

NONLINEAR AUTOREGRESSION IN THE THEORY OF SIGNAL COMPRESSION

Timo Koski

Luleå University of Technology, Department of Applied Mathematics
S-95187 Luleå, Sweden

Abstract. The non-linear autoregression

$$U_{n+1} = U_n - b \cdot \text{sign}(U_n) + m + \xi_n,$$

where ξ_n is a white noise, emerges both in delta modulation and in sigma-delta modulation, see [6, 10, 31]. We prove the existence of a unique invariant measure and some ergodic properties, and discuss a time continuous approximation of the process.

1. Introduction

The technique of digitization of waveforms (signal or source compression) leads in several situations to consideration of certain nonlinear stochastic difference equations, see [2, 12, 16]. In (single-loop) *sigma-delta modulation* (see [10, 31]) and in (ideal) *delta modulation* (see e.g. [6, 10]) an important role is played by the first order nonlinear autoregression

$$(1) \quad U_{n+1} = U_n - Q(U_n) + m + \xi_n, \quad n = 0, 1, \dots,$$

where ξ_n is a white noise sequence and

$$(2) \quad Q(u) := b \cdot \text{sign}(u) = \begin{cases} b & \text{if } u > 0 \\ -b & \text{if } u < 0, \end{cases}$$

with $b > 0$, i.e. $Q(u)$ is a one-bit quantizer characteristic.

An interesting development of an approximate time continuous model for (1) is given by P.W. Wong and R.M. Gray in [31] with a Gaussian white noise ξ_n . Their time continuous model is a Brownian motion with a drift considered (and modified) in the intervals between certain stopping times. In the present paper the properties of the process $\mathcal{U} = \{U_n \mid n = 0, 1, \dots\}$ in (1) are surveyed as a real valued, discrete time Markov process. The aspects of the theory of Markov

1991 Mathematics Subject Classification: Primary: 94A29; Secondary 60J20.

Research supported by the Air Force Office of Scientific Research Contract No. F49620 85C 0144.

processes required for this are found e.g. in [20, 21, 25], see also [26]. We shall briefly discuss another time continuous model, a stochastic differential equation based on the drift function of (1) when the noise ξ_n is Gaussian.

In time series analysis the type of processes determined by (1) is regarded as one of the so-called *SETAR(1)*-processes, where *SETAR(1)* stands for ‘self-excited threshold autoregressive of first order’. In our case there is one threshold, i.e. the origin. Results on *SETAR(1)*-processes closely resembling (1) are found e.g. in [1, 13, 24, 22, 30]. A comprehensive account of the theory of and the statistical modeling by nonlinear autoregressions is the monograph by H. Tong [27].

The paper is organized as follows. In Section 2 the states of the process \mathcal{U} are classified in terms of the theory in [19, 28, 29], i.e. it is shown under certain assumptions on the density of the noise ξ_n , see Assumption 2.1, that \mathcal{U} is aperiodic and φ -irreducible. A condition for stationarity, which is central in the theory of delta and sigma-delta modulators, is derived and stated in (6). In Section 3 the existence of a unique invariant measure and some ergodic properties are established evoking stochastic Lyapunov functions (mean drift conditions). In Section 4 a stochastic differential equation related to (1) by a mean square convergence argument is discussed. Further results on this topic are found in [16, 18]. The main proofs are given in the Appendix. The subsequent work [17] will contain a more detailed probabilistic analysis of sigma-delta modulation, in particular, of binary quantizer error.

2. Some elementary properties and classification of states

Assume for the moment that $\mathcal{U} = \{U_n \mid n = 0, 1, \dots\}$ is a stationary process. Then the mean value function is a constant $\mu = \mathbf{E}(U_{n+1}) = \mathbf{E}(U_n)$, and hence from (1)

$$(3) \quad \mathbf{E}(Q(U_n)) = m,$$

since we are, by convention (c.f. [11, p. 197]), assuming that a white noise has mean zero. Since the quantizer characteristic in (2) can be written as

$$(4) \quad Q(u) = 2b\mathbf{I}(u > 0) - b,$$

where $\mathbf{I}(u > 0)$ is the indicator function of the positive real line, equation (3) yields

$$(5) \quad \begin{cases} \mathbf{P}(U_n > 0) = (b + m)/2b \\ \mathbf{P}(U_n \leq 0) = (b - m)/2b \end{cases}$$

For this to make sense we must assume that $|m| \leq b$. The equality $|m| = b$ is clearly impossible, when the assumption $\mathbf{E}(\xi_n) = 0$ is imposed, unless the process

is constant. Excluding the trivial process $U_n = U_0$ we have a necessary condition for stationarity,

$$(6) \quad |m| < b.$$

For the conclusion in (6) we assumed just the existence of the (zero) mean for the distribution of the ξ_n 's. We need to add some more restrictive hypotheses.

