TÍTOL DEL TFC: Light coding and protocols for network-on-chip systems of low energy consumption

TITULACIÓ: Enginyeria Tècnica de Telecomunicació, especialitat Telemàtica

AUTOR: Pedro Antonio Vicente Núñez

DIRECTOR: Joan Bas

DATA: 13 de juny de 2014
Overview

Communication systems are constantly evolving and need to process increasingly larger amounts of data in a short time. It is in this field that Network-on-Chip architectures emerge as a viable model to support such traffic demands. However, integration levels in Networks-on-Chip (NoCs) are expected to increase significantly in the next years, accompanied with a reduction in supplied power. This forces NoC technologies to evolve and attempt to find solutions to the problems that are about to arise.

This thesis studies the effects of channel noise on NoC communications and aims to compare the performance of coding methods (with different error-protection levels) to that of plain, uncoded transmissions. This way, it is possible to determine the minimum noise levels that would make it preferable to not use coding methods on the messages and thus avoid the throughput drop they entail.

Furthermore, this thesis proposes the use of an alternative to the classic NoC: the Wireless Network-on-Chip (WNoC), which provides several advantages over its wired counterpart and has a wide array of possible improvements, also explained and analyzed in this thesis.

Accompanying the analysis is a series of simulations, achieved through the MATLAB platform, which aim to provide proof of the proposed solutions' viability.
Título: Codificación leve y protocolos para sistemas network-on-chip de bajo consumo energético

Autor: Pedro Antonio Vicente Núñez

Director: Joan Bas
Tutor: Luis Alonso Zárate

Data: 13 de junio de 2014

Resumen

Los sistemas de comunicación evolucionan constantemente y necesitan procesar cantidades de datos cada vez más elevadas en un corto periodo de tiempo. Es en este aspecto que las arquitecturas de Network-on-Chip emergen como un modelo viable para soportar estas demandas de tráfico. Sin embargo, está previsto que los niveles de integración en las networks-on-chip van a experimentar un aumento considerable en los próximos años, acompañado con una reducción en la potencia suministrada. Esto fuerza a las tecnologías de las NoC a evolucionar y a tratar de encontrar soluciones a los problemas que están surgiendo.

Esta tesis estudia los efectos de ruido de canal sobre las comunicaciones en NoC y tiene como objetivo comparar el rendimiento de métodos de codificación (con diferentes niveles de protección de errores) a los de transmisiones simples, sin código. De esta manera, es posible encontrar niveles mínimos de ruido que harían preferible no usar métodos de codificación en los mensajes y por consiguiente evitar la pérdida de throughput que éstos suponen.

Además, esta tesis propone el uso de una alternativa a las NoC clásicas: la Wireless Network-on-Chip (WNoC), que proporcionan diversas ventajas respecto a su contrapartida con cables y que dispone de una gran cantidad de posibles mejoras, también explicadas y analizadas en esta tesis.

Acompañando a los análisis hay una serie de simulaciones, creadas a partir de la plataforma MATLAB, que tienen como objetivo presentar pruebas de la viabilidad de las soluciones propuestas.
LIST OF ACRONYMS

NoC: Network-on-Chip
BCH: Bose-Chaudhuri-Hocquenghem
SNR: Signal to Noise Ratio
IP: Intellectual Property
NI: Network Interface
ECC: Error-Control Coding
ARQ: Automatic Repetition Request
HARQ: Hybrid Automatic Repetition Request
GF: Galois Field
BPSK: Binary Phase-Shift Keying
PSK: Phase-Shift Keying
BER: Bit Error Rate
WER: Word Error Rate
FET: Field-Effect Transistors
FLIT: Flow Control Unit
WNoC: Wireless Network-on-Chip
SIC: Serial Interference Cancellation
SA: Slotted-ALOHA
DSA: Diversity Slotted-ALOHA
XOR: Exclusive OR
IC: Integrated Circuits
1 Introduction

1.1 Motivation

Future communication systems such as the novel 5G or applications based on big data will need to process large amounts of data in a short time. In these scenarios, Network-on-Chip (NoC) architectures emerge as a viable model to support such traffic demands.

