since $(m+1)E\{u^2(k)\} > 0$. Using $E\{y^2(k)\} > \sigma_v^2$ produces

$$\Pr\left[nE\{y^2(k)\} - (E\{y^2(k)\} - \sigma_v^2)\sum_{i=1}^{n} a_i \geq 0\right] = \Pr\left[\sum_{i=1}^{n} a_i \leq \frac{nE\{y^2(k)\}}{E\{y^2(k)\} - \sigma_v^2}\right]. \quad (7)$$

Then, it results that

$$\Pr\left[\sum_{i=1}^{n} a_i \leq \frac{nE\{y^2(k)\}}{E\{y^2(k)\} - \sigma_v^2}\right] > \Pr\left[\sum_{i=1}^{n} a_i < n\right]. \quad (8)$$

From (6)–(8), we get

$$\Pr[\mathrm{tr}\,\mathbf{R}_{z\varphi} \geq 0] > \Pr\left[\sum_{i=1}^{n} a_i < n\right]. \quad (9)$$

Therefore, it follows that

$$\Pr[\mathrm{tr}\,\mathbf{R}_{z\varphi} < 0] < 1 - \Pr\left[\sum_{i=1}^{n} a_i < n\right]. \quad (10)$$

Now, we need to compute $\Pr\left[\sum_{i=1}^{n} a_i < n\right]$. We can express the denominator of the unknown system as

$$z^n + a_1 z^{n-1} + \cdots + a_n = \prod_{i=1}^{n}(z - c_i)$$

where c_is are real or complex number. Substituting $z = 1$, we get

$$1 + \sum_{i=1}^{n} a_i = \prod_{i=1}^{n}(1 - c_i).$$

Therefore, it follows that

$$\Pr\left[\sum_{i=1}^{n} a_i < n\right] = \Pr\left[\prod_{i=1}^{n}(1 - c_i) < n + 1\right]. \quad (11)$$

Let $c_i = \alpha_i + j\beta_i$, where $j^2 = -1$. Since c_i is inside the unit circle, $\alpha_i^2 + \beta_i^2 < 1$. Assume that c_is are mutually independent and uniformly distributed inside the unit circle. Then, if c_is are real, (11) can be written as

$$\Pr\left[\sum_{i=1}^{n} a_i < n\right] = \Pr\left[\prod_{i=1}^{n}(1 - \alpha_i) < n + 1\right] \quad (12)$$

where $|\alpha_i| < 1$. In the case where c_is are complex and n is even ($n = 2p$), (11) can be written as

$$\Pr\left[\sum_{i=1}^{n} a_i < n\right] = \Pr\left[\prod_{i=1}^{p}\left((1 - \alpha_i)^2 + \beta_i^2\right) < 2p + 1\right]. \quad (13)$$

Equations (12) and (13) can be computed using the Monte Carlo simulation. For the simulation, 10^5 samples are randomly chosen such that $\alpha_i^2 + \beta_i^2 < 1$, and each result is averaged over ten independent trials.

REFERENCES

[1] H.-N. Kim and W.-J. Song, "Unbiased equation-error adaptive IIR filtering based on monic normalization," *IEEE Signal Processing Lett.*, vol. 6, pp. 35–37, Feb. 1999.

[2] ——, "Adaptive IIR filtering with monic normalization: reduced-order approximation," *IEEE Signal Processing Lett.*, vol. 7, pp. 54–56, Mar. 2000.

[3] P. R. Regalia, "An unbiased equation error identifier and reduced-order approximations," *IEEE Trans. Signal Processing*, vol. 42, pp. 1397–1412, June 1994.

[4] T. Söderström, "Comments on adaptive IIR filtering with monic normalization," *IEEE Trans. Signal Processing*, vol. 48, pp. 892–894, Mar. 2000.

Erratum: Stochastic Resonance in Discrete Time Nonlinear AR(1) Models[1]

Steeve Zozor and Pierre-Olivier Amblard

***Index Terms*—Nonlinear autoregressive model, stochastic resonance.**

The model studied in the title paper is the following nonlinear AR(1) scheme:

$$\begin{cases} x_n = \Phi(x_{n-1}) + e_n \\ y_n = c\ \mathrm{sign}\,(x_n) \end{cases} \quad (1)$$

where the input is the sum of noise and a sine $e_n = b_n + \varepsilon_n$. b_n is assumed to a be zero-mean white noise of variance σ^2 and of probability density function (pdf) f_b; function Φ is taken to be bistable and odd. Furthermore, we assume that the signal-to-noise ratio (SNR) is very small, or $\varepsilon \ll \sigma$.

