EXPLOITING SIGNAL NONGAUSSIANITY AND NONLINEARITY FOR PERFORMANCE ASSESSMENT OF ADAPTIVE FILTERING ALGORITHMS: QUALITATIVE PERFORMANCE OF KALMAN FILTER

Mo Chen[1]*, Temujin Gautama*[2]*, Dragan Obradovic*[3]*, Jonathon Chambers*[4] *and Danilo Mandic*[1]

[1]Department of Electrical and Electronic Engineering
Imperial College London, Exhibition Road, London, SW7 2BT, U.K.
E-mail: {mo.chen, d.mandic}@imperial.ac.uk

[2]Philips Leuven, Interleuvenlaan 80, B-3001 Leuven, Belgium
E-mail: temujin.gautama@philips.com

[4]Centre of Digital Signal Processing, School of Engineering, Cardiff University
E-mail: ChambersJ@cardiff.ac.uk

[5]Siemens AG, Corporate Technology, Otto-Hahn-Ring 6, D-81739 Munich, Germany
E-mail: Dragan.Obradovic@siemens.com

ABSTRACT

A new framework for the assessment of the qualitative performance of Kalman filter is proposed. This is achieved by the recently proposed 'Delay Vector Variance' (DVV) method for the signal modality characterisation, which is based upon the local predictability in the phase space. It is shown that Kalman filter not only outperforms common linear and nonlinear filters in terms of quantitative performance but also achieves a better qualitative performance. A set of comprehensive simulations on representative data sets supports the analysis.

1. INTRODUCTION

Much research has been dedicated towards devising 'optimal' adaptive filtering architectures and algorithms, according to a certain pre-defined optimization criterion [1][2], most typically based on the mean square error (MSE). Kalman filter is such an optimal estimator within the framework of second-order statistics, and has been well established in state-space based estimation. This is achieved based on the knowledge of the system inputs and outputs, and based on a model of the relation between them [3]. To assess its performance, various performance measures have been proposed, usually based on some sort of signal to error power ratio [4].

This way, it is possible to introduce so-called *quantitative* performance indexes which reflect the 'goodness' of estimation (ratio between the powers of the signal, noise and error), however these criteria cannot provide insight into so-called *qualitative* performance of adaptive filters; to this end, open literature provides various approaches including those based on the preservation of the nature of a signal, perceptual quality estimation, and probability density function (PDF) matching. In fact, little has been known about the characterising the nature[1] of the signal before and after it is processed by an adaptive filter. The importance of such characterisation is two-fold. On one hand, it is desirable to verify the presence of underlying linear or nonlinear nature of the signal, before actual adaptive filters are employed. As a matter of fact, in the absence of nonlinearity, it is not preferable to use nonlinear adaptive filters despite of their ability to process linear signals [5], because they are known to be more computationally complex, slower converging and more sensitive to parameter perturbation than their linear counterparts [1][4].

On the other hand, since the1990s, research on 'signal modality characterisation' has been established and has just started finding its applications in signal processing. For instance, in the analysis of biomedical signals (heart rate variability (HRV), electro-encephalogram (EEG)), we need to assess the presence/absence of the nonlinear behavior within the signals, since this may convey essential information about the signal generation system. In addition, in biomedical and en-

[1]In this paper, we refer "nature" to four fundamental properties of a signal, that is linear, nonlinear, deterministic and stochastic signal behaviour, which will be addressed later in the following section.

vironmental applications, the nature of a signal conveys important information about the underlying signal generation mechanism:- 1). In electrocardiogram (ECG) and heart rate variability (HRV) signals, where the change in the nature provides an indication of health hazard [6]; 2). Similar phenomena have been reported in the analysis of air pollutants [7] and brain electrical activity [8]. Hence, in these cases, it is essential that during processing of such signals we not only optimise for the 'best' performance in terms of a certain quantitative performance criterion, but also that the filter preserves the basic signal modality (nature of the signal). If the nature of the signal has significantly changed after being processed (e.g. prediction within compression algorithms and denoising), the application of such filters will be greatly limited.