However, in the next years, the integration level of NoCs will increase by about 63%, whereas the supplied power will experience a reduction of about 18%. This evolution translates into a drop in both the reliability of the bus lines and NoCs’ throughput, due to the smaller separation between lines and the increase of functional block units per NoC. In order to overcome these drawbacks, solutions such as light coding based on Bose-Chaudhuri-Hocquenghem (BCH) codes and/or wireless access strategies with diversity in slot and signal interference cancellers may be used.

Another aspect to consider in WNoCs is the frequency band. In the case of the Teraherzt band, the antennas and communication range are very small and consequently become the perfect candidate for implementation in WNoCs. However, this work is based on the access and encoding schemes of WNoCs instead of the antennas. Therefore, reduced complexity techniques in order to implement the MAC layer of WNoCs will be proposed. In particular, the study of the real needs of encoding data from the latency point of view. In case of using coding, protecting data with BCH codes able to correct one and two incorrect bits since their decoding schemes are very easy. Moreover, the use of an interference canceller will also be presented, although its initial complexity could reduce its validity as a solution for WNoC. However, if the number of cores that transmit simultaneously is small, interference cancellation solutions could be an interesting approach.

This paper aims to study these solutions in both low and high Signal-to-Noise Ratio (SNR) contexts and propose a viable and effective way to implement them.

1.2 Objectives

This thesis has three main objectives:

1) To compare, with the use of simulations, the efficiency of coded and uncoded data at different noise levels in order to find the minimum SNR at which uncoded data transmissions start outperforming coded data transmissions.
2) To simulate a scenario where devices transmit data within a NoC, while testing the impact different elements such as SNR, codification and error correction bits have on the transmissions.

3) To analyze the efficacy of the ALOHA access scheme in NoCs and to propose tools that can be used to improve it, as well as testing the impact the elements described above have on the scenario.

1.3 Paper structure

This paper consists of six parts. In the first part, the introduction and the objectives of the paper have been presented. The second part describes the scenario to be simulated, explaining the different factors that have to be taken into account and the search for a solution to the problems that arise. The third part centers on access schemes, explaining the reason for their use and proposing the use of ALOHA. Part four collects the results from the simulations run according to the scenarios presented in the previous parts. Finally, the conclusions drawn from these results can be found in part five, followed by an environmental study and a look into future lines of work. To round the paper up, the annex contains the extended, full calculations done in order to obtain the results, as well as the source code used for the MATLAB simulations, making up part six.
2 Networks-on-Chip

2.1 Introduction

2.1.1 The Network-on-Chip paradigm

In order to process large amounts of data in a short period of time it is necessary to resort to parallel solutions. In this aspect, Network-on-Chip architectures emerge as the paradigmatic solution, capable of providing high throughput efficiently. However, NoCs suffers from scalability problems, as the latency that comes from communicating distant nodes increases with the number of cores. Moreover, in the timeframe of the next ten years, the International Technology Roadmap for Semiconductors (ITRS) has anticipated that the same FETs that are implemented in 24nm technology in 2015 will be implemented in 7.5nm in 2025. Moreover, in the same time-period, the supplied voltage of the FETs will be lowered from 0.83V to 0.68V. Consequently, it means the area of FET implementation will be reduced by about 63% whereas the supplied power will diminish around 18%. These facts will increase the cross-talk effect and decrease the throughput of the NoC. As a result, the NoC paradigm has to be revisited in order to support the target integration level of the next years.

![NoC structure](image_url)
These changes make more apparent than ever the need to optimize networks-on-chip. Ensuring the success of the transmissions while maintaining the supplied power as low as possible within the limits of the Signal-to-Noise Ratio of the channel becomes an enticing idea. This section focuses on analyzing these concepts and proposing solutions that allow us to optimize these networks in channels with a wide variety of noise levels.

