# Optimal Mean-Square Noise Benefits in Quantizer-Array Linear Estimation

Ashok Patel and Bart Kosko

***Abstract*—A new theorem shows that additive quantizer noise decreases the mean-squared error of threshold-array optimal and suboptimal linear estimators. The initial rate of this noise benefit improves as the number of threshold sensors or quantizers increases. The array sums the outputs of identical binary quantizers that receive the same random input signal. The theorem further shows that zero-symmetric uniform quantizer noise gives the fastest initial decrease in mean-squared error among all finite-variance zero-symmetric scale-family noise. These results apply to all bounded continuous signal densities and all zero-symmetric scale-family quantizer noise with finite variance.**

***Index Terms*—Noise benefit, quantizer array, scale-family noise, suprathreshold stochastic resonance.**

## I. Quantizer Noise Benefits in Parallel Threshold Arrays

The noise-benefit theorem below shows that small amounts of additive quantizer noise can reduce the initial rate of the mean-squared quantization error of an array of parallel connected threshold sensors or quantizers. Fig. 1 displays this array noise benefit for several types of quantizer noise. This result extends to the mean-square case recent results for Neyman–Pearson and maximum-likelihood quantization noise benefits in array correlation detectors [1]. These results are array or *collective* noise benefits and thus examples of what Stocks has called "suprathreshold SR" effects [2], [3].

Noise can benefit many signal processing systems if the user judiciously adds noise or exploits the noise already present in the system [2], [4]–[8]. This noise benefit or *stochastic resonance* (SR) effect requires some form of nonlinear signal processing [9], [10]. Then a small amount of noise improves the system performance while too much noise degrades it. Such noise benefits in threshold arrays can increase the mutual information between the array input and output [11], decrease the mean-squared error [12], [13] or increase the Fisher information in signal or parameter estimation [14], or improve the detection performance in signal detection or hypothesis testing [15]–[18].

Manuscript received May 10, 2010; revised June 21, 2010; accepted June 25, 2010. Date of publication July 19, 2010; date of current version November 04, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Aleksandar Dogandzic.

The authors are with the Signal and Image Processing Institute, Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-2564 USA (e-mail: kosko@usc.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2010.2059376

Each threshold element in the quantizer array receives the same random input signal $X$ but with independent additive symmetric scale-family noise $N_m$. The continuous signal random variable $X$ has bounded probability density function (pdf) $f_X$ and has a finite variance $V[X]$. The $m$th quantizer output $Y_m$ is 1 or 0 depending on whether the noisy input $X+N_m$ exceeds the quantization threshold $\theta$:

$$Y_m = \begin{cases} 1 & \text{if } X+N_m > \theta \\ 0 & \text{else} \end{cases}. \tag{1}$$

The overall array output $Y = \sum_{m=1}^{M} Y_m$ just sums the $M$ quantizer outputs. Such quantizer arrays can apply in stochastic analog-to-digital converters [12], [19], stochastic pooling networks [16], [20], or in digital data acquisition in noisy sensor networks that use binary sensors with independent and identically distributed (i.i.d.) noise [21], [22]. The independent quantizer noise variables $N_m$ may be present near the threshold sensors or the user may deliberately add them before quantization.

The noise variables $N_m$ are i.i.d. and have a zero-symmetric (thus zero-mean) scale-family [23] pdf $f_N(\sigma_N, n) = (1/\sigma_N) f_{\tilde{N}}(n/\sigma_N)$ with finite standard deviation $\sigma_N$. The standard noise variable $\tilde{N}$ of this family has unit variance as well as zero mean. So the symmetric scale-family cumulative distribution function (CDF) is $F_N(\sigma_N, n) = F_{\tilde{N}}(n/\sigma_N)$ if $F_{\tilde{N}}$ is the CDF of $\tilde{N}$. The noise $N_m$ can be discrete or continuous. Scale-family densities include many common pdfs such as the Gaussian and uniform. They do not include Poisson or binomial pdfs.

The overall array output $Y$ is a discrete random variable that takes values in $\{0, 1, \ldots, M\}$. The conditional quantizer outputs $Y_m|X=x$ are independent. The probability that $Y_m|X=x$ equals 1 is $1 - F_{\tilde{N}}((\theta - x)/\sigma_N)$ and that it equals 0 is $F_{\tilde{N}}((\theta - x)/\sigma_N)$. So $Y|X{=}x$ is a binomial random variable and thus has mean

$$E[Y|X=x] = M\left(1 - F_{\tilde{N}}\left(\frac{\theta - x}{\sigma_N}\right)\right) \tag{2}$$

and variance

$$V[Y|X=x] = M\left(1 - F_{\tilde{N}}\left(\frac{\theta - x}{\sigma_N}\right)\right) F_{\tilde{N}}\left(\frac{\theta - x}{\sigma_N}\right). \tag{3}$$

