A RAYLEIGH DOPPLER FREQUENCY ESTIMATOR DERIVED FROM MAXIMUM LIKELIHOOD THEORY*

Henrik HANSEN[1]   *Sofiène AFFES*[2]   *Paul MERMELSTEIN*[2]

[1]Dept. of Telecommunication, Tech. Univ. of Denmark, Bldg. 371, DK-2800 Lyngby, Denmark
[2]INRS-Télécommunications, Université du Quebec, 16, Place du Commerce, Ile-des-Sours, Verdun, Qc., H3E 1H6, Canada

ABSTRACT

Reliable estimates of Rayleigh Doppler frequency are useful for the optimization of adaptive multiple access wireless receivers. The adaptation parameters of such receivers are sensitive to the amount of Doppler and automatic reconfiguration to the speed of terminal movement can optimize cell capacities in low and high speed situations. We derive a Doppler frequency estimator using the maximum likelihood method and Jakes model [1] of a Rayleigh fading channel. This estimator requires an FFT and simple post-processing only. Its performance is verified through simulations and found to yield good results.

1. INTRODUCTION

The mobile channel is characterized by multipath propagation and the received signal manifests significant variations of the channel gain. The received signal over a propagation channel may be written as (complex base band representation)

$$r(t) = \alpha(t)e^{i\omega_0 t}s(t) + n(t) \qquad (1)$$

where the variations of α are modeled as Rayleigh fading, ω_0 is a frequency offset due to instability of the receiver and transmitter oscillators, $s(t)$ is the information bearing signal, and $n(t)$ is additive noise. ω_0 will change over time as a result of drift of oscillators and transmitter/receiver relative acceleration; however, its rate of change is much smaller than the symbol rate. We will, as a result, assume that this factor is cancelled already. On the other hand, the variations of $\alpha(t)$, caused by transmitter/receiver relative motion and reflections, can be very rapid because the wavelength of transmission is often small compared to the relative speed between receiver and transmitter. The Rayleigh fading characteristics of $\alpha(t)$ is a random process determined by one parameter only, the Doppler frequency f_D [1]. The amount of Doppler present affects the performance of mobile receivers, and the parameters used for demodulation can be optimized for one Doppler frequency only. The Spatio-Temporal Array-Receiver (STAR) is an example of receiver for CDMA which has been used as a reference herein [2]. STAR tracks the channel using an LMS-type approach and the performance of STAR is largely determined by the step-size used for tracking. Fig. 1 shows the BER performance of STAR as a function of the SNR for two step-size values and two different Doppler frequencies corresponding to a speed of 180km/h (e.g., in a train) and 5km/h (e.g., pedestrian), respectively. The optimal step-size for the fast mobile is $\mu = 0.5$; whereas, the optimal step-size for the slow mobile is $\mu = 0.1$. Using the wrong step-size in the two situations causes approximate degradations of $1 - 2dB$ and $3 - 4dB$, respectively. Clearly, knowledge of the Doppler frequency can lead to significant improvements of receiver performance. We therefore propose an estimator for the Doppler frequency derived from maximum likelihood (ML) theory, which to our knowledge has not been previously addressed. The estimator is very simple to implement in a digital environment, requiring only an FFT plus some simple post-processing. By simulation we show that the estimator provides good results.

2. BACKGROUND

We consider the Rayleigh fading channel and assume that there is no frequency offset. Rayleigh fading is caused by transmitter/receiver relative motion and multipath propagation. According to Jakes [1], the Rayleigh fading channel can be described by

$$\begin{aligned}
\alpha(t) &= \alpha(t; \omega_D, \phi_1, \cdots, \phi_L) \qquad (2)\\
&= \lim_{L\to\infty} \sqrt{\frac{E_0}{L}} \sum_{l=1}^{L} e^{i\phi_l} e^{i\omega_D \cos(\theta_l)t}, \qquad \theta_l = \frac{2\pi l}{L},
\end{aligned}$$

