Duplex and Triplex Memory: Which Is More Reliable?

Nitin H. Vaidya
Department of Computer Science
Texas A&M University
College Station, TX 77843-3112
E-mail: vaidya@cs.tamu.edu

Technical Report 94-025 February 1994

Abstract

A large number of choices exist when designing a reliable memory system. The choices range from simple replication to complex error control codes (ECC). Simple replication (without ECC) has a better performance as compared to systems using ECC, as ECC necessitate decoding, deteriorating the performance. Clearly, simple replication improves performance at the cost of increased redundancy. An intermediate solution is to use combination of replication and simple ECC. Similar approach has been used in some commercial systems [1].

This report compares reliability of memory systems formed using simple triplication (without ECC) with memory systems formed by duplicating memory modules that use ECC. Reliability of a system using duplication of memory modules, with codes capable of only error detection or codes only capable of single error correction, is shown to be worse than simple triplication. It is also shown that systems using duplication of memory modules, with codes capable of single error correction and double error detection can achieve better reliability than simple triplication.

1 Introduction

A large number of choices exist when designing a reliable memory system. The choices range from simple replication to complex error control codes (ECC). Simple replication (without ECC) has a better performance as compared to systems using ECC, as ECC necessitate decoding, deteriorating the performance. Clearly, simple replication improves performance at the cost of increased redundancy. An intermediate solution is to use combination of replication and simple ECC. Such schemes have been used in commercial systems [1].

This report compares reliability of memory systems formed by simple triplication (without ECC) with memory systems formed by duplicating memory modules that use ECC. Following the standard terminology, these two systems are referred to as *triplex* and *duplex* systems, respectively. Figure 1 illustrates the two systems under consideration.

The triplex system in Figure 1(a) is formed by simple triplication of memory modules that do not use any error control coding. The memory system output is obtained by bit-wise voting on the output of the three modules in the triplex system.

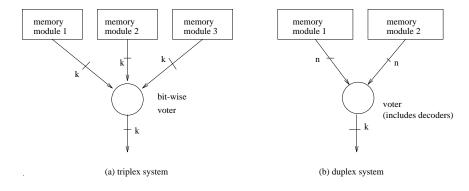


Figure 1: Triplex and Duplex Memory System Models

The duplex system in Figure 1(b) consists of two identical memory modules. Each memory module uses an (n, k) error control code. Encoded outputs of the two modules are available to a voter that can decode the outputs of the two memory modules. The exact function of the voter in duplex systems will be defined more precisely in the next section.

The objective here is to determine the minimum *capability* required in the ECC such that the duplex system can achieve a higher reliability than the triplex system. The

capability of an ECC is defined here by the number of errors it can correct and detect. A t_1 error correcting- d_1 error detecting code is said to be less capable than a t_2 -error correcting- d_2 error detecting code if (i) $t_1 < t_2$ or (ii) $t_1 = t_2$ and $d_1 < d_2$. The work presented in this
report is motivated by our previous work on modular redundant system reliability and safety
[4].

In this report, reliability of the triplex system is compared with duplex systems that use error control codes (ECC) of different capabilities:

- Error detection only.
- Single error correction (and no error detection, i.e. more than one error is assumed to result in erroneous decoding.).
- Single error correction and double error detection.

In the following, reliability of the triplex system is evaluated first, followed by evaluation of reliability of the three duplex systems and comparison with the triplex system.

For the reliability analysis, we use the independent symmetric error model [3]. It is assumed that each bit of a codeword in memory may become erroneous independently with probability p. In reality, probability p is expected to be quite small (of the order of 10^{-3} or less). Reliability of the voters is assumed to 1.

Each data word contains k bits. The ECC used in the duplex scheme is an (n, k) code for some $n \geq k$. The number of words in the memory system is denoted by W.

Definition 1 Reliability R_S of a memory system S is defined as the probability that all words in the memory can be accessed correctly.

