recognition computationally less efficient due to an increased number of machine states. This paper explores generalization of Hilbert transform that addresses symbolic analysis of noise-corrupted dynamical systems. In this context, theoretical results are derived based on the concepts of information theory. These results are validated on time series data, generated from a laboratory apparatus of nonlinear electronic systems.**

1. Introduction

HILBERT transform and the associated concept of analytic signals, introduced by Gabor [1], have been widely adopted for time-frequency analysis in diverse applications of signal processing. Hilbert transform [2] of a real-valued signal $x(t)$ is defined as:

$$\widetilde{x}(t) \triangleq \mathcal{H}[x](t) = \frac{1}{\pi}\int_{\mathbb{R}} \frac{x(\tau)}{t-\tau} d\tau \tag{1}$$

That is, $\widetilde{x}(t)$ is the convolution of $x(t)$ with $\frac{1}{\pi t}$ over $\mathbb{R} \triangleq (-\infty, \infty)$, which is represented in the Fourier domain as:

$$\widehat{\widetilde{x}}(\omega) = -i\ sgn(\omega)\ \widehat{x}(\omega) \tag{2}$$

where $\widehat{x}(\omega) \triangleq \mathcal{F}[x](\omega)$ and $sgn(\omega) \triangleq \begin{cases} +1 & \text{if } \omega > 0 \\ -1 & \text{if } \omega < 0 \end{cases}$

Given the Hilbert transform of a real-valued signal $x(t)$, the complex-valued analytic signal [2] is defined as:

$$\mathcal{X}(t) \triangleq x(t) + i\ \widetilde{x}(t) \tag{3}$$

★This work has been supported in part by the U.S. Army Research Office (ARO) under Grant No. W911NF-07-1-0376, by NASA under Cooperative Agreement No. NNX07AK49A, and by the U.S. Office of Naval Research under Grant No. N00014-08-1-380. Any opinions, findings and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the sponsoring agencies.

and the (real-valued) transfer function with input $\widehat{x}(\omega)$ and output $\widehat{\mathcal{X}}(\omega)$ is formulated as:

$$G(\omega) \triangleq \frac{\widehat{\mathcal{X}}(\omega)}{\widehat{x}(\omega)} = 1 + sgn(\omega) \tag{4}$$

Recently, Subbu and Ray [3] have reported an application of Hilbert transform for symbolic time series analysis of dynamical systems where the space of analytic signals, derived from real-valued time-series data, is partitioned for symbol sequence generation. This method, called analytic signal space partitioning (*ASSP*), is comparable or superior to other partitioning techniques, such as symbolic false nearest neighbor partitioning (*SFNNP*) [4] and wavelet-space partitioning (*WSP*) [5], in terms of performance, complexity and computation time. A major shortcoming of *SFNNP* is that the symbolic false neighbors rapidly grow in number for noisy data and may erroneously require a large symbol alphabet to capture pertinent information on the system dynamics. The wavelet transform largely alleviates these shortcomings and thus *WSP* is particulary effective for noisy data from high-dimensional dynamical systems. However, WSP has several other shortcomings such as identification of an appropriate basis function, selection of appropriate scales, and non-unique and lossy conversion of the two-dimensional scale-shift wavelet domain to a one-dimensional domain of scale-series sequences [5].

When applied to symbolic analysis in dynamical systems, *ASSP* is used to formulate a probabilistic finite-state model, called the *D*-Markov model [6], where the machine states are symbol blocks of depth *D*. For noisy systems, it is expected that modeling with a large *D* in the *D*-Markov machine would result in higher gain in information on the system dynamics. However, a large *D* increases the number of machine states, which in turn degrades computation efficiency (e.g., increased execution time and memory requirements) [7].

This paper introduces a generalization of the classical Hilbert transform to modify *ASSP* for application to noisy systems. The objective here is to partition the transformed signal space such that *D*-Markov machines can be constructed with a small *D* without significant loss of information for noisy signals. The key idea is to provide a

mathematical structure of the generalized Hilbert transform such that the low-frequency region is more heavily weighted than that in the classical Hilbert transform.

2. Generalization of Hilbert Transform

Lohmann et al. [8] introduced the concept of a generalized Hilbert transform in the fractional Fourier space instead of the conventional Fourier space; a discrete version of this generalized Hilbert transform was developed later [9]. For geophysical applications, Luo et al. [10] proposed another type of generalized Hilbert transform that is essentially the windowed version of traditional Hilbert transform. Our Generalization of Hilbert transform is different from previously published methods of generalization of Hilbert transform.