where ϕ_l are uniform random variables, E_0/T is the power of $\alpha(t)$, and $\omega_D = 2\pi f_D = 2\pi v f_c/c$ is the Doppler frequency, with v the relative speed, c the speed of light, and f_c the carrier frequency.
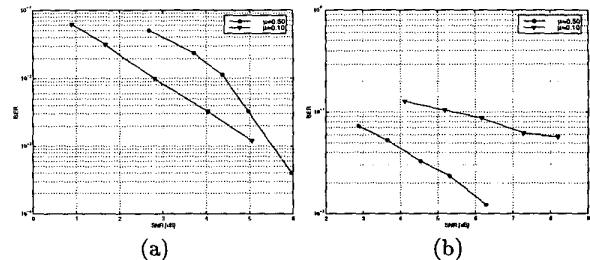

(a)   (b)

Figure 1. STAR performance in (a) slow Doppler, and (b) fast Doppler for various step-sizes.

The channel gain $\alpha(t)$ is the sum of independent waves with equal gains, independent phases, and different frequencies which are not equidistant. In practice, the number of waves, L, is limited; however, Jakes [1] shows that even small L (e.g. $L = 6$) can provide a very good approximation to the true Rayleigh fading case.

The temporal correlation of the envelope of Eq. 2 is a Bessel function of order zero, and the power spectrum is the Doppler spectrum given by

$$S(f) = \frac{1}{\sqrt{1 - (\frac{f}{f_D})^2}}, \quad |f| \leq f_D, \quad (3)$$

which is a function of the Doppler frequency only. Evidently, the knowledge of the of Doppler frequency gives all necessary information about the nature of the Rayleigh fading. Note, how this spectrum is the sum of unevenly spaced frequency components of equal strength (compare with Eq. 2).

3. THEORY OF THE PROPOSED ML ESTIMATOR

We consider the following sampled model:

$$\beta_k = \alpha_k + n_k \quad (4)$$

where $\alpha_k = \alpha(kT)$ is the channel parameter (Eq. 2), T is the sampling time, $n_k = n(kT)$ is noise (identification error) which is assumed as white with power σ_n^2, and $\beta_k = \beta(kT)$ is the noisy observation. In a receiver β_k can be identified by a pilot ($s(t)$ known in Eq. 1) or by tracking as is done in STAR [2]. With this model, the likelihood function is given by:

$$\Lambda(\Psi) = C \exp \left\{ \frac{-T}{2\sigma_l^2} \sum_{k=1}^{K} |\beta_k - \alpha_k(\Psi)|^2 \right\} \quad (5)$$

where K is the length of the observation interval, $\Psi = (\omega_D, \phi_{-N}, \cdots, \phi_N)$ is the unknown parameter vector, and C is an immaterial constant. The ML estimate is the parameter vector $\hat{\Psi}$ which optimizes the likelihood function. Joint optimization of the likelihood function with respect to the elements of Ψ is not straightforward because of the infinite number of unknown phases, ϕ_l. Even if we assume for instance $L = 6$, it requires joint optimization over an $L + 1 = 7$ dimensional space. In the following we therefore adopt the idea of [6] and average the likelihood function with respect to the phases.

The likelihood function can be written in the alternative form

$$\Lambda(\Psi) = \exp \left\{ \frac{-T}{2\sigma_n^2} \sum_{k=1}^{K} \left(|\beta_k|^2 + |\alpha_k(\Psi)|^2 - 2\Re[\beta_k \alpha_k(\Psi)'] \right) \right\} \quad (6)$$

where $\Re$ denotes real part. The first two terms of the sum are constants equal to the total powers of α_k and β_k, respectively, and the last term is the cross correlation between β_k and α_k. Only the latter is decisive for the maximization; therefore, maximization of the likelihood function is equivalent to the maximization of