Reliability of Triplex System

A data word contains k bits, therefore, the probability that a given word in the memory can be accessed correctly is

$$R_{triplex}^* = \left[(1-p)^3 + 3p(1-p)^2 \right]^k = (1-p)^{2k} (1+2p)^k \tag{1}$$

Therefore, reliability of the triplex system is given by $R_{triplex} = (R_{triplex}^*)^W$. Recall that W is the total number of words in memory system (i.e., each memory module contains W words).

2 Reliability of Duplex Systems

This section evaluates reliability of three different types of duplex systems and compares them with the triplex system.

2.1 Duplex system using error detecting codes

In this section we show that the reliability achieved by a duplex system, using an ECC for error detection only, is always less than a triplex system. The triplex system uses three replicas of each word, requiring 3k bits per word. The duplex system requires 2n bits per word. For a fair comparison, we consider only those (n,k) error control codes for which $2n \leq 3k$ or $n \leq 3k/2$.

Let the reliability of the duplex system under consideration here be denoted by $R_{duplex1}$. Each memory module in the duplex system uses an (n, k) error detecting code with $n \leq 3k/2$. Let P_u denote the probability that an undetected error occurs in a codeword of this code. The function of the voter in this system is as follows: When a word is to be read from the memory, the corresponding codewords from the two memory modules are provided as input to the voter. The voter decodes the two codewords to detect errors. If errors are detected in both codewords, then the voter does not produce any data. If exactly one word is detected to contain an error, then the other decoded codeword is produced as output. If neither codeword is detected to contain an error, then any one decoded codeword is produced as the output. This voter will maximize the reliability under the constraint that each codeword is to be used only for error detection (no error correction). With such a voter, the reliability is given by $R_{duplex1} = (R_{duplex1}^*)^W$, where

$$R_{duplex1}^* = (1-p)^{2n} + 2(1-P_u - (1-p)^n)(1-p)^n + \frac{1}{2}2(1-p)^n P_u$$

$$= (2 - P_u)(1 - p)^n - (1 - p)^{2n}$$
(2)

In the first expression above, the term $(1-p)^{2n}$ is the probability that both codewords are error-free. The term $2(1-P_u-(1-p)^n)(1-p)^n$ is the probability that one of the codewords contains a detectable error and the other codeword is error-free. The term $\frac{1}{2}2(1-p)^nP_u$ corresponds to the probability that one of the codewords contains an undetectable error, the other codeword is error-free and the voter chooses the error-free codeword. Note that in this situation, the voter output will be erroneous with probability $\frac{1}{2}$.

The theorem below states that triplex memory reliability is larger than that of a duplex system using ECC for error detection only.

Theorem 1 Given $0 and <math>n \le 3k/2$, $R_{duplex1}$ is always smaller than $R_{triplex}$ independent of the error detecting code used in the duplex system.

Proof: The proof is somewhat complex and is presented in Appendix A.1.

2.2 Duplex systems using single error correcting (SEC) codes

In this section, we assume that the error control code used in the duplex system can correct a single error and not detect any other errors. In other words, it is assumed that more than one error will result in incorrect decoding of this code. In the next section, we will consider a single error correcting and double error detecting code.

For the duplex system considered here, the voter function is as follows: The voter decodes the two codewords and corrects any errors that may be detected. Then, it outputs any one of the decoded codewords. This voter will maximize the reliability under the constraint that each codeword can be used only to correct a single error and that more than one error in a codeword causes erroneous decoding.