Assumption 2.1. *The random variables ξ_n have for each $n = 0, 1, \dots$, a density function $f_\xi(x)$ such that*

1. $f_\xi(x) > 0$ for all $x \in R$,
2. $f_\xi(x)$ is continuous for all $x \in R$,
3. $\mathbf{E}(\xi_n) = 0$,
4. $\mathbf{E}(\xi_n^2) := \sigma^2 < \infty$.

The continuity assumption 2.1.2 is actually needed for the first time in the proof of Proposition 3.1. From the assumed white noise property of the sequence ξ_n we have that \mathcal{U} is Markov with the state space R , and from (1) it follows that the n -step (temporally homogeneous) transition probability

$$(7) \quad \mathbf{P}^n(u; A) := \mathbf{P}^n(U_n \in A \mid U_0 = u)$$

can be expressed as

$$(8) \quad \mathbf{P}^n(u; A) = \int_A p_n(x \mid u) dx,$$

where A is a Borel set of the real line. In other words the n -step transition density, $p_n(x \mid u)$, of the process \mathcal{U} exists, and for all x and u

$$(9) \quad p_n(x \mid u) > 0.$$

In order to briefly motivate this we iterate (1) obtaining

$$(10) \quad U_n = u - \sum_{i=0}^{n-1} Q(U_i) + n \cdot m + \sum_{i=0}^{n-2} \xi_i + \xi_{n-1},$$

where $U_0 = u$. In view of the assumption that $\xi_0, \xi_1, \dots, \xi_{n-2}, \xi_{n-1}$ are independent and identically distributed with the density $f_\xi(x)$, it follows that

$$(11) \quad p_n(x \mid u) = \int_{R^{n-1}} f_\xi \left(x - u + \sum_{i=0}^{n-1} Q(u_i) - n \cdot m - \sum_{i=0}^{n-2} x_i \right) \prod_{i=0}^{n-2} f_\xi(x_i) dx_i,$$

which entails (9) under Assumption 2.1.1. (In (11) the sequence u_i is computed for any given sequence x_i using the rule (1), $u_0 = u$).

Some significant technical properties of the process \mathcal{U} follow readily from (9) by the results in [29] or [19, 20, 28]. In fact, due to (8) and (9) we have that for every $u \in R$

$$(12) \quad \sum_{n=1}^{\infty} 2^{-n} \mathbf{P}^n(u; A) > 0$$

if $\varphi(A) > 0$, where φ denotes the Lebesgue-measure. Equation (12) recapitulates one of many similar definitions of φ -irreducibility, see [19, 28, 29]. In our case any of them is easily verified. This means that every set in R with positive Lebesgue-measure is attainable by the process \mathcal{U} from any $u \in R$. Furthermore, (9) implies that \mathcal{U} is an aperiodic process, see [20, p. 20–21], i.e. there are no cyclical subsets in the state space. For convenience of reference we collect these observations in the next lemma:

Lemma 2.1. *Under Assumption 2.1.1 the process \mathcal{U} in (1) is φ -irreducible and aperiodic, φ standing for the Lebesgue measure.*

Proof. See e.g. [20, p. 11] or [28, p. 741].

3. Existence and uniqueness of the invariant measure, ergodic properties

3.1. The mean drift condition. Ergodic and recurrence properties of a Markov process are often conveniently established, if one can verify a set of *mean drift conditions* (also known as *Foster's conditions*) for the given process, see [4, 19, 26, 28, 29].

Mean drift conditions: There exist a non-negative function $V(\cdot)$, a compact set $\mathbf{K} \subseteq R$, and a number $k > 0$, such that

$$(13) \quad \mathbf{E}(V(U_{n+1}) \mid U_n = u) - V(u) < -k \quad \text{for } u \in \mathbf{K}^c,$$

where \mathbf{K}^c is the complement of \mathbf{K} , and for some fixed $B > 0$,

$$(14) \quad \mathbf{E}(V(U_{n+1}) \mid U_n = u) \leq B$$

for all $u \in \mathbf{K}$.

The function $V(\cdot)$ is also sometimes called a *stochastic Lyapunov function*.