Since signal b_n is iid, signal x_n is Markovian, but in the title paper, we considered that as a consequence, signal y_n is Markovian. This fact is *false* in the general case (see [1]). Due to the Markovian property of x_n, the study of the title paper is right until formula (2.8) (resp. Appendix A), but (2.9) and Appendix C are (slightly) erroneous.

In this erratum, we use simpler notations than that used in the paper, The sine is written here $\varepsilon_n = \varepsilon\Re e\{e^{-\imath(2\pi n\lambda_0+\varphi_0)}\}$. This change implies that the pdf of x_n is expressed as $f_n(x) = f_{\mathrm{wm}}(x) + \Re e\{f_{\mathrm{mod}}(x)\varepsilon e^{-\imath(2\pi n\lambda_0+\varphi_0)}\}$, and (2.6) is replaced by

$$\begin{cases} f_{\mathrm{wm}} = \mathcal{K}(f_{\mathrm{wm}}) \\ f_{\mathrm{mod}} = -\left(\mathcal{I} - \mathcal{K}e^{+2\imath\pi\lambda_0}\right)^{-1} \circ \mathcal{L}(f_{\mathrm{wm}}). \end{cases} \quad (2)$$

Furthermore, (2.7) can be written as $\chi(\lambda_0) = 2c\int_0^{+\infty} f_{\mathrm{mod}}(u)\,du$, and (2.8) is simply expressed as $\underline{p}(n) = 1/2\,[1 \quad 1]^t + \Re e\{(\chi(\lambda_0)/2c)\varepsilon e^{-\imath(2\pi n\lambda_0+\varphi_0)}\}\,[-1 \quad 1]^t$.

The correlation function of y_n is

$$\begin{aligned}\Gamma(n,\,q) = E[y_n y_{n+q}] = c^2(&\Pr[y_{n+q} = +c,\ y_n = +c] \\ &+ \Pr[y_{n+q} = -c,\ y_n = -c] \\ &- \Pr[y_{n+q} = +c,\ y_n = -c] \\ &- \Pr[y_{n+q} = -c,\ y_n = +c])\end{aligned}$$

Manuscript received December 7, 2000; revised January 16, 2001. The associate editor coordinating the review of this paper and approving it for publication was Editor-in-Chief Prof. Arye Nehorai.

The authors are with LIS-ENSIEG, Saint Martin d'Heres, France.

Publisher Item Identifier S 1053-587X(01)03342-6.

[1] *IEEE Trans. Signal Processing*, vol. 47, pp. 108–122, Jan. 1999.

but as y_n is not Markovian, we cannot write, as in the paper, that the q-step transition matrix is the product of the one-step transition matrices. Nevertheless, we can write, as an example, that for $q \geq 0$

$$\begin{aligned}
&\Pr[y_{n+q} = +c,\, y_n = +c] \\
&\quad = \Pr[x_{n+q} > 0,\, x_n > 0] \\
&\quad = \int_0^{+\infty} \int_0^{+\infty} f(x_{n+q},\, x_n)\, dx_{n+q}\, dx_n \\
&\quad = \int_0^{+\infty} \int_{\mathbb{R}^{q-1}} \int_0^{+\infty} f(x_{n+q},\, \ldots,\, x_n)\, dx_{n+q} \cdots dx_n
\end{aligned}$$

and so on for the other probabilities. At this point, we can use the Markov property of x_n, $f(x_{n+q}, \ldots, x_n) = f(x_{n+q}|_{x_{n+q-1}}) \cdots f(x_{n+1}|_{x_n}) f(x_n)$. As in the title paper, since $\varepsilon \ll \sigma$, we use a first-order Taylor expansion to write

$$\begin{aligned}
f\left(x_{n+k+1}|_{x_{n+k}}\right) &= f_b\left(x_{n+k+1} - \Phi(x_{n+k}) - \varepsilon_{n+k+1}\right) \\
&\approx f_b\left(x_{n+k+1} - \Phi(x_{n+k})\right) - \varepsilon_{n+k+1} \\
&\quad \times f_b'\left(x_{n+k+1} - \Phi(x_{n+k})\right).
\end{aligned}$$