### 2.1.2 Simulation tools

The tool used to perform all simulations in this thesis is the free software MATLAB. MATLAB a computing environment developed by MathWorks that allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other languages. All figures in the results section come from MATLAB’s plotting function.

### 2.2 Simulated scenario

We assume the traditional structure of a communication system is used, as featured in the following diagram:

![Diagram](image.png)

**Fig. 2.2 Classic communication model**

Once the original message is encoded, it passes through a modulator before being sent through the channel (which applies the noise effects on the message). Finally, the data is received, demodulated and decoded. In the simulations, once the whole process has been completed, the resulting message is compared to the original to check for errors in the transmission.

When a message is injected into the network, it is first segmented into packets, which are then divided into fixed-length flits, short for flow control units.

Every block’s function has been translated to code in the simulation tool (MATLAB) and the details of each block and the different variables in the scenario are explained in the following subsections.
2.2.1 Encoding

In this thesis, we aim to compare the performance of coded and uncoded data frames when transmissions occur from multiple NoC functional blocks to a single destination.

Encoding is a tool that aims to correct errors that may have happened during the transmission of the message. The code used in this thesis is Bose-Chaudhuri-Hocquenghem (BCH). It is capable of error correction by appending or prepending parity symbols to the original message. Among error-control techniques, it has been proven that Error-Control Coding (ECC), such as the use of BCH, is preferable to the use of other alternatives, such as Automatic Repetition request (ARQ) or Hybrid ARQ (HARQ) when time constraints are tight and link distances are short. In scenarios with low SNR, ARQ and HARQ entail the retransmission of a large amount of messages, which is not desirable.

Let \( m \) be the order of the primitive polynomial of the Galois Field (GF). Let \( n \) be the codeword’s length (original message with the appended parity symbols) and let \( t \) be the number of correctable errors. The original message’s length \( k \) will be:

\[
\begin{align*}
    n &= 2^m - 1 \\
    k &\geq n - t \cdot m
\end{align*}
\]  

For low values of \( t \), the following expression is valid:

\[
    k = n - t \cdot m
\]  

As such, the notation used for a message coded by using BCH, with length \( k \), codeword length \( n \) and error-correction capability \( t \) is \( BCH(n,k,t) \).

The following table shows some valid \([n,k]\) pairs, as well as their corresponding error-correction capability, \( t \):

**Table 2.1.** Some values of \([n,k]\) pairs and \( t \)

<table>
<thead>
<tr>
<th>( n )</th>
<th>( k )</th>
<th>( t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>31</td>
<td>26</td>
<td>1</td>
</tr>
</tbody>
</table>
This thesis focuses on analyzing the effects of encoding when the error-correction capability is 1 and 2 and the message lengths are low. Specifically, the simulations will use a codeword length of 15 as a paradigmatic case, and as a result, as table 2.1 shows, a message length of 11 and 7 for error-correction capabilities of 1 and 2 respectively. This can be noted as \( BCH(15,11,1) \) and \( BCH(15,7,2) \).

### 2.2.2 Modulation

The modulation used in the simulations is Binary Phase-Shift Keying (BPSK), the simplest and most robust form of PSK, and provides the highest resistance to noise. It is only capable of modulating 1 bit/symbol, which is the required transmission mode in order to transfer data using the buses.

![Constellation diagram for BPSK](\text{Fig. 2.3 Constellation diagram for BPSK})

### 2.2.3 Noise and error probability
It is safe to assume the noise in the channel is Gaussian noise. This is due to the fact that channel noise is a combination of effects from uncorrelated sources, the sum of which can be accurately approximated as statistical noise with a Gaussian (or normal) distribution of average voltage $V_N$ and variance $\sigma_N^2$. The error probability $\varepsilon$ (also called Bit Error Rate (BER)) in a flit can then be denoted as:

$$\varepsilon = Q\left(\frac{V_{dd}}{2\sigma_N}\right)$$  \hspace{1cm} (2.3)$$

Where $Q(\cdot)$ is the tail probability of the standard normal distribution and is equal to:

$$Q(x) = \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$$  \hspace{1cm} (2.4)$$

