Event probability analysis of capacity and throughput in wireless local area networks using distributed medium access control based on listen before talk

T. P. Spiller[*], T. A. Wilkinson[†] and G. Vanzini[‡]

Hewlett-Packard Laboratories, Bristol,

Filton Road, Stoke Gifford, Bristol BS34 8QZ, U.K.

August 19, 1999

Abstract

An analysis of event probabilities in wireless local area networks using distributed medium access control based on listen before talk is presented. More significantly it is shown how this can be used to calculate capacity or throughput for a given offered load for these networks. This in turn can be used to optimize system parameters for whatever performance aspect is of importance. The results are particularly applicable to systems designed for operation at 5 GHz. It is shown that for such systems with a usable system bandwidth of 70 Mbps in typical deployments the capacity or throughput are the order of 100 kbps/m^2 for reasonable packet failure rates.

[*]email: ts@hplb.hpl.hp.com

[†]email: taw@hplb.hpl.hp.com

[‡]email: gvanzini@hotmail.com

1 Introduction: The origins of the problems

Recently significant amounts of spectrum have been allocated at frequencies such as 5 GHz for unlicensed operation of wireless local area networks [1, 2]. Such networks invariably use some form of listen before talk in their medium access control protocols [3, 4]. The calculation of the performance of these networks is critical to understanding what types of applications they will support. This is the aim of the analysis presented in this paper.

The evaluation of the capacity is one of the most important aspects in the design of a wireless communication system. This concept gives a measure of how efficiently the spectral resource is being utilised and it is usually derived from some notion of spectral re-use. The re-use distance is simply the distance beyond which a part of the spectrum can be re-used or the distance separating two simultaneous users of the same part of the spectrum such that they are subject to negligible mutual interference. This re-use distance or radius can be used to define a re-use area in which this part of the spectrum can be used only once and this yields a spectral re-use in Hz/m^2. If this is combined with a spectral efficiency of the transmission scheme in bps/Hz this yields a capacity in bps/m^2 or more usefully in kbps/km^2 in wide area networks and Mbps/m^2 in local area networks.

The calculation of capacity is relatively straightforward for cellular telephony networks with star topologies and centralised medium access control. In this case the frequency re-use is structured in that cell size and re-use pattern can be fixed and the medium access control will limit the number of calls in a cell. The result is that the re-use interference is the aggregate of the significant co-channel interferers surrounding a cell. This simply manifests itself as background noise causing bit errors or packet failures. This is true of any FDMA or TDMA system but the situation is more complicated with a CDMA system. In a CDMA system the re-use is not as structured and the medium access involves statistical multiplexing but the calculation of the aggregate interference is similar and its effect the same.

The calculation of capacity is much more complicated for wireless local area networks that have mesh topologies and distributed medium access control invariably involving listen before talk. In this case the frequency re-use is totally unstructured and indeed is dependent upon the detailed mechanism of the medium access control. This is because interference measurement or clear channel assessment is central to the medium access control in listen before talk. Moreover, its function is not perfect because clear channel assessment is carried out at the transmitter wishing to initiate a communication, whereas collisions clearly occur at the receiver where the interference environment may be quite different. This can result in collisions from what is commonly known as the hidden node effect. An obvious way to avoid such collisions is to increase the sensitivity of the clear channel assessment so that all potential interferers that could cause collisions at the receiver can be heard by the transmitter. However, this has the effect of inhibiting communication over a larger area, perhaps unnecessarily. So there are two distinct effects of re-use interference: unwanted collisions and unnecessary deferrals. Hence, the concept of capacity based on re-use is not as useful as this must be tied to one single effect such as co-channel interference causing bit errors or packet failures. As collisions can cause packet failures it would seem sensible to calculate re-use based on collision probability if the capacity concept is applied to these networks.

In local area network systems it is much more usual to consider system performance in terms of throughput against offered load. However, these concepts were invented for wired communication systems where the capacity of the medium is bounded and there is no re-use aspect.

This paper presents an analysis of the event probabilities in these communication systems in an attempt to link these with the more familiar concepts of capacity and throughput against offered load. In section 2 the basic event probabilities are described. In section 3 the analysis is resolved into two basic problems. In sections 4 and 5 these two problems are analysed and the resulting probabilities are presented in section 5. In section 6 the results are interpreted and links are identified between these probabilities and the familiar concepts of capacity and throughput against offered load. This is done with the specific example of HIPERLAN. In section 7 conclusions are presented.

2 Event Probabilities

Suppose that a transmitter (labelled number 2) wishes to transmit a packet to an unspecified receiver in the presence of an on-going communication between transmitter 1 and receiver 0. There are four possible outcomes with probabilities p_1, p_2, p_3 and p_4. These are respectively defined as the probabilities of:

1. Transmitter 2 sensing transmitter 1 and deferring, correctly, because if it had transmitted it would have caused a collision at receiver 0.

2. Transmitter 2 sensing transmitter 1 and deferring, incorrectly, because if it had transmitted it would not have caused a collision at receiver 0.

3. Transmitter 2 not sensing transmitter 1 and thus not deferring, incorrectly, as when it transmits it causes a collision at receiver 0. (The collision is due to the fact that receiver 0 cannot distinguish between the signals from 1 and 2.)

4. Transmitter 2 not sensing transmitter 1 and thus not deferring, but correctly, as when it transmits it does not cause a collision at receiver 0.

Of the four events listed above 1 is desirable, 4 is acceptable (because of the correct outcome) and two are undesirable, 2 and 3. Of the two undesirable events it could be argued that 2 has a less destructive impact on performance than 3. A greater probability of event 2 means a lower system capacity in Mbps/m^2 because there are less simultaneous communications in a given area. A greater probability of 3 means more collisions and associated re-transmissions. Both of these things ultimately mean reduced capacity and increased latency. However, event 2 causes the deferral of only one communication. This is simply handled in the medium access control, resulting in minimal additional latency, whereas event 3 causes the possible destruction of two communications along with the far greater additional latency associated with the necessary re-transmissions.

It can be seen from the descriptions of the event probabilities above that these are formed from all four combinations of the outcomes of two events namely the probability of transmitter 2 sensing transmitter 1 and probability of transmitter 2 causing a collision with transmitter 1 at receiver 0. The following analysis details the solution of these distinct problems and the resulting event probabilities.

3 Two problems and some notation

Before attacking the real problem, the calculation of the probabilities p_1 to p_4 (which we will call "case B"), we address a very much simpler problem, ("case A"). This generates some definitions and other useful input (such as choices of system parameters) for the main calculations.

- Case A arises when there are just two communicators. Number 1 is defined to lie somewhere inside a circle 1, which has radius R_1, distributed randomly with a constant probability per unit area. Number 2 is similarly somewhere inside a circle 2 (of radius R_2), which is concentric with circle 1. In the model we employ in this work, the communicators are assumed to operate isotropically and over a finite range a. Thus, if they are simply further apart than the distance a, they are unable to make contact. There are two possible interpretations which can be placed on this *contact*—we outline these briefly before discussing the important issue of parameter values.

  1. Communication: If both parties operate transceivers (or one operates a transmitter and the other a receiver), their separation compared to a determines whether or not they can achieve two-way (one-way) communications. In this case a is clearly equal to the maximum *communication* distance, defined as D.
  2. Sensing: As will be relevant for our full problem, case B, it could be that

both parties are transmitters—both intent on contacting some other receiving party—which have the ability to sense other active transmitters. (This ability is a desirable feature as it enables the prevention of communication collisions, by a transmitter deferring if it is able to sense a nearby transmitter already at work.) In such a case, a will be the maximum *sensing* distance. This may well be larger than D, the communication distance, because the signal to noise ratio for reliable sensing (which depends on the sensing hardware) can be smaller than that for the actual demodulation of a signal.