Inserting this expansion in the last equation allows us to write

$$\begin{aligned}
&\Pr[y_{n+q} = +c,\, y_n = +c] \\
&\quad = P^{(1)}_{+,+}(n+q;\, n) + P^{(2)}_{+,+}(n+q;\, n) \\
&\qquad + P^{(3)}_{+,+}(n+q;\, n) + P^{(4)}_{+,+}(n+q;\, n)
\end{aligned}$$

where

$$\begin{aligned}
P^{(1)}_{+,+}(n+q;\, n) &= \int_0^{+\infty} \int_{\mathbb{R}^{q-1}} \int_0^{+\infty} \prod_{j=1}^{q} \\
&\quad f_b(x_{n+j} - \Phi(x_{n+j-1})) f_{\rm wm}(x_n) \\
&\quad \times dx_{n+q} \cdots dx_n \\
P^{(2)}_{+,+}(n+q;\, n) &= \int_0^{+\infty} \int_{\mathbb{R}^{q-1}} \int_0^{+\infty} \prod_{j=1}^{q} \\
&\quad f_b(x_{n+j} - \Phi(x_{n+j-1})) \\
&\quad \times \Re e\left\{ f_{\rm mod}(x_n) \varepsilon e^{-\imath(2\pi n \lambda_0 + \varphi_0)} \right\} \\
&\quad \times dx_{n+q} \cdots dx_n \\
P^{(3)}_{+,+}(n+q;\, n) &= -\sum_{k=1}^{q} \varepsilon_{n+k} \int_0^{+\infty} \int_{\mathbb{R}^{q-1}} \int_0^{+\infty} \prod_{\substack{j=1 \\ j \neq k}}^{q} \\
&\quad f_b(x_{n+j} - \Phi(x_{n+j-1})) \\
&\quad \times f_b'(x_{n+k} - \Phi(x_{n+k-1})) \\
&\quad \times f_{\rm wm}(x_n)\, dx_{n+q} \cdots dx_n \\
P^{(4)}_{+,+}(n+q;\, n) &= -\sum_{k=1}^{q} \varepsilon_{n+k} \int_0^{+\infty} \int_{\mathbb{R}^{q-1}} \int_0^{+\infty} \prod_{j=1}^{q} \\
&\quad f_b(x_{n+j} - \Phi(x_{n+j-1})) \\
&\quad \times \Re e\left\{ f_{\rm mod}(x_n) \varepsilon e^{-\imath(2\pi n \lambda_0 + \varphi_0)} \right\} \\
&\quad \times dx_{n+q} \cdots dx_n
\end{aligned}$$

and so on for the other probabilities. Then using recursion (A.4) on the n-fold iterated kernel of $\mathcal{K}$, $N_n(x,\, y)$, these four terms can be written in simpler form. Using the symmetrical properties of $f_{\rm wm}$, $f_{\rm mod}$ (see Appendix A of the paper) and the obvious centro-symmetry of the n-fold iterated kernels $N_n(x,\, y)$, it can easily be seen that

$$\begin{aligned}
&P^{(2)}_{+,+}(n+q;\, n) + P^{(2)}_{+,+}(n+q;\, n) \\
&\quad = P^{(2)}_{+,+}(n+q;\, n) + P^{(2)}_{+,+}(n+q;\, n) \\
&\quad = P^{(3)}_{+,+}(n+q;\, n) + P^{(3)}_{+,+}(n+q;\, n) \\
&\quad = P^{(3)}_{+,+}(n+q;\, n) + P^{(3)}_{+,+}(n+q;\, n) = 0.
\end{aligned}$$

Hence, the only terms to contribute to the correlation are the terms 1 and 4. Interested by the zero-cycle correlation $\Gamma(q) = \langle \Gamma(n,\, q) \rangle_n$, we finally obtain the correlation

$$\Gamma(q) = \Gamma_{\rm wm}(q) + \Gamma_{\rm mod}(q) \tag{3}$$

where the terms are, respectively, the correlation of the output in the absence of the sine and the contribution of the sine to this correlation

$$\begin{cases}
\Gamma_{\rm wm}(q) = 2c^2 \displaystyle\int_0^{+\infty} \int_0^{+\infty} (N_q(u,\, v) - N_q(u,\, -v)) \\
\qquad f_{\rm wm}(v)\, du\, dv \\
\Gamma_{\rm mod}(q) = \dfrac{c^2 \varepsilon^2}{2} \Re e \left\{ \displaystyle\sum_{k=1}^{q} e^{2\imath\pi k \lambda_0} \int_0^{+\infty} \int_{\mathbb{R}^2} \int_0^{+\infty} \right. \\
\qquad (N_{q-k}(u,\, v) - N_{q-k}(u,\, -v)) \\
\qquad (N_k(w,\, t) + N_k(w,\, -t)) \\
\qquad \left. f_b'(v - \Phi(w)) f_{\rm mod}(t)\, du\, dv\, dw\, dt \right\}
\end{cases} \tag{4}$$

for $q \geq 0$ (and the correlation is even). This result still looks like the linear response theory (LRT) result (see [2] and [3]).