Let us define a generalized Hilbert transform as: $\mathcal{H}^\alpha$ of a real-valued signal $x(t)$ as the convolution:

$$\widetilde{x}^\alpha(t) \triangleq \mathcal{H}^\alpha[x](t) = x(t) * \left(\frac{sgn(t)}{\pi\ |t|^\alpha}\right) \quad \text{for } \alpha \in (0,1] \quad (5)$$

It is shown in the sequel that, as $\alpha \uparrow 1$ (i.e., the values of α form an increasing sequence of positive real numbers with the limit equal to 1), $\mathcal{H}^\alpha$ converges to $\mathcal{H}$, where $\mathcal{H}$ is the classical Hilbert transform defined in Eq. (1); that is, $\mathcal{H}^1 \equiv \mathcal{H}$.

Two lemmas are presented, which are necessary for derivation of the main results in the Fourier space.

Lemma 2.1:

$$\int_{-\infty}^{\infty} \frac{e^{-i\omega t}}{\pi|t|^\alpha} sgn(t)dt = -i\ sgn(\omega)\frac{2}{\pi}\frac{\Gamma(1-\alpha)}{|\omega|^{1-\alpha}} sin\left(\frac{\pi}{2}(1-\alpha)\right) \quad (6)$$

where $\alpha \in (0,1)$; and $\Gamma(1-\alpha) \triangleq \int_0^\infty \frac{e^{-y}}{y^\alpha} dy$.

Lemma 2.2: As $\alpha \uparrow 1$, the integral $\int_{-\infty}^{\infty} \frac{e^{-i\omega t}}{\pi|t|^\alpha} sgn(t)dt \rightarrow -i\ sgn(\omega)$, i.e.,

$$\lim_{\alpha\uparrow 1} \Gamma(1-\alpha)\ \left(\frac{2}{\pi} sin\frac{\pi}{2}\ (1-\alpha)\right) = 1. \quad (7)$$

Proofs of the above two lemmas are provided in Appendix A.

Taking Fourier transform of the convolution in Eq. (5) and an application of Lemma 2.1 yield:

$$\begin{aligned} \widehat{\widetilde{x}}^\alpha(\omega) &= \mathcal{F}\left(x(t) * \frac{sgn(t)}{\pi|t|^\alpha}\right) \\ &= \mathcal{F}(x(t)) \cdot \mathcal{F}\left(\frac{sgn(t)}{\pi|t|^\alpha}\right) \\ &= -i\ sgn(\omega)\ \frac{2}{\pi}\ \frac{\hat{x}(\omega)\Gamma(1-\alpha)}{|\omega|^{1-\alpha}} \cdot \\ &\qquad \sin\left(\frac{\pi}{2}(1-\alpha)\right) \end{aligned} \quad (8)$$

Since $\Gamma(1-\alpha) < \infty$ for $\alpha \in (0,1)$, the generalized Hilbert transform $\widetilde{x}^\alpha(t)$ can be evaluated by taking the inverse Fourier transform of $\widehat{\widetilde{x}}^\alpha(\omega)$.

The above formulation shows that reduced α puts more weight on the low frequency part of the signal $x(t)$ and hence more effectively attenuates the high-frequency noise than the classical Hilbert transform. Following Lemma 2.2, as $\alpha \uparrow 1$, Fourier transform of the signal $\frac{sgn(t)}{\pi|t|^\alpha}$ converges to $-i\ sgn(\omega)$. This leads to the fact, that as $\alpha \uparrow 1$, $\mathcal{H}^\alpha$ converges to $\mathcal{H}$, where $\mathcal{H}$ is the classical Hilbert transform defined in Eq. (1).

Analogous to the analytic signal in Eq. (3), the (complex-valued) generalized analytic signal of the real-valued signal $x(t)$ is defined as:

$$\mathcal{X}^\alpha(t) \triangleq x(t) + i\ \widetilde{x}^\alpha(t) \quad (9)$$

and the (real-valued) transfer function with input $\widehat{x}(\omega)$ and output $\widehat{\mathcal{X}}^\alpha(\omega)$ is formulated as:

$$G^\alpha(\omega) \triangleq \frac{\widehat{\mathcal{X}}^\alpha(\omega)}{\widehat{x}(\omega)} = 1 + sgn(\omega)\ \left(\frac{\omega_0(\alpha)}{|\omega|}\right)^{(1-\alpha)} \quad (10)$$

where $\omega_0(\alpha) \triangleq \left(\frac{2}{\pi\ \Gamma(1-\alpha)} sin\left(\frac{\pi}{2}(1-\alpha)\right)\right)^{\frac{1}{1-\alpha}}$.

