N92-24129

Sloppy-Slotted ALOHA

Stewart N. Crozier

Department of Communications, Communications Research Centre, DSAT,
PO Box 11490, Station H, 3701 Carling Ave., Ottawa, Canada,
K2H 8S2, Ph: (613)-998-9262, Fax: 990-0316

ABSTRACT

Random access signaling, which allows slotted packets to spill over into adjacent slots, is investigated. It is shown that sloppy-slotted ALOHA can always provide higher throughput than conventional slotted ALOHA. The degree of improvement depends on the timing error distribution. Throughput performance is presented for Gaussian timing error distributions, modified to include timing error corrections. A general channel capacity lower bound, independent of the specific timing error distribution, is also presented.

1.0 INTRODUCTION

The random access channels for the North American MSAT system are likely to involve some form of slotted ALOHA signaling. A potential problem is the large guard times which may be required to ensure that packets stay correctly slotted. A variation in distance of 6000 km from mobile terminal to satellite, results in a two way propagation delay variation on the order of 40 msec. As an example, with 192 transmission bits per packet, a guard time of 40 msec corresponds to half a packet length for 2400 bps, and a full packet length for 4800 bps. The throughput reduction resulting directly from the use of non-zero guard times is well understood. If the guard time is narrower than the width of the timing error distribution, then packets transmitted in adjacent slots will occasionally collide. The throughput reduction, caused by adjacent packet collisions, is usually assumed to be small, or forced to be negligible by choosing a sufficiently large guard time. It is shown that this is not the best strategy, and that the optimum guard time is often much narrower than the width of the timing error distribution. Random access signaling, which allows slotted packets to spill over into adjacent slots, is denoted as Sloppy-Slotted ALOHA.

The throughput and channel capacity performance of classical unslotted ALOHA [1], and classical slotted ALOHA, including non-zero guard times [2, 3, 4], is reviewed. The corresponding performance measures for sloppy-slotted ALOHA are derived. Performance results are presented for Gaussian timing error distributions, modified to include a fraction of traffic with corrected timing.

A convenient and general channel capacity lower bound, independent of the specific timing error distribution, is also presented. This lower bound is particularly useful for designing signaling systems where most users are expected to have accurate timing, but a few users could have very large timing errors, and the type and width of the timing error distribution is unknown.

2.0 CLASSICAL ALOHA PERFORMANCE

2.1 Unslotted ALOHA

The throughput performance of classical unslotted ALOHA is given by the well known formula [3, 4]

$$S = G\,e^{-2G} \tag{1}$$

where S is the normalized channel throughput in packets per packet length, and G is the normalized channel traffic, or offered traffic, also in packets per packet length. The channel capacity, C, is defined as the maximum channel throughput achievable, and is found by differentiating (1) with respect to G and equating to zero. The result is

$$C = \max[S] = \frac{1}{2e} \simeq 18.4\,\% \tag{2}$$

and occurs for $G = 0.5$. Unslotted ALOHA does not require guard times since there are no slots to guard.

2.2 Slotted ALOHA

The throughput performance for classical slotted ALOHA, which assumes that the required guard time is negligible, is given by [3, 4]

$$S = G e^{-G} \tag{3}$$

where S and G are as defined above. The channel capacity for this case is

$$C = \max[S] = \frac{1}{e} \simeq 36.8\,\% \tag{4}$$

and occurs for $G = 1.0$. The channel capacity is twice that of unslotted ALOHA. The above result holds only if it is assumed that packets transmitted in adjacent slots never collide, and that the necessary guard time is negligible. This is not the case in practice, and non-zero guard times will be required. This is especially true for the MSAT system, which will involve time-varying propagaton delays with large delay differences.