We focus on noise benefits in terms of the mean-squared error (MSE) $E\left[(X - \hat{X}(Y))^2\right]$ between the input signal $X$ and its linear estimators $\hat{X}(Y) = aY + b$. The SR effect or noise benefit occurs if an increase in the quantizer noise intensity $\sigma_N$ decreases the MSE for the same input signal $X$. We define the SR effect as an *initial SR effect* if there exists some $h > 0$ such that $MSE(\sigma_N) < MSE(0)$ for all $\sigma_N \in (0, h)$. We define the

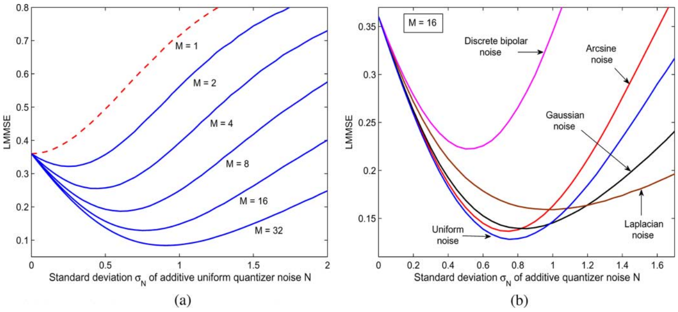


Fig. 1. Initial SR noise benefits in the mean-squared error (MSE) of the quantizer-array linear minimum MSE (LMMSE) estimator (4) for a standard Gaussian input signal $X$. Each curve shows the LMMSE where sample statistics estimated the population statistics in (5) using $10^5$ random-sample realizations. (a) Initial SR effects for zero-mean uniform quantizer noise. The solid lines show that the LMMSE decreases at first as the quantizer noise intensity $\sigma_N$ increases. The dashed line shows that the SR effect does not occur if $M = 1$ as part (a) of the theorem predicts. The solid lines also show that the rate of the initial SR effect $\lim_{\sigma_N \to 0}(dLMMSE(\sigma_N)/d\sigma_N)$ improves as the number $M$ of quantizers increases as part (b) of the theorem predicts. $M = 32$ quantizers gave 0.085 as the MSE minimum. It thus gave a 76% decrease over the noiseless MSE of 0.36. (b) compares the initial rates of the LMMSE SR effect in the quantizer-array LMMSE estimator for different types of symmetric quantizer noise when $M = 16$. Zero-symmetric uniform noise gave the fastest initial decrease in MSE as part (c) of the theorem predicts. Zero-symmetric discrete bipolar noise gave the smallest SR effect.

initial rate of the SR effect as the rate of the SR effect in the small-noise limit: $\lim_{\sigma_N \to 0}(dMSE(\sigma_N)/d\sigma_N)$.

## II. MSE Noise Benefits in Quantizer-Array Linear Estimation

The MSE noise-benefit theorem uses only linear estimators $\hat{X}(Y)$ of $X$ based on the output $Y$ of the parallel quantizer array. The conditional expectation $E[X|Y]$ is the minimum mean-squared-error (MMSE) estimator [2] as well as the Bayesian estimator for a squared-error loss measure. It can have a complicated form because it averages against the conditional pdf $f(x|y)$. The conditional pdf $f(x|y)$ may be hard to compute or may not even be known. The related linear minimum mean-squared-error (LMMSE) estimator $\hat{X}_{LMMSE}$ is suboptimal in general but requires only the population covariance $Cov[X,Y]$, variance $V[Y]$, and expected values $E[X]$ and $E[Y]$ [24]:

$$\hat{X}_{LMMSE} = a^* Y + b^* \tag{4}$$

with coefficients $a^* = Cov[X,Y]/V[Y]$ and $b^* = E[X] - a^* E[Y]$. We use the sum $Y$ of quantizer outputs $Y_m$ instead of their vector $\mathbf{Y} = [Y_1, \ldots, Y_M]^T$. This causes no loss of information because both give the same LMMSE estimate of $X$ from the orthogonality principle since all the quantizers receive the same input signal and use i.i.d. quantizer noise. The LMMSE term $E_{X,Y}\left[(X - \hat{X}_{LMMSE}(Y))^2\right]$ is [24]

$$LMMSE = V[X] - \frac{Cov^2[X,Y]}{V[Y]}. \tag{5}$$

The proof of the MSE noise-benefit theorem uses two lemmas. The first lemma gives the initial rate of the LMMSE SR effect. Its proof is in the Appendix. All the results assume that the signal pdf $f_X$ is bounded: there exists a constant $B$ such that $f_X(x) \le B$ for all $x$.