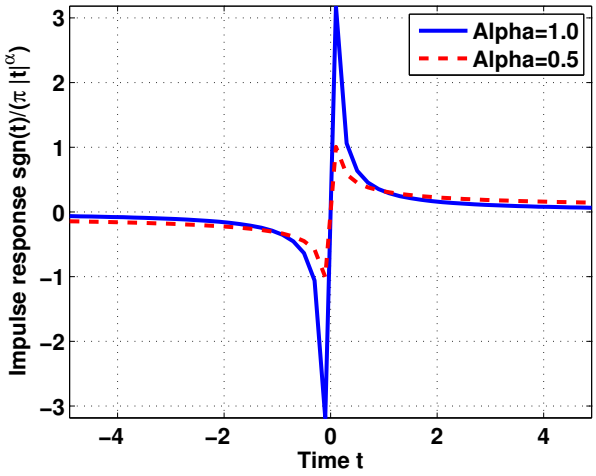


Fig. 1. Impulse response $\left(\frac{sgn(t)}{\pi|t|^\alpha}\right)$ of the generalized Hilbert transform

Remark 2.1: For $\alpha = 1$, it follows from Eq. (5) that the real-valued signal $x(t)$ is convoluted with $\frac{1}{\pi t}$. The implication is that the effects of memory in the signal $x(t)$ reduce as fast as $\frac{1}{\pi|t|}$. As α is decreased, the tail of the impulse response of the generalized Hilbert transform $\widetilde{x}^\alpha(t)$ becomes increasingly fat as seen in Fig. 1. Hence, for $0 < \alpha < 1$, the generalized analytic signal $\mathcal{X}^\alpha(t)$ captures more (low-frequency) information from time series data than that for $\alpha = 1$.

Remark 2.2: Fourier transform of a real-valued signal does not contain any additional information beyond what is

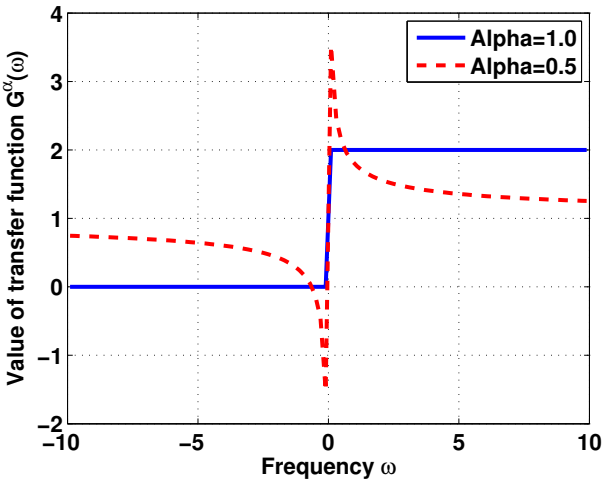


Fig. 2. Transfer function $G^{\alpha}(\omega)$ of the generalized analytic signal

provided by the positive frequency components, because of the symmetry of its spectrum. Therefore, in the construction of an analytic signal in Eq. (3) and its transfer function in Eq. (4), Hilbert transform removes the negative frequency components while doubling the positive frequency components. For $\alpha < 1$, it follows from Fig. 2 that the negative frequency components of the transfer function $G^{\alpha}(\omega)$ of a generalized analytic signal are no longer zero. Therefore, the generalized analytic signal in Eq. (9) is not an analytic signal in the sense of Gabor [1] for $\alpha < 1$. However, the transfer functions of both analytic and generalized analytic signals are real-valued almost everywhere in the range $\omega \in \mathbb{R}$. The phase of the (real-valued) transfer function $G^{\alpha}(\omega)$ is either 0 or $-\pi$ as explained below.

- The phase of $G(\omega)$ (i.e., $G^{\alpha}(\omega)$ for $\alpha = 1$) is 0 radians in the frequency range $(0, \infty)$. Its magnitude in the negative frequency range $(-\infty, 0)$ is identically equal to 0; therefore, the phase in this range is inconsequential.
- For $0 < \alpha < 1$, the phase of $G^{\alpha}(\omega)$ is $-\pi$ radians in the frequency range $(-\omega_0(\alpha), 0)$, where ω_0 is defined in Eq. (10), and is 0 radians in the frequency range $\big(-\infty, -\omega_0(\alpha)\big) \bigcup (0, \infty)$.