$$\Lambda_1(\Psi) = \exp \left\{ \frac{T}{\sigma_n^2} \Re \left[\sum_{k=1}^{K} \beta_k \alpha_k(\Psi) \right] \right\}, \quad (7)$$

where we can substitute Eq. 2 into Eq. 7 to obtain

$$\Lambda_1(\Psi) = \exp \left\{ \frac{-T\sqrt{E_0}}{\sigma_n^2 \sqrt{L}} \Re \left[\sum_{k=1}^{K} \beta_k \sum_{l=1}^{L} e^{i\phi_l} e^{i\omega_D kT \cos(\theta_l)} \right] \right\}$$

$$= \prod_{l=1}^{L} \exp \left\{ C \sum_{k=1}^{K} \Re \left[\beta_k e^{-i\phi_l} e^{-i\omega_D kT \cos(\theta_l)} \right] \right\} \quad (8)$$

$$= \prod_{l=1}^{L} \lambda_l.$$

All phases, ϕ_j, are mutually uncorrelated and therefore

$$E_{\phi_1, \cdots, \phi_L} \{ \Lambda_1(\Psi) \} = \prod_{l=1}^{L} E_{\phi_l} \{ \lambda_l \} \quad (9)$$

where $E_v\{\cdot\}$ is the average operator with respect to the parameters v. For one of the right hand side terms we find

$$E_{\phi_l} \{ \lambda_l \} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp \left(C \sum_{k=1}^{K} \Re \{ \beta_k e^{-i\phi_l} e^{-i\omega_D kT \cos(\theta_l)} \} \right) d\phi_l. \quad (10)$$

This integral cannot be solved. Therefore we linearize the argument by using the Taylor series of the exponential to arrive at

$$E_{\phi_l} \{ \lambda_l \} =$$

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ 1 + \frac{1}{2} \left(C \sum_{k=1}^{K} \Re \{ \beta_k e^{-i\phi_l} e^{-i\omega_D kT \cos(\theta_l)} \} \right)^2 \right.$$
$$\left. + \sum_{m=2}^{\infty} \frac{1}{2m!} \left(C \sum_{k=1}^{K} \Re \{ \beta_k e^{-i\phi_l} e^{-i\omega_D kT \cos(\theta_l)} \} \right)^{2m} \right\} d\phi_l, \quad (11)$$

because odd terms vanish as a result of the averaging. In Appendix A we show that the SNR of the second term is $7.8 dB$ better than the remainder. As a result, we neglect the higher order terms and arrive at the simpler expression

$$E_{\phi_l} \{ \lambda_l \} \simeq C_1 + \frac{C^2}{4\pi} \int_{-\pi}^{\pi} \left(\sum_{k=1}^{K} \Re \{ \beta_k e^{-i\phi_l} e^{-i\omega_D kT \cos(\theta_l)} \} \right)^2 d\phi_l. \quad (12)$$

Taking the square, products between $\cos(\cdot)$ and $\sin(\cdot)$ vanish after the sum (in the limit $K \to \infty$) or as a result of the integration, leaving

$$E_{\phi_l} \{ \lambda_l \} \simeq C_1 + \frac{C^2}{4\pi} \sum_{k_1=1}^{K} \sum_{k_2=1}^{K} \int_{-\pi}^{\pi} \quad (13)$$

$$\{ \beta_{k_1}^R \beta_{k_2}^R \cos^2(\phi_l) \cos(\omega k_1 T \cos(\theta_l)) \cos(\omega k_2 T \cos(\theta_l))$$
$$+ \beta_{k_1}^R \beta_{k_2}^R \sin^2(\phi_l) \sin(\omega k_1 T \cos(\theta_l)) \sin(\omega k_2 T \cos(\theta_l))$$
$$+ \beta_{k_1}^I \beta_{k_2}^I \cos^2(\phi_l) \sin(\omega k_1 T \cos(\theta_l)) \sin(\omega k_2 T \cos(\theta_l))$$
$$+ \beta_{k_1}^I \beta_{k_2}^I \sin^2(\phi_l) \cos(\omega k_1 T \cos(\theta_l)) \cos(\omega k_2 T \cos(\theta_l)) \} d\phi_l,$$