2.3 Slotted ALOHA with Non-Zero Guard Times

The analysis of classical slotted ALOHA with non-zero guard times is identical to that with zero guard times, provided the traffic statistics are presented in terms of packets per slot, instead of packets per packet length. The result is

$$S' = G' e^{-G'} \tag{5}$$

where S' is the channel throughput in packets per slot, and G' is the offered traffic, also in packets per slot. When the packet and slot lengths are the same, corresponding to zero guard time, S' and G' are equal to the normalized traffic parameters S and G, respectively, and equations (5) and (3) are equivalent. If the guard time is not zero, then the slot length is given by

$$\tau_s = \tau_p + \tau_g = (1 + g)\,\tau_p \tag{6}$$

where τ_p is the packet length in units of time, τ_g is the guard time, and

$$g = \frac{\tau_g}{\tau_p} \tag{7}$$

is the normalized guard time, measured in packet lengths. The normalized traffic parameters, S and G, are given by

$$S = S' \frac{\tau_p}{\tau_s} = \frac{S'}{1 + g} \tag{8}$$

$$G = G' \frac{\tau_p}{\tau_s} = \frac{G'}{1 + g} \tag{9}$$

Combining (5), (8), and (9), gives the result

$$S = G e^{-(1+g)G} \tag{10}$$

The channel capacity is given by

$$C = \max[S] = \frac{1}{e(1+g)} \tag{11}$$

and occurs for $G = 1/(1+g)$. Figure 1 shows the channel capacity as a function of the normalized guard time, g. As the guard time approaches one full packet length, the capacity degrades to that of unslotted ALOHA.

3.0 SLOPPY-SLOTTED ALOHA

3.1 Throughput and Capacity

The previous conventional slotted results are based on the assumption that packets always fall within their intended slots. It was shown that the performance of slotted ALOHA is poor, approaching that of unslotted ALOHA, if the required guard time is on the order of a packet length. Reducing the guard time results in the following two effects: (a) It increases the potential channel capacity by increasing the number of slots available per unit time, and (b) It introduces the possibility of adjacent packet collisions, which in turn will reduce the channel capacity. Finding the optimum guard time obviously involves a trade-off between these two effects. The approximate throughput performance, with non-zero guard times, and with the possibility

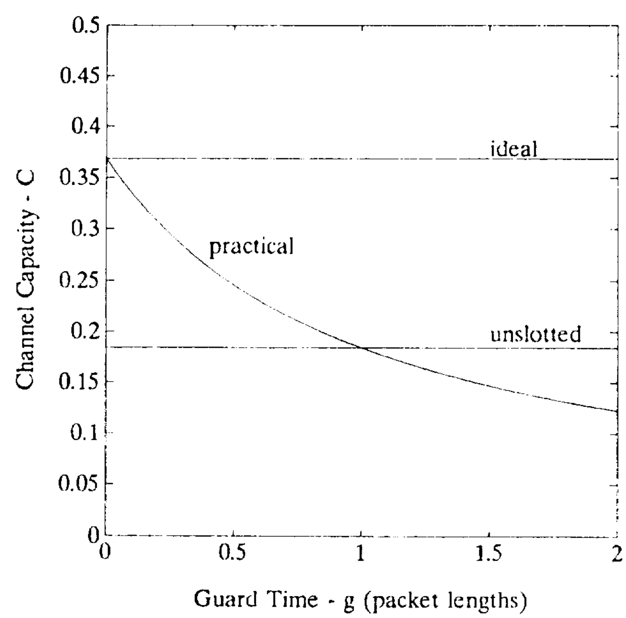

Fig. 1: Channel capacity of conventional slotted ALOHA with non-zero guard times.

of adjacent packet collisions, is derived in Appendix A. The result is

$$S \simeq G\, e^{-(1+g)(1+2p)G} \tag{12}$$

where, S, G, and g are as defined previously, and p is the probability that 2 packets transmitted in adjacent slots collide. The channel capacity evaluates to

$$C \simeq \frac{1}{e(1+g)(1+2p)} \tag{13}$$

and occurs for $G = 1/[(1+g)(1+2p)]$. Comparing (2) and (13), it is seen that the capacity of sloppy-slotted ALOHA is higher than the capacity of unslotted ALOHA only if $(1+g)(1+2p) < 2$. Ideally one would like to keep both g and p as close to zero as possible. One must be traded off against the other, however, since p is a function of g and the timing error distribution.

