A Lower Bound on the Probability of a Finite Union of Events with Applications*

Hongyan Kuai, *Fady Alajaji* and *Glen Takahara*

Mathematics and Engineering
Department of Mathematics and Statistics
Queen's University
Kingston, Ontario, Canada K7L 3N6

Abstract

A new lower bound on the probability of $\mathbf{P(A_1 \cup \cdots \cup A_N)}$ is established in terms of only the individual event probabilities $\mathbf{P(A_i)}$'s and the pairwise event probabilities $\mathbf{P(A_i \cap A_j)}$'s. This bound is shown to be always at least as good as two similar lower bounds, one by de Caen (1997) and the other by Dawson and Sankoff (1967). Numerical examples for the computation of this inequality are also provided. Finally, the application of this result to the symbol error probability of an uncoded communication system used in conjunction with M-ary Phase-Shift Keying (M-PSK) modulation and maximum a posteriori (MAP) decoding is examined.

1 Introduction

We establish a new lower bound on the probability of the union of a finite family of events in terms of only the individual and pairwise event probabilities. We also demonstrate that this inequality is always at least as good as a recent bound by de Caen [4] that uses the same information. We illustrate the bound by means of several numerical examples. Finally, we examine the application of this bound to the probability of symbol error of non-uniform coherent M-PSK signaling in the presence of additive white Gaussian noise.

2 Main Results

Consider a finite family of events $A_1, A_2, \ldots, A_N$ in a finite[1] probability space (Ω, P), where N is a fixed positive integer. For each $x \in \Omega$, let $p(x) \triangleq P(\{x\})$, and let the degree of x – denoted by $deg(x)$ – be the number of A_i's that contain x. Define

$$B_i(k) \triangleq \{x \in A_i : \deg(x) = k\}$$

and

$$a_i(k) \triangleq P(B_i(k)),$$

where $i = 1, 2, \ldots, N$ and $k = 1, 2, \ldots, N$. We obtain the following lemma.

Lemma 1

$$P\left(\bigcup_{i=1}^{N} A_i\right) = \sum_{i=1}^{N}\sum_{k=1}^{N}\frac{a_i(k)}{k}.$$

Proof: Cf [6].

This brings us to our main result.

Theorem 1

$$P\left(\bigcup_{i=1}^{N} A_i\right) \geq \sum_{i=1}^{N}\left(\frac{\theta_i P(A_i)^2}{\sum_{j=1}^{N} P(A_i \cap A_j) + (1-\theta_i)P(A_i)} + \frac{(1-\theta_i)P(A_i)^2}{\sum_{j=1}^{N} P(A_i \cap A_j) - \theta_i P(A_i)}\right), \quad (1)$$

where

$$\theta_i \triangleq \frac{\beta_i}{\alpha_i} - \left\lfloor \frac{\beta_i}{\alpha_i} \right\rfloor,$$

$$\alpha_i \triangleq \sum_{k=1}^{N} a_i(k) = P(A_i),$$

and

$$\beta_i \triangleq \sum_{k=1}^{N}(k-1)a_i(k) = \sum_{j:j\neq i} P(A_i \cap A_j).$$

Proof: Cf [6].

*This work was supported in part by NSERC and TRIO.

[1] For a general probability space, the problem can be directly reduced to the finite case since there are only finitely many Boolean atoms specified by the A_i's [4].

3 Comparison with Other Bounds

In a recent work [4], de Caen also presented a lower bound on $P(\cup_{i=1}^N A_i)$ in terms of the $P(A_i)$'s and the $P(A_i \cap A_j)$'s.

Lemma 2 (de Caen [4]) Let $A_1, A_2, \ldots, A_N$ be any finite family of events in a probability space (Ω, P). Then

$$P\left(\bigcup_{i=1}^N A_i\right) \geq \sum_{i=1}^N \frac{P(A_i)^2}{\sum_{j=1}^N P(A_i \cap A_j)}. \qquad (2)$$

We demonstrate that our new bound is *always* at least as good as de Caen's bound.