To that end, we provide an qualitative assessment of qualitative performance of Kalman filter, supported by quantitative performance measures (prediction) together with some illustrative examples highlighting the need to take into account the nature of processed signals. This is supported by comprehensive simulations on linear and nonlinear benchmark signals.

2. "DELAY VECTOR VARIANCE" (DVV) METHOD AND NOVEL DVV SCATTER DIAGRAM

Before we introduce the 'Delay Vector Variance' method, there is a need to give definitions of some notions which will be frequently referred to in the remainder of the paper.

2.1. Background

Following the approaches from the physics literature, we introduce the following definitions [9]:-

1. Linear Signal (strict definition) – A linear signal is generated by a linear time-invariant system, driven by white Gaussian noise;
2. Linear Signal (commonly adopted) – Definition 1) is relaxed somewhat by allowing the distribution of the signal to deviate from the Gaussian one, which can be interpreted as a linear signal from 1), measured by a static, monotonic, and possibly nonlinear observation function;
3. Nonlinear Signal – A signal that cannot be generated in the above way is considered nonlinear;
4. Deterministic (predictable) Signal – A signal is considered deterministic if it can be precisely described by a set of equations;
5. Stochastic Signal – A signal that is not deterministic.

For simplicity, we shall refer to 'nature' as the first two signal properties from the above.

2.2. "Delay Vector Variance" (DVV) method

Several methods for detecting nonlinear nature of a signal have been proposed over the past few years, which include the *'Deterministic versus Stochastic'* (DVS) plots [10], the Correlation Exponent, and 'δ-ε' method [11]. The recently introduced DVV method [12] is able to simultaneously examines both the nonlinear/linear and deterministic/stochastic nature of a signal, which can be summarised as follows: For an given embedding dimension m:

- Generate delay vectors (DVs): $\mathbf{x}(k) = [x_{k-m}, \dots, x_{k-1}]^T$ and the corresponding target x_k,
- The mean μ_d and standard deviation σ_d are computed over all pairwise Euclidean distances between DVs, $\|\mathbf{x}(i) - \mathbf{x}(j)\|$ $(i \neq j)$,
- The sets $\Omega_k(r_d)$ are generated such that $\Omega_k(r_d) = \{\mathbf{x}(i) | \|\mathbf{x}(k) - \mathbf{x}(i)\| \leq r_d\}$, *i.e.*, sets which consist of all DVs that lie closer to $\mathbf{x}(k)$ than a certain distance r_d, taken from the interval [max$\{0, \mu_d - n_d\sigma_d\}$; $\mu_d + n_d\sigma_d$], where n_d is a parameter controlling the span over which to perform the DVV analysis,
- For every set $\Omega_k(r_d)$, the variance of the corresponding targets, $\sigma_k^2(r_d)$, is computed. The average over all sets $\Omega_k(r_d)$, normalised by the variance of the time series, σ_x^2, yields the 'target variance', or DVV values, $\sigma^{*2}(r_d)$:

$$\sigma^{*2}(r_d) = \frac{\frac{1}{N}\sum_{k=1}^{N} \sigma_k^2(r_d)}{\sigma_x^2} \quad (1)$$

As r_d increases, the target variance smoothly converges to unity. This is because all DVs start to belong to the same universal set, and the variance of targets is equal to the variance of the time series.

As a result of the standardisation of the distance axis, the resulting 'DVV plot' (target variance, $\sigma^{*2}(r_d)$ as a function of the standarised[2], $\frac{r_d - \mu_d}{\sigma_d}$ distance), are easy to interpret, as illustrated in Figure 1(a) and Figure 1(b). The minimal target variance, *e.g.*, the lowest point of the curve, is a measure for the amount of noise which is present in the time series. The presence of a strong deterministic component will lead to small target variances for small spans. At the extreme right, the DVV plots smoothly converge to unity, since for maximum spans, all DVs belong to the same set, and the variance of the targets is equal to the variance of the time series.