We show next that the function

$$(15) \quad V(x) := x^2$$

has the properties required in (13) and (14) under the condition (6). In fact, using (1) with (15) we have

$$\begin{aligned}
 \mathbf{E}(V(U_{n+1}) | U_n = u) - V(u) &= 2u\mathbf{E}(-Q(U_n) + \xi_n | U_n = u) + 2um \\
 (16) \qquad \qquad \qquad &+ \mathbf{E}((-Q(U_n) + m + \xi_n)^2 | U_n = u) \\
 &\leq 2u(m - Q(u)) + 4(b^2 + m^2 + \sigma^2),
 \end{aligned}$$

where $\mathbf{E}(\xi_n^2) = \sigma^2$, by Assumption 2.1.4.

Setting

$$(17) \qquad \qquad \varrho(u) := 2u(m - Q(u)) + 4(b^2 + m^2 + \sigma^2),$$

we have, in view of (6), that

$$(18) \qquad \qquad \varrho(u) \downarrow -\infty \quad \text{as} \quad |u| \uparrow +\infty.$$

This means that the set \mathbf{K} in the mean drift conditions above can be taken as $\mathbf{K} = \{|u| \leq R_0\}$ for a sufficiently large R_0 . These observations verify the requirements (13)–(14) and will be used in the proof of existence of a unique invariant measure, see Proposition 3.1.

3.2. Invariant measure, ergodic properties. An invariant measure, denoted by π , is a (probability) measure satisfying

$$(19) \qquad \qquad \pi(A) = \int_{\mathbf{R}} \pi(dx) \mathbf{P}^1(x; A),$$

for any Borel set A .

Proposition 3.1. *Suppose the density of the white noise f_ξ satisfies the conditions in Assumption 2.1. Then there exists a unique invariant measure π for the process \mathcal{U} defined in (1) if and only if (6) holds.*

The proof of this proposition is given in Appendix.

In terms of the classification of Markov processes introduced in [29], we say that the process \mathcal{U} is a *positively recurrent* process. From the existence of an invariant measure together with aperiodicity and φ -irreducibility we obtain the following results.

Proposition 3.2. *The process \mathcal{U} is metrically transitive, if U_0 is distributed according to the invariant measure $\pi(\cdot)$.*

Proof. We show first that the invariant measure is equivalent to the Lebesgue measure φ . By definition (19) and by (8) we have

$$\pi(A) = \int_A \left(\int_R p_1(y | x) \pi(dx) \right) \varphi(dy)$$

since Fubini's theorem applies to the kernel $p_1(y | x)$. Hence $\pi(\cdot)$ is absolutely continuous with respect to the Lebesgue measure.

Conversely, if $\varphi(A) > 0$, then $\mathbf{P}^1(x, A) > 0$ by (9). Thus, (19) entails that $\pi(A) > 0$. Hence we have established the asserted equivalence between $\varphi(\cdot)$ and $\pi(\cdot)$.

Suppose now that A is an *invariant set* in the sense that

$$(20) \quad \mathbf{P}^1(x, A) = 1 \quad \text{for almost all } x \in A \text{ and } \pi(A) \neq 0.$$

Hence, for almost all $x \in A$

$$\mathbf{P}^1(x, A^c) = 0,$$

which in its turn implies $\pi(A^c) = 0$, since $\mathbf{P}^1(x, \cdot)$, $\varphi(\cdot)$ and $\pi(\cdot)$ are equivalent measures. Thus $\pi(A) = 1$, and any invariant set has $\pi(\cdot)$ -measure 1 or 0. This establishes the assertion as claimed.

The definition of an invariant set and that of metrical transitivity invoked in (20) can be translated into the identically named concepts in ergodic theory, see [9, p. 207–208] or [23, p. 42] and vice versa, as shown in [25, p. 95–97]. In fact, for any transition probability $\mathbf{P}^1(\cdot; \cdot)$ and invariant measure $\pi(\cdot)$, a probability measure \mathbf{P}_π can be constructed on the space of all real valued sequences with nonnegative indices, denoted by $R^{\mathbb{Z}^+}$, so that \mathbf{P}_π is invariant under the shift transformation on that space. Next we prove an ergodic theorem using these remarks.

Proposition 3.3. *For all $u \in R$ and all real valued functions $\psi(\cdot)$, integrable w.r.t. the invariant measure π , we have*

$$(21) \quad \lim_{n \rightarrow +\infty} \frac{1}{n+1} \sum_{i=0}^n \psi(U_i) = \int_{-\infty}^{\infty} \psi(x) \pi(dx),$$

where $U_0 = u$, and the convergence holds for every u .