In both of these interpretations of case A, what is desired is the probability of successful contact, as a function of R_1, R_2 and a. It is clear that this is invariant under a scaling of all the lengths. It is therefore convenient to use dimensionless lengths; the appropriate choice here will be to divide everything by R_1.

The choices for all the parameters in this study are clearly important if our example results are to be realistic and useful. We therefore choose parameter values for the two case A scenarios with reference to real communications systems.

1. Communication: The maximum communication distance D (which equals the contact range a for this case) is set by the transmit power at 1m (which depends on antenna gains/losses) $P_t(1)$, the noise power in the system bandwidth (which depends on the transmission rate and the modulation scheme) P_n and the signal to noise ratio for reliable demodulation (which also depends on the modulation scheme) of n dB. Assuming inverse square path loss of the signal, all these quantities are related by $10\log_{10}\left(\frac{P_t(1)}{P_n D^2}\right) = n$. The signal to noise ratio for reliable demodulation obviously depends on many details of the link design and implementation. However, given that our analysis is aimed at high rate systems, it is appropriate to select an example system design such as HIPERLAN (High Performance Radio Local Area Network) [3]. This system has a transmission rate of 23.5 Mbps with GMSK modulation and BCH error-correction coding. HIPERLAN implementations typically use decision feedback equalization to enable the high transmission rate in dispersive multipath propagation channels. For such implementations the system typically requires a signal to noise ratio of 17 dB for reliable demodulation. Other systems using alternative methods of dealing with the propagation channel, such as multicarrier modulation, require similar signal to noise ratios for reliable demodulation. An appropriate choice for our work is therefore $n = 17$ dB [4].

2. Sensing: The relationship between the contact range for this case, a, and the signal to noise ratio for reliable *sensing*, m, is analogous to the communication case, so $10\log_{10}\left(\frac{P_t(1)}{P_n a^2}\right) = m$. Thus the relationship here between a and D is set simply by the difference between n and m, with $a = D\, 10^{\frac{(n-m)}{20}}$. The choice of m is somewhat more arbitrary than that of n. For efficient

operation of listen before talk protocols, it is desirable if the signal to noise ratio for reliable sensing is lower than that for reliable demodulation. This is achieved in a number of ways. Simply averaging over a period of time or number of bit periods will give a more reliable measure of signal strength for sensing. For example, detecting or correlating to a synchronization sequence at the beginning of a transmission will give an increase in signal strength equivalent to the the length of the sequence. Consequently, sensing reliability can be traded against sensing time. In HIPERLAN the sensing threshold is adaptive to allow the system to adjust to interference environments [3]. However, a reasonable choice of the signal to noise ratio for reliable sensing is 6 dB lower than that for reliable demodulation. We therefore take $n - m = 6$ dB, which gives $a = 2.00\ D$.

- The full collision problem, case B, arises if the first two communicators are both considered to be transmitters (with the ability to sense other transmitters). Transmitter 1 would like to signal to a third person 0, who is also somewhere inside circle 1. However, this receiver 0 is unable to distinguish the signal from transmitter 1 above that of transmitter 2 if the *ratio* of the distances $0 \to 2$ and $0 \to 1$ is less than a certain dimensionless quantity $b > 1$. A very simple estimate for b, given the discussions above, follows by assuming that transmitter 2 simply acts as *noise* to the communication between receiver 0 and transmitter 1. In this case, given a signal to noise ratio for demodulation of n dB, it follows that $b = 10^{\frac{n}{20}}$. From our choice of $n = 17$ dB, this gives $b = 7.08$. Clearly this simple estimate neglects the actual noise present. Nevertheless, for the parameters which interest us, the constant b approximation is not at all bad, and we shall employ it in our analysis. (The justification for this will be given at the end of the next section, after we have addressed the two communicator problem of case A.) In the three communicator problem (case B), two separate questions can be asked: (i) Whether or not transmitter 2 is aware of transmitter 1 (and might thus be able to avoid simultaneous use of the broadcast channel); (ii) Whether or not the receiver 0 can distinguish signals from 1 and 2 if they both transmit. Four possible joint answers exist; the probabilities of these four possibilities are the desired p_1 to p_4 listed in sec. 2.

4 Case A: Two communicators

The mathematical problem to which this reduces is the following:

Take two concentric circles with radii R_1 and R_2. Pick a point $\vec{x}_1$ in circle 1 and a point $\vec{x}_2$ in circle 2 (at random and with uniform weight). For a given a, what is the probability that $|\vec{x}_1 - \vec{x}_2| > a$?

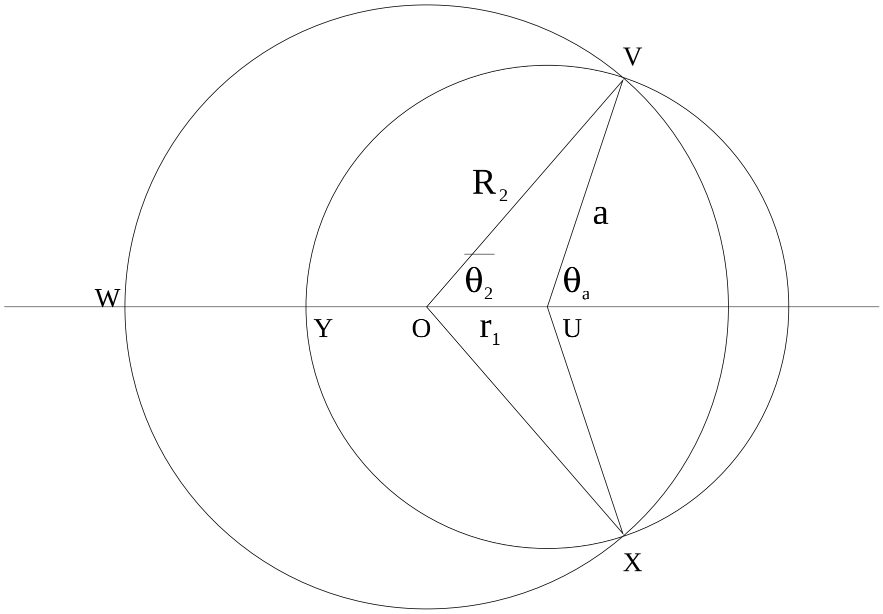

Figure 1: The geometric arrangement relevant for case 1. The variable r_1 (which ranges from 0 to 1 in units of R_1) is the length OU. The maximum contact distance a is UV. R_2, the radius of the circle constraining transmitter 2, is OV.

This is illustrated in fig. 1. For a chosen $\vec{x}_1$ of magnitude r_1, the points $\vec{x}_2$ which fit the bill lie in the moon shaped area YVWXY. This area, A, may be evaluated from $A(r_1, R_2, a) = OVWXO - UVYXU + OVUXO = R_2^2\bar{\theta}_2 - a^2\theta_a + 2OVUO$. Using the standard results [7]

$$a^2 = R_2^2 + r_1^2 - 2r_1R_2\cos\bar{\theta}_2 \tag{1}$$

$$R_2^2 = a^2 + r_1^2 - 2r_1 a\cos(\pi - \theta_a) \tag{2}$$

$$OVUO = (s(s-a)(s-r_1)(s-R_2))^{1/2}\ , \tag{3}$$

with $s = \frac{1}{2}(a + r_1 + R_2)$, the required area is

$$\begin{aligned} A(r_1, R_2, a) = {} & R_2^2 \arccos\left(\frac{a^2 - r_1^2 - R_2^2}{2r_1R_2}\right) - a^2 \arccos\left(\frac{a^2 + r_1^2 - R_2^2}{2r_1 a}\right) \\ & + \frac{1}{2}\left(2R_2^2a^2 + 2R_2^2r_1^2 + 2r_1^2a^2 - a^4 - r_1^4 - R_2^4\right)^{1/2}\ . \end{aligned} \tag{4}$$