Lemma 3 Let $A_1, A_2, \ldots, A_N$ be any finite family of events in a probability space (Ω, P). Then

$$\sum_{i=1}^N \left(\frac{\theta_i P(A_i)^2}{\sum_{j=1}^N P(A_i \cap A_j) + (1-\theta_i)P(A_i)} + \frac{(1-\theta_i)P(A_i)^2}{\sum_{j=1}^N P(A_i \cap A_j) - \theta_i P(A_i)} \right) \geq \sum_{i=1}^N \frac{P(A_i)^2}{\sum_{j=1}^N P(A_i \cap A_j)},$$

where

$$\theta_i \triangleq \frac{\beta_i}{\alpha_i} - \left\lfloor \frac{\beta_i}{\alpha_i} \right\rfloor.$$

Proof: Cf [6].

We next prove that our bound is also *always* at least as good as the Dawson-Sankoff bound [3].

Lemma 4 (Dawson-Sankoff [3])
Let $A_1, A_2, \ldots, A_N$ be any finite family of events in a probability space (Ω, P). Then

$$P\left(\bigcup_{i=1}^N A_i\right) \geq \frac{\theta S_1^2}{(2-\theta)S_1 + 2S_2} + \frac{(1-\theta)S_1^2}{(1-\theta)S_1 + 2S_2}, \qquad (3)$$

where

$$S_1 \triangleq \sum_{i=1}^N P(A_i),$$

$$S_2 \triangleq \sum_{i=1}^N \sum_{j=1}^{i-1} P(A_i \cap A_j),$$

and

$$\theta \triangleq \frac{2S_2}{S_1} - \left\lfloor \frac{2S_2}{S_1} \right\rfloor.$$

Lemma 5 Let $A_1, A_2, \ldots, A_N$ be any finite family of events in a probability space (Ω, P). Then (1) is always sharper than (3); i.e.,

$$\sum_{i=1}^N \left(\frac{\theta_i P(A_i)^2}{\sum_{j=1}^N P(A_i \cap A_j) + (1-\theta_i)P(A_i)} + \frac{(1-\theta_i)P(A_i)^2}{\sum_{j=1}^N P(A_i \cap A_j) - \theta_i P(A_i)} \right) \geq \frac{\theta S_1^2}{(2-\theta)S_1 + 2S_2} + \frac{(1-\theta)S_1^2}{(1-\theta)S_1 + 2S_2}.$$

Proof: Cf [6].

4 Numerical Examples

Example 1 We first give an example in which our proposed bound is tight. Let $3|n$ (n is a multiple of 3) and

$$A_i = \begin{cases} \{\frac{3i-1}{2}, \frac{3i+1}{2}\}, & \text{if } i \text{ is odd,} \\ \{\frac{3i}{2} - 1, \frac{3i}{2}\}, & \text{if } i \text{ is even,} \end{cases}$$

where $1 \leq i \leq \frac{2n}{3}$. Then $A_i \cap A_j \neq \phi$ if and only if $\lceil \frac{i}{2} \rceil = \lceil \frac{j}{2} \rceil$. If the points are uniformly distributed with probability $\frac{1}{n}$, then

$$P(A_i) = \frac{2}{n},$$

$$\sum_{j:j\neq i} P(A_i \cap A_j) = \sum_{j\neq i:\lceil \frac{i}{2} \rceil = \lceil \frac{j}{2} \rceil} P(A_i \cap A_j) = \frac{1}{n},$$

and

$$\theta_i = \frac{1}{2}.$$

Clearly

$$P\left(\bigcup_{i=1}^{\frac{2n}{3}} A_i\right) = 1.$$

(1) gives

$$\sum_{i=1}^{\frac{2n}{3}} \left(\frac{\frac{1}{2}(\frac{2}{n})^2}{\frac{3}{n} + \frac{1}{2}\frac{2}{n}} + \frac{\frac{1}{2}(\frac{2}{n})^2}{\frac{3}{n} - \frac{1}{2}\frac{2}{n}} \right) = \sum_{i=1}^{\frac{2n}{3}} \frac{3}{2n} = 1.$$