[2]Note that we use the term 'standarised' in the statistical sense, namely as having zero mean and unit variance.

2.3. Signal nature Characterisation: DVV Scatter Diagram

In the following step, the linear or nonlinear nature of the time series is examined by performing DVV analyses on both the original and a number[3] of surrogate[4] time series, using the optimal embedding dimension of the original time series. Due to the standardisation of the distance axis, these plots can be conveniently combined within a *scatter diagram*, where the horizontal axis corresponds to the DVV plot of the original time series, and the vertical to that of the surrogate time series. If the surrogate time series yield DVV plots similar to that of the original time series, as illustrated by Figure 1(a), the DVV scatter diagram coincides with the bisector line, and the original time series is judged to be linear, as shown in Figure 1(c) (for the linear signal (2)). If not, as illustrated by Figure 1(b), the DVV scatter diagram will deviate from the bisector line and the original time series is judged to be nonlinear, as depicted in Figure 1(d) (for the nonlinear signal (3)).

To illustrate the usage of DVV scatter diagram, consider a linear signal (AR(4)), given by [4]

$$\begin{aligned} x(k) &= 1.79\,x(k-1) - 1.85\,x(k-2) + 1.27\,x(k-3) \\ &\quad - 0.41\,x(k-4) + n(k) \end{aligned} \tag{2}$$

and a nonlinear signal (a Narendra Model Three realisation), given by [14]

$$\begin{aligned} z(k) &= \frac{z(k-1)}{1+z^2(k-1)} + r^3(k) \\ r(k) &= 1.79\,r(k-1) - 1.85\,r(k-2) + 1.27\,r(k-3) \\ &\quad - 0.41\,r(k-4) + n(k) \end{aligned} \tag{3}$$

where $\{n(k)\}$ is white Gaussian noise $n(k) \in \mathcal{N}(0,1)$.

If the surrogate time series yield DVV plots similar to that of the original time series, as illustrated by Figure 1(a), the DVV scatter diagram coincides with the bisector line, and the original time series is judged to be linear, as shown in Figure 1(c) (for the linear signal (2)). If not, as illustrated by Figure 1(b), the DVV scatter diagram will deviate from the bisector line and the original time series is judged to be nonlinear, as depicted in Figure 1(d) (for the nonlinear signal (3)).

2.4. Qualitative and Quantitative Performance Analysis

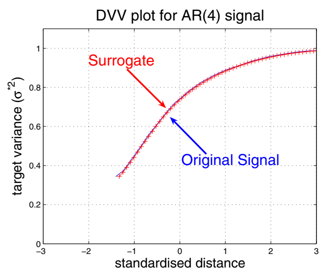
Notice that the prediction is the core of adaptive algorithms in machine learning applications, both supervised and blind. To assess the *quantitative* performance of learning algorithms, it is convenient to use the standard one-step forward prediction gain R_p, defined as [15]

$$R_p = 10\log_{10}\left(\frac{\hat{\sigma}_s^2}{\hat{\sigma}_e^2}\right)[dB] \tag{4}$$

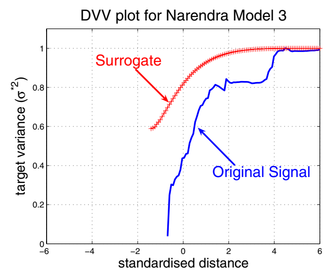
which is a logarithmic ratio between the estimated signal variance $\hat{\sigma}_s^2$ and estimated prediction error variance $\hat{\sigma}_e^2$.

On the other hand, to assess the *qualitative* performance, that is, *a possible change in the signal nature introduced by a filter*, we compare DVV scatter diagrams of the outputs of the learning algorithms with those of the original signal. The target variances Eq. (1) for the predicted signal and its surrogates are obtained by performing the DVV test on the predicted signal. For robustness, these steps are repeated 100 times and DVV scatter diagrams are obtained by plotting the averaged target variance against that for its surrogates.