3.2 Gaussian Timing Error Distribution

Consider a Gaussian timing error distribution with a standard deviation of d packet lengths. The derivation in Appendix A is accurate for $d \leq 0.25$. With $d = 0.25$, a timing error of half a packet length or more, in either direction, will occur with a probability of 4.5%. The channel capacity is evaluated in Appendix B, and is given by

$$C \simeq \frac{1}{e(1+g)\left[1 + 2Q\left[\frac{g}{\sqrt{2}\,d}\right]\right]} \tag{14}$$

where Q[x] is the area under the tail of the normal distribution from x to infinity. Figure 2 shows the channel capacity versus guard time, with the standard deviation of the timing error distribution as a parameter. Typically, one might choose g to be many standard deviations, to keep the number of miss-slotted packets small. For example, one would expect 4.5% of all packets to be miss-slotted with $g = 4d$. The optimum guard time is defined as the guard time which maximizes the channel capacity. For very wide timing error distributions, the optimum guard time is seen to be closer to $g = 2d$, which corresponds to over 30% of all packets being miss-slotted. The optimum guard times, and corresponding optimum channel capacities, are presented with the results in the next section.

3.3 Gaussian Distribution with Corrections

A Gaussian timing error distribution, modified to include timing corrections, is now considered. Fraction q of all transmitted packets are assumed to have uncorrected, Gaussianly distributed timing errors, with a standard deviation of d packet lengths. The other transmitted packets, fraction (1-q), are assumed to have perfect timing. The channel capacity is evaluated in Appendix B. The result is

$$C \simeq \frac{1}{e(1+g)\left[1+4q(1-q)\,Q\left[\frac{g}{d}\right]+2q^2\,Q\left[\frac{g}{\sqrt{2}\,d}\right]\right]} \tag{15}$$

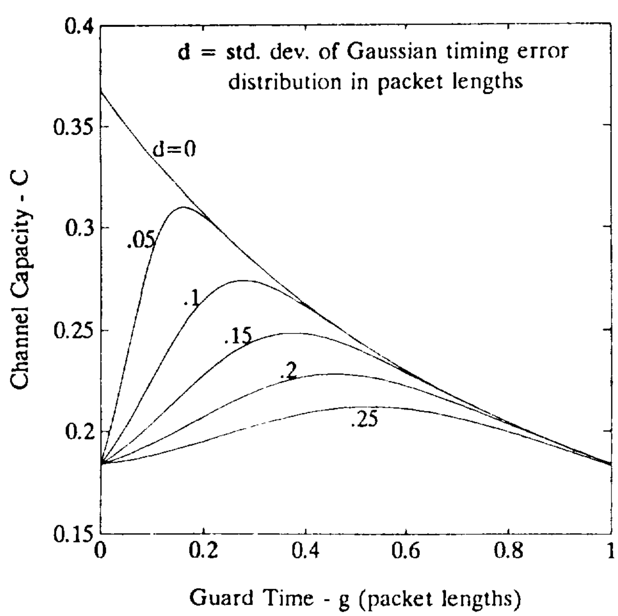

Fig. 2: Channel capacity of sloppy-slotted ALOHA versus guard time, with Gaussian timing error distribution.