*Lemma 1:* The initial rate of the SR effect in the LMMSE is

$$\begin{aligned}\lim_{\sigma_N \to 0} & \frac{dLMMSE}{d\sigma_N} \\ &= \lim_{\sigma_N \to 0} \frac{dV[Y]}{d\sigma_N} \\ &\quad \times \frac{\left[\int_\theta^\infty xMf_X(x)dx - E[X]M(1-F_X(\theta))\right]^2}{\left[M^2(1-F_X(\theta))F_X(\theta)\right]^2}\end{aligned} \tag{6}$$

if the bounded signal pdf $f_X$ is continuous at $\theta$.

Lemma 2 involves the linear mean-squared error (LMSE) term $E_{X,Y}\left[(X - \hat{X}_{LMSE}(Y))^2\right]$ of a simple but suboptimal linear estimator $\hat{X}_{LMSE}(Y)$:

$$\hat{X}_{LMSE}(Y) = \tilde{c}(2Y - M) \tag{7}$$

where $\tilde{c} = c/M$ for $M$ quantizer-array outputs. This estimator is useful if we know only that the input signal lies in the range $[-c, c]$ for some constant $c>0$.

Lemma 2 uses Lemma 1 and shows that the initial rate of the LMSE and LMMSE SR effects are proportional. Its proof is also in the Appendix.

*Lemma 2:*

$$\lim_{\sigma_N \to 0} \frac{dLMSE}{d\sigma_N} = K \lim_{\sigma_N \to 0} \frac{dLMMSE}{d\sigma_N} \tag{8}$$

for $K = 4\tilde{c}^2 \left( \left[M^2(1-F_X(\theta))F_X(\theta)\right]^2 / \left[\int_\theta^\infty xMf_X(x)dx - E[X]M(1-F_X(\theta))\right]^2 \right)$ if the bounded signal pdf $f_X$ is continuous at $\theta$.

The MSE noise-benefit theorem states that it takes more than one quantizer to produce the initial SR effect in the LMMSE or LMSE estimators and that the rate of the initial SR effect increases as the number $M$ of quantizers increases. It further states that uniform quantizer noise gives the fastest initial decrease in MSE among all finite-variance symmetric scale-family quantizer noise. The theorem holds for both optimal and suboptimal linear estimators. These results on the initial rate of noise benefits can apply in practice when the user has only a small amount of noise to enhance preferred signals. They resemble but differ from the array SR results in [1], [17] that hold only for quantizer-array correlation detectors based on Neyman–Pearson or maximum-likelihood detection in infinite-variance symmetric alpha-stable channel noise or in generalized-Gaussian channel noise. We note that the array-SR result in [18] can hold even for $M = 1$ quantizer when the performance measure is error probability. We also note that the result in [25] gives the optimal standard deviation of the quantizer noise for the suboptimal estimator (7) if the input and quantizer noise are both Gaussian or both uniform. But we do not know the optimal quantizer noise for either the LMSE or the LMMSE even if the input signal is Gaussian or uniform.

*MSE Noise Benefit Theorem:* Suppose that the $M$ threshold sensors or quantizers in (1) receive the same input signal $X$ with bounded signal density $f_X$ that is positive and continuous at the threshold $\theta$. Suppose that the additive quantizer noise $N_m$ in (1) is i.i.d. zero-symmetric scale-family noise with finite variance. Then

a) $M>1$ is necessary and sufficient for the initial LMMSE and LMSE SR effects.
b) $M_2$ quantizers give a faster initial decrease in MSE than $M_1$ quantizers give if $M_2 > M_1$.
c) Zero-mean uniform noise is optimal: it gives the fastest initial decrease in MSE among all zero-symmetric scale-family quantizer noise with finite variance.

Fig. 1 shows simulation instances of the theorem for the LMMSE estimator (4). $M = 32$ quantizers in Fig. 1(a) gave the minimal LMMSE of 0.085. It thus gave a 76% decrease over the noiseless LMMSE of 0.36. Fig. 1(b) shows that zero-symmetric uniform noise gave the maximal rate of the initial SR effect as part (c) the theorem predicts. Fig. 1(b) also shows that arcsine noise had a larger MSE than did uniform quantizer noise even though arcsine noise is asymptotically (large $M$) optimal if the performance measure is mutual information [26].