3. Test Results and Validation

The concept of generalized Hilbert transform is tested and validated by symbolic analysis of time series data, generated from the same apparatus of nonlinear electronic systems reported in the earlier publication [3]. The symbol sequence, constructed from time series data, is passed through a fixed structure *D*-Markov machine [6] to compute the state-transition matrices, called Π-matrices, for two values of the depth parameter, $D = 1$ and $D = 2$. Performance of the two *D*-Markov representations for each partition, corresponding to different values of the parameter α, is compared in terms of the mutual information [11] and the associated information gain. The procedure is delineated below.

A. Collection of Time Series Data

The nonlinear active electronic system in the test apparatus emulates the forced Duffing equation:

$$\frac{d^2x}{dt^2} + \beta\frac{dx}{dt} + x(t) + x^3(t) = Acos(\omega t) \qquad (11)$$

Having the system parameters set to $\beta = 0.24$, $A = 22.0$, and $\omega = 5.0$, time series data of the variable $x(t)$ were collected from the electronic system apparatus. These data sets do not contain any substantial noise because the laboratory apparatus is carefully designed to shield spurious signals and noise. Therefore, to emulate the effects of noise in the time series data, additive first-order colored Gaussian noise was injected to the collected time series data to investigate the effects of signal-to-noise ratio (SNR). The profile of a typical signal, contaminated with 10 db additive Gaussian noise (i.e., $SNR = 10$), is shown in Fig. 3.

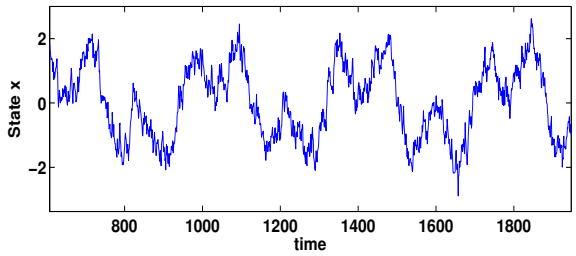


Fig. 3. Signal contaminated with 10 db additive colored Gaussian noise

B. Construction of the Transformed Phase Space

Let the real-valued noisy time-series data $x(t)$ contain N data points. Upon generalized Hilbert transformation of this data sequence, a complex-valued generalized analytic signal $\mathcal{X}^{\alpha}(t)$ is constructed. Similar to the procedure described in [3], $\mathcal{X}^{\alpha}(t)$ is represented as a one-dimensional trajectory in the two-dimensional pseudo-phase space. Let Ω be a compact region in the pseudo-phase space, which encloses the trajectory of N such data points.

C. Partitioning and Symbol Generation

The next task is to partition Ω into finitely many mutually exclusive and exhaustive segments, where each segment is labelled with a symbol or letter. The partitioning is based on the magnitude and phase of the complex-valued signal $\mathcal{X}^{\alpha}(t)$ as well as the density of data points in these segments, following the procedure described in [3]. Each point in the partitioned data set is represented by a pair of symbols; one belonging to the alphabet Σ_R based on the magnitude (i.e., in the radial direction) and the other belonging to the alphabet Σ_A based on the phase (i.e., in the angular direction). In this way, the complex-valued signal $\mathcal{X}^{\alpha}(t)$ is partitioned into a symbol sequence by associating each pair of symbols to a single symbol belonging to an alphabet Σ that is defined as:

$$\Sigma \triangleq \big\{(\sigma_i, \sigma_j) \in \Sigma_R \times \Sigma_A\big\} \text{ and } |\Sigma| = |\Sigma_R||\Sigma_A| \qquad (12)$$

The results presented in this paper are generated with $|\Sigma_R| = 8$ in the radial direction and $|\Sigma_A| = 5$ in the angular direction, i.e., $|\Sigma| = 40$.

D. State Transition Matrices

The symbol sequence is now used to construct *D*-Markov machine models [6]. The assumption of statistical stationarity of the symbol sequence is implicit in the construction of Markov models. In this paper, Markov chain models of depth $D = 1$ and $D = 2$ have been constructed.

Modeling of the symbolic process as a $(D = 1)$ Markov chain involves evaluation of the Π^1 matrix, where the ij^{th} matrix element π^1_{ij} is defined as the probability that $(n+1)^{th}$ state is i given that the n^{th} state was j, i.e.,

$$\pi^1_{ij} \triangleq P\left(q^1_{n+1} = i \mid q^1_n = j\right) \tag{13}$$

where q_k is the state at discrete time instant k. Evidently, the size of the Π matrix is $|\Sigma| \times |\Sigma|$, where $|\Sigma|$ is the number of symbols in the alphabet Σ.