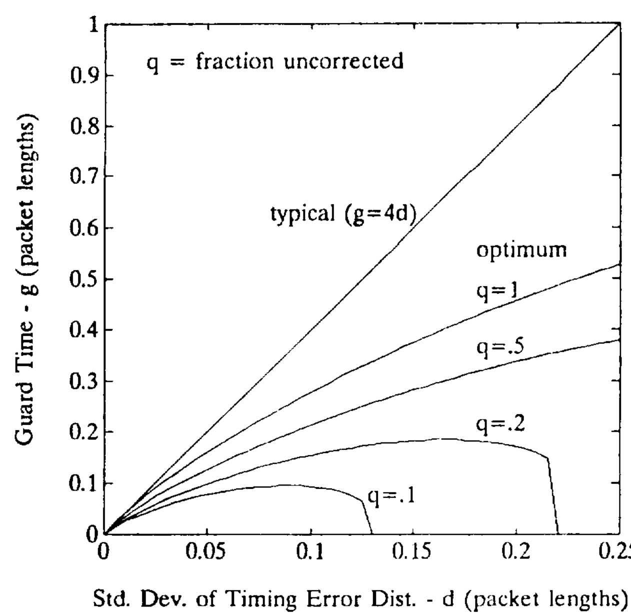

Fig. 3: Optimum guard time versus standard deviation of Gaussian timing error distribution, modified to include timing corrections.

Figure 3 shows the optimum guard times versus standard deviation, d, with fraction uncorrected, q, as a parameter. This figure was obtained using numerical methods on (15), to find the guard time which maximized the capacity for each d and q. Note that, with $q < 0.5$, as the width of the timing error distribution becomes very large, the optimum guard time jumps back to zero. Figure 4 shows the corresponding optimum channel capacities.

3.4 Channel Capacity Lower Bound

A lower bound on channel capacity is derived in Appendix C. The lower bound is general in that it is independent of the specific type and width of the timing error distribution. All that is required is the probability of being miss-slotted, m, given a specific guard time, g. The result is

$$C \geq \frac{1}{e(1+g)[1 + 2m - m^2]} \tag{16}$$

This lower bound is plotted versus guard time in Figure 5, with m as a parameter. As one might expect, the lower bound predicts ideal slotted performance with g=0 and m=0, and ideal unslotted performance with g=0 and m=1. The greater the guard time the poorer the bound.

This lower bound is useful for designing systems where corrected timing is possible, but not all packets will have corrected timing, and the timing error distribution without corrections is not well known, or time-varying. For example, it may be known that, on average, the timing of 95% of all packets can be corrected to within 10% of a packet length, and that the timing for the remaining 5% is very unpredictable. It can be seen, from the lower bound plotted in Figure 5, with g=0.1 and m=0.05, that the channel capacity is at least 30%.

4.0 CONCLUSIONS

A tight approximation for the throughput and channel capacity, with sloppy-slotted ALOHA signaling, was derived. Performance results were presented for Gaussian timing error distributions, modified to include timing corrections. The results show that sloppy-slotted ALOHA can always provide higher throughput than conventional slotted ALOHA. The degree of improvement depends on the specific timing error distribution. The greatest improvement is for wide timing error distributions, with the optimum guard time often being close to zero.

A convenient and general channel capacity lower bound, independent of the specific timing error distribution, was also presented. This lower bound is particularly useful for designing signaling systems where most users are expected to have accurate timing, but a few users could have very large timing

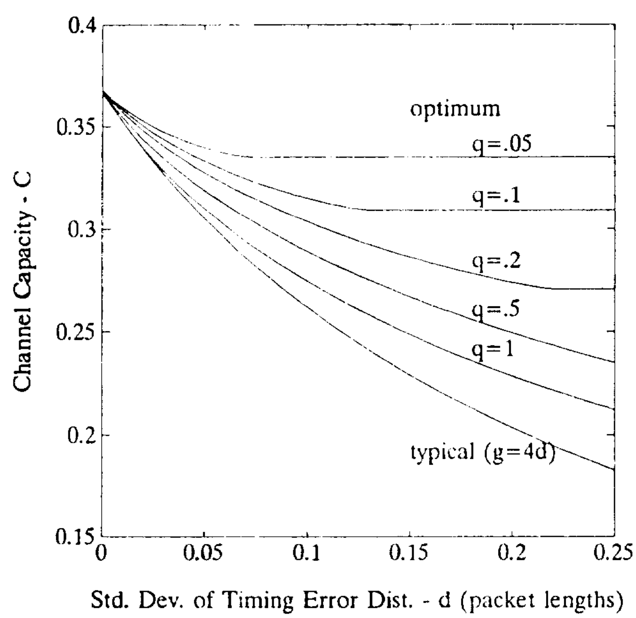

Fig. 4: Optimum channel capacity versus standard deviation of Gaussian timing error distribution, modified to include timing corrections.