Proof. Let us consider $\mathbf{x} = (x_0, x_1, \dots)$, a generic element of the sequence space $R^{\mathbb{Z}^+}$ referred to above. Let, furthermore, τ be the *shift transformation*, i.e. $\tau(\mathbf{x}) = (x_1, x_2, \dots)$. Let also $\varphi(\mathbf{x}) := x_0$. Then we can take $\psi(U_i) = \psi(\varphi(\tau^i(\mathbf{x})))$, where τ^i denotes the i th successive application of the shift transformation. We know, by Proposition 3.2 and by the argument preceding the proposition under consideration, that τ is ergodic. Hence the ergodic theorem, see e.g. [23, p. 44] or [25, Corollary 5, p. 44], implies (21) for almost every u . Using the equivalence of φ and π , established in the proof of Proposition 3.2, we obtain the convergence in (21) for every u .

Example 3.4 (Decoder in sigma-delta modulation). The process \mathcal{U} in (1) plays the role of the encoder state in what is known as a (*single loop*) *sigma-delta modulator*, see [8, 10, 31]. The actual output (i.e. the encoded or quantized sequence) is $\{Q(U_n) \mid n = 0, 1, \dots\}$. The *decoder* output, denoted by y_n , of the sigma-delta modulator is given by

$$y_n := \frac{1}{N} \sum_{i=0}^{N-1} Q(U_{n-i})$$

The number N is known as the *oversampling ratio*, see [8] for discussion of this concept. In the stationary state we have by (3) and (21) that for any n and for N large

$$(22) \quad y_n \approx m.$$

Since the second moment of $Q(U_{n-i})$ is equal to b^2 and the process \mathcal{U} is metrically transitive, the approximate equality in (22) is valid also in mean square, (see [11, p. 178]). Thus, if the input sequence to the sigma-delta modulator is taken as $(m + \xi_n)$, the decoder output converges for higher oversampling ratios N to the dc-level of the input, which is a reasonable property for a quantizer of this type. There are, however, simpler ways of obtaining (22) than proving a full ergodic theorem.

4. Remarks on a time continuous model

Rewriting (1) as

$$U_{n+1} - U_n = -Q(U_n) + m + \xi_n$$

makes it natural, at least in the case of a white Gaussian input ξ_n , to consider its approximation by means of the stochastic differential equation (SDE)

$$(23) \quad dv(t) = -Q(v(t)) dt + m \cdot dt + \sigma \cdot dw(t),$$

where $w(t)$ is a standard Wiener process. The SDE (23) is known in stochastic control theory as a 'bang-bang'-feedback system, c.f. [3]. The existence of a strong solution to (23) is well known and easily established. Techniques for computing the transition densities of (23) are found in [14]. A straightforward solution to the associated stationary Fokker-Planck equation (see [15, p. 231]) is obtained as

$$f_v(u) = \frac{(b^2 - m^2)}{\sigma^2 \cdot b} e^{\frac{2}{\sigma^2} \int^u \mu(x) dx},$$

where $\mu(u) := -Q(u) + m$ denotes the drift function of the SDE in (23). In other words $f_v(u)$ is the invariant density of the process defined in (23). Our attention to this SDE is additionally justified by the fact that for any $t > 0$

$$\mathbf{P}\{Q(v(t)) = -b\} = \frac{(b^2 - m^2)}{\sigma^2 \cdot b} \int_{-\infty}^0 e^{\frac{2}{\sigma^2}(b+m)u} du = (b - m)/2b,$$

which agrees with (5).

Finally we outline a justification for the approximation of (1) by the SDE in (23). We set first

$$\Delta t_n := t_n - t_{n-1},$$

and for $t_{n-1} \leq t < t_n$,

$$\bar{U}_t = \bar{U}_{n-1} - b \cdot \text{sign}(\bar{U}_{n-1})(t - t_{n-1}) + m \cdot (t - t_{n-1}) + \sigma(w(t_n) - w(t_{n-1})),$$

where $\bar{U}_{t_n} := \bar{U}_n \equiv U_n$. Evidently this defines a time continuous interpolation of \mathcal{U} . Then we have

Proposition 4.1. *If $\Delta t_n \rightarrow 0$ with $n \rightarrow +\infty$, then*

$$E\{|v_{t_n} - \bar{U}_n|^2\} \rightarrow 0.$$

Proof. This follows by using Ito's lemma on $F(v_{t_n}, \bar{U}_n) := (v_{t_n} - \bar{U}_n)^2$, by certain upper bounds and by Bellman–Gronwall's lemma. The straightforward details are omitted.