Modeling the symbolic process as a $(D = 2)$ Markov chain involves evaluation of a 3-dimensional matrix, where the ijk^{th} matrix element π^2_{ijk} is defined as:

$$\pi^2_{ijk} \triangleq P\left(q^2_{n+2} = i | q^2_{n+1} = j, q^2_n = k\right) \tag{14}$$

and size of the (sparse) Π^2 matrix is $|\Sigma| \times |\Sigma| \times |\Sigma|$.

Remark 3.1: Elements of both Π^1 and Π^2 matrices are estimated by conditional frequency count and their convergence requires a symbol sequence of sufficient length. This aspect has been discussed in [12][7] and is referred to as the stopping rule that assigns a bound on the length of the symbol sequence for parameter identification of the stochastic matrices Π^1 and Π^2.

E. Computation of Mutual Information

Effectiveness of generalized Hilbert transform for Markov model construction has been examined from an information theoretic perspective [11]. The rationale is that, in a noise-corrupted system, higher values of mutual information imply less uncertainties in the symbol sequence. The mutual information $\mathcal{I}$ is expressed in terms of entropy $\mathcal{S}$ for both $(D = 1)$ and $(D = 2)$ Markov chains in the following set of equations:

$$\mathcal{I}\left(q_{n+3}; q_{n+2}\right) \triangleq \mathcal{S}\left(q_{n+3}\right) - \mathcal{S}\left(q_{n+3}|q_{n+2}\right) \tag{15}$$

$$\mathcal{S}\left(q_{n+3}\right) \triangleq -\sum_{\ell=1}^{|\Sigma|} P\left(q_{n+3} = \ell\right) \quad logP\left(q_{n+3} = \ell\right) \tag{16}$$

Usage of maximum entropy partitioning [5] for symbol generation yields: $\mathcal{S}\left(q_{n+3}\right) = log\left(|\Sigma|\right)$.

$$\mathcal{S}\left(q_{n+3}|q_{n+2}\right) \triangleq \sum_{\ell=1}^{|\Sigma|} P\left(q_{n+2} = \ell\right) \; \mathcal{S}\left(q_{n+3}|q_{n+2} = \ell\right) \tag{17}$$

where

$$\mathcal{S}\left(q_{n+3}|q_{n+2} = \ell\right) = -\sum_{j=1}^{|\Sigma|} P\left(q_{n+3} = j \mid q_{n+2} = \ell\right) \cdot logP\left(q_{n+3} = j|q_{n+2} = \ell\right) \tag{18}$$

$$\mathcal{I}\left(q_{n+3}; q_{n+2}, q_{n+1}\right) \triangleq \mathcal{S}\left(q_{n+3}\right) - \mathcal{S}\left(q_{n+3}|q_{n+2}, q_{n+1}\right) \tag{19}$$

$$\mathcal{S}\left(q_{n+3}|q_{n+2}, q_{n+1}\right) \triangleq -\sum_{i=1}^{|\Sigma|}\sum_{j=1}^{|\Sigma|} P\left(q_{n+2} = i, q_{n+2} = j\right) \cdot \mathcal{S}\left(q_{n+3}|q_{n+2} = i, q_{n+1} = j\right) \tag{20}$$

where

$$\mathcal{S}\left(q_{n+3}|q_{n+2} = i, q_{n+1} = j\right) = -\sum_{\ell=1}^{|\Sigma|} P\left(q_{n+3} = \ell|q_{n+2} = i, q_{n+1} = j\right) \cdot logP\left(q_{n+3} = \ell|q_{n+2} = i, q_{n+j} = j\right) \tag{21}$$

Based on Eq. (13) and Eqs. (15) to (18), the mutual information $\mathcal{I}\left(q_{n+3}; q_{n+2}\right)$ is calculated from the Π^1 matrix. Similarly, based on Eq. (14) and Eqs. (19) to (21), $\mathcal{I}\left(q_{n+3}; q_{n+2}, q_{n+1}\right)$ is calculated from the Π^2 matrix. Then, information gain (abbreviated as $\mathcal{I}_\mathcal{G}$) with $D = 2$ instead of $D = 1$ in the Markov chain construction is defined as:

$$\mathcal{I}_\mathcal{G} \triangleq \mathcal{I}\left(q_{n+3}; q_{n+2}, q_{n+1}\right) - \mathcal{I}\left(q_{n+3}; q_{n+2}\right) \tag{22}$$

F. Pertinent Results

This subsection presents test and validation of the concept of generalized Hilbert transform based on the time series data collected from a laboratory apparatus of nonlinear electronic systems. The test results are interpreted in terms of mutual information for $(D = 1)$ and $(D = 2)$ Markov chains for noise-contaminated data for different values of SNR and the parameter α (see Eq. (5)). The pertinent results on mutual information and information gain are presented in Fig. 4 and Fig. 5, respectively. Although results are shown only for $SNR = \infty, 10, 4$ and 0, several other experiments with intermediate values of SNR between ∞ and 0 were performed, which show the same trend.