To progress in this sense, consider the very particular case $\Phi = c$ sign. In this case, it can easily be shown that

$$\begin{cases}
\Gamma_{\rm wm}(q) = c^2 \beta^q \\
\Gamma_{\rm mod}(q) = \dfrac{\varepsilon^2 |\chi(\lambda_0)|^2}{2} (\cos(2\pi q \lambda_0) - \beta^q) \\
\qquad \approx \dfrac{\varepsilon^2 |\chi(\lambda_0)|^2}{2} \cos(2\pi q \lambda_0)
\end{cases} \tag{5}$$

where $\beta = F_b(c) - F_b(-c)$, and $\chi(\lambda_0) = (2c\, f_b(c))/(1 - \beta e^{+2\imath\pi\lambda_0})$ (the term in β^q of the modulation part is rejected in the "noise only" part and then neglected as it is of order ε^2). These expressions are exactly that given in the paper; this is due to the fact that in this particular case, y_n *is* Markovian ($y_n = c$ sign$(y_{n-1} + e_n)$). Let us insist on the fact that in the particular case where $\Phi = c$ sign, all the results of the paper are right; hence, part IV is entirely right.

Let us come back to the general case. Using the LRT, comforted by the last result, the validity domain of the LRT (since ε is small enough compared with σ; see [2] and [3]); also comforted by many simulated results, we obtain $\Gamma(q) \approx \Gamma_{\rm wm}(q) + \langle E[y_{n+q}] E[y_n] \rangle_n$, that is

$$\Gamma(q) \approx \Gamma_{\rm wm}(q) + \frac{\varepsilon^2 |\chi(\lambda_0)|^2}{2} \cos(2\pi q \lambda_0) \tag{6}$$

where

$$\begin{cases} \Gamma_{\rm wm}(q) = 2c^2 \int_0^{+\infty} \int_0^{+\infty} (N_q(u, v) - N_q(u, -v)) \\ \qquad f_{\rm wm}(v)\, du\, dv \\ \chi(\lambda) = -2c \int_0^{+\infty} \left(\mathcal{I} - \mathcal{K} e^{+2\imath\pi\lambda}\right)^{-1} \circ \mathcal{L}(f_{\rm wm})(u)\, du \end{cases} \tag{7}$$

[where $f_{\rm wm} = \mathcal{K}(f_{\rm wm})$ is unique] that correct (2.11)–(2.13). In the SETAR case, these expressions degenerate

$$\begin{cases} \Gamma_{\rm wm}(q) = 2c^2 \left(\underline{1}_2 - \underline{1}_1\right)^t \underline{\underline{K}}^q \underline{\alpha}_2 \\ \chi(\lambda) = -2c\underline{1}_2 \left(\underline{\underline{I}} - \underline{\underline{K}} e^{2\imath\pi\lambda}\right)^{-1} \underline{\underline{L}}\, \underline{\alpha}. \end{cases} \tag{8}$$

This result corrects part III and Appendix E.

Unsurprisingly, with these correct expressions, doing the same work that is done in Appendix D leads to exactly the same conclusions. The local SNR tends toward 0 when the noise amplitude σ tends to 0 or to infinity. The conclusion that SR exists in discrete time holds.

Furthermore, these results comfort us about the fact that to find an optimal system Φ for a given noise pdf is very complicated. We hope to use SR for detection purposes, and our current investigations have led us to use the system $\Phi = c$ sign. This choice is reinforced.

ACKNOWLEDGMENT

The authors would like to thank Prof. A. Ferrari from University of Nice, Sophia-Antipolis, France, who pointed out the mistake. The authors would also like to thank Dr. A. Bulsara from the Space and Naval Warfare Systems Center of San Diego for the intensive discussion on this point and for references [4]–[9] concerning the master equation for non-Markovian processes. Finally, the authors would like to thank Prof. M. Dykman from the Physics and Astronomy Department of Michigan State University for the discussion about LRT and its validity domain.

REFERENCES

[1] J. Ledoux, *On Weak Lumpability of Denumerable Markov Chains*. Rennes, France: IRISA–INRIA, Feb. 1994.