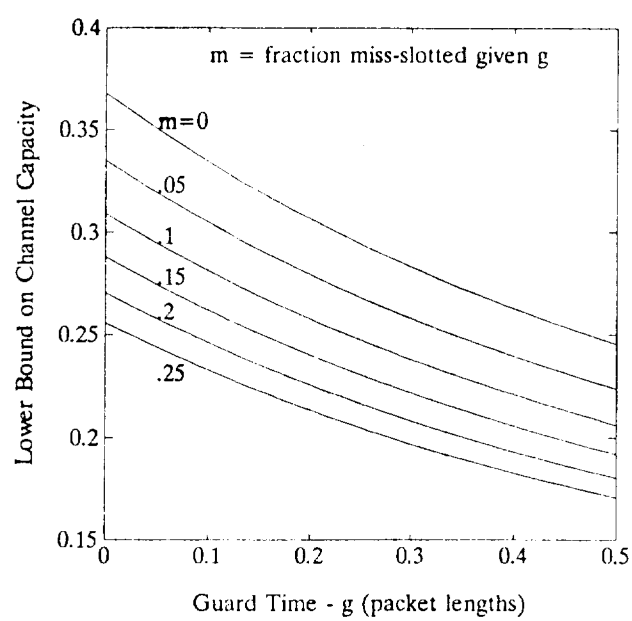

Fig. 5: Channel capacity lower bound versus guard time, with fraction of miss-slotted packets as a parameter.

errors, and the type and width of the timing error distribution is unknown.

Appendix A: Analysis of Sloppy-Slotted ALOHA

The following definitions are used:
N = number of users
M = middle (or co-slot) collision
L = left collision
R = right collision
$\bar{x}$ = no x, or not x (eg. $\bar{M}$ = no middle collision)
P(x) = probability of x
$P(x \cap y)$ = probability of x and y
$P(x|y)$ = probability of x given y

Then S'/N = average throughput in packets per slot per user = probability of a successful packet per slot per user, and G'/N = average offered traffic in packets per slot per user = probability of a transmission per slot per user. It follows that

$$\frac{S'}{N} = \frac{G'}{N} P(\bar{M} \cap \bar{L} \cap \bar{R}) \qquad (A.1)$$

Multiplying through by N, and using the fact that M is independent of L and R, gives

$$S' = G' \, P(\bar{M}) \, P(\bar{L} \cap \bar{R}) \qquad (A.2)$$

M is independent of L and R because P(M) does not depend on the timing error distribution. This is because a second co-slot transmission is always assumed to cause a middle collision. L and R are not independent. Knowing that a left collision has occurred reduces the probability of a right collision, since it is more likely that the packet of interest has shifted left than right. Expressed mathematically,

$$P(R|L) \le P(R) \qquad (A.3)$$

Note that

$$\begin{aligned} P(\bar{L} \cap \bar{R}) &= 1 - P(L) - P(R) + P(L \cap R) \\ &\ge 1 - P(L) - P(R) \\ &= [1-P(L)]\,[1-P(R)] - P(L)\,P(R) \\ &= P(\bar{L})\,P(\bar{R}) - P(L)\,P(R) \qquad (A.4) \end{aligned}$$

Also, it follows from (A.3), that

$$\begin{aligned} P(\bar{L} \cap \bar{R}) &= P(\bar{L})\,P(\bar{R}|\bar{L}) \\ &\le P(\bar{L})\,P(\bar{R}) \qquad (A.5) \end{aligned}$$