The information gain is always a positive quantity as seen in Eq. (22). In other words, there is always a non-negative gain in information upon increasing the depth of the Markov chain model. Pertinent inferences, drawn from these results, are presented below.

1) Mutual information increases with decrease in α irrespective of D and SNR as seen in Fig. 4.
2) Information gain $\mathcal{I}_\mathcal{G}$ (see Eq. (22) and Fig. 5) is minimal for $SNR \to \infty$ (i.e., for the signal with no

noise injection). Therefore, D=1 Markov chain should be adequate with $ASSP$ using conventional Hilbert transform (see Eq. 1) for low-noise signals.

3) As SNR is decreased (i.e., percentage of additive noise is increased), information gain $\mathcal{I}_{\mathcal{G}}$ increases for all values of α in the range of 1.0 down to about 0.2. As α is decreased, information gain decreases as seen in Fig. 5. Therefore, even for a considerable amount of noise, a smaller value of α should be able to achieve noise attenuation and thus allow usage of $D = 1$ in D-Markov machines.
4) Results for a pathological case with $SNR \to 0$, (i.e., complete noise capture of the signal) in Fig. 4 and Fig. 5 show similar trends as above. The crossing of the information gain curves in Fig. 5 at low values of α (e.g., $\alpha \leq 0.2$) could possibly be attributed to the effects of coarse graining [13] due to symbol generation.

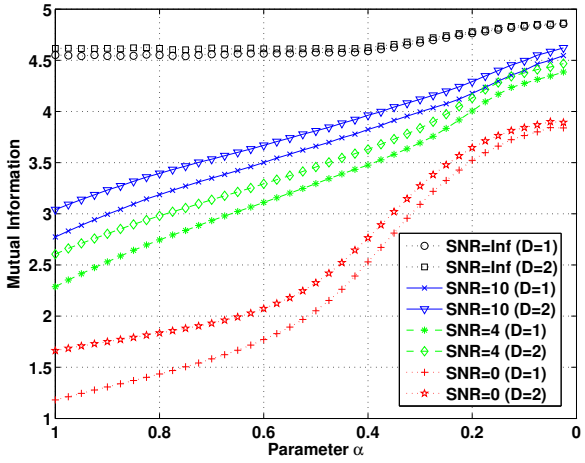


Fig. 4. Profiles of mutual information

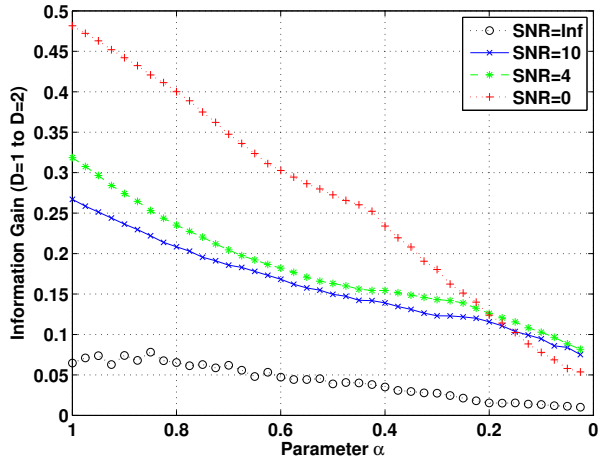


Fig. 5. Profiles of information gain

4. Summary, Conclusion and Future Work

This section summarizes the specific contributions of the paper highlighting the pertinent conclusions. Future research in several ares has also been recommended.

A. Summary

This paper formulates a generalization of the classical Hilbert transform along with mathematical proofs. The proposed scheme of generalized Hilbert transform is shown to be potentially useful for symbolic time series analysis of noise-corrupted dynamical systems. The proposed concept of noise reduction via generalization of Hilbert transform is tested and validated on experimental data collected from a laboratory test apparatus.

B. Pertinent Conclusions

The following conclusions are drawn from the test and validation results presented in the previous section.

- Generalized Hilbert transform with a smaller value of parameter α is capable of extracting more information from a data sequence irrespective of the depth of the D-Markov machine chosen for modeling.
- Information gain for a larger depth D reduces with smaller values of the parameter α.
- By selecting small values of the parameter α in the generalized Hilbert transform, it is possible to avoid using a computationally expensive larger depth D without loss of significant information.