Let the reliability of the duplex system being considered in this section be denoted by $R_{duplex2}$. A given word can be accessed correctly when the two codewords contain at most one error each. In the case where one of the codewords has at most one error and the other

codeword contains more than one error, there is a 50% chance that the correct information will be obtained (recollect that multiple errors in a codeword are not detected). When both codewords contain more than one error, correct information cannot be obtained. Therefore, $R_{duplex2} = (R_{duplex2}^*)^W$ where

$$R_{duplex2}^* = ((1-p)^n + np(1-p)^{n-1})^2 + \frac{1}{2}2((1-p)^n + np(1-p)^{n-1}) (1 - (1-p)^n - np(1-p)^{n-1})$$
$$= (1-p)^n + np(1-p)^{n-1}$$

The above expression is identical to the reliability that would be obtained if just one memory module with a single error correcting code were used (instead of two). This implies that when the error control code is *only* capable of correcting a single error, it does not help to use more than one memory module. Therefore, for this system, we impose a weaker constraint on n that $n \leq 3k$, instead of $n \leq 3k/2$.

Theorem 2 Given $0 and <math>n \le 3k$, $R_{duplex2}$ is always smaller than $R_{triplex}$ independent of the single error correcting code used in the duplex system.

Proof: The proof is presented in Appendix A.2.

Although the result stated above is proved for 0 , we conjecture that it holds true when <math>0 . In practice, <math>p is much smaller that 1/3, therefore, the above result is adequate for real applications.

2.3 Duplex systems using SEC-DED codes

This section shows that a duplex system using a single error correcting and double error detecting (SEC-DED) code can achieve reliability better than a triplex system. This is demonstrated with the help of examples.

Assume that the voter for duplex system using SEC-DED code functions as follows: It decodes the codeword from one of the memory modules and if zero or one error is detected in this codeword, the decoded codeword is produced as the output. If two errors are detected in this codeword, then the second codeword is decoded. In this case, the second decoded codeword is produced as output if it is detected to contain at most one error.

Let the reliability of the duplex system being considered here be denoted by $R_{duplex3}$. Then, $R_{duplex3} = (R_{duplex3}^*)^W$ where,

$$R_{duplex3}^* = (1-p)^n + np(1-p)^{n-1} + \binom{n}{2} p^2 (1-p)^{n-2} \left((1-p)^n + np(1-p)^{n-1} \right)$$

Unlike the results presented in Theorems 1 and 2, in this case, the duplex system can achieve a better reliability that the triplex system. We illustrate this with an example. Assume that the error control code used in the duplex system is a (n, k) code obtained by (possibly) shortening the distance-4 extended Hamming code [3]. For k = 16 and k = 32, Figures 2 and 3 plot the unreliability (i.e. 1-reliability) for duplex and triplex systems as a function of p. For the SEC-DED code, n is equal to 22 when k = 16 and 39 when k = 32. From the unreliability plots, it can be seen that for sufficiently small values of p the reliability of the duplex system is larger than that of the triplex system.

For $R_{duplex3}$ to be larger than $R_{triplex}$ p needs to be small enough. For example, in case of k = 32, p must be smaller than 9.05×10^{-3} . There are two aspects to this issue: (a) In practice, given realistic failure rates, the value of p is likely to be small enough to meet this bound. (b) Secondly, the duplex system with SEC-DED code uses much fewer bits than 3k. It should be easy to construct a single error correcting-triple error detecting code with n much less than 3k/2. The duplex system using this code would achieve reliability higher than the triplex system for values of p larger than those for the SEC-DED code.

Observe that for small p the slope of the unreliability curves is very large. This implies that, for a given value of p, the difference between the unreliabilities of duplex and triplex systems is significant (even though the curves are very close to each other).

The objective here was to demonstrate that a duplex system with a SEC-DED code

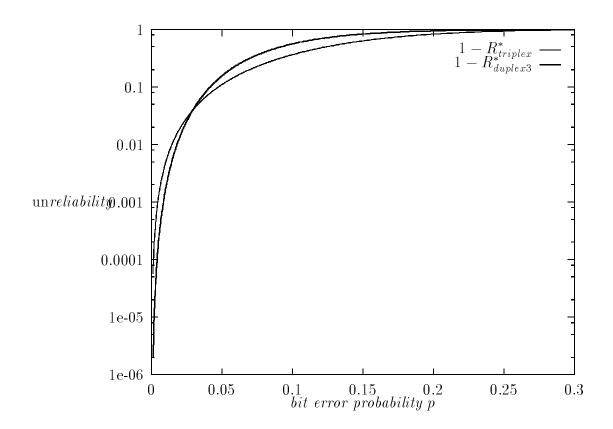


Figure 2: Comparison of $R_{duplex3}^*$ and $R_{triplex}^*$ for k = 16 and n = 22

can achieve reliability higher than the triplex system. We have shown this to be true provided the error probability p is small enough.