*Proof:* We need prove only claims (a)–(c) for the initial rate of the LMSE SR effect $(\lim_{\sigma_N\to 0}(dLMSE/d\sigma_N))$ because the initial rate of the LMSE and LMMSE SR effects are proportional by Lemma 2. So we first find the limit $\lim_{\sigma_N\to 0}(dLMSE/d\sigma_N)$. Equation (44) in the Appendix implies that $\lim_{\sigma_N\to 0}(dLMSE/d\sigma_N)= \lim_{\sigma_N\to 0}(d\,\tilde{c}^2 E[(2Y-M)^2]/d\sigma_N)$ because of (7) and (37). Total expectation gives

$$\begin{aligned}
&\tilde{c}^2 E[(2Y-M)^2]\\
&\quad= \tilde{c}^2 E_X\left[E[(2Y-M)^2|X]\right] && (9)\\
&\quad= \tilde{c}^2 E_X\left[4E[Y^2|X]-4ME[Y|X]+E[M^2|X]\right] && (10)\\
&\quad= 4\tilde{c}^2 E_X\left[V[Y|X]+E^2[Y|X]-ME[Y|X]\right]\\
&\qquad+\tilde{c}^2M^2 && (11)\\
&\quad= \tilde{c}^2M^2+4\tilde{c}^2\\
&\qquad\times E_X\left[M\left(1-F_{\tilde{N}}\left(\frac{\theta-X}{\sigma_N}\right)\right)F_{\tilde{N}}\left(\frac{\theta-X}{\sigma_N}\right)\right.\\
&\qquad\quad+M^2\left(1-F_{\tilde{N}}\left(\frac{\theta-X}{\sigma_N}\right)\right)^2\\
&\qquad\quad\left.-M^2\left(1-F_{\tilde{N}}\left(\frac{\theta-X}{\sigma_N}\right)\right)\right] && (12)
\end{aligned}$$

because of (2) and (3). Then the distributional derivative [27] of (12) with respect to $\sigma_N$ is

$$\begin{aligned}
&\frac{dLMSE}{d\sigma_N}\\
&=4\tilde{c}^2\frac{d}{d\sigma_N}E\left[M(1-M)\left[1-F_{\tilde{N}}\left(\frac{\theta-X}{\sigma_N}\right)\right]F_{\tilde{N}}\left(\frac{\theta-X}{\sigma_N}\right)\right] && (13)\\
&=4\tilde{c}^2M(1-M)\frac{d}{d\sigma_N}\\
&\int_{-\infty}^{\infty}\left(1-F_{\tilde{N}}\left(\frac{\theta-x}{\sigma_N}\right)\right)F_{\tilde{N}}\left(\frac{\theta-x}{\sigma_N}\right)f_X(x)dx. && (14)
\end{aligned}$$

This allows [27] us to interchange the order of differentiation and integration in (14):

$$\begin{aligned}
\frac{dLMSE}{d\sigma_N} &= 4\tilde{c}^2M(M-1)\\
&\quad\times\int_{-\infty}^{\infty}\left[1-2F_{\tilde{N}}\left(\frac{\theta-x}{\sigma_N}\right)\right]\\
&\quad\times\frac{\theta-x}{\sigma_N^2}f_{\tilde{N}}\left(\frac{\theta-x}{\sigma_N}\right)f_X(x)dx && (15)\\
&=4\tilde{c}^2M(M-1)\\
&\quad\times\int_{-\infty}^{\infty}[1-2F_{\tilde{N}}(n)]nf_{\tilde{N}}(n)f_X(\theta-\sigma_N n)dn && (16)
\end{aligned}$$

if we substitute $((\theta-x)/\sigma_N)=n$ in (15). Then we can fix $n$ when we take the $\sigma_N$-limit:

$$\begin{aligned}
&\lim_{\sigma_N\to 0}\frac{dLMSE}{d\sigma_N}\\
&\quad=\lim_{\sigma_N\to 0}4\tilde{c}^2M(M-1)\\
&\qquad\times\int_{-\infty}^{\infty}[1-2F_{\tilde{N}}(n)]nf_X(\theta-\sigma_N n)f_{\tilde{N}}(n)dn && (17)\\
&\quad=4\tilde{c}^2M(M-1)\\
&\qquad\times\int_{-\infty}^{\infty}[1-2F_{\tilde{N}}(n)]n\lim_{\sigma_N\to 0}f_X(\theta-\sigma_N n)f_{\tilde{N}}(n)dn && (18)\\
&\quad=4\tilde{c}^2M(M-1)f_X(\theta)\int_{-\infty}^{\infty}[1-2F_{\tilde{N}}(n)]nf_{\tilde{N}}(n)dn. && (19)
\end{aligned}$$

Equality (18) holds by the dominated convergence theorem [28] because $|[1-2F_{\tilde{N}}(n)]nf_X(\theta-\sigma_N n)|\leq|n|B$ for bound $B$ on $f_X$ and because $E[|\tilde{N}|]<\infty$ since $\tilde{N}$ has finite variance. Equality (19) holds because $f_X$ is continuous at $\theta$ and thus $\lim_{\sigma_N\to 0}f_X(\theta-\sigma_N n)=f_X(\theta)$.