As all lengths are dimensionless and are defined in units of R_1, the variable r_1 ranges from 0 to 1. The required probability is a function of R_2 and a and follows from averaging $A(r_1, R_2, a)$, normalized by πR_2^2, over all positions of $\vec{x}_1$ inside circle 1. Angular symmetry leaves only an integral over r_1 for the communication failure probability:

$$P(R_2, a)_{|\vec{x}_1 - \vec{x}_2| > a} = \int_0^1 dr_1\ 2r_1 \frac{A(r_1, R_2, a)}{\pi R_2^2}\ . \tag{5}$$

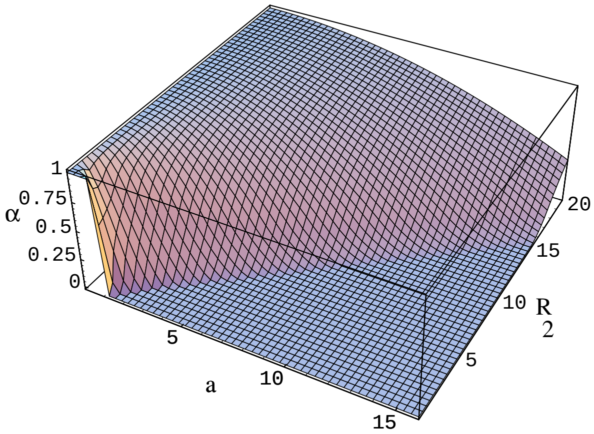

Figure 2: Plot of $\alpha(r_1, R_2, a) = A(r_1, R_2, a)/\pi R_2^2$, given by eq. (4), as a function of a and R_2, for $r_1 = 1$.

The complementary probability, the probability of successful communication between the two, follows from

$$P(R_2, a)_{|\vec{x}_1 - \vec{x}_2| \leq a} = 1 - P(R_2, a)_{|\vec{x}_1 - \vec{x}_2| > a} \; . \tag{6}$$

Fig. 2 shows a plot of $A(r_1, R_2, a)/(\pi R_2^2)$ as a function of R_2 and a, for a fixed value of $r_1 = 1.0$. Clearly, although fig. 1 shows the two circles intersecting, this is not always the case.

- If $a - r_1 > R_2$, the circle of radius R_2 lies entirely within the circle of radius a and so $A = 0$.
- If $a + r_1 < R_2$, the opposite is true and so $A/(\pi R_2^2) = (1 - (a/R_2)^2)$.

Both of these cases are covered[1] in parts of fig. 2.

[1] A very convenient trick can be employed for the evaluation of the probabilities. If $\arccos(x)$ is defined to lie in the range 0 to π for $|x| \leq 1$, the real part to satisfy $\Re \arccos(1 + x) = 0$ for $x > 0$ and $\Re \arccos(-1 - x) = \pi$ for $x > 0$, then, by taking the real part of eq. (4), $A(r_1, R_2, a)$ can be defined for *all* eventualities from the one formula. Consequently, the integrals in the probabilities (5) and (6) do not have to be broken up into pieces by hand. This does not matter too much here. However, it proves invaluable if such integrals are nested inside others, as happens in the next case B.

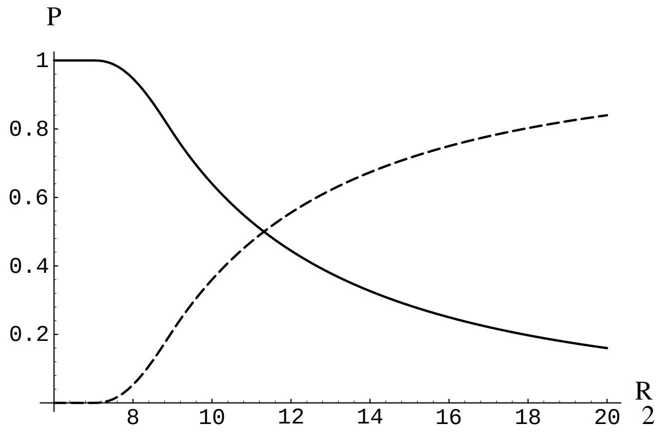

Figure 3: Plots of $P(R_2, a)_{|\vec{x}_1 - \vec{x}_2| \leq a}$ (solid) and $P(R_2, a)_{|\vec{x}_1 - \vec{x}_2| > a}$ (dashes) as functions of R_2 for $a = 8.0$. See eqs. (5) and (6).

Fig. 3 shows a plot of the two probabilities as functions of R_2 for a fixed value of $a = 8.0$. Clearly, as expected, $P(R_2, a)_{|\vec{x}_1 - \vec{x}_2| \leq a} = 1$ for $R_2 < a$ and $P(R_2, a)_{|\vec{x}_1 - \vec{x}_2| \leq a} = (a/R_2)^2$ for $R_2 > a + 1$.

Our results for this two communicator situation can be used to justify the constant discrimination ratio b approximation employed in our next (three communicator) example. For receiver 0 and transmitter 1 both randomly distributed inside circle 1, it is clear that the maximum possible value of their separation x_{01} is 2 (in units of R_1). To consider the validity of the constant b approximation, we require the probability distribution, $G(x_{01})$, over the separation x_{01}, from zero up to its maximum value. This follows simply from our analysis thus far. The total accumulated probability from zero up to a separation x_{01} is given by eq. (6) with some label changes and setting $R_2 = 1$ (as the receiver and transmitter are both distributed over a unit radius circle). The probability distribution is just the derivative of this; thus

$$G(x_{01}) = \frac{\partial}{\partial x_{01}} \left(1 - \frac{1}{\pi} \int_0^1 dr_1 \, 2r_1 A(r_1, 1, x_{01}) \right) \; . \tag{7}$$

This is shown (as the dashed plot) in fig. 4.

Next we need the actual variation of b over this range. In the introduction we estimated this assuming that transmitter 2 played the role of noise for communication between transmitter 1 and receiver 0. However, this neglected the actual noise power

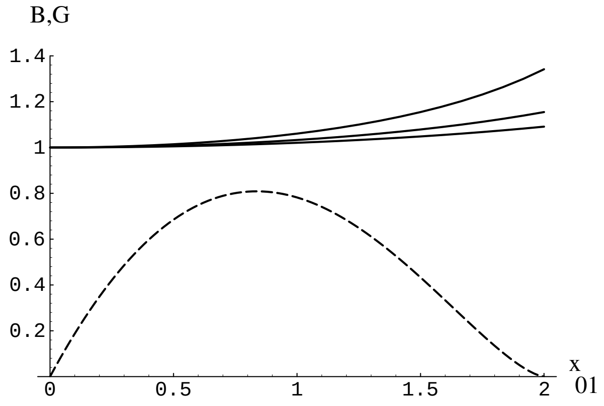

Figure 4: The variation of $B \equiv b/10^{\frac{n}{20}}$ with x_{01}, for—uppermost trace first—the cases $D =$ 3, 4 and 5, as given by eq. (9). Also shown (dashed) is the probability distribution for the distance x_{01}, $G(x_{01})$ given by eq. (7).

also present. Adding this to the signal power from transmitter 2, then in fact successful demodulation can only occur if

$$10 \log_{10} \left(\frac{\frac{P_t(1)}{x_{01}^2}}{\frac{P_t}{x_{02}^2} + P_n} \right) \geq n \; , \tag{8}$$

where x_{02} is the distance from 0 to 2. Identifying $b = \frac{x_{01}}{x_{02}}$ for the equality case, rearrangement gives

$$b = \frac{10^{\frac{n}{20}}}{\left(1 - \frac{x_{01}^2}{D^2}\right)^{1/2}} \; , \tag{9}$$

so there is now a dependence on the separation of 0 and 1, relative to the maximum communication distance D. This is also illustrated in fig. 4, for the physically sensible choices[2] of $D = 3$, 4 and 5. It is clear that b rises away from the constant approximation as x_{01} approaches its maximum—this is where the actual noise added to the transmitter 2 power begins to bite. However, even for the worst case $D = 3$, given the region where $G(x_{01})$ carries the most weight then it is not a bad approximation to simply take b to equal its limiting ($D \to \infty$) value of $10^{\frac{n}{20}}$. We shall employ this approximation in what follows.