where the integration leaves a simple scaling by $\frac{1}{2}$ for all terms and we can therefore write

$$E_{\phi_l} \{ \lambda_l \} \simeq C_1 + \frac{C^2}{8\pi} \left| \sum_{k=1}^{K} \beta_k \exp(-i\omega kT \cos(\theta_l)) \right|^2. \quad (14)$$
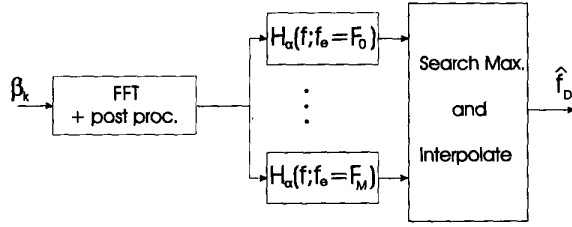
Figure 2. Implementation of Eq. 17.

The latter term is recognized as the power of β_k at the frequency $\omega_D \cos(\theta_n)$, that is $S_\beta(\omega_D \cos(\theta_n))$. Generalizing the result, we get the averaged likelihood function (Eq. 9)

$$\tilde{\Lambda}(\omega_D) = C_2 \prod_{l=1}^{L} [1 + C_3\ S_\beta(\omega_D \cos(\theta_l))], \qquad (15)$$

and as a final step we take the logarithm utilizing $\ln(1+x) \simeq x$ for x small (which is the case in the limit), to get the log-likelihood function (ignoring the immaterial constant):

$$\Lambda_L(\omega_D) = \sum_{l=1}^{L} S_\beta(\omega_D \cos(\theta_l)). \qquad (16)$$

The ML Doppler frequency estimate $\hat{f}_D$ can therefore be obtained from:

$$\hat{f}_D = \arg \max_{f_e} \left\{ \int_{-f_e}^{f_e} S_\beta(f) S_\alpha(f; f_e) df \right\} \qquad (17)$$

where $S_\alpha(f; f_e)$ is the Rayleigh spectrum with $f_D = f_e$ (see Eq. 3). From Eq. 17, the ML estimate is found by computing the spectrum of β, and sweeping f_e over the continuous range of Doppler frequencies considered to find the largest scalar product between $S_\beta(f)$ and $S_\alpha(f; f_e)$. The f_e which maximizes the scalar product is the Doppler frequency estimate, $\hat{f}_D$.

Since the continuous sweep of f_e is not possible in practice, we provide a practical implementation of Eq. 17 in the following section.

4. DIGITAL IMPLEMENTATION

Fig. 2 illustrates a simplified model of an implementation of Eq. 17. The model consists of three major modules, spectrum computation, filtering, and maximum search followed by interpolation.

First, the input is processed in blocks and the DFT is computed using an N-point FFT. From the FFT the amplitude spectrum or the power spectrum is computed. The raw spectrum is computed although a window function or other means may be used to reduce the variance. We also consider the case where the nominal value of the noise is known, which allows us to subtract the (expected) noise bias from the computed spectrum.