Appendix: Proof of Proposition 3.1

In this Appendix we present a proof of the existence of an invariant measure i.e. Proposition 3.1. The difficulties posed by the threshold in the quantizer map $Q(\cdot)$ are made explicit in this particular proof, c.f. [7].

By [28, Theorem 4.2, Theorem 4.3] or [19], for a φ -irreducible process the conditions (13)–(14) and the weak continuity or the *Feller condition* would imply the existence of a unique invariant measure (see (19) above). The weak continuity condition is formulated as follows.

Feller condition: The function $g(x)$ defined by

$$g(x) := \int_{\mathcal{R}} l(y) p_1(y | x) dy$$

is a continuous function for every bounded measurable $l(\cdot)$.

In the present case we have, however,

$$(24) \quad p_1(y | x) = f_\xi(y - (x - Q(x) + m))$$

and the required continuity need not hold for every $l(\cdot)$ with the assumed properties. It turns out on the other hand that we are able to prove the following lemma on *tightness* using (13) with $V(x) = x^2$, and that this will entail directly the desired conclusions about the invariant measure.

Lemma A.1. *Let (6) and Assumptions 2.1.1, 2.1.3, and 2.1.4 hold. Then the measures $\pi_n(A)$, defined on the Borel sets $A \subset R$ by*

$$(25) \quad \pi_n(A) := \frac{1}{n} \sum_{i=0}^{n-1} F_i(A),$$

are tight, where for $i = 1, 2, \dots$

$$(26) \quad F_i(A) := \int_R F_{i-1}(dx) \mathbf{P}^1(x; A)$$

with $\mathbf{P}^1(x; A)$ defined in (7) and assuming that F_0 is a distribution function such that $\int_{-\infty}^{\infty} x^2 F_0(dx) < \infty$.

Proof. Let us consider the function $\varrho(\cdot)$ defined in (17) as

$$(27) \quad \varrho(u) := 2u(m - Q(u)) + 4(b^2 + m^2 + \sigma^2).$$

The function $\varrho(\cdot)$ is linear for u in either half-axis and it decreases to $-\infty$, under the assumption (6), as $|u|$ increases. We set

$$\mathbf{K} := \{|u| \leq R_0\},$$

where R_0 is so large that

$$\sup_{u \in \mathbf{K}^c} \varrho(u) < 0.$$

In addition we set

$$(28) \quad -k := -k(R_0) := \sup_{u \in \mathbf{K}^c} \varrho(u), \quad M := \sup_{u \in \mathbf{K}} \varrho(u).$$

Thus, it follows that

$$\mathbf{E}(\varrho(U_j) | U_0) \leq -k(R_0) \mathbf{P}(|U_j| > R_0 | U_0) + M \mathbf{P}(|U_j| \leq R_0 | U_0).$$

Rearranging this inequality we obtain

$$(29) \quad k(R_0) \mathbf{P}(|U_j| > R_0 | U_0) \leq M \mathbf{P}(|U_j| \leq R_0 | U_0) - \mathbf{E}(\varrho(U_j) | U_0).$$

It follows from the definition of $\varrho(\cdot)$ (cf. (27)) combined with the inequality (16) that

$$(30) \quad \begin{aligned} -\mathbf{E}(\varrho(U_j) | U_0) &\leq -\mathbf{E}\{\mathbf{E}(V(U_{j+1}) - V(U_j) | U_j) | U_0\} \\ &= -\mathbf{E}\{V(U_{j+1}) - V(U_j) | U_0\}. \end{aligned}$$

Summing the inequalities (29)–(30) gives

$$k(R_0) \frac{1}{n} \sum_{j=0}^{n-1} \mathbf{P}(|U_j| > R_0 \mid U_0) \leq M - \mathbf{E}(V(U_n) \mid U_0) + V(U_0).$$