[2]Receiver 0 and transmitter 1 want to be able to communicate wherever they are in circle 1. This immediately constrains $D \geq 2$, and it makes sense to operate somewhat above the absolute minimum.

A remark is in order about our choice of normalization. In our analysis we render lengths dimensionless by using units of the radius of the circle in which receiver 0 and transmitter 1 are confined, R_1, rather than, for example, the maximum communication distance, D. As the "cell size" R_1 is a physical parameter, which will be determined in any real situation,[3] it is appropriate to measure other lengths in terms of this. For some set R_1, equipment must be chosen with a constraint on D—at bare minimum D must be twice R_1 so 0 and 1 can communicate wherever they are in circle 1. Fig. 4 shows that with this constraint, when D is appreciably greater than R_1, it induces little variation of the important quantity b (which represents the distance ratio for the signal to noise or carrier to interference limit). We thus approximate b as a constant parameter. The other important parameter which determines the probabilities is the maximum *sensing* distance a, (which is a multiple of D). We therefore analyse our probability results in terms of b, a and the circle radius R_2, the two lengths expressed in units of the "cell size" R_1.

5 Case B: Two transmitters and a receiver

The mathematical problem to which this reduces is the following:

Take two concentric circles with radii R_1 and R_2. Pick points $\vec{x}_0$ and $\vec{x}_1$ in circle 1 and a point $\vec{x}_2$ in circle 2 (at random and with uniform weight). For a given a and b, calculate the four probabilities:

$$\begin{aligned}
P(R_2,a,b)_{|\vec{x}_1-\vec{x}_2|\le a\ ,\ \frac{|\vec{x}_2-\vec{x}_0|}{|\vec{x}_1-\vec{x}_0|}\le b} &\equiv p_1(R_2,a,b)\\
P(R_2,a,b)_{|\vec{x}_1-\vec{x}_2|\le a\ ,\ \frac{|\vec{x}_2-\vec{x}_0|}{|\vec{x}_1-\vec{x}_0|}> b} &\equiv p_2(R_2,a,b)\\
P(R_2,a,b)_{|\vec{x}_1-\vec{x}_2|> a\ ,\ \frac{|\vec{x}_2-\vec{x}_0|}{|\vec{x}_1-\vec{x}_0|}\le b} &\equiv p_3(R_2,a,b)\\
P(R_2,a,b)_{|\vec{x}_1-\vec{x}_2|> a\ ,\ \frac{|\vec{x}_2-\vec{x}_0|}{|\vec{x}_1-\vec{x}_0|}> b} &\equiv p_4(R_2,a,b)\ .
\end{aligned} \tag{10}$$

These correspond to the four situations given in the introduction. The probabilities are clearly exhaustive, so that

$$\sum_{i=1}^{4} p_i(R_2,a,b)\ =\ 1 \tag{11}$$

independent of R_2, a and b.

The strategy for the evaluation of, say, p_1 proceeds as follows:

- **(1)** Choose $\vec{x}_1$ in circle 1 and $\vec{x}_2$ in circle 2, such as shown in fig. 5.

[3] In the case of an office environment, which we discuss later, it is the size of an individual's workspace.

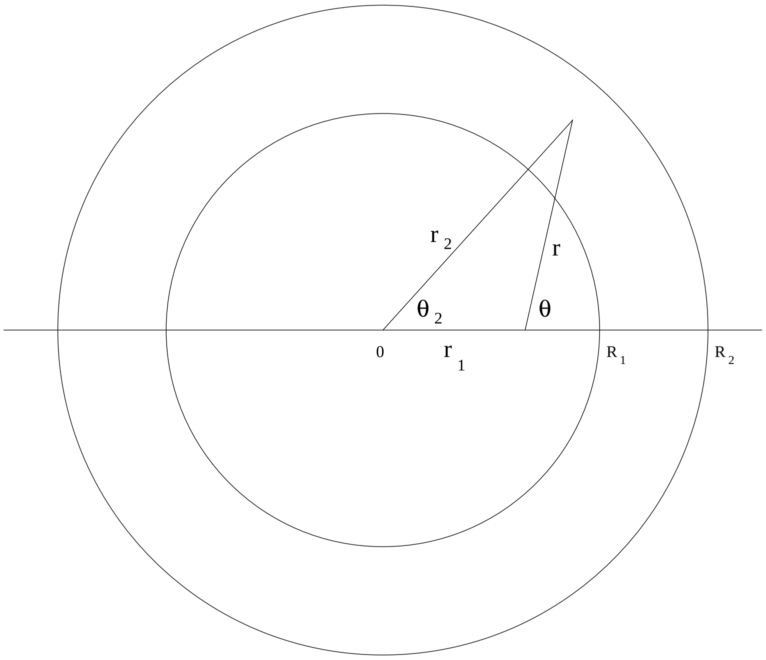

Figure 5: The geometric arrangement relevant for case B. The distance between points $\vec{x}_1$ and $\vec{x}_2$ is defined to be r.

- **(2)** Calculate the probability of choosing $\vec{x}_0$ in circle 1 which satisfies $|\vec{x}_2 - \vec{x}_0| \le b\,|\vec{x}_1 - \vec{x}_0|$ for the fixed $\vec{x}_1$ and $\vec{x}_2$.
- **(3)** Average this over all $\vec{x}_2$ in circle 2 which satisfy $|\vec{x}_1 - \vec{x}_2| \le a$, for fixed $\vec{x}_1$.
- **(4)** Average this over all $\vec{x}_1$ in circle 1.

The detailed calculations behind steps **(2)**-**(4)** are:

(2) First, consider the two fixed points $\vec{x}_1$ and $\vec{x}_2$, defined to be separated by a distance r. The points $\vec{x}_0$ which satisfy the equality $|\vec{x}_2 - \vec{x}_0| = b\,|\vec{x}_1 - \vec{x}_0|$ lie on a circle of radius

$$r_0 \equiv \frac{rb}{b^2 - 1} \tag{12}$$

centred a distance $\frac{r}{b^2-1}$ from point $\vec{x}_1$, as shown in fig. 6. This can be seen by defining the cartesian co-ordinates u and w as in fig. 6; the boundary of interest then gives the standard form for a circle,

$$\left(u + \frac{r}{b^2 - 1}\right)^2 + w^2 = \frac{r^2 b^2}{(b^2 - 1)^2}\,. \tag{13}$$

Points which satisfy the inequality $|\vec{x}_2 - \vec{x}_0| \le b\,|\vec{x}_1 - \vec{x}_0|$ lie on or outside this circle.

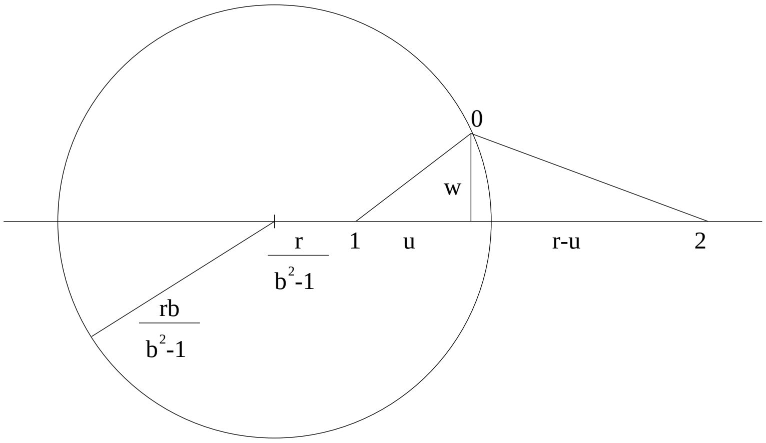

Figure 6: The locus of points defined by the equality $|\vec{x}_2 - \vec{x}_0| = b\,|\vec{x}_1 - \vec{x}_0|$. Cartesian co-ordinates u and w define the position relative to point $\vec{x}_1$.

