lants, require large numbers of samples for reliable estimation and hence are not practical to use in this case.

Conclusions: MFR coupled with CMA provides a direct method of blind channel estimation in a symbol-spaced configuration without explicit use of higher-order statistics of the channel output.

© IEE 2003
22 May 2003
Electronics Letters Online No: 20030807
DOI: 10.1049/el:20030807

B. Baykal (*Department of Electrical and Electronic Engineering, Middle East Technical University, Ankara 06531, Turkey*)

E-mail: buyurman@metu.edu.tr

A.G. Constantinides (*Department of Electrical and Electronic Engineering, Imperial College of Science, Technology and Medicine, London SW7, United Kingdom*)

References

1 TONG, L., XU, G.H., and KAILATH, T.: 'Blind identification and equalization based on 2nd-order statistics—a time-domain approach', *IEEE Trans. Inf. Theory*, 1994, **40**, pp. 340–349
2 LABAT, J., MACCHI, O., and LAOT, C.: 'Adaptive decision feedback equalization: can you skip the training period?', *IEEE Trans. Commun.*, 1998, **46**, pp. 921–930
3 SEO, B.S., LEE, J.H., and LEE, C.W.: 'Approach to blind decision feedback equalisation', *Electron. Lett.*, 1996, **32**, pp. 1639–1640
4 FORNEY, G.D.: 'Maximum-likelihood sequence estimation of digital sequences in presence of intersymbol interference', *IEEE Trans. Inf. Theory*, 1972, **18**, pp. 363–378
5 TANRIKULU, O., BAYKAL, B., CONSTANTINIDES, A.G., and CHAMBERS, J.: 'Soft constraint satisfaction (SCS) blind channel equalisation algorithm', *Int. J. Adapt. Control Signal Process.*, 1998, **12**, pp. 117–134
6 ZENG, H.H., TONG, L., and JOHNSON, C.R.: 'Relationships between the constant modulus and Wiener receivers', *IEEE Trans. Inf. Theory*, 1998, **44**, pp. 1523–1538

Stochastic resonance and data processing inequality

M.D. McDonnell, N.G. Stocks, C.E.M. Pearce and D. Abbott

The data processing inequality of information theory proves that no more information can be obtained out of a set of data than was there to begin with. However, many papers in the field of stochastic resonance report signal to noise ratio gains in some nonlinear systems due to the addition of noise. Such an observation appears on the surface to contradict the data processing inequality. It is demonstrated that the data processing inequality is upheld for the case of a periodic input signals.

Introduction: The data processing inequality (DPI) states that given random variables X, Y and Z that form a Markov chain in the order $X \to Y \to Z$, then the mutual information between X and Y is greater than or equal to the mutual information between X and Z [1]. That is $I(X; Y) \geq I(X; Z)$. In other words, no signal processing on Y can increase the information that Y contains about X. However, in the field of stochastic resonance (SR) [2, 3] many papers report that it is possible to obtain a signal to noise ratio (SNR) gain in some nonlinear systems by the addition of noise [4, 5]. To a casual observer, this would seem to contradict the DPI.

Background: An SNR gain is not in itself a remarkable thing; SNR gains are routinely obtained by filtering. However in the SR literature, the reported SNR gains are said to be due to stochastic resonance, rather than a deliberately designed filter, which is why the SNR gains are taken to be quite remarkable. SR is the term given to the phenomenon where the optimal output of a nonlinear system occurs for nonzero noise in that system. It was at first thought to occur only in bistable dynamical systems, generally driven by a periodic input signal and broadband noise. The work on such systems showed that the ratio of the output power at the input frequency to the background noise spectral density at the same frequency could be maximised by a nonzero value of noise intensity.

It has been proven using linear response theory that for the case of stationary Gaussian noise and a signal that is small compared to the noise, that for nonlinear systems the SNR gain must be less than or equal to unity, and that hence no SNR gain can be induced by utilising the SR effect [6]. Once this fact was established, researchers still hoping to be able to find systems in which SNR gains due to noise could occur, turned their attention to situations not covered by the proof, that is, the case of a signal that is not small compared to the noise, or broadband signals or non-Gaussian noise.

For broadband or aperiodic input signals, SNR is not appropriate, and methods such as cross-correlation [2] and mutual information [7] are the most commonly used to show that SR can occur for nonperiodic signals. However to compare the detectability at the output to the input for broadband input signals, a measure analogous to the input–output SNR ratio for periodic input signals is required. One such possible measure is channel capacity. This measure has been used by Chapeau-Blondeau [4], who stated that comparing the channel capacity at the input and output for an aperiodic input signal was analogous to a comparison of the input and output SNRs for periodic input signals. His work indicated that it was possible for the channel capacity at the output to exceed that at the input. This appears *prima facie* to be a clear contradiction of the DPI, and we resolve this in the following.