However (2) gives

$$\sum_{i=1}^{\frac{2n}{3}} \frac{(\frac{2}{n})^2}{\frac{3}{n}} = \sum_{i=1}^{\frac{2n}{3}} \frac{4}{3n} = \frac{8}{9}.$$

Thus, in this case, (1) is stronger than (2).

Example 2 We next consider several systems and compare our bound to the de Caen and Dawson-Sankoff bounds. The different systems are described in Tables 1-4. The lower bounds for each system are computed in Table 5. It can be clearly observed from Table 5 that the new bound ((1)) is sharper than the de Caen ((2)) bound and the Dawson-Sankoff bound ((3)).

Observation: de Caen's bound is tight (i.e., (2) is an equality) if and only if the degrees $deg(x)$ are constant on each A_i [4]; this condition includes the case where all the events are disjoint. Since (1) is stronger than (2), we conclude that the above condition is only a sufficient (but *not* necessary, cf. Example 1) condition for the tightness of (1).

5 Applications to Communication Systems

In [8], Séguin employed de Caen's inequality (given by (2)) to derive a lower bound on the probability of error for M-ary signals derived from a binary linear code. He assumed an additive white Gaussian noise (AWGN) channel with a maximum likelihood (ML) decoding criterion. He showed that de Caen's bound converges to the union upper bound as the signal-to-noise ratio (SNR) increases to infinity, and that it provides fairly good results at low SNR.

We similarly apply the new lower bound as well as the De Caen and the Dawson-Sankoff lower bounds to estimate the symbol error probability of a non-uniform M-PSK modulation system. More specifically, the problem formulation is as follows. We consider a *non-uniform*[2] independent and identically distributed (i.i.d.) binary source $\{X_i\}$, with $P\{X = 0\} = p$, that is transmitted via M-PSK modulation (with Gray mapping) over an AWGN channel [7]. The source stream is grouped in blocks of $\log_2 M$ bits which are each subsequently mapped to an M-PSK signal for transmission over the channel. At the receiver, optimal maximum a posteriori decoding (MAP) is performed in estimating the transmitted M-ary signal.

The computation of all three lower bounds to the probability of symbol error P_E for $M = 8$ and $p = 0.5$, 0.7 and 0.9 are displayed in terms of the SNR E_b/N_0, where E_b is the energy per information bit, in Figures 1, 2 and 3 respectively. Note that when $p = 0.5$, then MAP decoding reduces to ML decoding. The chosen values of E_b/N_0 correspond to a very noisy channel environment (e.g. $E_b/N_0 \leq 6$ dB). In Figures 1-3, two additional upper bounds are also calculated: the union upper bound and an upper bound by Kounias [5] which uses the individual and pairwise error probabilities. More specifically, given events $A_1, A_2, \ldots, A_N$, Kounias' upper bound is

$$P\left(\bigcup_{i=1}^{N} A_i\right) \leq \sum_{i=1}^{N} P(A_i) - \max_{k=1,2,\ldots,N} \sum_{i=1, i\neq k}^{N} P(A_i \cap A_k). \quad (4)$$

We first observe from Figures 1-3 that all bounds start to converge at $E_b/N_0 = 6$ dB, and that all three lower bounds are fairly good over the entire considered range of SNR. The best improvement of (1) over bounds (2) and (3) occurs when the channel is very noisy and $p = 0.5$ (i.e., when the source is totally uniform). Furthermore, we remark that when the source is very redundant ($p = 0.9$), our lower bound (1) and Kounias' upper bound (4) are very close for all values of SNR, thus providing a very good estimate of the exact error probability.