Averaging over F_0 entails, as we are assuming that $\mathbf{E}(V(U_0)) = \int_{-\infty}^{\infty} x^2 F_0(dx) < +\infty$, that

$$(31) \quad k(R_0) \frac{1}{n} \sum_{j=0}^{n-1} \mathbf{P}(|U_j| > R_0) \leq M + \mathbf{E}(V(U_0)),$$

since $-V(U_n) \leq 0$. By the definitions in (25)–(26), as $\mathbf{P}(|U_j| > R_0) = F_j(\mathbf{K}^c)$, we can rewrite (31) as

$$(32) \quad \pi_n(\mathbf{K}^c) \leq \frac{1}{k(R_0)} \left(M + \mathbf{E}(V(U_0)) \right),$$

since $k(R_0) > 0$ by (28). Furthermore, by (28) and the properties of the function $\varrho(\cdot)$, $k(R_0) \rightarrow +\infty$ as R_0 becomes larger, whereas M stays constant for increasing R_0 . Hence, the inequality (32) indicates that there is a compact set G , like \mathbf{K} for R_0 , large enough satisfying

$$\pi_n(G) \geq 1 - \varepsilon,$$

for any $\varepsilon > 0$ and all n . This establishes the tightness of the sequence π_n , as claimed.

Proof of Proposition 3.1. The necessity of the condition (6) has already been discussed in Section 2. We shall establish the sufficiency of (6) under the stated assumptions on $f_\xi(\cdot)$. The basic outline of the argument is based on and, in fact, completes a similar proof in [7, p. 831–835].

Let us first define a linear operator \mathbf{T} on the space of all the probability distributions $F(\cdot)$ on R by

$$(33) \quad (\mathbf{T}F)(A) := \int_R F(dx) \mathbf{P}^1(x; A),$$

where A is a Borel set of R . It follows by (25) and (26) that

$$(34) \quad \begin{aligned} (\mathbf{T}\pi_n)(A) &= \frac{1}{n} \sum_{i=0}^{n-1} \int_R F_i(dx) \mathbf{P}^1(x; A) \\ &= \frac{1}{n} \sum_{i=0}^{n-1} F_{i+1}(A) \\ &= \frac{1}{n} \sum_{i=0}^{n-1} F_i(A) + \frac{1}{n} (F_n(A) - F_0(A)) \\ &= \pi_n(A) + \frac{1}{n} (F_n(A) - F_0(A)). \end{aligned}$$

Hence we have for all Borel sets A , not just for continuity sets (a set A is called a π -continuity set, if its boundary has π -measure zero, see e.g. [5, p. 303]), that

$$(35) \quad |(\mathbf{T}\pi_n)(A) - \pi_n(A)| \leq \frac{1}{n} |F_n(A) - F_0(A)| \leq \frac{1}{n}.$$

Next, the condition (6) entails by Lemma A.1 that the measures π_n are tight. Hence there exists a subsequence π_{n_k} converging to a probability measure π in the sense of weak convergence, denoted by $\pi_{n_k} \implies \pi$, or equivalently, for all π -continuity sets A

$$(36) \quad \pi_{n_k}(A) \rightarrow \pi(A).$$

We shall now prove that

$$(37) \quad (\mathbf{T}\pi_{n_k})(A) \rightarrow (\mathbf{T}\pi)(A)$$

holds for all Borel sets A . The difficulty is that $\pi_{n_k} \implies \pi$ does imply, by any standard characterization of weak convergence [5, p. 303], that

$$(38) \quad \int_R \pi_{n_k}(dx) \varphi(x) \rightarrow \int_R \pi(dx) \varphi(x)$$

for any bounded continuous function φ . This would be the desired result (37) provided the choice $\varphi(x) := \mathbf{P}^1(x; A)$, A any Borel set, would give a continuous function of x . In the present situation, as already pointed out (see (8), (24)), we do not have this property.

We can, however, obtain (38) for $\varphi(x) := \mathbf{P}^1(x; A)$ by means of the following argumentation. What we need is the fact that origin, or the threshold of the quantizer in (1), is the only discontinuity of the function $\varphi(x) := \mathbf{P}^1(x; A)$. In order to be assured of this we have imposed the continuity condition in Assumption 2.1.2.

Let us define for an arbitrary $\delta > 0$

$$I_\delta(x) := \begin{cases} 1 & \text{for } |x| \geq \delta, \\ |x|/\delta & \text{for } |x| < \delta. \end{cases}$$

The function $I_\delta(\cdot)$ is bounded and continuous, and so is

$$\varphi_\delta(x) := I_\delta(x) \mathbf{P}^1(x; A),$$

for an arbitrary Borel set $A \subset R$. Next,

$$(39) \quad \left| \int_R \pi_{n_k}(dx) \mathbf{P}^1(x; A) - \int_R \pi_{n_k}(dx) \varphi_\delta(x) \right| \leq \pi_{n_k}(A_\delta),$$

where $A_\delta := \{|x| < \delta\}$, and similarly

$$(40) \quad \left| \int_R \pi(dx) \mathbf{P}^1(x; A) - \int_R \pi(dx) \varphi_\delta(x) \right| \leq \pi(A_\delta).$$