With equal traffic and timing error statistics for all users, the symmetry of the problem forces P(L) = P(R), even if the timing error distribution is not symmetric. From (A.4) and (A.5), the following lower and upper bounds on $P(\bar{L} \cap \bar{R})$ are obtained

$$P(\bar{L})^2 - P(L)^2 \le P(\bar{L} \cap \bar{R}) \le P(\bar{L})^2 \qquad (A.6)$$

These bounds differ only by $P(L)^2$. Typically, in the operating region of interest, P(L) is fairly small, so that $P(L)^2$ is a very small second order effect. Even for P(L) as high as 10%, the bounds are only 1% apart. Thus, $\bar{L}$ and $\bar{R}$ are approximately independent in the operating region of interest. Using the upper bound of (A.6) as an approximation, (A.2) simplifies to

$$S' \simeq G' \, P(\bar{M}) \, P(\bar{L})^2 \qquad (A.7)$$

For N users, $P(\bar{M})$ is given by

$$P(\bar{M}) = \left[1 - \frac{G'}{N} \right]^{N-1} \qquad (A.8)$$

Taking the limit as N approaches infinity gives

$$P(\bar{M}) = e^{-G'} \qquad (A.9)$$

For N users, $P(\bar{L})$ can be approximated by

$$\begin{aligned} P(\bar{L}) \simeq\ & P(0 \text{ left adjacent packets}) \qquad (A.10) \\ & + P(1 \text{ left adjacent packet}) \times P(\text{ no overlap}) \\ & + P(2 \text{ left adjacent packets}) \times P(\text{ no overlap})^2 \\ & + \ldots \\ =\ & \left[1 - \frac{G'}{N}\right]^N \\ & + \binom{N}{1}\left[\frac{G'}{N}\right]^1 \left[1 - \frac{G'}{N}\right]^{N-1} (1-p) \qquad (A.11) \\ & + \binom{N}{2}\left[\frac{G'}{N}\right]^2 \left[1 - \frac{G'}{N}\right]^{N-2} (1-p)^2 \\ & + \ldots \end{aligned}$$

where

$$\binom{N}{n} = \frac{N!}{(N-n)!\,n!} \qquad (A.12)$$

is the number of "N choose n" combinations, and p is the probability that 2 packets transmitted in adjacent slots overlap (collide). The approximation is only for the high order terms, and is due to the independence assumption for the probability of no overlap when 2 or more packets are transmitted in the same adjacent slot. The independence assumption provides a very tight lower bound in this case, provided p is small (e.g. less than 10%). Taking the limit as N approaches infinity gives

$$\begin{aligned} P(\bar{L}) &\simeq e^{-G'}\left[1 + G'(1-p) + \frac{1}{2!}\left[G'(1-p)\right]^2 + \ldots\right] \\ &= e^{-G'} e^{G'(1-p)} \\ &= e^{-pG'} \qquad (A.13) \end{aligned}$$

Substituting (A.9) and (A.13) into (A.7) yields (as N approaches infinity)

$$S' \simeq G' e^{-(1+2p)G'} \qquad (A.14)$$

Performance can be presented in terms of the normalized throughput, S, and offered traffic, G, measured in packets per packet length, by accounting for the non-zero guard times, as in Section 2.3. The result is equation (12).

Appendix B: Gaussian Distribution with Corrections

Let u and v represent the timing errors, measured in packet lengths, for the first and second of two adjacent packets, respectively. With probability (1-q), the timing is correct, and with probability q, the timing is Gaussianly distributed with standard deviation d. The probability that the two adjacent packets collide is given by

$$p = P(u > v+g) \qquad (B.1)$$

$$= \iint_A f_u(x)\, f_v(y)\, dx\, dy \qquad (B.2)$$

$$= \int_{L_x} (1\text{-}q) f_u(x)\, dx + \int_{L_y} (1\text{-}q) f_v(y)\, dy$$

$$+ \iint_{A\text{-}L_x\text{-}L_y} f_u(x) f_v(y)\, dx\, dy \qquad (B.3)$$

$$= 2q(1-q)Q\left[\frac{g}{d}\right] + q^2 Q\left[\frac{g}{\sqrt{2}\,d}\right] \qquad (B.4)$$

where $f_u(x)$ and $f_v(y)$ represent the probability density functions for u and v respectively, L_x and L_y are the infinite half-lines shown on the x and y axes in Figure 6, and A is the shaded half-plane. Substituting (B.4) into (13) gives (15), and letting q=1 gives (14).