6 Conclusion

A new lower bound on the probability of a finite union of events was proven in terms of only the individual and pairwise event probabilities. This bound was shown to be always sharper than two similar inequalities by de Caen and by Dawson and Sankoff. The goodness of this bound was illustrated via numerical examples and in the estimation of the error probability of non-uniform M-PSK signaling over very noisy AWGN communication channels.

References

[1] F. Alajaji, N. Phamdo and T. Fuja, "Channel Codes That Exploit the Residual Redundancy in CELP-Encoded Speech," *IEEE Trans. on Speech and Audio Processing*, Vol. 4, No. 5, pp. 325-336, September 1996.

[2] P. Burlina and F. Alajaji, "An Error Resilient Scheme for Image Transmission Over Noisy Channels with Memory," *IEEE Trans. on Image Processing*, Vol. 7, No. 4, pp. 593-600, April 1998.

[3] D. A. Dawson and D. Sankoff, "An Inequality for Probabilities," *Proc. Amer. Math. Soc.*, vol. 18, pp. 504-507, 1967.

[2]The justification for the non-uniformity assumption of the source is as follows. In many practical data compression schemes such as image or speech coding, after some transformation, the transform coefficients are turned into bit streams (binary source) [1, 2]. Due to the suboptimality of the compression algorithm, the bit stream often exhibits some redundancy. This embedded residual redundancy can be characterized by modeling the bitstream as an i.i.d. non-uniform process or as a Markov process [1, 2].

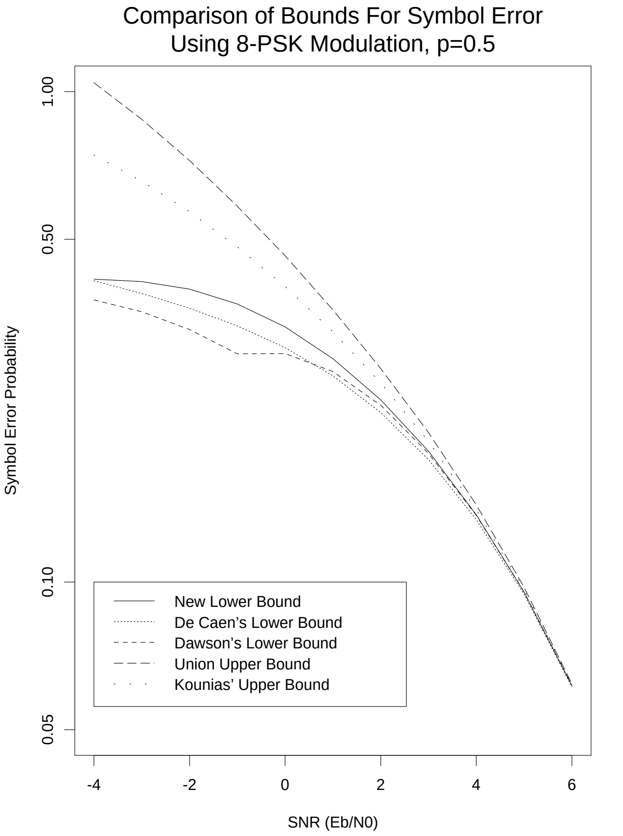

Figure 1: Bounds for P_E using 8-PSK modulation and $p = 0.5$ (ML decoding).

[4] D. de Caen, "A Lower Bound on the Probability of a Union," *Discrete Mathematics*, vol. 169, pp. 217-220, 1997.

[5] E. G. Kounias, "Bounds on the Probability of a Union, with Applications," *Annals of Mathematical Statistics,* Vol. 39, No. 6, pp. 2154-2158, 1968.

[6] H. Kuai, F. Alajaji and G. Takahara, "A Lower Bound on the Probability of a Finite Union of Events," submitted for publication to *Discrete Mathematics*, February 1998.

[7] J. G. Proakis, *Digital Communications,* McGraw-Hill, New York, 1995.