Thus, those which satisfy the inequality *and* lie within circle 1 are given by the area shown in fig. 7. Using units of R_1, this area is given by $A(y, 1, r_0)$ where A is defined by eq. (4), r_0 by eq. (12) and y is shown in fig. 7. This area must be divided by π, the area of circle 1, to turn it into a probability. Length y is given by the cosine rule [7] as

$$y^2 = \frac{r^2}{(b^2-1)^2} + r_1^2 - \frac{2r_1 r\cos\theta}{(b^2-1)}\;. \tag{14}$$

(3) In order to average $A(y, 1, r_0)/\pi$ over the allowed positions of $\vec{x}_2$ for fixed $\vec{x}_1$, the lengths y and r_0 must be given in terms of r_1, r_2 and θ_2. From

$$r = (r_1^2 + r_2^2 - 2r_1 r_2\cos\theta_2)^{1/2},$$

$$r\cos\theta = r_2\cos\theta_2 - r_1$$

and the above formulae for r_0 and y, we obtain

$$y = \frac{(r_1^2 b^4 + r_2^2 - 2r_1 r_2 b^2\cos\theta_2)^{1/2}}{(b^2-1)} \tag{15}$$

and

$$r_0 = \frac{b(r_1^2 + r_2^2 - 2r_1 r_2\cos\theta_2)^{1/2}}{(b^2-1)}\;. \tag{16}$$

For the case p_1, the average over points $\vec{x}_2$ which satisfy $|\vec{x}_1 - \vec{x}_2| \le a$ corresponds to an average over the area defined in fig. 8. For $0 \le \theta_2 < \bar{\theta_2}$ the range of integration for r_2 is $0 \to R_2$ and for $\bar{\theta_2} \le \theta_2 \le \pi$ it is $0 \to \bar{r_2}$, where $\bar{r_2}$ is shown in fig. 8. The cosine rule gives

$$\bar{r_2} = r_1\cos\theta_2 + (a^2 - r_1^2\sin^2\theta_2)^{1/2}\;. \tag{17}$$

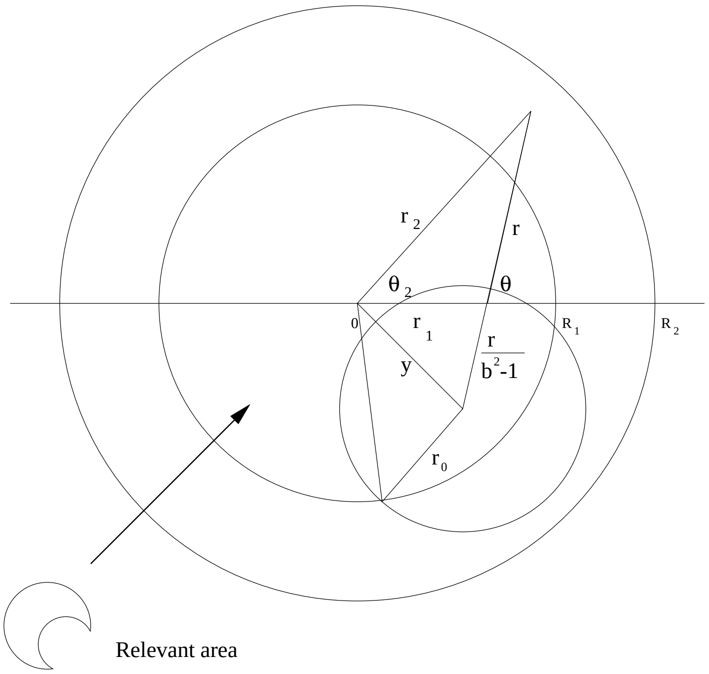

Figure 7: The relevant area of points which satisfy $|\vec{x}_2 - \vec{x}_0| \ \leq \ b\,|\vec{x}_1 - \vec{x}_0|$ and lie within circle 1.

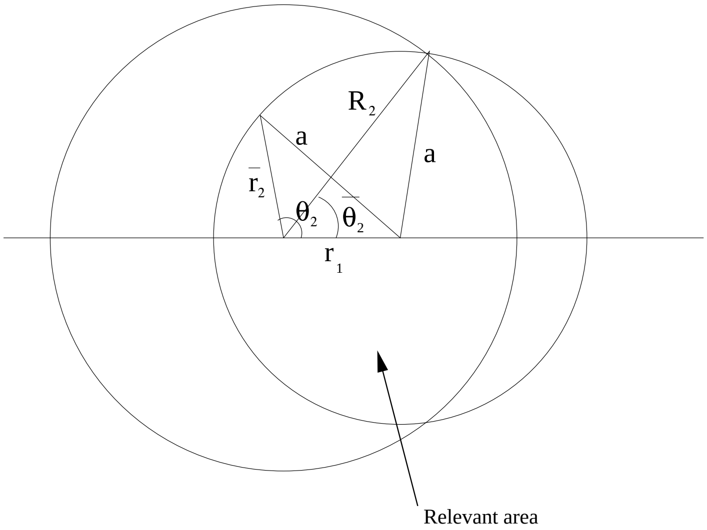

Figure 8: The relevant integration region for points $\vec{x}_2$ which satisfy $|\vec{x}_1 - \vec{x}_2| \le a$. This is the appropriate region for probability p_1.

$\bar{\theta_2}$ is given by eq. (1). Division by πR_2^2 normalizes this integration to give a probability.

(4) The final integration over all $\vec{x}_1$ in circle 1 is easy. There is angular symmetry which simply leaves an integral $0 \to 1$ for r_1.

The calculations for the other probabilities, p_2, p_3 and p_4 are very similar. For p_2 the relevant area in fig. 7 is the other piece of circle 1; this is integrated over the same ranges as for p_1. p_3 uses the same relevant area of fig. 7 as p_1, but it is integrated over the other piece of circle 2 in fig. 8. (The final integral over circle 1 is still the same.) p_4 uses the same fig. 8 area as for p_2, but this integrated over the same ranges as for p_3.

The four probability results are therefore given by

$$p_1(R_2, a, b) = \frac{4}{\pi^2 R_2^2} \int_0^1 dr_1\, r_1 \left[\int_0^{\bar{\theta_2}} d\theta_2 \int_0^{R_2} dr_2\, r_2 + \int_{\bar{\theta_2}}^{\pi} d\theta_2 \int_0^{\bar{r_2}} dr_2\, r_2 \right] A(y, 1, r_0)\ , \tag{18}$$

$$p_2(R_2, a, b) = \frac{4}{\pi^2 R_2^2} \int_0^1 dr_1 r_1 \left[\int_0^{\bar{\theta_2}} d\theta_2 \int_0^{R_2} dr_2\, r_2 + \int_{\bar{\theta_2}}^{\pi} d\theta_2 \int_0^{\bar{r_2}} dr_2\, r_2 \right] (\pi - A(y, 1, r_0))\ , \tag{19}$$

$$p_3(R_2, a, b) = \frac{4}{\pi^2 R_2^2} \int_0^1 dr_1\, r_1 \int_{\bar{\theta_2}}^{\pi} d\theta_2 \int_{\bar{r_2}}^{R_2} dr_2\, r_2\, A(y, 1, r_0) \tag{20}$$

and

$$p_4(R_2, a, b) = \frac{4}{\pi^2 R_2^2} \int_0^1 dr_1\, r_1 \int_{\bar{\theta_2}}^{\pi} d\theta_2 \int_{\bar{r_2}}^{R_2} dr_2\, r_2\, (\pi - A(y, 1, r_0))\ . \tag{21}$$