3 Summary

This report compares reliability of memory systems formed using simple triplication (without ECC) with memory systems formed by duplicating memory modules that use ECC. Reliability of a system using duplication of memory modules is shown to be always worse than simple triplication if the ECC used in the duplex system is capable of only error detection or only single error correction. It is also shown that if the ECC is capable of single

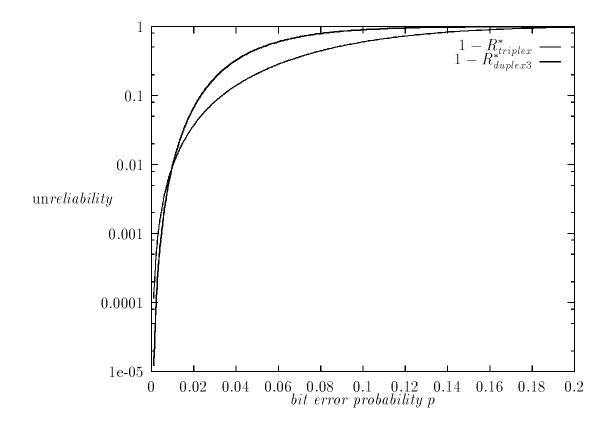


Figure 3: Comparison of $R_{duplex3}^*$ and $R_{triplex}^*$ for k=32 and n=39

error correction as well as double error detection, then the duplex system achieves higher reliability than the triplex system for small values of p.

From the results presented in this report it can be concluded that for the duplex system to be able to achieve higher reliability than the triplex system the ECC must at least be capable of single error correction and double error detection.

Acknowledgements

The author thanks Prasad Padmanabhan for helping with the proof of Lemma 1 in the appendix.

A Appendix

A.1 Proof of Theorem 1

To be able to prove the theorem, we first prove the following lemma.

Lemma 1 $(1-p)^{2k-n}(1+2p)^k + (1-p)^n > 2$ provided $0 , <math>n \ge 3$ and $k+1 \le n \le 3k/2$.

Proof: Assume that $0 , <math>n \ge 3$ and $k + 1 \le n \le 3k/2$. Let function $g(k, n, p) = (1 - p)^{2k-n}(1 + 2p)^k + (1 - p)^n$. Then,

$$\frac{\partial g}{\partial k} = \frac{(1-p)^{2k}(1+2p)^k}{(1-p)^n} \ln[(1-p)^2(1+2p)].$$

As $0 , <math>0 < (1-p)^2(1+2p) < 1$, and it follows that $\frac{\partial g}{\partial k} < 0$. Thus, g is a monotonically decreasing function of k. Therefore, we will choose the largest possible value of k, i.e., k = n - 1, and show that g is larger than 2 for this value of k. When k = n - 1, we have $g(n-1,n,p) = (1-p)^{n-2}(1+2p)^{n-1} + (1-p)^n = [(1-p)(1+2p)]^{n-2}(1+2p) + (1-p)^n$. Let function f(n,p) = g(n-1,n,p). Thus, our goal now is to prove that f(n,p) > 2. Now,

$$\frac{\partial f}{\partial n} = (1-p)^{n-2}(1+2p)^{n-1}\ln[(1-p)(1+2p)] + (1-p)^n\ln(1-p).$$

To find the extrema of f, we set $\frac{\partial f}{\partial n} = 0$. This implies that, at the extrema (i.e. minima or maxima),

$$(1+2p)^{n-1} = -(1-p)^2 \frac{\ln(1-p)}{\ln[(1-p)(1+2p)]}.$$

This equation can hold for only one real value of n. Thus, there exists only one extrema of f with respect to n. Looking at f(n,p) it is clear that by increasing n, f(n,p) can be made arbitrarily large.¹ Therefore, the above extrema must be a minima. Let the minima occur at $n = n^*$. Two cases can occur: (a) $n^* > 3$ and (b) $n^* \le 3$. We consider the two cases separately.