a) The limit (19) implies that the initial rate of the SR effect $\lim_{\sigma_N\to 0}(dLMSE/d\sigma_N)=0$ if $M=1$. The integral of (19) is negative because $[1-2F_{\tilde{N}}(n)]n$ is non-positive since $\tilde{N}$ is a symmetric random variable. So $\lim_{\sigma_N\to 0}(dLMSE/d\sigma_N)<0$ and thus the MSE decreases initially if $M>1$. This proves part (a).
b) The limit (19) also implies that the initial rate of the LMSE SR effect $\lim_{\sigma_N\to 0}(dLMSE/d\sigma_N)$ improves (increases the initial decrease in the MSE) as the number $M$ of quantizers increases. This proves part (b).
c) Fix the input signal $X$ and the number $M$ of quantizers and choose the symmetric scale-family quantizer noise $N$. The integral in (19) is non-positive because $[1-2F_{\tilde{N}}(n)]n$ is non-positive and because $f_{\tilde{N}}$ is a pdf. So we want to find the symmetric standard (zero-mean and unit-variance) pdf $f_{\tilde{N}}$ of the zero-mean unit-variance scale-family quantizer noise $\tilde{N}$ that minimizes the expectation $\int_{-\infty}^{\infty}[1-2F_{\tilde{N}}(n)]nf_{\tilde{N}}(n)dn$. But this is equivalent to maximizing

$$\int_{-\infty}^{\infty}|[1-2F_{\tilde{N}}(n)]nf_{\tilde{N}}(n)|dn. \qquad (20)$$

Suppose first that the quantizer noise $\tilde{N}$ is a symmetric discrete random variable on the sample space $\Omega_{\tilde{N}} = \{n_\ell \in \mathbf{R} : \ell \in \mathcal{L} \subseteq \mathbb{Z}\}$ where $n_0 = 0$. Suppose also that $n_{-\ell} = -n_\ell$ and that $n_\ell > 0$ for all $\ell \in \mathbb{N} = \{1, 2, 3, \ldots\}$. Let $P_{\tilde{N}}(n_\ell) = P(\tilde{N} = n_\ell) = (p_\ell/2)$ denote the pdf of this symmetric standard discrete quantizer noise $\tilde{N}$ where $P_{\tilde{N}}(n_{-\ell}) = (p_\ell/2) = P_{\tilde{N}}(n_\ell)$ for all $\ell \in \mathbb{N}$ and $P_{\tilde{N}}(n_0) = p_0$. Then the finite variance of $\tilde{N}$ lets us replace (20) with the appropriate convergent series $\sum_{\ell=-\infty}^{\infty} |[1 - 2F_{\tilde{N}}(n_\ell)]n_\ell P_{\tilde{N}}(n_\ell)|$. Write

$$\sum_{\ell=-\infty}^{\infty} |[1 - 2F_{\tilde{N}}(n_\ell)]n_\ell P_{\tilde{N}}(n)| = 2\sum_{\ell=1}^{\infty} [F_{\tilde{N}}(n_\ell) - F_{\tilde{N}}(-n_\ell)]n_\ell P_{\tilde{N}}(n_\ell) \quad (21)$$

because $E[\tilde{N}] = 0$, $n_{-\ell} = -n_\ell$, and $P_{\tilde{N}}(-n_\ell) = P_{\tilde{N}}(n_\ell)$ and because $F_{\tilde{N}}$ is a CDF and so $F_{\tilde{N}}(n_\ell) - F_{\tilde{N}}(-n_\ell)$ is nonnegative. The derivation in the Appendix of [1] shows that

$$2\sum_{\ell=1}^{\infty} [F_{\tilde{N}}(n_\ell) - F_{\tilde{N}}(-n_\ell)]n_\ell P_{\tilde{N}}(n_\ell) \leq \frac{1}{\sqrt{3}}. \quad (22)$$

So (22) gives the upper bound $1/\sqrt{3}$ on (21) for any discrete symmetric standard quantizer noise $\tilde{N}$.