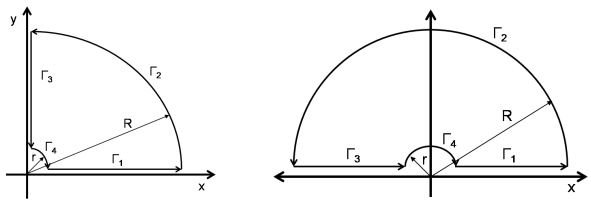
C. Recommendations for Future Research

The proposed method of noise attenuation via generalization of Hilbert transform is potentially useful for symbolic time series analysis of noise-corrupted dynamical systems. The future work should be directed toward advancement of the theory of partitioning as well as on application to different real-life uncertain systems. Examples include sensor networks that require on-board real-time analysis of noisy signals with very low computation capacity.

Appendix A. Proofs of Lemmas

This appendix presents the proofs of Lemma 2.1 and Lemma 2.2 that are stated in Section 2.

A. Proof of Lemma 2.1



(a) $0 < \alpha < 1$ (b) $\alpha = 1$

Fig. 6. Contours of integration paths for generalized Hilbert transform

Proof: The lemma is proved by integration around a quarter circular contour as shown in Fig. 6(a). Following Cauchy residual theorem, if z is a complex number, then

$$\int_{\Gamma} \frac{e^{iz}}{z^{\alpha}} dz = 0 \tag{23}$$

where the closed contour Γ contains no residues and consists of the paths Γ_1, Γ_2, Γ_3, and Γ_4. As $R \to \infty$ and $r \to 0$, the integral around Γ_2 and the integral around Γ_4 goes to 0, respectively. Therefore,

$$\int_{\Gamma_1} \frac{e^{iz}}{z^{\alpha}} dz + \int_{\Gamma_3} \frac{e^{iz}}{z^{\alpha}} dz = 0 \tag{24}$$

$$\Rightarrow \int_0^{\infty} \frac{e^{ix}}{x^{\alpha}} dx + \int_{\infty}^0 \frac{e^{-y}}{(iy)^{\alpha}} dy = 0 \tag{25}$$

The following two equations are derived from Eq. (25).

$$\int_0^{\infty} \frac{e^{ix}}{x^{\alpha}} dx = e^{i\frac{\pi}{2}(1-\alpha)} \Gamma(1-\alpha) \tag{26}$$

$$\int_0^{\infty} \frac{e^{-ix}}{x^{\alpha}} dx = e^{-i\frac{\pi}{2}(1-\alpha)} \Gamma(1-\alpha) \tag{27}$$

Now,

$$\int_{\infty}^{\infty} \frac{e^{-i\omega t}}{|t|^{\alpha}} sgn(t) dt = \int_0^{\infty} \frac{e^{-i\omega t}}{t^{\alpha}} dt - \int_{-\infty}^0 \frac{e^{-i\omega t}}{|t|^{\alpha}} dt \tag{28}$$

For $\omega > 0$, it follows from Eq. (27) that

$$\int_0^{\infty} \frac{e^{-i\omega t}}{t^{\alpha}} dt = \frac{\Gamma(1-\alpha)}{|\omega|^{1-\alpha}} e^{-i\frac{\pi}{2}(1-\alpha)} \tag{29}$$

and using Eq. (26)

$$\int_{-\infty}^0 \frac{e^{-i\omega t}}{|t|^{\alpha}} dt = \frac{\Gamma(1-\alpha)}{|\omega|^{1-\alpha}} e^{i\frac{\pi}{2}(1-\alpha)} \tag{30}$$

Similarly, for $\omega < 0$

$$\int_0^{\infty} \frac{e^{-i\omega t}}{t^{\alpha}} dt = \frac{\Gamma(1-\alpha)}{|\omega|^{1-\alpha}} e^{i\frac{\pi}{2}(1-\alpha)} \tag{31}$$

$$\int_{-\infty}^0 \frac{e^{-i\omega t}}{|t|^{\alpha}} dt = \frac{\Gamma(1-\alpha)}{|\omega|^{1-\alpha}} e^{-i\frac{\pi}{2}(1-\alpha)} \tag{32}$$