Observe that, for 0 , <math>(1-p)(1+2p) = 1 + p(1-2p) > 1.

Case (a) $n^* > 3$: In this case, we have

$$(1+2p)^{n^*-1} = -(1-p)^2 \frac{\ln(1-p)}{\ln[(1-p)(1+2p)]}$$
(3)

and $f(n,p) \ge f(n^*,p)$. Our goal now is to prove that $f(n^*,p) > 2$. From Equation 3, we get

$$(1-p)^{n^*} = -(1-p)^{n^*-2}(1+2p)^{n^*-1}\frac{\ln[(1-p)(1+2p)]}{\ln(1-p)}.$$

Substituting this expression into $f(n^*, p)$, we get

$$f(n^*, p) = (1 - p)^{n^* - 2} (1 + 2p)^{n^* - 1} - (1 - p)^{n^* - 2} (1 + 2p)^{n^* - 1} \frac{\ln[(1 - p)(1 + 2p)]}{\ln(1 - p)}$$

$$= (1 - p)^{n^* - 2} (1 + 2p)^{n^* - 1} \frac{\ln(1 + 2p)}{-\ln(1 - p)}$$

$$= [(1 - p)(1 + 2p)]^{n^* - 2} (1 + 2p) \frac{\ln(1 + 2p)}{-\ln(1 - p)}$$

$$> (1 + 2p) \frac{\ln(1 + 2p)}{-\ln(1 - p)}, \quad \text{because } n^* > 3 \text{ and } (1 - p)(1 + 2p) > 1 \text{ for } 0$$

Define function $h(p) = (1+2p)\ln(1+2p) + 2\ln(1-p)$. h(p) > 0 implies that $(1+2p)\frac{\ln(1+2p)}{-\ln(1-p)} > 2$ which in turn (by the above inequality) implies that $f(n^*, p) > 2$. Therefore, our goal now is to prove that h(p) > 0.

Note that h(0) = 0 and h(1/2) = 0. Also function h is differentiable in [0, 1/2]. Therefore, by Rolle's theorem [2], at least one extrema (maxima or minima) exists between p = 0 and p = 1/2. Now, $\frac{dh}{dp} = 2 + 2\ln(1+2p) - 2/(1-p)$ and $\frac{d^2h}{dp^2} = \frac{4}{1+2p} - \frac{2}{(1-p)^2}$. Note that for p = 0, $\frac{dh}{dp} = 0$ and $\frac{d^2h}{dp^2} > 0$. Thus, h has a minima p = 0 and at least one maxima in [0,1/2] (by Rolle's theorem). Let the maxima occur at p_{max} . As $p = p_{max}$ is a maxima, $\frac{d^2h}{dp^2}$ must be negative at p_{max} . $\frac{d^2h}{dp^2}$ is a decreasing function of p, therefore, it will remain negative for $p > p_{max}$. This implies that in the interval $(p_{max}, 1/2)$, no minima exists. This

in turn implies that between 0 and 1/2, there exists only one maxima and no minima. As h(0) = h(1/2) = 0, it implies that h(p) > 0 for $0 . This implies that <math>f(n^*, p) > 2$. Now,

$$2 < f(n^*, p) \le f(n, p) \le g(k, n, p).$$

Therefore, $g(k, n, p) = (1 - p)^{2k-n} (1 + 2p)^k + (1 - p)^n > 2$.

