[7] S. Haykin, *Adaptive Filter Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1991.
[8] G. J. Bierman, *Factorization Method for Discrete Sequential Estimation.* New York: Academic, 1977.
[9] P. Eykhoff, *Trends and Progress in System Identification.* New York: Pergamon, 1981.
[10] G. H. Golub and C. F. Van Loan, *Matrix Computations,* 2nd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1989.

The Barankin Bound and Threshold Behavior in Frequency Estimation

Luc Knockaert

***Abstract*— This correspondence presents the Barankin bound as a fundamental statistical tool for the understanding of the threshold effect associated with the estimation of the frequency of a sinusoid in additive white Gaussian noise. It is shown that the threshold effect takes hold whenever the Barankin bound departs significantly from the Cramer–Rao bound. In terms of the signal-to-noise ratio (SNR) and the data length T, the quantity SNR $\times T/\ln T$ is shown to be a good indicator for deciding whether the SNR is above threshold or not.**

I. Introduction

The problem of estimating the frequency of a sinusoid in additive white Gaussian noise is one of considerable interest. In most cases [1], [2], the maximum likelihood (ML) procedure is utilized to obtain what one expects to be a sufficiently unbiased and efficient estimator of the frequency. Due to the nonlinear nature of the frequency estimation problem, the so-called *threshold effect* [2], [3] takes hold whenever the SNR drops below a critical data-length dependent level $\text{SNR}_c(T)$. The threshold effect can be characterized by an almost instant and drastic deterioration of the frequency estimator variance with respect to the Cramer–Rao bound (CRB) below this critical SNR level. In [2], the threshold effect was related to the existence of highly probable outliers that were fairly removed from the exact frequency, and in [3], a more technical device related to the phase-locked loop was proposed to explain the phenomenon. The aim of this correspondence is to provide a more fundamental approach to the understanding of the threshold effect. Our starting point is the fact that the CRB, although being the best lower bound in the linear Gaussian case, is a less appropriate tool when dealing with nonlinear problems such as frequency estimation. For nonlinear problems, the Barankin bound (BRB) [4], [5] is a stronger lower bound for the variances of unbiased estimators, including the CRB as a limit case. The threshold effect region can therefore, in the sense of Barankin, be defined as the region where the BRB suddenly departs from the CRB. This is fully exploited in the sequel, resulting in a simple indicator quantity for threshold behavior in frequency estimation.

Manuscript received March 13, 1996; revised March 27, 1997. The associate editor coordinating the review of this paper and approving it for publication was Dr. Ananthram Swami.

The author is with the Department of Information Technology, Intec, Gent, Belgium (e-mail: knokaert@intec.rug.ac.be).

Publisher Item Identifier S 1053-587X(97)06438-6.

II. The Barankin Bound

The simplest form of the BRB for the estimation of a scalar real parameter a can be stated as follows [5]. Let $p(\mathbf{x}|a)$ be the probability density of the vector $\mathbf{x}$, given a. Let h be a real number independent of $\mathbf{x}$ such that $a+h$ ranges over all possible values of a. Then for any unbiased estimator $\xi(\mathbf{x})$, we have

$$\text{var}\,(\xi) \geq \text{BRB} \tag{1}$$

where

$$\text{BRB} = \sup_h \frac{h^2}{\displaystyle\int \frac{p(\mathbf{x}|a+h)^2}{p(\mathbf{x}|a)} d\mathbf{x} - 1} \geq \text{CRB} \tag{2}$$

and the CRB is given by

$$\text{CRB} = \lim_{h\to 0} \frac{h^2}{\displaystyle\int \frac{p(\mathbf{x}|a+h)^2}{p(\mathbf{x}|a)} d\mathbf{x} - 1} = \frac{1}{\text{var}\left[\dfrac{\partial \ln p(\mathbf{x}|a)}{\partial a}\right]}. \tag{3}$$

To avoid theoretical complications, we assume that the integral in the denominator of (2) exists and that the support of $p(\mathbf{x}|a)$ and its partial derivative with respect to a is R for almost all a. A natural way to measure the deviation of the BRB from the CRB is the ratio

$$Q = \frac{\text{BRB}}{\text{CRB}} = \sup_h V(h) \tag{4}$$

where

$$V(h) = \frac{h^2\,\text{var}\left[\dfrac{\partial \ln\, p(\mathbf{x}|a)}{\partial a}\right]}{\displaystyle\int \frac{p(\mathbf{x}|a+h)^2}{p(\mathbf{x}|a)} d\mathbf{x} - 1}. \tag{5}$$