The computed spectrum is applied to $M + 1$ filters each representing a different Doppler frequency spanning the range 0 to $1/(2T)$ ($H_\alpha = S_\alpha$ or $H_\alpha = \sqrt{S_\alpha}$ in Fig. 2). In the simulations the number of filters is equal to $N/2 + 1$
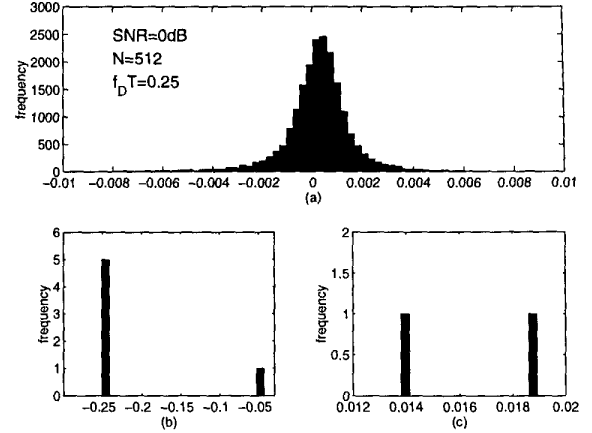

Figure 3. Histogram of frequency estimation errors.

unless otherwise is specified. Finally, the maximum filter output is determined and the estimate is refined by cubic interpolation over the maximum and its two neighbors.

We denote versions which rely on the power spectra as Power Biased (PB) for the case where no knowledge of the noise power is assumed, and Power UnBiased (PUB) when the noise power is assumed known and subtracted from the computed spectrum. Similarly we use AB and AUB for the amplitude spectrum based versions.

5. SIMULATION

Fig. 3 shows a histogram of the Doppler frequency estimation errors, $\hat{f}_D - f_D$, when PB is used with SNR=0dB and $N = 512$. In each run, the Doppler frequency f_D to be estimated is chosen randomly in $[f_D^n T - 1/(2M); f_D^n T + 1/(2M)]$ where $f_D^n T = 0.25$ is the nominal normalized frequency and $M + 1 = N/2 + 1$ is the number of filters. The main lobe of (a) resembles a normal distribution, which agrees with the fact that the estimation errors for ML estimators are normal in the limit. However, the estimates are biased. The ML theory promises unbiased estimates in the limit only. Plots (b) and (c) show outliers which are estimates which fall outside the main lobe. This phenomenon is well known for single-tone frequency estimators as well [5].

In Fig. 4 the percentage of outliers for different N is depicted for the four versions at an SNR of $-3dB$ and $f_D^n T = 0.25$. The PB and PUB both have a quite high percentage of outliers at low N ($4 - 7\%$) and are therefore useless for estimation. However, the probability of an outlier is reduced significantly with larger N and already at $N = 512$ the probability is less than 0.2%. The AUB model performs better than the PUB and PB and offers an improvement ranging from $2 - 10dB$ over the versions based on the power spectra. In contrast, the AB is useless for all N at this level of interference! Therefore, the amplitude spectra approach is only lucrative provided that the noise power is known. The AB model will not be considered in what follows. Fig. 5 shows the Mean Square Error (MSE) of the estimation errors (outliers removed). Again, the PB and PUB have similar performance. The AUB method is slightly superior in this case as well. All methods have
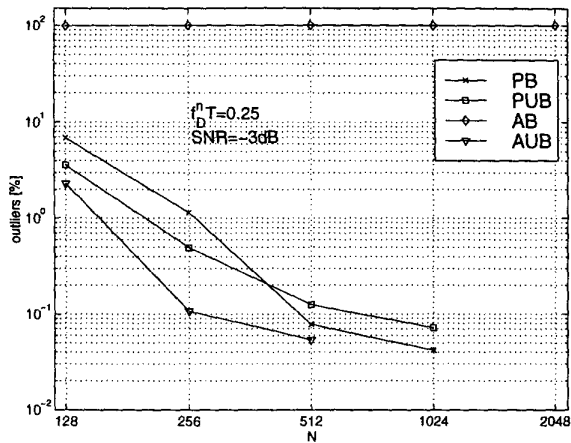
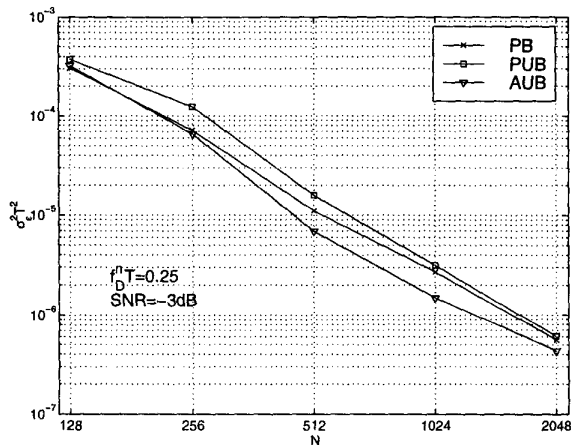384