Case (b) $n^* \leq 3$: Observe that the range of interest for parameter n is $n \geq 3$. If n^* is no larger than 3, then in the range of interest, function f(n,p) will be minimized at n=3, i.e. $f(n,p) \geq f(3,p)$. Therefore, our goal in this case is to prove that f(3,p) > 2. Now,

$$f(3,p) = (1-p)(1+2p)^{2} + (1-p)^{3}$$

$$= 2+3p^{2}-5p^{3} = 2+p^{2}(3-5p)$$

$$> 2 \quad \text{because } 0$$

This implies that f(n,p) > 2. As $f(n,p) \le g(k,n,p)$, we have $g(k,n,p) = (1-p)^{2k-n}(1+2p)^k + (1-p)^n > 2$.

Theorem 1 can be proved now. If n=k, then it is clear that all errors in a codeword will be undetected. In other words, P_u in Equation 2 is $1-(1-p)^n$. It can easily seen that, in this case, $R_{triplex} > R_{duplex1}$. Now we assume that $n \ge k+1$ or $k \le n-1$. Also, to ensure that the duplex system does not use more bits than the triplex system, we impose the constraint that $2n \le 3k$ (equivalently, $k \ge 2n/3$ or $n \le 3k/2$). When k = 1, n must be at least 2, and this case contradicts the condition $3k \le 2n$. Therefore, in the following, we assume $k \ge 2$ which in turn implies that $n \ge 3$. To summarize, we have $n \ge 3$, $k \ge 2$, $k+1 \le n \le 3k/2$ and 0 . Under these conditions, the result proved in Lemma 1 is applicable. Therefore, we have

$$(1-p)^{2k-n}(1+2p)^k + (1-p)^n > 2$$

$$\Rightarrow (1-p)^{2k}(1+2p)^k + (1-p)^{2n} > 2(1-p)^n$$

$$\Rightarrow (1-p)^{2k}(1+2p)^k + (1-p)^{2n} > (2-P_u)(1-p)^n, \text{ because } P_u \ge 0$$

$$\Rightarrow (1-p)^{2k}(1+2p)^k > (2-P_u)(1-p)^n - (1-p)^{2n}$$

$$\Rightarrow R^*_{triplex} > R^*_{duplex1} \text{ by Equations 1 and 2}$$

$$\Rightarrow R_{triplex} > R_{duplex1}$$

A.2 Proof of Theorem 2

The number of checkbits in the (n, k) code is r = n - k. It is not possible to design a single error correcting code with just one checkbit. Therefore $r \geq 2$. Also, it is not possible to design a single error correcting code for k > 1 with r = 2. When k = 1, the triplex system essentially implements a single error correcting code using a total of 3 bits. Therefore $R_{duplex2}$ with k = 1 and r = 2 is identical to $R_{triplex}$ with k = 1.

For $k \leq 4$, r may be equal to 3. For k > 4, r must be at least 4 for any single error correcting code. We consider the case of $r \geq 4$ first followed by r = 3.

Case 1: $r \ge 4$, 0 : To prove the theorem, we first derive three inequalities.

$$(1-p)^{k}(1+2p)^{k} = [1+p(1-2p)]^{k}$$

$$= \sum_{i=0}^{k} {k \choose i} p^{i} (1-2p)^{i}$$

$$\Rightarrow (1-p)^{k} (1+2p)^{k} \ge 1 + kp(1-2p)$$
(4)

$$1 = (p + (1 - p))^{r} \quad \text{where } r = n - k$$

$$= \sum_{i=0}^{r} {r \choose i} p^{i} (1 - p)^{r-i}$$

$$\Rightarrow 1 > (1 - p)^{r} + rp(1 - p)^{r-1} \quad \text{as } r \ge 4$$
(5)