Model: Consider a system where a signal, s, is subject to independent additive random noise, n, to form another random signal, $y = s + n$. The signal y is then subjected to a nonlinear function, g, to give a final random signal, $z = g(y)$. Since we have $z = g(y)$, z is conditionally independent of s, our model forms a Markov chain and therefore the DPI applies. We consider the case of s consisting of a random binary pulse train. Hence, our model is similar to those reported to show SNR gains for aperiodic input signals [4].

Example: Let the input s take on values $\pm s_v$. The output, z is one of three values (s_v, $-s_v$ and 0) and is determined by two thresholds, $\pm\theta$ such that $g(y)$ is given by

$$z = g(y) = \begin{cases} s_v & \text{if } y = s + n > \theta \\ -s_v & \text{if } y = s + n < -\theta \\ 0 & \text{otherwise} \end{cases}$$

where an output value of zero indicates the complete erasure of an input value, rather than its corruption.

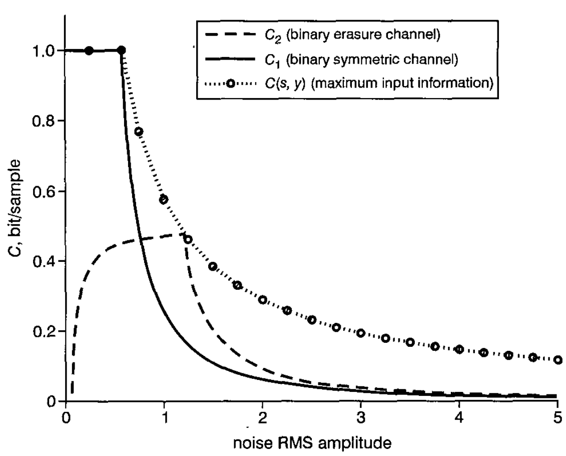

Fig. 1 *Channel capacity against RMS noise amplitude for three cases*

If we have $\pm\theta = 0$, and a probability of error given the input, $p_e = p(z = s_v | x = -s_v) = p(z = -s_v | x = s_v)$ then the channel is a binary symmetric channel, for which it is known that channel capacity occurs when $p(\pm s_v) = 0.5$ and is given by $C_1 = 1 + p_e \log_2(p_e) + (1 - p_e)\log_2(1 - p_e)$ [1]. If the noise in the channel has an even probability density then $p_e = F_n(s_v)$, where F_n is the cumulative distribution function of the noise.

For the more general case of $\theta > 0$, the channel can be considered as a binary erasure channel with error. This is a symmetric memoryless discrete channel provided the noise has an even probability density function. Hence capacity is also achieved when $p(\pm s_v) = 0.5$. Formulas were derived in [4] for the channel capacity for such a channel, which we have denoted as C_2 and plotted in Fig. 1 against the RMS value of a uniform noise distribution when $\theta = \pm 1.1$. It also shows the capacity of the binary symmetric channel, C_1, for the same noise. Note that the maximum value of the capacity for the binary erasure channel corresponds to a nonzero value of noise. As noted in [4], this shows that SR can occur in the channel capacity measure.

However, in [4] C_1 is considered to be the maximum mutual information at the input of the binary erasure channel, and C_2 to be the maximum mutual information at the output of the binary erasure channel. When the RMS noise amplitude is such that $C_2 > C_1$, this is interpreted to mean that the capacity at the output of the binary erasure channel is greater than the capacity at the input. However C_1 is the channel capacity of the binary symmetric channel, not the erasure channel, and hence cannot be considered as the maximum input mutual information to the binary erasure channel. Thus [4] has shown only that for the values of RMS noise where the ratio $C_2/C_1 > 1$, more information can be obtained about s by the binary erasure channel than for the binary symmetric channel.

The actual capacity at the input, which we denote as $C(s, y)$, should be taken as the maximum of the mutual information between the input signal, s, and the input signal plus noise, $y = s + n$, over all probability distributions of the input symbols, s. Owing to symmetry, this maximum occurs when $p(\pm s_v) = 0.5$ and we can derive $C(x, y)$ as

$$C(s,y) = \begin{cases} 1 & \sigma_n < 2 \\ \dfrac{2}{\sigma_n} & \sigma_n \geq 2 \end{cases}$$

where $\sigma_n/12$ is the variance of the uniform noise. $C(s, y)$ is plotted in Fig. 1, where it can be seen that both C_1 and C_2 are less than or equal to $C(s, y)$ and that hence the DPI holds. For the binary symmetric channel, where there is a noise induced maximum in the capacity, this maximum can be interpreted to mean that a certain nonzero value of noise can minimise the information lost in the channel.

The majority of studies indicating the possibility of SNR gains due to stochastic resonance make use of periodic input signals in the large signal domain (i.e. signal and noise outside the validity of linear response theory) and SNR measures (for example [5]). However it has previously been pointed out [8] that for such circumstances, the SNR measure discards some statistical information due to the nonlinearity involved. For example, if knowledge of the shape of the input signal is required at the output, rather than simply a reduced SNR for the fundamental frequency, then any possible SNR gain is irrelevant, and information theoretic or correlation based measures must be used. As shown in this Letter the DPI applies, and although an SNR gain might occur for the amplitude of the input frequency, that does not mean that more information about the input signal is available at the output.