Appendix C: Channel Capacity Lower Bound

Using definitions similar to those in Appendix B, the probability that the two adjacent packets collide can be upper bounded as follows:

$$p = P(u > v+g) \qquad (C.1)$$

$$= \iint_A f_u(x)\, f_v(y)\, dx\, dy \qquad (C.2)$$

$$\le \iint_{A+B+C} f_u(x)\, f_v(y)\, dx\, dy \qquad (C.3)$$

$$= \tfrac{1}{2} - \iint_D f_u(x)\, f_v(y)\, dx\, dy \qquad (C.4)$$

$$= \tfrac{1}{2} - \tfrac{1}{2}(1-m)^2 \qquad (C.5)$$

$$= m - \tfrac{1}{2}m^2 \qquad (C.6)$$

where A, B, C, and D represent the regions shown in Figure 7, and the probability of u or v being outside the box is m. Substituting (C.6) into (13) gives (16).

References

[1] N. Abramson, "The ALOHA System - Another Alternative for Computer Communications", Proceedings, Fall Joint Computer Conference, 1970.

[2] L. Roberts, "ALOHA Packet System With and Without Slots and Capture", Computer Communications Review, April, 1975.

[3] R. Rosner, *Distributed Telecommunications Networks, via Satellites and Packet Switching*, Wadsworth, Inc., 1982.

[4] W. Stallings, *Data and Computer Communications*, Macmillan Publishing Company, New York, 1985.

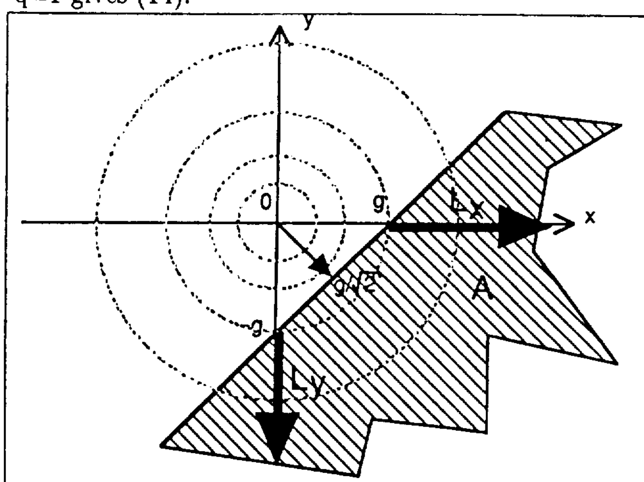

Fig. 6: Joint timing error distribution (Gaussian).

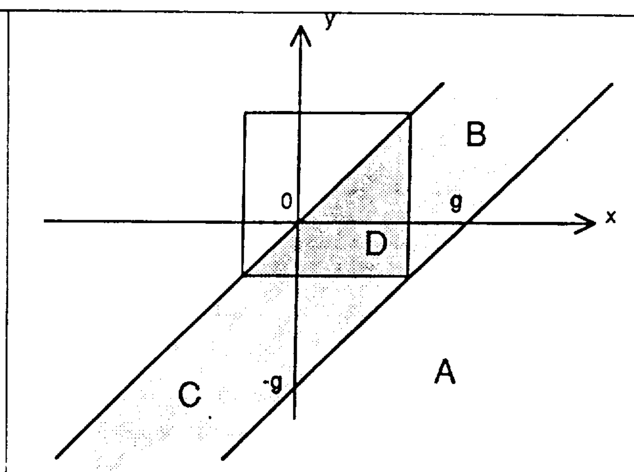

Fig. 7: Joint timing error distribution (General).