When $Q=1$, then the supremum in (5) is obtained for $h=0$, and in that case, we say that there is no Barankin threshold effect. This does not imply that there is no threshold effect whatsoever since there exist still stronger bounds than the above BRB [4], [5]. When $Q>1$, then the supremum is obtained for $h \neq 0$, and in that case, there surely exists a threshold effect since the BRB and, hence, the variance of the estimator then depart from the CRB.

To show what this means in practice, we apply this to a simple but frequently occurring nonlinear problem. Let the observed data vector $\mathbf{x}$ be given by

$$\mathbf{x} = \mathbf{f}(a) + \mathbf{n} \tag{6}$$

where $\mathbf{n}$ is $N(0, \sigma^2\mathbf{I})$ Gaussian noise, and $\mathbf{f}(a)$ is a function, in general nonlinear, mapping the parameter a into the data space.

After some elementary calculations, we obtain

$$V(h) = \frac{h^2|\mathbf{f}'(a)|^2}{\sigma^2(e^{|\mathbf{f}(a+h)-\mathbf{f}(a)|^2/\sigma^2} - 1)}. \tag{7}$$

Note that when the problem is linear, i.e., when $\mathbf{f}'(a)$ is a constant vector, $V(h)$ is a strictly decreasing function of h^2, which implies that $Q=1$. This is easily understood since linear problems in additive Gaussian noise never exhibit a threshold effect. Note also that when we have M independent realizations of the same process, the above formula remains valid after replacing σ^2 with σ^2/M.

III. Application to Frequency Estimation

Consider the single tone frequency estimation problem

$$z(t) = Ae^{j(at+\alpha)} + \nu(t) \qquad t = 0, 1, \cdots, T-1. \tag{8}$$

The $\nu(t)$ are white complex independent Gaussian random variables with the same noise variance σ^2, and the amplitude $A > 0$ and initial phase α are assumed known. The estimation of the angular frequency $-\pi \le a < \pi$ is of primary interest.

In the terms of the preceding section, we have a $2T$-dimensional data space

$$\begin{aligned} x_k &= \Re\{z(k-1)\} \\ x_{k+T} &= \Im\{z(k-1)\} \qquad k = 1, 2, \cdots, T \end{aligned} \tag{9}$$

and

$$\begin{aligned} f_k(a) &= A\cos[(k-1)a+\alpha] \\ f_{k+T}(a) &= A\sin[(k-1)a+\alpha] \qquad k = 1, 2, \cdots, T. \end{aligned} \tag{10}$$

In addition

$$\begin{aligned} f_k'(a) &= -A(k-1)\sin[(k-1)a+\alpha] \\ f_{k+T}'(a) &= A(k-1)\cos[(k-1)a+\alpha] \\ & k = 1, 2, \cdots, T \end{aligned} \tag{11}$$

and

$$\begin{aligned} |\mathbf{f}'(a)|^2 &= \sum_{k=1}^{T}\{|f_k'(a)|^2 + |f_{k+T}'(a)|^2\} \\ &= A^2\sum_{k=1}^{T}(k-1)^2 = A^2\,\frac{T}{6}\,(2T-1)(T-1). \end{aligned} \tag{12}$$

In the same vein, we have

$$\begin{aligned} |\mathbf{f}(a+h) - \mathbf{f}(a)|^2 &= 2A^2\left[T - 1 - \sum_{k=1}^{T-1}\cos(kh)\right] \\ &= A^2\left[2T - 1 - \frac{\sin(T-\frac{1}{2})h}{\sin\dfrac{h}{2}}\right] \end{aligned} \tag{13}$$

where the explicit result for the sum of cosinuses can be found in [6, p. 73]. Note that there is no dependence on the initial phase α. The objective function $V(h)$ can be written as

$$V(h) = \frac{\dfrac{\Gamma h^2 T(T-1)}{6}}{\exp\left\{\Gamma\left[1 - \dfrac{\sin(T-\frac{1}{2})h}{(2T-1)\sin\dfrac{h}{2}}\right]\right\} - 1} \tag{14}$$