In these expressions, $\bar{\theta_2}$ is given by eq. (1), $\bar{r_2}$ by eq. (17) and A, as a function of

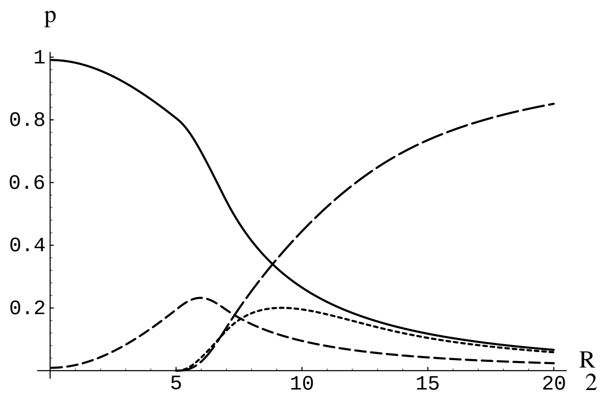

Figure 9: Plots of the four probabilities p_1 (solid), p_2 (medium dashes), p_3 (small dashes) and p_4 (large dashes), as functions of R_2 for $a = 6.0$ and $b = 7.08$. See eqs. (18)-(21).

it arguments, by eq. (4). The relationships (15) and (16) must be used to convert $A(y, 1, r_0)$ to a function of the integration variables.

Examples of these four probabilities are plotted in figs. 9, 10 and 11. They are shown as functions of R_2, for fixed values of $b = 7.08$ and $a = 6$, 8 and 10. The value of b was calculated in the introduction, as was the relationship $a = 2D$ for sensing. We have already argued that $D = 3$, 4 and 5 are sensible choices for D; hence the values of a. Every point used to make up each of the four curves shown was computed from a 3-dimensional numerical integration of the appropriate expression taken from (18)-(21). The plots demonstrate some generic features; for $R_2 \leq (a-1)$, $p_3 = p_4 = 0$ and for $R_2 \to \infty$, $p_4 \to 1$ and $p_1, p_2, p_3 \to 0$.

It is tempting to suppose that for $R_2 \to 0$, $p_1 \to 1$ and $p_2, p_3, p_4 \to 0$. However, this is not generally true; it only holds for $b \gg 1$. For $b \not\gg 1$ (such as in the numerical example of $b = 7.08$ considered here) there is a small calculable correction. The limit $r_2 \to 0$ can be taken in eqs. (15) and (16) to give $y = \frac{r_1 b^2}{(b^2-1)}$ and $r_0 = \frac{r_1 b}{(b^2-1)}$. For $r_2 = 0$, the centre of the circle with radius r_0 lies along the radial direction defined by r_1. (See fig. 7.) The relevant area for integration to evaluate p_1 is still $A(y, 1, r_0)$ with the simplified y and r_0 and there is now only an integral over r_1 to be done. Thus

$$p_1(0, a, b) = \frac{2}{\pi} \int_0^1 dr_1\, r_1\, A(y, 1, r_0) \tag{22}$$

and $p_2(0, a, b) = 1 - p_1(0, a, b)$. For $b = 7.08$ these give $p_1(0, a, 7.08) = 0.991$ and $p_2(0, a, 7.08) = 0.009$, in good agreement with the data computed for figs. 9-11. This

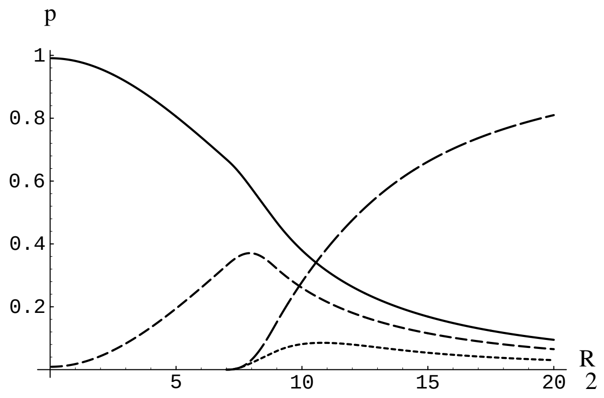

Figure 10: Plots of the four probabilities p_1 (solid), p_2 (medium dashes), p_3 (small dashes) and p_4 (large dashes), as functions of R_2 for $a = 8.0$ and $b = 7.08$. See eqs. (18)-(21).

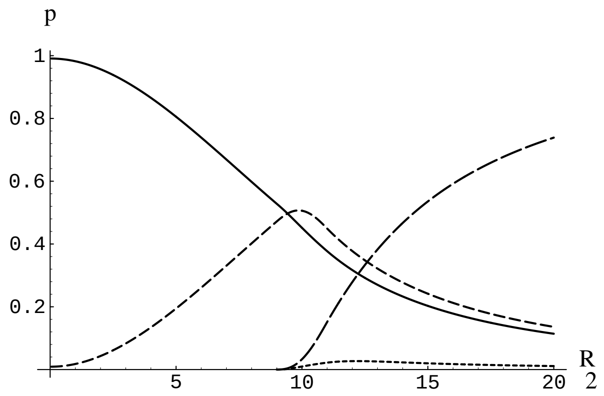

Figure 11: Plots of the four probabilities p_1 (solid), p_2 (medium dashes), p_3 (small dashes) and p_4 (large dashes), as functions of R_2 for $a = 10.0$ and $b = 7.08$. See eqs. (18)-(21).

provides a useful consistency check on the more involved numerical calculations.

From figs. 9-11 it is apparent that the probabilities of the two undesirable events, p_2 and p_3, exhibit maxima as functions of R_2, for the given a and b. This is to be expected since their definitions (see eqs. (10)) effectively contain competing inequalities. The R_2 values at which these occur can be found by setting the derivatives

$$\begin{aligned}
&\frac{\partial p_2(R_2,a,b)}{\partial R_2} = \\
&\quad \frac{-8}{\pi^2 R_2^3}\int_0^1 dr_1\, r_1 \left[\int_0^{\bar{\theta_2}} d\theta_2 \int_0^{R_2} dr_2\, r_2 + \int_{\bar{\theta_2}}^{\pi} d\theta_2 \int_0^{\bar{r_2}} dr_2\, r_2\right] (\pi - A(y,1,r_0)) \\
&\quad + \frac{4}{\pi^2 R_2^2}\int_0^1 dr_1\, r_1 \int_0^{\bar{\theta_2}} d\theta_2\, R_2\, (\pi - A(\bar{y},1,\bar{r_0})) \\
&\quad + \frac{4}{\pi^2 R_2^2}\int_0^1 dr_1\, r_1 \frac{\partial\bar{\theta_2}}{\partial R_2}\left[\int_0^{R_2} dr_2\, r_2 - \int_0^{\hat{r_2}} dr_2\, r_2\right] (\pi - A(\hat{y},1,\hat{r_0}))
\end{aligned} \tag{23}$$

and

$$\begin{aligned}
\frac{\partial p_3(R_2,a,b)}{\partial R_2} = {} & \frac{-8}{\pi^2 R_2^3}\int_0^1 dr_1\, r_1 \int_{\bar{\theta_2}}^{\pi} d\theta_2 \int_{\bar{r_2}}^{R_2} dr_2\, r_2\, A(y,1,r_0) \\
& + \frac{4}{\pi^2 R_2^2}\int_0^1 dr_1\, r_1 \int_{\bar{\theta_2}}^{\pi} d\theta_2\, R_2\, A(\bar{y},1,\bar{r_0}) \\
& - \frac{4}{\pi^2 R_2^2}\int_0^1 dr_1\, r_1 \frac{\partial\bar{\theta_2}}{\partial R_2}\int_{\hat{r_2}}^{R_2} dr_2\, r_2\, A(\hat{y},1,\hat{r_0})
\end{aligned} \tag{24}$$