[2] M. I. Dykman, H. Haken, G. Hu, D. G. Luchinsky, R. Mannella, P. V. E. McClintock, C. Z. Ning, N. D. Stein, and N. G. Stocks, "Linear response theory in stochastic resonance," *Phys. Lett. A*, vol. 180, pp. 332–336, Sept. 1993.

[3] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 3rd ed. Oxford, U.K.: Butterworth Heinemann, 1980, pt. 1, vol. 5.

[4] P. Hänggi and P. Talkner, "First-passage time problems for non-Markovian processes," *Phys. Rev. A*, vol. 32, no. 3, pp. 1934–1937, Sept. 1985.

[5] B. Carmeli and A. Nitzan, "Non-Markovian theory of activated rate processes. I. Formalism," *J. Chem. Phys.*, vol. 79, no. 1, pp. 393–404, July 1983.

[6] ——, "Non-Markovian theory of activated rate processes. II. Thermal desorption," *Israeli J. Chem.*, vol. 22, no. 4, pp. 360–364, 1982.

[7] ——, "Non-Markovian theory of activated rate processes. III. Bridging between the Kramers limits," *Phys. Rev. A*, vol. 29, no. 3, pp. 1481–1495, Mar. 1984.

[8] ——, "Non-Markovian theory of activated rate processes. IV. The double well model," *J. Chem. Phys.*, vol. 80, no. 8, pp. 3596–3605, Apr. 1984.

[9] ——, "Non-Markovian theory of activated rate processes. V. External periodic forces in the low-friction limit, resp.," *Phys. Rev. A*, vol. 32, no. 4, pp. 2439–2454, Oct. 1985.

Comment: On the Unnecessary Assumption of Statistical Independence Between Reference Signal and Filter Weights in Feedforward Adaptive Systems

J. Minkoff

In [1], I presented a formulation in which the time-evolving weight-iteration equation for random signals is derived without the necessity of invoking the usual unsatisfactory assumption that is customarily made, namely, that the weights $\mathbf{W}$ and the reference signal $\mathbf{X}$ in the weight-iteration equation are statistically independent. I neglected, however, to give a physical argument for it. This is, that in this derivation [see (16)–(18) in the paper], it is not necessary for $\mathbf{W}$ to be independent of $\mathbf{X}$ but only of $\mathbf{XX}^\dagger$, which does not contain the phase information of $\mathbf{X}$. That is, the off-diagonal terms of $\mathbf{XX}^\dagger$ contain only phase differences, which could be produced by an infinite number of different, arbitrary $\mathbf{X}$s. Thus, considering, for example, an echo-canceler application, the all-important information concerning echo delay, which is essential for proper convergence of $\mathbf{W}$, is arrived at in the convergence process by iteratively cross-correlating the reference signal $\mathbf{X}$—not $\mathbf{XX}^\dagger$—with the iterated error signal. The iterated phase differences between $\mathbf{X}$ and the error signals serve to drive the phase of $\mathbf{W}$ to its ultimate converged value. For this purpose, $\mathbf{XX}^\dagger$ is of no use, and declaring $\mathbf{W}$ to be independent of the phase of $\mathbf{XX}^\dagger$ is therefore not equivalent to declaring it to be independent of the phase of $\mathbf{X}$. The term $\mathbf{XX}^\dagger$ does contain information about the magnitude of $\mathbf{X}$, which must be considered in setting the value of the step-size parameter μ, but the adaptive-filter operation is statistically independent of the magnitude of $\mathbf{X}$; if the strength of the echo is proportional to $|\mathbf{X}|$, so is the filter input, and the corresponding output that cancels it. In fact, as is well known, in order to avoid excessive fluctuations during and after convergence, it is customary to normalize the step-size parameter μ with respect to input reference-signal power, which removes all information about $|\mathbf{X}|$. And again, since $\mathbf{XX}^\dagger$ does not contain the phase information necessary for proper convergence, it is therefore possible to write

$$\overline{\mathbf{WXX}^\dagger} = \overline{\mathbf{W}}\ \overline{\mathbf{XX}^\dagger}.$$

By means of the same arguments, one can of course write an equivalent expression for time functions.

REFERENCES

[1] J. Minkoff, "Operation of multichannel feedforward adaptive systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 2993–3004, Dec. 1997.

Manuscript received November 16, 2000; revised February 5, 2001. The associate editor coordinating the review of this paper and approving it for publication was Dr. Alle-Jan van der Veen.

The author is with ITT Aerospace Communications, Clifton, NJ 07014 USA USA (e-mail: john.minkoff@itt.com).

Publisher Item Identifier S 1053-587X(01)03341-4.