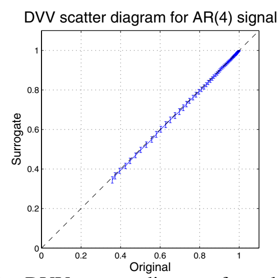
For the assessment of qualitative performance, if the considered filters yield high prediction gain (R_p), the quantitative performance of the filters is judged to be 'good'. As for the qualitative performance as explained above, the more similar the DVV scatter diagram for the filtered signal is to that for the original signal, the better the qualitative performance of the considered filter.



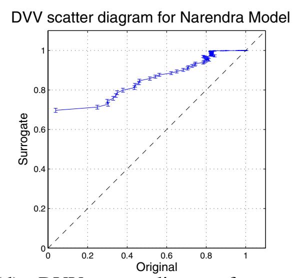
(a) DVV plot for a linear signal (AR(4) signal). The line with crosses denotes the DVV plot for the average of 25 iAAFT-based surrogate while the solid line denotes that for the original signal.



(b) DVV plot for a nonlinear signal (Narendra Model 3). The line with crosses denotes the DVV plot for the average of 25 iAAFT-based surrogate while the solid line denotes that for the original signal.



(c) DVV scatter diagram for a linear signal (AR(4) signal). Error bars denote the standard deviation of the target variances of surrogates.



(d) DVV scatter diagram for a nonlinear signal (Narendra Model 3). Error bars denote the standard deviation of the target variances of surrogates.

Fig. 1. Nonlinear and deterministic nature of signals.

---

[3] In all of our simulations, we choose to generate 25 surrogate each time as increasing the number of surrogate will not improve experiment results but increase the computational complexity.

[4] Simply speaking, surrogate data, or "surrogate" for short, are artificially generated randomised linear version of the original data. For more information on surrogate data, please refer to [13][9].

3. EXPERIMENTAL RESULTS

For rigour, the qualitative assessment of Kalman filter was performed in comparison with that of a linear filter trained by the standard least-mean-square (LMS) algorithm and a nonlinear feedforward perceptron trained by the nonlinear version of LMS, *e.g.*, nonlinear gradient descent (NGD) algorithm [4].

Figure 2 illustrates the qualitative and quantitative performance comparison of three different adaptive filters on the prediction of the linear benchmark signal (2). From the Figure, all the filters were able to preserve the linear nature of the processed signal, as illustrated by the fact that all the DVV scatter diagrams lie on the bisector line. However, in terms of quantitative performance, Kalman filter outperformed the other two adaptive filters with a noticeable increase in the prediction gain.

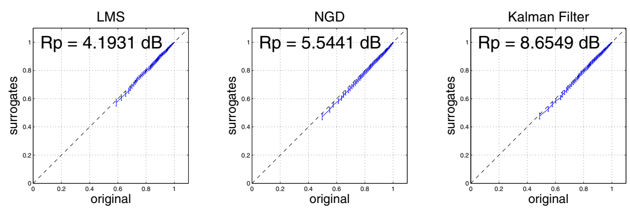


Fig. 2. Quantitative and qualitative performance comparison of three different filters on the prediction of the linear benchmark signal (2): the linear filter trained by LMS algorithm (leftmost diagram), the nonlinear feedforward perceptron trained by NGD algorithm (middle diagram) and the Kalman filter (rightmost diagram).

Figure 3 illustrates a similar experiment performed on prediction of the benchmark nonlinear signal (3). From Figure 3, Kalman filter not only had the best quantitative performance in terms of prediction gain but also has the best qualitative performance in terms of signal nature preservation, judged by the fact that the DVV scatter diagram of the prediction (solid line) was closest to that of the original signal (dotted line) out of all the thee diagrams.