where

$$\Gamma = (2T-1)\,\frac{A^2}{\sigma^2}. \tag{15}$$

In order to find $Q = \sup_h V(h)$, we need to know the range of h. Since $V(h)$ is an even function of h and since $a + h$ has to range over all possible values of a modulo 2π, we can take $0 \le h \le \pi$, and hence

$$Q = \sup_{0\le h\le\pi} V(h). \tag{16}$$

In order not to have immediate threshold behavior, we must require that $h = 0$ is at least a local maximum of the function $V(h)$. For h sufficiently small, we have

$$\begin{aligned} V(h) \approx 1 - h^2\,\frac{T(T-1)}{12}\left\{\Gamma - \left[\frac{3}{5} - \frac{1}{5T(T-1)}\right]\right\} \\ + O(h^4). \end{aligned} \tag{17}$$

For $h = 0$ to be a local maximum, a necessary condition is therefore

$$\Gamma > \tfrac{3}{5} - \frac{1}{5T(T-1)}. \tag{18}$$

The above inequality is always satisfied if we simply take $\Gamma > \frac{3}{5}$. Since the SNR (not in decibels) is defined as [2], [3] SNR $= A^2/2\sigma^2$, this implies that we should have

$$\text{SNR} \times T > \tfrac{3}{20}. \tag{19}$$

When (18) is not satisfied, $h = 0$ is a local minimum of $V(h)$, and Q will certainly always largely exceed unity. This means that the threshold effect will always be active in that case, resulting in large variances for any unbiased estimator of the angular frequency and thereby making efficient frequency estimation almost impossible. This confirms partly, but not completely, as we shall see in the sequel, the conclusions of [3], where the indicator quantity $3\sigma^2/A^2T$ was utilized, which quantity should therefore be smaller than 10. Note also that when we have M independent data realizations at our disposition, (19) remains valid after replacing T with MT, which the total number of data points.

On the other hand, when (18) is satisfied, the function $V(h)$ drops sharply in the vicinity of $h = 0$ and afterwards behaves as a parabola proportional to h^2 that is slightly modulated due to the presence of the factor $\sin[(T-0.5)h]/(2T-1)\sin(h/2)$. Hence, it appears that the local maximum closest to $h = \pi$ of the latter factor determines the global behavior of $V(h)$ in the vicinity of π. For T odd, this happens when $h = \pi$, and for T even, this happens when $h \approx \pi - \pi/(T-0.5)$. This is illustrated in Fig. 1, where $V(h)$ is plotted for SNR $\times\, T = 2.6, 2.4$, and $T = 33$. This allows us to infer that

$$Q \approx \max\left\{1, \frac{\dfrac{\Gamma\pi^2 T(T-1)}{6}}{\exp\dfrac{\Gamma(2T-2)}{(2T-1)} - 1}\right\} \tag{20}$$

where the approximation is exact for T odd. Hence, we conclude that we are certainly in the threshold region when

$$\frac{\dfrac{\Gamma\pi^2 T(T-1)}{6}}{\exp\dfrac{\Gamma(2T-2)}{\{(2T-1)\}} - 1} > 1 \tag{21}$$

since in that case, no unbiased estimator can achieve the exact CRB. On the other hand, we have chances of not being in the threshold region when

$$\frac{\dfrac{\Gamma\pi^2 T(T-1)}{6}}{\exp\dfrac{\Gamma(2T-2)}{\{(2T-1)\}} - 1} \ll 1. \tag{22}$$

Defining $\gamma = \Gamma(2T-2)/(2T-1)$, (21) can be written as

$$\frac{e^\gamma - 1}{\gamma} < (2T-1)T\,\frac{\pi^2}{12}. \tag{23}$$

Taking advantage of the fact that

$$e^{\gamma/2} \le \frac{e^\gamma - 1}{\gamma} \quad \forall\gamma \ge 0 \tag{24}$$