Suppose next that the quantizer noise pdf $f_N$ is continuous. The Cauchy-Schwarz inequality [29] and the finite variance of $\tilde{N}$ give

$$\begin{aligned}
\int_{-\infty}^{\infty} &|[1 - 2F_{\tilde{N}}(n)]n f_{\tilde{N}}(n)|\, dn \\
&\leq \left[\int_{-\infty}^{\infty} [1 - 2F_{\tilde{N}}(n)]^2 f_{\tilde{N}}(n) dn\right]^{1/2} \\
&\quad \times \left[\int_{-\infty}^{\infty} n^2 f_{\tilde{N}}(n) dn\right]^{1/2} && (23) \\
&= \left[\int_{-\infty}^{\infty} [1 - 2F_{\tilde{N}}(n)]^2 f_{\tilde{N}}(n) dn\right]^{1/2} \\
&\quad \text{because } E[\tilde{N}^2] = 1 && (24) \\
&= \left[E\left[[1 - 2F_{\tilde{N}}(\tilde{N})]^2\right]\right]^{1/2} && (25) \\
&= \left[E\left[(1 - 2U)^2\right]\right]^{1/2} && (26) \\
&= \frac{1}{\sqrt{3}} && (27)
\end{aligned}$$

because the transformed random variable $U = F_{\tilde{N}}(\tilde{N})$ is uniform in $[0, 1]$ since $F_{\tilde{N}}$ is continuous when $\tilde{N}$ is continuous [23]. Inequality (23) becomes an equality if and only if $F_{\tilde{N}}$ obeys $[1 - 2F_{\tilde{N}}(n)]^2 = \kappa n^2$ for some constant $\kappa$ on the support of $f_{\tilde{N}}$ [29]. Then $F_{\tilde{N}}(n) = (1/2) + (\sqrt{\kappa} n/2)$ for all $n \in [(-1/\sqrt{\kappa}), (1/\sqrt{\kappa})]$ for $\kappa = 1/3$ because $F_{\tilde{N}}$ is the CDF of standard quantizer noise. The same CDF implies that $\tilde{N}$ is uniformly distributed in $[-\sqrt{3}, \sqrt{3}]$. So symmetric uniform quantizer noise achieves the upper bound $1/\sqrt{3}$ of (22) and (27). Then zero-mean uniform noise gives the maximal rate of the initial SR effect among all zero-symmetric scale-family quantizer noise with finite variance. □

## Appendix
## Proof of Lemmas

*Proof of Lemma 1:* Equation (5) implies that

$$\begin{aligned}
\lim_{\sigma_N \to 0} &\frac{dLMMSE}{d\sigma_N} \\
&= -\frac{\lim_{\sigma_N \to 0} V[Y] 2Cov[X, Y] \frac{dCov[XY]}{d\sigma_N}}{\lim_{\sigma_N \to 0} V^2[Y]} \\
&\quad - \frac{\lim_{\sigma_N \to 0} Cov^2[X, Y] \frac{dV[Y]}{d\sigma_N}}{\lim_{\sigma_N \to 0} V^2[Y]} && (28)
\end{aligned}$$

where $(d/d\sigma_N)$ is the distributional derivative [27] with respect to the quantizer-noise standard deviation $\sigma_N$.

The first term in the numerator of (28) is zero because

$$\begin{aligned}
\lim_{\sigma_N \to 0} \frac{dCov[X, Y]}{d\sigma_N} &= \lim_{\sigma_N \to 0} \frac{dE[XY]}{d\sigma_N} \\
&\quad - E[X] \lim_{\sigma_N \to 0} \frac{dE[Y]}{d\sigma_N} && (29) \\
&= 0. && (30)
\end{aligned}$$

Equation (30) follows because both terms on the right side of (29) are zero:

$$\begin{aligned}
\lim_{\sigma_N \to 0} &\frac{dE[XY]}{d\sigma_N} \\
&= \lim_{\sigma_N \to 0} \frac{dE\left[XE[Y|X]\right]}{d\sigma_N} \text{ by total expectation} && (31) \\
&= \lim_{\sigma_N \to 0} \frac{d}{d\sigma_N} \int_{-\infty}^{\infty} xM(1 - F_{\tilde{N}}\left(\frac{\theta - x}{\sigma_N}\right)) f_X(x) dx && (32) \\
&= \lim_{\sigma_N \to 0} \int_{-\infty}^{\infty} xM \frac{\theta - x}{\sigma_N^2} f_{\tilde{N}}\left(\frac{\theta - x}{\sigma_N}\right) f_X(x) dx && (33)
\end{aligned}$$

because the distributional derivative with respect to $\sigma_N$ allows [27] us to interchange the order of integration and differentiation in (33). Then putting $((\theta - x)/\sigma_N) = n$ in (33) gives

$$\begin{aligned}
\lim_{\sigma_N \to 0} &\frac{dE[XY]}{d\sigma_N} \\
&= \lim_{\sigma_N \to 0} M \int_{-\infty}^{\infty} (\theta - n\sigma_N) n f_X(\theta - n\sigma_N) f_{\tilde{N}}(n) dn && (34) \\
&= M \int_{-\infty}^{\infty} \lim_{\sigma_N \to 0} (\theta - n\sigma_N) n f_X(\theta - n\sigma_N) f_{\tilde{N}}(n) dn && (35) \\
&= M \int_{-\infty}^{\infty} \theta n f_X(\theta) f_{\tilde{N}}(n) dn \\
&\quad \text{because } f_X \text{ is continuous at } \theta && (36) \\
&= 0. && (37)
\end{aligned}$$