As in the proof of Proposition 3.2 we see that every $F_i(\cdot)$ is actually equivalent to the Lebesgue-measure φ . Hence, every π_{n_k} has the same property. As $\varphi(A_\delta) = \delta$, we have for some function $\mathcal{E}(\delta)$ that

$$\pi_{n_k}(A_\delta) \leq \mathcal{E}(\delta),$$

where $\mathcal{E}(\delta) \rightarrow 0$ as $\delta \downarrow 0$. Furthermore, A_δ is an open set and the standard portmanteau theorem, see e.g. [5, p. 303], entails that

$$\pi(A_\delta) \leq \liminf_{k \rightarrow +\infty} \pi_{n_k}(A_\delta) < \mathcal{E}(\delta).$$

Hence, for any Borel set A

$$(41) \quad \begin{aligned} & \left| \int_R \pi_{n_k}(dx) \mathbf{P}^1(x; A) - \int_R \pi(dx) \mathbf{P}^1(x; A) \right| \\ & \leq \left| \int_R \pi_{n_k}(dx) \mathbf{P}^1(x; A) - \int_R \pi_{n_k}(dx) \varphi_\delta(x) \right| \\ & \quad + \left| \int_R \pi_{n_k}(dx) \varphi_\delta(x) - \int_R \pi(dx) \varphi_\delta(x) \right| \\ & \quad + \left| \int_R \pi(dx) \varphi_\delta(x) - \int_R \pi(dx) \mathbf{P}^1(x; A) \right|. \end{aligned}$$

By (39) and (40) we obtain an upper bound for the right hand side of (41), yielding

$$(42) \quad \begin{aligned} & \left| \int_R \pi_{n_k}(dx) \mathbf{P}^1(x; A) - \int_R \pi(dx) \mathbf{P}^1(x; A) \right| \\ & \leq \pi_{n_k}(A_\delta) + \left| \int_R \pi_{n_k}(dx) \varphi_\delta(x) - \int_R \pi(dx) \varphi_\delta(x) \right| + \pi(A_\delta). \end{aligned}$$

On the right hand side of (42) the sum of the two probability terms is smaller than or equal to $2\mathcal{E}(\delta)$. The middle expression can be made arbitrarily small by letting k become sufficiently large by weak convergence, since (38) holds for $\varphi_\delta(x)$. Note that k and δ can be treated independently. Hence we have actually proved that

$$(43) \quad (\mathbf{T}\pi_{n_k})(A) \rightarrow (\mathbf{T}\pi)(A)$$

for any Borel set A . We know, in addition, that $\mathbf{T}\pi$ is a probability measure, which we denote by μ . The inequality (35) entails now that $\pi_{n_k}(A)$ converges to $\mu(A)$

for every Borel set A . Hence, we have $\pi_{n_k} \implies \mu$ by the standard characterization of weak convergence used above. Thus the proof of (37) above can be repeated by using μ instead of π as the limiting measure in the weak convergence arguments. This means that both $\pi_{n_k}(A) \rightarrow \mu(A)$ and $(\mathbf{T}\pi_{n_k})(A) \rightarrow (\mathbf{T}\mu)(A)$ or, in view of (35) that $\mu(A) = (\mathbf{T}\mu)(A)$ for any Borel set A , which by definition (see (19)) means that there exists an invariant measure. As any φ -irreducible, aperiodic Markov-process, c.f. Lemma 2.1, has at most one invariant measure, the assertion about uniqueness follows, finishing the proof.

Acknowledgments. The author thanks Professor S. Cambanis for many useful discussions. This work was done during the author's stay at the Center for Stochastic Processes at the University of North Carolina at Chapel Hill.