[8] G. E. Séguin, "A Lower Bound on the Error Probability for Signals in White Gaussian Noise," *IEEE Transactions on Information Theory*, submitted for publication, October 1996.

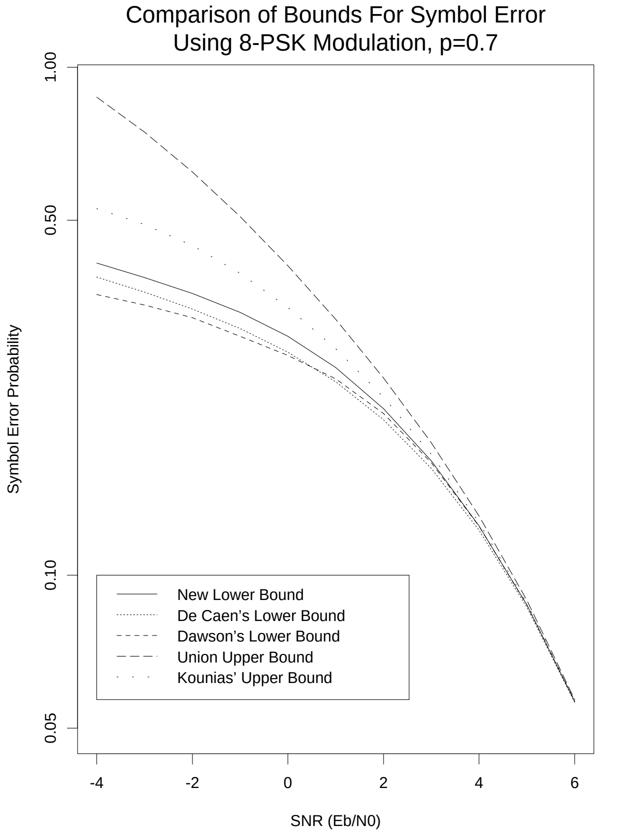

Figure 2: Bounds for P_E using 8-PSK modulation and $p = 0.7$ (MAP decoding).

x	$p(x)$	A_1	A_2	A_3	A_4	A_5	A_6
x_0	0.012	$\times$		$\times$		$\times$	
x_1	0.022		$\times$		$\times$		$\times$
x_2	0.023	$\times$		$\times$		$\times$	
x_3	0.033		$\times$				
x_4	0.034	$\times$				$\times$	$\times$
x_5	0.044		$\times$	$\times$		$\times$	
x_6	0.045		$\times$			$\times$	$\times$
x_7	0.055		$\times$	$\times$	$\times$		$\times$
x_8	0.056	$\times$		$\times$			
x_9	0.066				$\times$	$\times$	
x_{10}	0.067		$\times$		$\times$	$\times$	
x_{11}	0.077		$\times$		$\times$		
x_{12}	0.078	$\times$			$\times$		$\times$
x_{13}	0.088		$\times$				
x_{14}	0.089	$\times$		$\times$		$\times$	$\times$

Table 1: Description of System I with $N = 6$ and $|\cup_{i=1}^{N} A_i| = 15$. An $\times$ in the (i, j)'th entry indicates that outcome $x_i \in A_j$.