equal to zero. Here $\bar{\theta_2}$ is given by eq. (1), $\bar{r_2}$ is given by eq. (17) and the definitions

$$\bar{y} \equiv \frac{(r_1^2 b^4 + R_2^2 - 2r_1 R_2 b^2 \cos\theta_2)^{1/2}}{(b^2-1)}\ , \tag{25}$$

$$\bar{r_0} \equiv \frac{b(r_1^2 + R_2^2 - 2r_1 R_2 \cos\theta_2)^{1/2}}{(b^2-1)}\ , \tag{26}$$

$$\hat{y} \equiv \frac{(r_1^2 b^4 + r_2^2 - 2r_1 r_2 b^2 \cos\bar{\theta_2})^{1/2}}{(b^2-1)}\ , \tag{27}$$

$$\hat{r_0} \equiv \frac{b(r_1^2 + r_2^2 - 2r_1 r_2 \cos\bar{\theta_2})^{1/2}}{(b^2-1)} \tag{28}$$

and

$$\hat{r_2} \equiv r_1 \cos\bar{\theta_2} + (a^2 - r_1^2 \sin^2\bar{\theta_2})^{1/2} \tag{29}$$

apply. The other quantity which arises is

$$\frac{\partial\bar{\theta_2}}{\partial R_2} = \frac{r_1^2 - R_2^2 - a^2}{R_2(2R_2^2 a^2 + 2R_2^2 r_1^2 + 2r_1^2 a^2 - a^4 - r_1^4 - R_2^4)^{1/2}}\ . \tag{30}$$

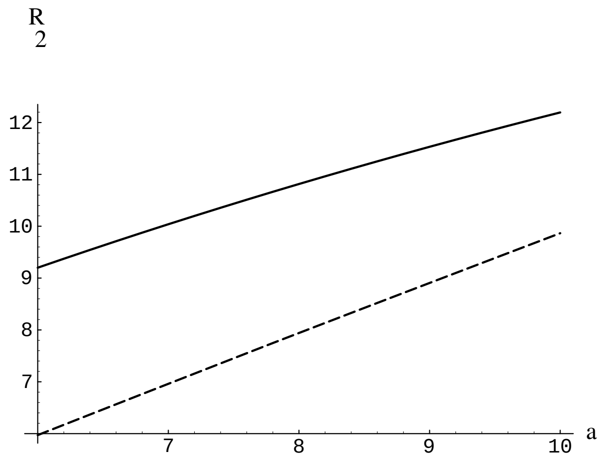

Figure 12: Plots of the R_2 positions of the maxima in p_2 (dashes) and p_3 (solid), as a function of a for $b = 7.08$.

Setting the probability derivatives, (23) and (24), equal to zero results in rather formidable equations. Nevertheless, they can be solved numerically to yield the locations of the maxima. Fig. 12 illustrates this, giving the R_2 locations of the peaks in p_2 and p_3, as a function of a for a fixed value of b. The actual values of p_2 and p_3 at their maxima can be found by feeding the (R_2, a, b) data back into eqs. (19) and (20). These results are shown in fig. 13. Although these calculations are again lengthy, they are in fact considerably shorter than a *run* of calculations to find one of the $p_i(R_2, a, b)$ as a function of R_2 for fixed a and b (as in figs. 9-11). Thus, if the positions and heights of the peaks are of primary interest, there exists a handy short cut for obtaining these, compared to the calculations required for the complete pictures of figs. 9-11.

6 Interpretation of results

The behaviour of the four probabilities as shown in figs. 9, 10 and 11 is as would be expected. They also agree well with Monte-Carlo simulations that were carried out for validation purposes.

Probability p_1 increases and probability p_2 decreases as R_2 increases. This is because as the circle containing the interfering transmitter 2 increases, the likelihood of it being able to carry out a simultaneous communication gets greater. However, this is not possible unless R_2 exceeds $(a - 1)$ and hence p_3 and p_4 are zero until this

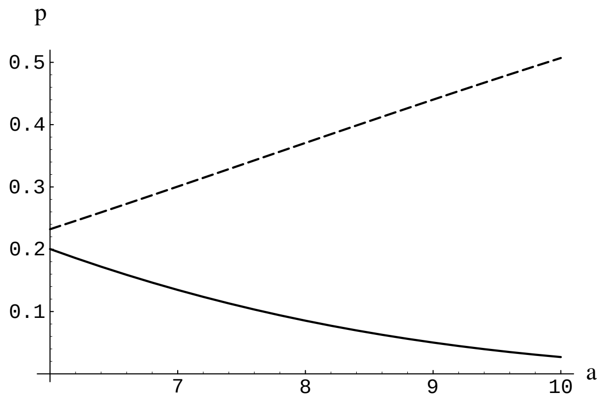

Figure 13: Plots of the values at their maxima of p_2 (dashes) and p_3 (solid), as a function of a for $b = 7.08$.

point is reached. Beyond here, p_3 and p_4 increase, as there is an increased probability of transmitter 2 not sensing transmitter 1, with positive and negative interference results. Ultimately, as R_2 tends to infinity, the likelihood is that transmitter 2 will not sense transmitter 1 and it will not interfere; hence p_3 tends to zero but p_4 tends to unity.

As stated earlier, both p_2 and p_3 are of interest to us, p_3 more so because its effect is more destructive. Both of these probabilities reach maxima and then tend to zero. The positions and values of these maxima, which are plotted in figs. 12 and 13 respectively, are of considerable interest to us. Practical probability estimates for multi-user environments, which we address shortly, require the large R_2 behaviour. A very useful approximation may be employed here. This makes use of the fact that, beyond their maxima, p_3 and p_2 decay like $1/R_2^2$ and so for this regime may be approximated as

$$p_i \approx p_{mi} \frac{R_{2mi}^2}{R_2^2} \tag{31}$$

for $i = 2, 3$. Here p_{mi} is the maximum value of the probability and R_{2mi} is the value of R_2 at which it is attained—these are given in figs. 12 and 13. Although deferrals of transmission which avoid a collision are correct as far as protocol is concerned, they also limit the actual throughput in a system. (These are described by scenario 1 of the four given in section 2.) Since for large R_2 the probability p_1 also decays like $1/R_2^2$, eq. (31) also applies for $i = 1$, with p_{m1} the value of p_1 at R_{2m1}, any value in the inverse-square law regime.

Thus far we have only considered the probability of a *single* transmitter in circle 2 interfering with an on-going communication between a transmitter and a receiver in circle 1. In reality, of course, there may be *many* transmitters within circle 2, sensing the channel prior to initiating a communication. Hence the *full* probability of a particular transmitter avoiding a collision or deferring when it is active is the product of the probabilities for it avoiding each of the others which is active[4] in circle 2. As an example (which can easily be scaled to other densities), let us assume that the density of interfering transmitters is uniform and equal to that of transmitter 1, i.e. the density is $1/\pi$ (in units of R_1)—there is one transmitter per "cell" area everywhere. In this case there are ηR_2^2 active transmitters inside a circle of radius R_2, where η is the probability of an individual being active. The full probability for successful transmission is therefore $p_{f4} \equiv p_4^{\eta R_2^2}$. Since $p_4 = 1 - p_1 - p_2 - p_3$ and assuming that the cut-off R_2 is large enough to apply eq. (31), this approximates to

$$p_{f4} = \exp\left[-\eta \sum_{i=1}^{3} p_{mi} R_{2mi}^2\right] \equiv \exp\left[-\eta K\right] \ . \qquad (32)$$

Provided that the argument is small this can be written as

$$p_{f4} = 1 - \eta \sum_{i=1}^{3} p_{mi} R_{2mi}^2 \ . \qquad (33)$$

In this limit the subtracted terms can be interpreted as the full probabilities for each of the three non-transmission scenarios (1-3 in section 2). These results are independent of the cut-off, so it is important to note that they hold *even if the cut-off is infinite*. As long as the transmitter *density* is uniform (and we have a decent estimate of its value), the results hold even if there is no limit to the actual number of potentially interfering transmitters. This is a very useful result. It is also interesting to note that p_{f4} decays exponentially with increasing η.