Equality (35) holds by the dominated convergence theorem [28] because $|(\theta - n\sigma_N) n f_X(\theta - n\sigma_N)| \leq (|\theta| + |n|)|n| B$ for bound $B$ on $f_X$ when $\sigma_N \leq 1$ and because $E[|\tilde{N}|] < \infty$ since $\tilde{N}$ has finite variance. Equality (37) holds because the quantizer noise $\tilde{N}$ has zero mean. A similar argument gives

$$\lim_{\sigma_N \to 0} \frac{dE[Y]}{\sigma_N} = 0. \quad (38)$$

So $\lim_{\sigma_N \to 0}(dCov[X,Y]/d\sigma_N) = 0$ and thus the first term in the numerator of (28) is zero.

The other two limiting terms in (28) expand as

$$\lim_{\sigma_N \to 0} Cov^2[X,Y] = \lim_{\sigma_N \to 0} (E[XY] - E[X]E[Y])^2 \tag{39}$$

$$= \left( \lim_{\sigma_N \to 0} E[XY] - E[X] \lim_{\sigma_N \to 0} E[Y] \right)^2 \tag{40}$$

$$= \left[ \int_\theta^\infty x M f_X(x) dx - E[X] M (1 - F_X(\theta)) \right]^2 \tag{41}$$

and

$$\lim_{\sigma_N \to 0} V^2[Y] = \left( \lim_{\sigma_N \to 0} V[Y] \right)^2 = [M^2 (1 - F_X(\theta)) F_X(\theta)]^2. \tag{42}$$

(41) and (42) follow because $Y$ is a binary random variable in the absence of quantizer noise $N$ and thus $Y = M$ if $X > \theta$ and $Y = 0$ else. Putting (30), (41), and (42) in (28) gives (6). □

*Proof of Lemma 2:* The linear mean-squared error $E\left[(X - \hat{X}_{LMSE}(Y))^2\right]$ or LMSE of $\hat{X}_{LMSE}(Y)$ is

$$LMSE = E\left[(X - \tilde{c}(2Y - M))^2\right] \text{ by (7)} \tag{43}$$

$$= E[X^2] - 4\tilde{c}E[XY] + 2\tilde{c}ME[X] + \tilde{c}^2 E[(2Y - M)^2] \tag{44}$$

$$= E[X^2] - 4\tilde{c}E[XY] + 2\tilde{c}ME[X] + 4\tilde{c}^2 \left( V[Y] + E^2[Y] - ME[Y] \right) + \tilde{c}^2 M^2. \tag{45}$$

Then (45) implies that the initial rate of the LMSE SR effect is

$$\lim_{\sigma_N \to 0} \frac{dLMSE}{d\sigma_N} = -4\tilde{c} \lim_{\sigma_N \to 0} \frac{dE[XY]}{\sigma_N} + 4\tilde{c}^2 \lim_{\sigma_N \to 0} \frac{dV[Y]}{\sigma_N} + 4\tilde{c}^2 \lim_{\sigma_N \to 0} \left[ (2E[Y] - M) \frac{dE[Y]}{\sigma_N} \right] \tag{46}$$

$$= 4\tilde{c}^2 \lim_{\sigma_N \to 0} \frac{dV[Y]}{\sigma_N} \text{ because of (37) and (38)} \tag{47}$$

$$= 4\tilde{c}^2 \lim_{\sigma_N \to 0} \frac{dLMMSE}{d\sigma_N} \times \frac{\left[M^2(1 - F_X(\theta))F_X(\theta)\right]^2}{\left[\int_\theta^\infty x M f_X(x) dx - E[X] M (1 - F_X(\theta))\right]^2} \tag{48}$$

because of (6). This proves (8). □

## REFERENCES


[1] A. Patel and B. Kosko, "Noise benefits in quantization-array correlation detection and watermark decoding," *IEEE Trans. Signal Process.*, to be published.

[2] M. D. McDonnell, N. G. Stocks, C. E. M. Pearce, and D. Abbott, *Stochastic Resonance: From Suprathreshold Stochastic Resonance to Stochastic Signal Quantization*. Cambridge, U.K.: Cambridge University Press, 2008.

[3] N. G. Stocks, "Suprathreshold stochastic resonance in multilevel threshold systems," *Phys. Rev. Lett.*, vol. 84, no. 11, pp. 2310–2313, Mar. 2000.

[4] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, "Stochastic resonance," *Rev. Mod. Phys.*, vol. 70, no. 1, pp. 223–287, Jan. 1998.