Figure 4. Prob. of an outlier as a function of N.

Figure 5. MSE performance as a function of N.

Figure 6. MSE performance as a function of SNR.

Figure 7. MSE performance as a function of M.

in common that the improvement is around $6dB$ when N is doubled[1] and therefore increasing N provides significant improvement.

In the following we concentrate on the MSE performance, and all curves supplied have an outlier probability less than 1%.

Fig. 6 shows the MSE performance with $f_D^n T = 0.25$ and $N = 512$ as a function of the SNR. The performance of the PB and PUB versions is similar: at low SNR the MSE reduces significantly with increasing SNR, but at good SNR the curves flatten out. This is also the case with the AUB model; however, this version provides a significantly better performance at good SNR where it flattens out around an MSE of 10^{-6} - approximately $4dB$ better than the models based on the power spectra. Unlike the estimation of a cisoid in noise [4] the MSE will approach zero as SNR

---
[1]The Cramer-Rao lower bound for the estimation of the location of a single tone in white noise [4] improves by $9dB$; but the observed tone-bandwidth is reduced by two each time N is doubled; however, in this case the bandwidth is constant which may account for the difference of $3dB$.
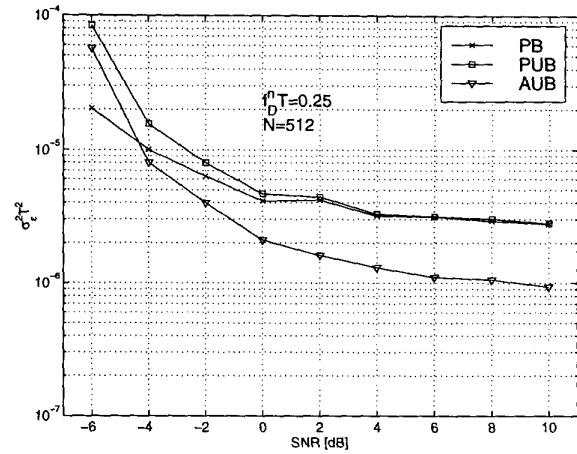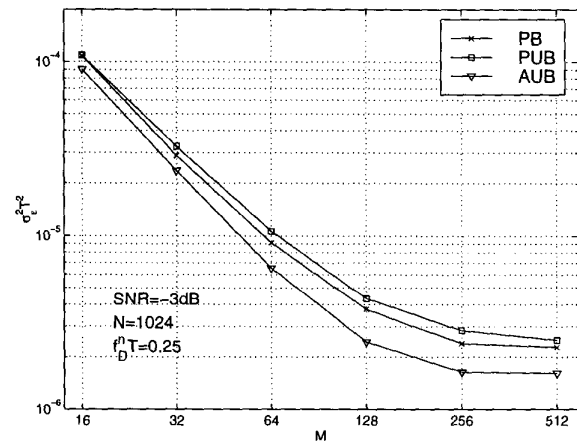
approaches infinity; here, however, the underlying process to be estimated is random itself and even in the absence of noise this randomness will introduce uncertainty, which explains the noise floor inherent in the figure.