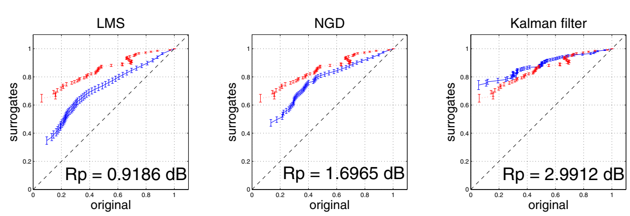


Fig. 3. Quantitative and qualitative performance comparison of three different filters on the prediction of the nonlinear benchmark signal (3). The dotted lines in the above diagrams stand for DVV scatter diagram for the original signal, while the solid lines for DVV scatter diagrams for one-step ahead prediction.

4. CONCLUSION

A novel framework for assessing the qualitative performance of the Kalman filter has been proposed. This has been achieved based upon the recently introduced ‘delay vector variance’ (DVV) method for characterising the nonlinearities present in the original signal. It has been shown that the Kalman filter not only has a better quantitative performance in terms of higher prediction gain, but also is able to better preserve the nature of the processed signal, as compared to standard linear and nonlinear filters. Simulations on both the linear and nonlinear benchmark signals support the findings.

5. REFERENCES

[1] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice Hall, 1994.

[2] B. Widrow, J. McCool, and M. Ball, “The complex lms algorithm,” *Proceedings of the IEEE*, vol. 63, pp. 712–720, 1975.

[3] R. E. Kalman, “New approach to linear filtering and prediction problems,” *Transactions of the ASME-Journal of Basic Engineering*, vol. 82, pp. 35–45, 1960.

[4] D. P. Mandic and J. A. Chambers, *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability*, John Wiley & Sons, 2001.

[5] T. Adali, “Why a nonlinear solution for a linear problem?,” *in Proceedings IEEE Workshop on Neural Networks for Signal Processing (NNSP), Madison, WI*, pp. 157–165, 1999.

[6] C. S. Poon and C. K. Merrill, “Decrease of cardiac chaos in congestive heart failure,” *Nature*, vol. 389, pp. 492–495, 1997.

[7] S. R. Dorling, R. J. Foxall, D. P. Mandic, and G. C. Cawley, “Maximum likelihood cost functions for neural network models of air quality data,” *Atomospheric Environment*, vol. 37, pp. 3435–3443, 2003.

[8] T. Gautama, D. P. Mandic, and M. M. Van Hulle, “Indications of nonlinear structures in brain electrical activity,” *Phys. Rev. E*, vol. 67, pp. 046204–1–046204–5, 2003.

[9] T. Schreiber and A. Schmitz, “Surrogate time series,” *Physica D*, vol. 142, pp. 346–382, 2000.

[10] M. C. Casdagli and A. S. Weigend, “Exploring the continuum between deterministic and stochastic modeling,” in *Time Series Prediction: Forecasting the Future and Understanding the Past*, A. S. Weigend and N. A. Gershenfeld, Eds., pp. 347–367. Reading, MA: Addison-Wesley, 1994.

[11] D. Kaplan, “Exceptional events as evidence for determinism,” *Physica D*, vol. 73, no. 1, pp. 38–48, 1994.

[12] T. Gautama, D. P. Mandic, and M. M. Van Hulle, “The delay vector variance method for detecting determinism and nonlinearity in time series,” *Physica D*, vol. 190, no. 3-4, pp. 167–176, 2004.

[13] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. Farmer, “Testing for nonlinearity in time series: the method of surrogate data,” *Physica D*, vol. 58, pp. 77–94, 1992.

[14] K.S. Narendra and K. Parthasarathy, “Identification and control of dynamical systems using neural networks,” *IEEE Trans. Neural Networks*, vol. 1, no. 1, pp. 4–27, 1990.

[15] S. Haykin and L. Li, “Nonlinear adaptive prediction of nonstationary signals,” *IEEE Transactions on Signal Processing*, vol. 43, no. 2, pp. 526–535, 1995.