[5] B. Kosko, *Noise*. New York: Viking/Penguin, 2006.

[6] A. Patel and B. Kosko, "Stochastic resonance in continuous and spiking neuron models with Levy noise," *IEEE Trans. Neural Netw.*, vol. 19, pp. 1993–2008, Dec. 2008.

[7] H. Chen, P. K. Varshney, S. M. Kay, and J. H. Michels, "Noise enhanced nonparametric detection," *IEEE Trans. Inform. Theory*, vol. 55, pp. 499–506, Feb. 2009.

[8] A. Patel and B. Kosko, "Optimal noise benefits in Neyman–Pearson and inequality-constrained statistical signal detection," *IEEE Trans. Signal Process.*, vol. 57, pp. 1655–1669, 2009.

[9] A. R. Bulsara and L. Gammaitoni, "Tuning in to noise," *Phys. Today*, vol. 49, pp. 39–47, Mar. 1996.

[10] M. D. McDonnell and D. Abbott, "What is stochastic resonance? definitions, misconceptions, debates, and its relevance to biology," *PLoS Comput. Biol.*, vol. 5, no. 5, pp. e1000348–e1000348, May 2009.

[11] N. G. Stocks, "Information transmission in parallel threshold arrays: Suprathreshold stochastic resonance," *Phys. Rev. E*, vol. 63, no. 4, pp. 041114–041123, Mar. 2001.

[12] M. McDonnell, N. Stocks, C. Pearce, and D. Abbott, "Analog to digital conversion using suprathreshold stochastic resonance," in *Proc. SPIE*, 2005, vol. 5649, pp. 75–84.

[13] D. Rousseau and F. Chapeau-Blondeau, "Noise-improved bayesian estimation with arrays of one-bit quantizers," *IEEE Trans. Instrum. Meas.*, vol. 56, pp. 2658–2652, Dec. 2007.

[14] F. Chapeau-Blondeau, S. Blanchard, and D. Rousseau, "Fisher information and noise-aided power estimation from one-bit quantizers," *Digital Signal Process.*, vol. 18, no. 3, pp. 434–443, 2008.

[15] D. Rousseau, G. V. Anand, and F. Chapeau-Blondeau, "Noise-enhanced nonlinear detector to improve signal detection in non-Gaussian noise," *Signal Process.*, vol. 86, no. 11, pp. 3456–3465, Nov. 2006.

[16] S. Zozor, P.-O. Amblard, and C. D. Duchêne, "On pooling networks and fluctuation in suboptimal detection framework," *Fluct. Noise Lett.*, vol. 7, no. 1, pp. L39–L60, 2007.

[17] A. Patel and B. Kosko, "Quantizer noise benefits in nonlinear signal detection with alpha-stable channel noise," in *IEEE Int. Conf. Acoustics, Speech and Signal Processing, ICASSP 2009*, 2009, pp. 3269–3272.

[18] A. Patel and B. Kosko, "Error-probability noise benefits in threshold neural signal detection," *Neural Netw.*, vol. 22, no. 57–6, pp. 697–706, Jul./Aug. 2009.

[19] T. Nguyen, "Robust data-optimized stochastic analog-to-digital converters," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2735–2740, Jun. 2007.

[20] M. McDonnell, P.-O. Amblard, and N. Stocks, "Stochastic pooling networks," *J. Statist. Mech.: Theory Applicat.*, pp. P01012–P01012, 2009.

[21] T. Nguyen, D. Nguyen, H. Liu, and D. A. Tran, "Stochastic binary sensor networks for noisy environments," *Int. J. Sensor Netw.*, vol. 2, no. 5/6, pp. 414–427, 2007.

[22] Z.-Q. Luo, "Universal decentralized estimation in a bandwidth constrained sensor network," *IEEE Trans. Inform. Theory*, vol. 51, pp. 2210–2219, 2005.

[23] G. Casella and R. Berger, *Statistical Inference*. Westford, MA: Duxbury Resource Center, 2001.

[24] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[25] M. D. McDonnell, D. Abbott, and C. E. M. Pearce, "An analysis of noise enhanced information transmission in an array of comparators," *Microelectron. J.*, vol. 33, no. 12, pp. 1079–1089, 2002.

[26] M. D. McDonnell, N. G. Stocks, and D. Abbott, "Optimal stimulus and noise distributions for information transmission via suprathreshold stochastic resonance," *Phys. Rev. E*, vol. 75, no. 6, pp. 061105–061105, 2007.

[27] D. Jones, *The Theory of Generalized Functions*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1982.

[28] P. Billingsley, *Probability and Measure*, 3rd ed. New York: Wiley-Interscience, 1995.

[29] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*, 2nd ed. New York: Wiley-Interscience, 1999.