It follows from Eqs. (29), (30), (31), and (32) that

$$\int_{-\infty}^{\infty} \frac{e^{(-i\omega t)}}{|t|^{\alpha}} sgn(t) dt = \begin{cases} \frac{\Gamma(1-\alpha)}{|\omega|^{1-\alpha}} [e^{-i\frac{\pi}{2}(1-\alpha)} - e^{i\frac{\pi}{2}(1-\alpha)}] & \text{if } \omega > 0 \\ \frac{\Gamma(1-\alpha)}{|\omega|^{1-\alpha}} [e^{i\frac{\pi}{2}(1-\alpha)} - e^{-i\frac{\pi}{2}(1-\alpha)}] & \text{if } \omega < 0 \end{cases} \tag{33}$$

$$\Rightarrow \int_{-\infty}^{\infty} \frac{e^{(-i\omega t)}}{\pi |t|^{\alpha}} sgn(t) dt = -i \frac{2}{\pi} sgn(\omega) \frac{\Gamma(1-\alpha)}{|\omega|^{1-\alpha}} sin\big(\frac{\pi}{2}(1-\alpha)\big) \tag{34}$$

■

Given $\Gamma(z) = \int_0^{\infty} t^{(z-1)} e^{-t} dt$ for $Re(z) > 0$, it follows that $0 < \Gamma(1-\alpha) < \infty$ for $\alpha \in (0,1)$.

B. Proof of Lemma 2.2

Proof: Let $\theta \triangleq (1-\alpha)$ and $Z \triangleq \frac{2}{\pi} sin\big(\frac{\pi\theta}{2}\big)\Gamma(\theta)$. Then, for $\theta \in (0,1)$, it follows that

$$\begin{aligned} \lim_{\theta \to 0} Z &= \lim_{\theta \to 0} \frac{\Gamma(\theta) 2 sin\frac{\pi\theta}{2} cos\frac{\pi\theta}{2}}{\pi cos\frac{\pi\theta}{2}} \\ &= \frac{\lim_{\theta \to 0} \Gamma(\theta) sin(\pi\theta)}{\pi \lim_{\theta \to 0} cos\frac{\pi\theta}{2}} \\ &= \frac{\lim_{\theta \to 0} \frac{\pi}{\Gamma(1-\theta)}}{\pi} = 1 \end{aligned} \tag{35}$$

because $\forall z \in \mathbb{C}, \Gamma(z)\Gamma(1-z) = \frac{\pi}{sin(\pi z)}$ and $\Gamma(1) = 1$. ■

It follows from the proofs of the two lemmas that the integration path changes from a quarter circular contour to a half circular contour, as seen in Fig. 6(a) and Fig. 6(b), as $\alpha \uparrow 1$.

REFERENCES

[1] D. Gabor, "Theory of communications," *J. Inst. Electrical Engineering*, vol. 93, p. 429457, 1946.

[2] L. Cohen, *Time-Frequency Analysis*. Prentice Hall PTR, 1995.

[3] A. Subbu and A. Ray, "Space partitioning via hilbert transform for symbolic time series analysis," *Applied Physics Letters*, vol. 92, p. 084107, February 2008.

[4] M. Buhl and M. Kennel, "Statistically relaxing to generating partitions for observed time-series data," *Physical Review E*, vol. 71, no. 4, p. 046213, 2005.

[5] V. Rajagopalan and A. Ray, "Symbolic time series analysis via wavelet-based partitioning," *Signal Processing*, vol. 86, no. 11, pp. 3309–3320, 2006.

[6] A. Ray, "Symbolic dynamic analysis of complex systems for anomaly detection," *Signal Processing*, vol. 84, no. 7, pp. 1115–1130, 2004.

[7] C. Rao, A. Ray, S. Sarkar, and M. Yasar, "Review and comparative evaluation of symbolic dynamic filtering for detection of anomaly patterns," *Signal, Image and Video Processing*, 2008. DOI 10.1007/s11760-008-0061-8.

[8] A. Lohmann, D. Mendlovic, and Z. Zalevsky, "Fractional hilbert transform," *Opt. Lett.*, vol. 21, no. 4, p. 281283, 1996.

[9] S.-C. Pei and M.-H. Yeh, "Discrete fractional hilbert transform," *IEEE Transactions on Circuits and SystemsII: Analog and Digital Signal Processing*, vol. 47, no. 11, 2000.

[10] Y. Luo, S. Al-dossary, M. Mahroon, and M. Alfaraj, "Generalized hilbert transform and its applications in geophysics," *The Leading Edge*, March 2003.

[11] T. Cover and J. Thomas, *Elements of Information Theory*. New York, NY, USA: John Wiley and Sons, Inc., 1991.

[12] A. Ray, "Signed real measure of regular languages for discrete-event supervisory control," *Int. J. Control*, vol. 78, no. 12, pp. 949–967, 2005.

[13] C. Beck and F. Schlögl, *Thermodynamics of chaotic systems: an introduction*. Cambridge University Press, United Kingdom, 1993.