Conclusions: Although SNR gains have been shown to exist for periodic input signals outside the linear response region, for random noisy aperiodic signals the DPI holds and the addition of more noise to a noisy signal cannot be of benefit as far as obtaining a mutual information gain is concerned. This result gives further weight to arguments that the SNR measure used for the case of periodic signals is inadequate, and that situations where SNR gains occur are limited in their potential usefulness.

Acknowledgments: This work was funded by the Australian Research Council and the Leverhulme Trust, grant number F/00 215/J. We also warmly thank F. Chapeau-Blondeau of the Université d'Angers, France, for useful discussions.

© IEE 2003

13 May 2003

Electronics Letters Online No: 20030792

DOI: 10.1049/el:20030792

M.D. McDonnell and D. Abbott (*Centre for Biomedical Engineering (CBME) and School of Electrical & Electronic Engineering, The University of Adelaide, SA 5005, Australia*)

E-mail: mmcdonne@eleceng.adelaide.edu.au

N.G. Stocks (*School of Engineering, The University of Warwick, Coventry CV4 7AL, United Kingdom*)

C.E.M. Pearce (*School of Applied Mathematics, The University of Adelaide, SA 5005, Australia*)

References

1 COVER, T.M., and THOMAS, J.A.: 'Elements of information theory' (John Wiley and Sons, 1991)

2 HARMER, G.P., DAVIS, B.R., and ABBOTT, D.: 'A review of stochastic resonance: circuits and measurement', *IEEE Trans. Instrum. Meas.*, 2002, **51**, (2), pp. 299–309

3 GODIVIER, X., ROJAS-VARELA, J., and CHAPEAU-BLONDEAU, F.: 'Noise-assisted signal transmission via stochastic resonance in a diode nonlinearity', *Electron. Lett.*, 1997, **33**, (20), pp. 1666–1668

4 CHAPEAU-BLONDEAU, F.: 'Periodic and aperiodic stochastic resonance with output signal-to-noise ratio exceeding that at the input', *Int. J. Bifurcation Chaos*, 1999, **9**, (1), pp. 267–272

5 GINGL, Z., MAKRA, P., and VAJTAI, R.: 'High signal-to-noise ratio gain by stochastic resonance in a double well', *Fluctuation and Noise Letters*, 2002, **1**, (3), pp. L181–L188

6 DYKMAN, M.I., LUCHINSKY, D.G., MANNELLA, R., MCCLINTOCK, P.V.E., STEIN, N.D., and STOCKS, N.G.: 'Stochastic resonance in perspective', *Il Nuovo Cimento*, 1995, **17**, (7–8), pp. 661–683

7 DEWEESE, M., and BIALEK, W.: 'Information flow in sensory neurons', *Nuovo Cimento D*, 1995, **17**, (7–8), pp. 733–741

8 ROBINSON, J.W.C., RUNG, J., BULSARA, A.R., and INCHIOSA, M.E.: 'General measures for signal-noise separation in nonlinear dynamical systems', *Phys. Rev. E*, 2001, **63**, (1), p. 0111071

Speech modelling by model-order reduction: SNR behaviour

L. Mitiche, B. Derras and A.B.H. Adamou-Mitiche

Using model reduction, a new approach for low-order speech modelling is presented. The process starts with a relatively high-order (full-order) autoregressive (AR) model obtained by some classical methods. The AR model is then reduced using the state projection method. The model reduction yields a reduced-order autoregressive moving-average model which interestingly preserves the key properties of the original model such as stability, minimality, and phase minimality. SNR behaviour are investigated. To illustrate the performance and the effectiveness of the proposed approach, computer simulations are conducted on practical speech segments.

Introduction: The purpose of any speech model synthesis is to determine the model parameters corresponding to a pre-selected order. However, the complexity of the synthesised model should be as low as possible while preserving the key properties of the speech being modelled. Classically, speech modelling is a very well-known process that can be accomplished using one of several AR methods, such as Levinson algorithm, maximum entropy etc. [1]. This usually leads to some models with over-estimated orders, i.e. with high complexity. To overcome this inherent problem, a model reduction approach is suggested. Starting with a full-order speech model, synthesised by one of the forementioned AR techniques, a reduced-order model is then obtained via some state-space projection scheme, in which the fundamental properties of the original model are preserved. To start the process assume that the original model has the following discrete-time state-space representation

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned} \tag{1}$$

where, $u(k)$, $x(k)$, and $y(k)$ are, respectively, the input (excitation), the n-dimensional state vector, and the output of the underlying model. The state-space parameters A, B, C, and D are matrices of dimensions $n \times n$, $n \times 1$, $1 \times n$, and 1×1, respectively. The purpose of any