References

- [1] ANDÉL, J., I. NETUKA, and K. ZUÁRA: On threshold autoregressive processes. - *Kybernetika* (Prague), 20, 1984, 89–106.
- [2] ARNSTEIN, D.S.: Quantization error in predictive coders. - *IEEE Trans. Comm.* 23, 1975, 423–429.
- [3] BALAKRISHNAN, A.V.: On stochastic bang bang control. - *Appl. Math. Optim.* 6, 1980, 91–96.
- [4] CHAN, K.S., and H. TONG: On the use of the deterministic Lyapunov function for the ergodicity of stochastic difference equations. - *Adv. in Appl. Probab.* 27, 1985, 666–678.
- [5] DUDLEY, R.M.: Real analysis and probability. - Wadsworth & Brooks/Cole Mathematics Series, Pacific Grove, 1989.
- [6] FINE, T.L.: The response of a particular nonlinear system with feedback to each of two random processes. - *IEEE Trans. Inform. Theory* 14, 1968, 255–264.
- [7] GERSHO, A.: Stochastic stability of delta modulation. - *Bell System Technical Journal* 51, 1972, 821–841.
- [8] GRAY, R.M.: Oversampled sigma-delta modulation. - *IEEE Trans. Comm.* 35, 1987, 481–489.
- [9] GRAY, R.M.: Probability, random processes and ergodic properties. - Springer-Verlag, New York, 1988.
- [10] GRAY, R.M.: Source coding theory. - Kluwer Academic Publishers, Boston, 1990.
- [11] GRAY, R.M., and L.D. DAVISSON: Random processes. A mathematical approach for engineers. - Prentice-Hall, Englewood-Cliffs, 1986.
- [12] HAYASHI, A.: Differential pulse code modulation of the Wiener process. - *IEEE Trans. Comm.* 26, 1978, 881–887.
- [13] HÖGNÄS, G.: Nonlinear autoregressive processes. - Reports on Computer Science & Mathematics. Åbo Akademi Ser. A 74, Åbo, Finland, 1989.
- [14] KARATZAS, I., and S.E. SHREVE: Trivariate density of Brownian motion, its local and occupation times, with application to stochastic control. - *Ann. Probab.* 12, 1984, 819–828.
- [15] KARLIN, S., and H.M. TAYLOR: A second course in stochastic processes. [Second edition]. - Academic Press, New York, 1981.
- [16] KOSKI, T.: A diffusion approximation for the quantization error in delta modulation with a Gauss Markov signal. - In *Limit Theorems in Probability and Statistics*, edited by P. Révész. D. Reidel Publ. Co., Dordrecht, 1990, 305–325.

- [17] KOSKI, T.: Statistics of the binary quantizer error in sigma-delta modulation with i.i.d. Gaussian input. - To appear.
- [18] KOSKI, T., and S. CAMBANIS: On the statistics of the error in predictive coding for stationary Ornstein-Uhlenbeck processes. - IEEE Trans. Inform. Theory (to appear).
- [19] LASLETT, G.M., D.B. POLLARD, and R.L. TWEEDIE: Techniques for establishing ergodic and recurrence properties of continuous-valued Markov chains. - Naval Res. Logist. Quart. 25, 1978, 455-472.
- [20] NUMMELIN, E.: General irreducible Markov chains and nonnegative operators. - Cambridge University Press, Cambridge, 1984.
- [21] OREY, S.: Limit theorems for Markov chain transition probabilities. - Van Nostrand Reinhold Co., London, 1971.
- [22] PEMBERTON, J.: Piecewise constant autoregressive models for univariate time series. - Technical Report MCS-90-04, Department of Mathematics and Computer Science, University of Salford, Salford, United Kingdom, 1990.
- [23] PETERSEN, K.: Ergodic theory. - Cambridge University Press, Cambridge, 1989.
- [24] PETRUCELLI, J.D., and S.W. WOOLFORD: A threshold AR(1) model. - J. Appl. Probab. 21, 1984, 270-286.
- [25] ROSENBLATT, M.: Markov processes, structure and asymptotic behavior. - Springer-Verlag, New York, 1971.
- [26] TJØSTHEIM, D.: Non-linear time series and Markov chains. - Adv. in Appl. Probab. 22, 1990, 587-611.
- [27] TONG, H.: Non-linear time series. A dynamical system approach. - Oxford Science Publications, Clarendon Press, Oxford, 1990.
- [28] TWEEDIE, R.L.: Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space. - Stochastic Process. Appl. 3, 1975, 385-403.
- [29] TWEEDIE, R.L.: Criteria for classifying general Markov chains. - Adv. in Appl. Probab. 8, 1976, 737-771.
- [30] VERRIEST, E.I., and A.H. HADDAD: Linear Markov approximations of piecewise linear stochastic systems. - Stochastic Anal. Appl. 5, 1987, 213-244.
- [31] WONG, P.W., and R.M. GRAY: Sigma-delta modulation with i.i.d. Gaussian inputs. - IEEE Trans. Inform. Theory 36, 1990, 784-778.