Clearly, the AUB version provides better performance than does the PB and PUB. However, it requires the knowledge of the noise power. If this value is not known accurately, one can use the PB scheme, which provides almost as good results. As an example (Fig. 6), if $1/T = 10kHz$ we can use $N = 512$ and obtain a standard deviation of only $20Hz$ with SNR= $0dB$ using the PB scheme, which is acceptable in most cases.

Finally, Fig. 7 shows the MSE performance as a function of the number of filters (M) used for the estimation with SNR= $-3dB$, $f_D^n T = 0.25$, and $N = 1024$. As expected, the performance is better with $M = N/2 = 512$. When M becomes less than $N/2$ the performance suffers. However, reducing M by four exacts a penalty of about $2dB$ only, and hereafter the penalty increases by about $5dB$ each time M is reduced by two. If complexity is of concern, one can reduce M significantly and still obtain useful estimates.

6. CONCLUSION

We have proposed an estimator for the Doppler frequency based on the ML principle. Using the model of Jakes [1] to describe the Rayleigh fading, the frequency estimator is derived from the ML-theory using approximations which are very good at high SNR. Based on the derived ML estimator, a model suitable for digital implementation is presented and verified through simulations. The simulations reveal that implementation of the proposed ML-estimator can be significantly simplified with slight degradation in performance. The estimator can be used in mobile receivers which can take advantage of the available Doppler estimate to optimize their performance. It can be also used to improve the frequency offset estimate relative to current techniques [3].

Appendix A - Letting R_{2m} denote term number $2m$ in Eq. 11, we can express term $2m + 2$ as

$$R_{2m+2} = \frac{(s + n)^2}{(2m + 2)(2m + 1)} R_{2m} \qquad (18)$$

$$= \frac{(s^2 + n^2 + 2sn)}{(2m + 2)(2m + 1)} R_{2m}$$

where s is the signal component and n is the noise component of the argument of the exponential. Without loss of generality, we assume that the signal component is $s = 1$ and the power of the noise, n, is σ^2. It can be shown that R_{2m+2} and R_{2m} are uncorrelated and therefore the power of the signal component , P_{2m+2}^s, and the power of the noise component, P_{2m+2}^n becomes

$$P_{2m+2}^s = \frac{1}{((2m + 2)(2m + 1))^2} P_{2m}^s \qquad (19)$$

and

$$P_{2m+2}^n = \frac{7\sigma^2}{((2m + 2)(2m + 1))^2} P_{2m}^n \qquad (20)$$

respectively. Comparing (19) and (20) it becomes evident that the SNR of R_{2m+2} is 7 times lower than R_{2m}. Summing over all terms $4, 6, 8 \cdots$ the total SNR is $10 log_{10}(6) = 7.8dB$ below the SNR of term 2.

REFERENCES

[1] W.C. Jakes, Ed., *Microwave Mobile Communications*, John Wiley & Sons, 1974.

[2] S. Affes and P. Mermelstein, "A new receiver structure for asynchronous CDMA: STAR - the spatio-temporal array receiver", *IEEE J. on Selected Areas in Comm.*, Vol. 16, No. 8, pp. 1411-1422, October 1998.

[3] W-Y. Kuo, M. P. Fitz, "Frequency offset compensation of pilot symbol assisted modulation in frequency flat fading" *IEEE Trans. on Communications*, Vol. 45, No. 11, pp. 1412-1416, Nov. 1997.

[4] D. C. Rife, R. R. Boorstyn, "Single-tone parameter estimation from discrete-time observations", *IEEE Trans. on Information Theory*, Vol. IT-20, No. 5, pp. 591-598, Sep. 1974.

[5] S. Oestreich, U. Appel, "Robust frequency estimation in the presence of outliers and noise", *Fifth ASSP Workshop on Spectrum Estimation and Modeling*, pp. 40-44, 1990.

[6] F. M. Gardner, "Frequency detectors for digital demodulators via maximum-likelihood derivation", ESTEC Contract No. 8022/88/NL/DG, European Space Agency, Jun. 1990.