As a practical example, let us consider a wireless LAN deployment in an office environment with one device per individual workspace and each work space having an area of πR_1^2, π in our units. If we assume that there are seven different channels in the system bandwidth and that, over a given time period—say the duration of a packet—about 5% of the transmitters are trying to send a packet, then this gives us $\eta = 0.007$ for a given channel. (This is the activity fraction $x = 0.05$ of a transmitter divided by the number of channels $k = 7$, so $\eta = x/k$.) Relevant data from figs. 9-13 is summarized in table 1. It can be seen from these data that a modest variation in the sensing distance a (6 to 10) causes a drastic variation in the collision probability p_{m3} (0.20 to 0.027). The collision probability that a system can tolerate is very much dependent on the traffic that it is carrying. However, systems are designed to cope well with a packet failure rate as high as 10assuming that these are recovered by re-transmissions to protect the higher layers in the protocol. So the choice of parameters

[4] We assume that they each satisfy our random distribution assumption.

Table 1: Data from figs. 9-13 and calculated values of K (see eq. 32). R_{2m1}, which can be any point in the inverse-square law regime for p_1, is set at 10 for all three cases.

a	6	8	10
p_{m1}	0.26	0.38	0.45
p_{m2}	0.23	0.37	0.50
p_{m3}	0.20	0.085	0.027
R_{2m1}	10	10	10
R_{2m2}	6.0	7.9	9.9
R_{2m3}	9.2	10.8	12.2
K	51	71	98

($a = 8$) is appropriate from an integrity perspective, but we must look at what this means in terms of capacity or throughput. For a sensing distance of $a = 8$ and taking data from table 1), an estimate for the successful transmission probability is $p_{f4} \approx 0.61$. Results for other values of a and also for $\eta = 0.014$ (twice the activity, $x = 0.1$) are given in tables 2 and 3. We stress that these probabilities follow solely as a result of the imperfections in the listen-before-talk over the wireless channel. For example, the calculations of p_{f4} do not take into account the additional collisions, which may occur through the contention resolution mechanism in the medium access control, because it is not the intention of this general study to consider such specific aspects of system design.

If in the example scenario introduced above the usable channel bandwidth is $C = 10$ Mbps per channel and there are seven channels in the band, then this system has a usable bandwidth of 70 Mbps. These round figures are representative of what a system like HIPERLAN [3] might achieve in user data rate after the overheads have been taken into account.

To calculate the system capacity we must estimate the re-use for this resource based on the distance beyond which a simultaneous transmission on the same channel has no additional effect on the collision probability p_{m3}. For the area of radius $R_{2m3} = 11$ and in more absolute terms, if R_1 is ~ 1 m then the system capacity is around $c = 0.19$ Mbps/m^2. To calculate the system throughput for an given offered load we must consider the probability of successful transmission p_{m4} given a number of active transmitters using the system bandwidth. With an activity of $x = 0.05$, the load applied by the users is $l = 0.16$ Mbps/m^2. Given the success probability $p_{f4} = 0.61$, the actual throughput is $t = 0.098$ Mbps/m^2, a fraction 0.51 of the capacity. The analogous results for other values of a and x are given in tables 2 and 3. Although the load l exceeds the capacity c in some cases, the throughput t never does because of the transmission failures.

Table 2: Model system response for transmitter activity fraction $x = 0.05$ spread over $k = 7$ channels (giving $\eta = 0.007$) each of usable bandwidth $C = 10$ Mbps. $R_1 = 1$ m.

a	6	8	10
p_{f4}	0.70	0.61	0.50
c (Mbps/m^2)	0.26	0.19	0.15
l (Mbps/m^2)	0.16	0.16	0.16
t (Mbps/m^2)	0.11	0.098	0.080
t/c	0.43	0.51	0.53

Table 3: Model system response for transmitter activity fraction $x = 0.1$ spread over $k = 7$ channels (giving $\eta = 0.014$) each of usable bandwidth $C = 10$ Mbps. $R_1 = 1$ m.

a	6	8	10
p_{f4}	0.49	0.37	0.25
c (Mbps/m^2)	0.26	0.19	0.15
l (Mbps/m^2)	0.32	0.32	0.32
t (Mbps/m^2)	0.16	0.12	0.080
t/c	0.60	0.62	0.53

Clearly this analysis contains no dynamics, in that no allowance has been made for the response of individual terminals to a deferral or a collision. In a dynamical system the activity fraction x would change in response to transmission failures according to the system protocol. Nevertheless the simple estimates presented here are representative of the equilibrium behaviour of dynamical systems whose parameters fall in the appropriate regime and protocols will accept throughputs and success probabilities of this order.

The above example is only an illustration of how this analysis can be used to investigate the system capacity/throughput and integrity trade-off, for selection of optimum system parameters. Nevertheless, this scenario is representative of systems that are currently being designed to operate at 5 GHz [3, 4]. Although the calculations are approximate, they clearly illustrate that a significant packet failure rate will exist through the imperfections of the listen before talk alone. This can only be reduced by increasing the sensing distance and significantly reducing the system capacity.

7 Conclusions

We have presented an analysis of event probabilities for wireless local area networks using distributed medium access control based on listen before talk. More significantly, we have shown how these can be used to calculate capacity or throughput for a given offered load for these networks. This in turn can be used to optimize system parameters for whatever performance aspect is of importance. The results agreed well with Monte-Carlo simulations carried out. The results are particularly applicable to systems designed for operation at 5 GHz [3, 4]. It is shown that for such systems with a usable system bandwidth of 70 Mbps in typical deployments, the capacity or throughput are of the order of 100 kbps/m^2 for reasonable packet failure rates.

Future work will consider the enhancement of the analysis to consider RTS/CTS (Request-To-Send Clear-To-Send) medium access control and the introduction of power control and/or adaptive sensing thresholds to improve system performance.

References

[1] FCC Report and Order in the matter of Amendment to the Commission's Rules to provide for Operation of Unlicensed NII Devices in the 5 GHz frequency range, (January 1997).

[2] CEPT Recommendation T/R 22-06, "Harmonised Radio Frequency Bands for High Performance Radio Local Area Networks (HIPERLANs) in the 5 GHz and 17 GHz Frequency Range," (Madrid 1992, revised Nicosia 1994).

[3] ETSI Radio Equipment and Systems (RES): High Performance Radio Local Area Network (HIPERLAN) Type 1, Functional Specification ETS 300 652.

[4] Draft Supplement to STANDARD [for] Information Technology- Telecommunications and information exchange between systems- Local and metropolitan area networks- Specific Requirements- Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: High Speed Physical Layer in the 5 GHz Band.

[5] Y. Sun, A. Nix and J. P. McGeehan, "HIPERLAN Performance Analysis with Dual Antenna Diversity and Decision Feedback Equalization," *IEEE VTC*, 1996, pp. 1549-1553.

[6] J. Tellado, E. Khayata and J. M. Cioffi, "Equalizing GMSK for High Data Rate Wireless LANs," *PIMRC*, 1995, pp. 16-20.

[7] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Section 4.3.148, p.79 (Dover 1972).