equation (23) can be strenghtened to

$$\gamma < 2\ln\left[(2T-1)T\,\frac{\pi^2}{12}\right]. \tag{25}$$

For sufficiently large T, this means that we are in the threshold region when

$$\text{SNR} \times \frac{T}{\ln T} < 1. \tag{26}$$

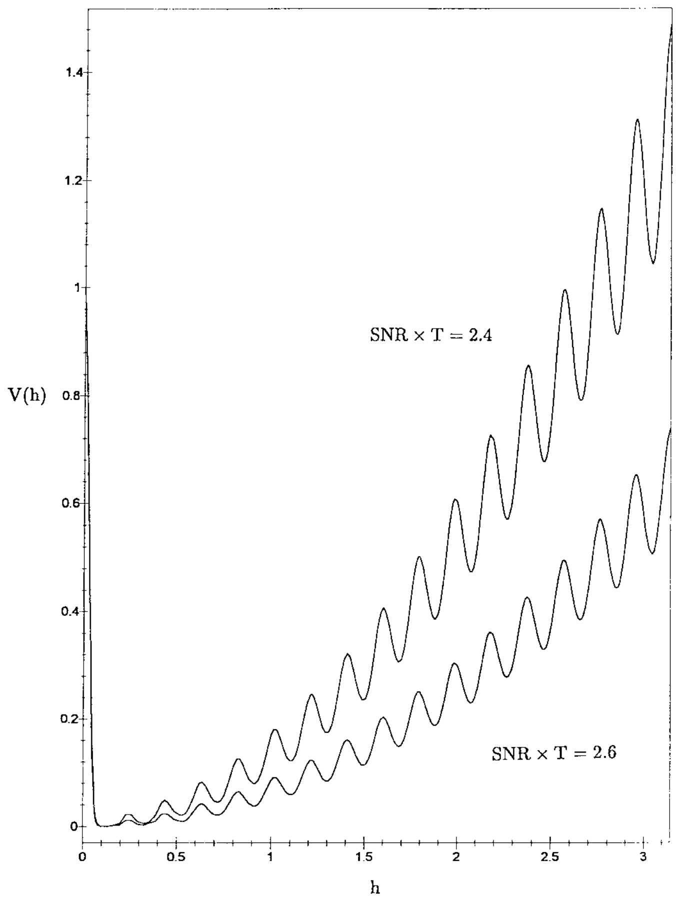

Fig. 1. $V(h)$ for SNR × $T = 2.6, 2.4$ and $T = 33$.

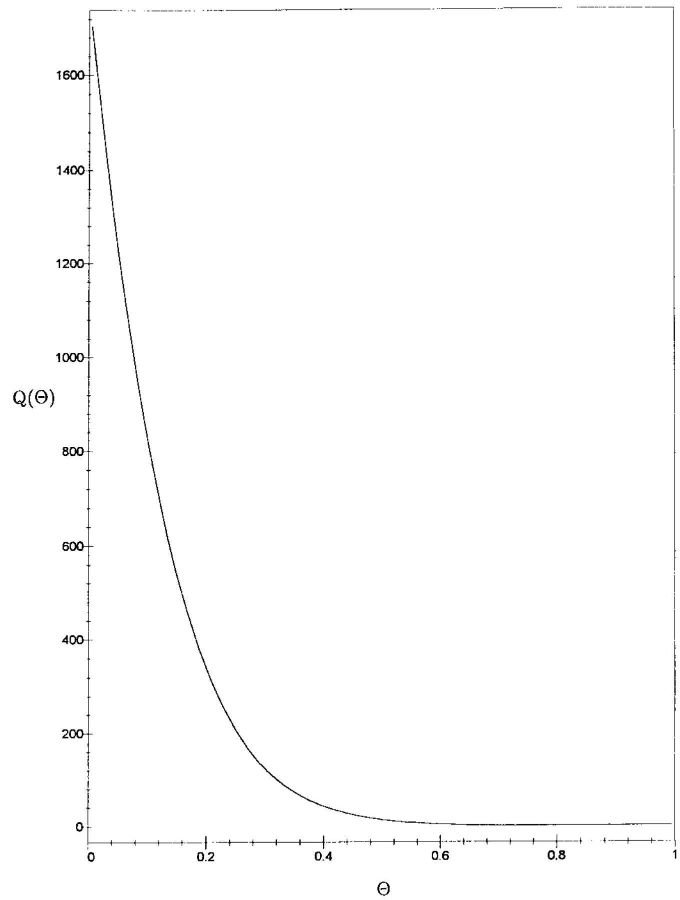

Fig. 2. Q as a function of Θ for $T = 33$.