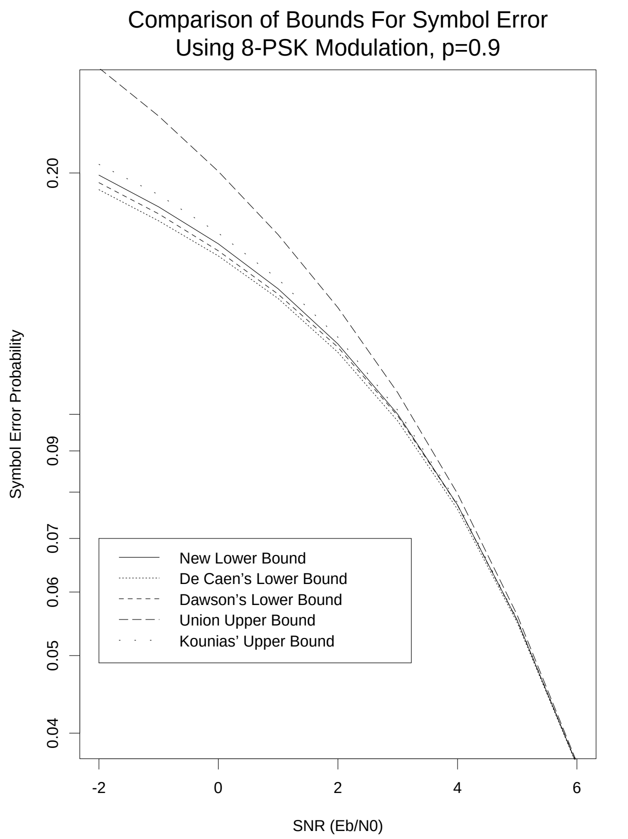

Figure 3: Bounds for P_E using 8-PSK modulation and $p = 0.9$ (MAP decoding).

x	$p(x)$	A_1	A_2	A_3	A_4	A_5	A_6
x_0	0.023	×		×		×	
x_1	0.034		×		×		
x_2	0.045	×		×		×	
x_3	0.056		×				
x_4	0.067	×				×	×
x_5	0.078		×	×		×	
x_6	0.067		×			×	×
x_7	0.056			×	×		×
x_8	0.045	×		×			
x_9	0.038				×	×	
x_{10}	0.011		×		×	×	
x_{11}	0.022		×				
x_{12}	0.033	×			×		×
x_{13}	0.044		×				
x_{14}	0.055	×		×		×	×

Table 2: Description of System II with $N = 6$ and $|\cup_{i=1}^{N} A_i| = 15$. An $\times$ in the (i, j)'th entry indicates that outcome $x_i \in A_j$.

x	$p(x)$	A_1	A_2	A_3	A_4	A_5	A_6
x_0	0.012	×		×		×	
x_1	0.022		×		×		
x_2	0.023	×		×		×	
x_3	0.033		×				
x_4	0.034	×				×	×
x_5	0.044		×	×		×	
x_6	0.045		×			×	×
x_7	0.055			×	×		×
x_8	0.056	×		×			
x_9	0.066				×	×	
x_{10}	0.067		×		×	×	
x_{11}	0.077		×				
x_{12}	0.078	×			×		×
x_{13}	0.088		×				
x_{14}	0.089	×		×		×	×

Table 3: Description of System III with $N = 6$ and $|\cup_{i=1}^{N} A_i| = 15$. An $\times$ in the (i, j)'th entry indicates that outcome $x_i \in A_j$.

x	$p(x)$	A_1	A_2	A_3	A_4	A_5	A_6	A_7
x_0	0.0329			×				
x_1	0.1076	×	×	×				×
x_2	0.0599					×		
x_3	0.1108			×		×		
x_4	0.0420		×					
x_5	0.0055		×	×				×
x_6	0.0508					×	×	×
x_7	0.1142	×				×		
x_8	0.0480						×	×
x_9	0.0235						×	×
x_{10}	0.0676	×	×					×
x_{11}	0.0295		×		×			
x_{12}	0.0441	×		×			×	
x_{13}	0.1265	×			×		×	
x_{14}	0.1058				×	×		×

Table 4: Description of System IV with $N = 7$ and $|\cup_{i=1}^{N} A_i| = 15$. An $\times$ in the (i, j)'th entry indicates that outcome $x_i \in A_j$.

Syst.	$P(\cup_i A_i)$	(2)	(3)	(1)
I	0.7890	0.7087	0.7007	0.7247
II	0.6740	0.6154	0.6150	0.6227
III	0.7890	0.7048	0.6933	0.7222
IV	0.9689	0.8759	0.8881	0.8911

Table 5: Bounds for the different systems.