When $0 , <math>1 - 2p > (1 - p)^3$. Also, $(1 - p)^3 > (1 - p)^i$ for $i \ge 4$. Therefore, for $r \ge 4$,

$$1 - 2p > (1 - p)^{r - 1}.$$

This implies that

$$kp(1-2p) > kp(1-p)^{r-1}$$
 (6)

By replacing the two terms on right hand side of Equation 4 by right hand sides of Equations 5 and 6, respectively, we get

$$(1-p)^k (1+2p)^k > (1-p)^r + rp(1-p)^{r-1} + kp(1-p)^{r-1}$$

$$\Rightarrow (1-p)^k (1+2p)^k > (1-p)^r + np(1-p)^{r-1} \quad \text{as } n = k+r$$

Multiplying both sides by $(1-p)^k$ and replacing n=k+r, we get

$$(1-p)^{2k}(1+2p)^k > (1-p)^n + np(1-p)^{n-1}$$

$$\Rightarrow R^*_{triplex} > R^*_{duplex2}$$

$$\Rightarrow R_{triplex} > R_{duplex2}$$

Case 2: r = 3, 0 : As discussed earlier, <math>r = 3 implies that k can at most be 4. If k = 1, then n = k + r = 4, which is larger than 3k = 3. Therefore, we consider only k = 2, 3, 4. We consider each value of k separately. Note that

$$\frac{R_{duplex2}^*}{R_{triplex}^*} = \frac{(1-p)^n + np(1-p)^{n-1}}{(1-p)^{2k}(1+2p)^k} = \frac{(1-p)^r + np(1-p)^{r-1}}{(1-p)^k(1+2p)^k}$$

(i) r = 3, k = 2: In this case, n = 5 and

$$\frac{R_{duplex2}^*}{R_{triplex}^*} = \frac{(1-p)^3 + 5p(1-p)^2}{(1-p)^2(1+2p)^2} = \frac{1+4p}{1+4p+4p^2} < 1.$$

Therefore, $R_{duplex2}^* < R_{triplex}^*$.

(ii) $r=3,\ k=3$: In this case, n=6. By following similar steps as above, we get

$$\frac{R_{duplex2}^*}{R_{triplex}^*} = \frac{1+5p}{(1-p)(1+2p)^3} = \frac{1+5p}{(1-p)(1+6p+12p^2+8p^3)}$$

$$= \frac{1+5p}{1+5p+6p^2-4p^3-8p^4} = \frac{1+5p}{1+5p+2p^2(3-2p-4p^2)}$$
< 1, because $3-2p-4p^2 > 0$ when $0 .$

Therefore, $R_{duplex2}^* < R_{triplex}^*$.

(iii) r=3, k=4: In this case, n=7. By following similar steps as above, we get

$$\frac{R_{duplex2}^*}{R_{triplex}^*} = \frac{1+6p}{(1-p)^2(1+2p)^4} = \frac{1+6p}{1+6p+9p^2-8p^3-24p^4+16p^6}$$

$$= \frac{1+6p}{1+6p+p^2(9-8p-24p^2)+16p^6}$$
< 1, because $9-8p-24p^2>0$ when $0< p<1/3$.

Therefore, $R_{duplex2}^* < R_{triplex}^*$.

Thus, in case 2, $R_{duplex2}^* < R_{triplex}^*$. This implies that $R_{duplex2} < R_{triplex}$.

References

- [1] P. A. Bernstein, "Sequoia: A fault-tolerant tightly coupled multiprocessor for transaction processing," *Computer*, pp. 37–45, February 1988.
- [2] W. E. Boyce and R. C. DiPrima, Calculus. John Wiley & Sons, Inc., 1988.
- [3] T. R. N. Rao and E. Fujiwara, Error-Control Coding for Computer Systems. Prentice-Hall, 1989.
- [4] N. H. Vaidya and D. K. Pradhan, "Fault-tolerant design strategies for high reliability and safety," *IEEE Trans. Computers*, vol. 42, pp. 1195–1206, October 1993.