This is further illustrated in Fig. 2, where Q is plotted as a function of Θ for fixed $T = 33$. By contrast, the property of being outside the threshold region may be described by the approximate inequality

$$\text{SNR} \times \frac{T}{\ln T} \gg 1. \tag{27}$$

Hence, it appears that at the level of the Barankin bound, the indicator quantity

$$\Theta = \text{SNR} \times \frac{T}{\ln T} \tag{28}$$

is crucial to the understanding of threshold behavior.

Based on the data in [3], we construct Table I, which pertains to the onset of the threshold effect for the maximum likelihood estimator of the angular frequency.

It is seen that the indicator quantity SNR × T is not a very good one since it clearly forms an increasing sequence and is therefore not likely to possess an upper bound. On the other hand, the quantity Θ forms a decreasing sequence and, therefore, always has an upper bound. Since $T/\ln T$ is an increasing function of T whenever $T > e$, the condition, say, $\Theta \geq 70$, will therefore always pull the ML estimator out of the threshold region. The reason why the Θ sequence is not approximately constant may be explained by the biasedness that typically affects ML estimators, especially when the number of data points is small [7, p. 426]. In the above context, the author would like to thank one of the reviewers for pointing out that in [8], a hybrid Barankin–Bhattacharryya bound was developed, with the possibility of incorporating biasedness corrections and a generalization to the multiple harmonics problem.

TABLE I

T	SNR(dB)	Θ	SNR × T
32	0	9.2	32.0
64	-2.75	8.2	34.0
128	-5.5	7.4	36.1
256	-8.25	6.9	38.3
512	-11	6.5	40.7

IV. Conclusion

A fundamental approach to the understanding of the threshold effect in frequency estimation, based on the Barankin bound, has been proposed. It is shown that threshold behavior is a typical nonlinear effect due to the departure of the Barankin bound from the Cramer–Rao bound at low SNR levels. A simple indicator quantity, with better behavior than the one proposed in [3], to characterize the onset of the threshold effect has been derived. The problem of including a possible biasedness corrector in the Barankin bound in the sense of [8], in order to come up with an even better indicator quantity, is the subject of ongoing research.

References

[1] T. J. Abatzoglou, “A fast maximum likelihood algorithm for frequency estimation of a sinusoid based on Newton’s method,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 77–89, Feb. 1985.

[2] D. C. Rife and R. R. Boorstyn, “Single tone parameter estimation from discrete-time observations,” *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 591–598, Sept. 1974.

[3] B. James, B. D. O. Anderson, and R. C. Williamson, "Characterization of threshold for single tone maximum likelihood frequency estimation," *IEEE Trans. Signal Processing,* vol. 43, pp. 817–821, Apr. 1995.
[4] E. W. Barankin, "Locally best unbiased estimates," *Ann. Math. Stat.,* vol. 20, pp. 477–501, 1949.
[5] H. L. Van Trees, *Detection, Estimation and Modulation Theory, Part I.* New York: Wiley, 1971.
[6] G. Pólya and G. Szegö, *Problems and Theorems in Analysis II.* New York: Springer, 1977.
[7] E. L. Lehmann, *Theory of Point Estimation.* Belmont, CA: Wadsworth, 1991.
[8] J. S. Abel, "A bound on mean-square-estimate error," *IEEE Trans. Inform. Theory,* vol. 39, pp. 1675–1680, Sept. 1993.

Full-Duplex Fast Estimation of Echo and Channel Responses in the Presence of Frequency Offsets in Both Far Echo and Far Signal

Weiping Li, Xixian Chen, Yi Wang, and Nobuhiro Miki

***Abstract*— In a previous publication, we proposed a full-duplex fast training procedure for simultaneously estimating echo and channel responses. It reduces the tap-setting time to half of that required by the traditional half-duplex fast training schemes. However, its estimation accuracy may be degraded by the frequency offsets in both far echo and far signal that are caused by the analog carrier network. In this correspondence, we extend the previous work and develop a new algorithm that can compensate for the phase rotation components introduced by these frequency offsets. The performance of the new method is analyzed in terms of mean-square error. Simulation results are presented to confirm the analysis.**

I. Introduction

The problem of reducing the initialization time of echo cancellers and channel equalizers in modems is extremely important for high-speed full-duplex data transmission over two-wire lines. A wide variety of approaches to this problem have been proposed and evaluated [1]–[6], although most of them were operated in the half-duplex transmission mode during the startup period. In our recent publication [7], we proposed a full-duplex fast training procedure for simultaneously estimating the echo and channel responses. Its novelty was that the echo cancellers and the channel equalizers at both ends can be trained simultaneously, rather than separately. The new method reduces the tap-setting time to half of that required by the traditional half-duplex fast training schemes. However, its estimation accuracy may be degraded by the frequency offsets in both the far echo and the far signal that are caused by up- and down-frequency shifts in the analog carrier network, where the modulators and the demodulators are not exactly matched.

Manuscript received February 19, 1996; revised February 19, 1997. The associate editor coordinating the review of this paper and approving it for publication was Prof. Georgios B. Giannakis.

W. Li is with the Training Center, Beijing University of Posts and Telecommunications, Beijing, China.

X. Chen is with the Department of Telecommunication Engineering, Beijing University of Posts and Telecommunications, Beijing, China.

Y. Wang is with the Department of Computer Science, University of New Brunswick, Fredericton, N.B., Canada.

N. Miki is with the Department of Electronics and Information Engineering, Hokkaido University, Sapporo, Japan.

Publisher Item Identifier S 1053-587X(97)06443-X.

When passing through an analog carrier system, the data signal is first modulated to the frequencies suitable for transmission over the network and then demodulated to the baseband signal at the end of transmission. Since the modulator and the demodulator, which are distantly located, are not synchronized in frequency, the demodulated baseband signal may be distorted by the frequency offset (difference) between them. This frequency offset can go as high as 7 Hz and is commonly present in the far signal. The far echo is the interfering signal that loops back to the modem through the carrier system. Since it traverses the carrier system twice and the same oscillator is often used for the modulator and the demodulator colocated, it is to be expected that frequency offset will be less prevalent in the far echo than in the far signal. However, frequency offsets with values of 1 Hz or less have been observed in the far echoes on international calls [8].

In this correspondence, we will extend the previous work [7] and develop a new algorithm that can compensate for the frequency offsets in both the far echo and the far signal. In the mathematical modeling of the received samples' vector, we discover that these frequency offsets have the effects of premultiplying the near-end and far-end data matrices by the diagonal phase rotation matrices associated with the far echo and the far signal, respectively. By constructing the appropriate estimation matrix, the far-signal component embedded in the received samples' vector can be effectively canceled when estimating the echo response. After obtaining the estimate of the echo response, the echo component embedded in the received samples' vector can be canceled by its synthesized counterpart. As a result, we obtain the estimate of the far-signal vector. The channel response can thus be calculated based on this far-signal estimate by using the appropriate pseudoinverse matrix.

This paper is organized as follows. In Section II, we first discuss the mathematical modeling of the received samples' vector in the presence of the frequency offsets in both the far echo and the far signal. Then, we derive the algorithm for estimating the echo response. Subsequently, in Section III, the algorithm for estimating the channel response is developed. In Section IV, we present the analytical and simulation results. Finally, conclusions are drawn in Section V.

II. Estimation of the Echo Response

By denoting the near echo, the far echo, the remote signal, and the channel noise as $e_N(t)$, $e_F(t)$, $s(t)$, and $v(t)$, respectively, the symbol rate $(1/T)$ sampled signal received at one end can be expressed as

$$x(n) = e_N(n) + e_F(n) + s(n) + v(n) \tag{1}$$

where we assume that the transmitters and receivers at both ends have exactly the same symbol rates, and the echo path and the transmission channel have finite complex impulse responses. Let $\{a(n)\}$ and $\{b(n)\}$ denote the data sequences transmitted by end A and end B, respectively, let $g(i)$ and $h(i)$ denote the sampled impulse responses of the echo path and the channel, respectively, let N_1, N_2, and M denote the spans of the near echo, the far echo, and the channel response, respectively, let d_e and d_f denote the transmission delays of the far echo and the far signal, respectively, and let f_1 and f_2 denote the frequency offsets in the far echo and the far signal, respectively.