Applying binomial distributions, the error probability of coded data (also called Word Error Rate (WER)) is, assuming we use $BCH(k,n,t)$:

$$\rho_{CD} = 1 - \sum_{q=0}^{t} \binom{n}{q} \varepsilon^q (1 - \varepsilon)^{n-q}$$  \hspace{1cm} (2.5)$$

In the case of uncoded data, as there is no error correction, a much simpler equation equation can be derived:

$$\rho_{UD} = 1 - (1 - \varepsilon)^n$$  \hspace{1cm} (2.6)$$

### 2.2.4 Signal-to-Noise Ratio

The following figure shows the different values of energy that appear throughout the classic communications model of figure 2.2:
Fig. 2.4 Energies corresponding to every segment

Since energy is not lost or gained within the system, the total energy in every step has to remain consistent. This means the total symbol energy \( E_s \) after the modulator should be the same as the total encoded bit energy before the modulator \( E_{enc} \) and the same as the total information bit energy \( E_{b \inf} \) before the encoder.

However, in this thesis we contemplate two scenarios with differing assumptions towards the energy distribution for coded data:

1) Considering the energy per bit is the same when dealing with coded and uncoded data:

\[
\begin{align*}
E_{enc} &= E_{b \inf} \\
E_s &= m \cdot E_{enc} \\
E_s &= m \cdot E_{b \inf}
\end{align*}
\]

Since the modulation used is BPSK, every symbol represents one bit and therefore the value of \( m \) is \( m = 1 \).

\[
E_s = E_{enc} = E_{b \inf}
\]

This means the SNR value is the same in the cases of coded data \( SNR_{CD} \) and uncoded data \( SNR_{UD} \).

\[
SNR = \frac{E_s}{N_0}
\]

\[
SNR_{CD} = SNR_{UD} = \frac{E_s}{N_0} = \frac{E_{b \inf}}{N_0}
\] (2.7)
Taking into account code rate. This refers to the fact that, despite having a codeword length of \( n \), coded words’ useful bits length is actually \( k \). This can be expressed the following way:

\[
\begin{align*}
n \cdot E_{b \text{ enc}} &= k \cdot E_{b \text{ inf}} \\
E_{b \text{ enc}} &= \frac{k}{n} E_{b \text{ inf}} \\
E_s &= m \cdot E_{b \text{ enc}} \\
E_s &= \frac{k}{n} \cdot m \cdot E_{b \text{ inf}}
\end{align*}
\]

Since the modulation used is BPSK, every symbol represents one bit and therefore the value of \( m \) is \( m = 1 \).

\[
E_s = \frac{k}{n} E_{b \text{ inf}} \tag{2.8}
\]

This means the SNR value is different in the cases of coded and uncoded data, as is the bit error rate.

The general formulas are:

\[
\begin{align*}
SNR &= \frac{E_s}{N_0} \tag{2.9} \\
\varepsilon &= Q(\sqrt{SNR}) \tag{2.10}
\end{align*}
\]

In the case of uncoded data, it can be noted as:

\[
\begin{align*}
SNR_{UD} &= \frac{E_{b \text{ inf}}}{N_0} \tag{2.11} \\
\varepsilon_{UD} &= Q(\sqrt{SNR_{UD}}) \tag{2.12}
\end{align*}
\]

And in the case of coded data, it can be noted as:
2.2.5 Voltage swing

Voltage swing refers to the supplied voltage at the input of the signal. This value can generally be lowered or increased depending on the needs of the channel (a channel with high noise levels would require a higher voltage swing so that the power of the signal stays above that of the noise).

The following expression determines the voltage swing value:

\[
\frac{E_s}{N_0} = \frac{v_s^2}{2\sigma_N^2}
\]

\[
v_s = \sqrt{\frac{2\sigma_N}{E_s/N_0}}
\]

(2.15)

Where \(v_s\) stands for voltage swing and \(\sigma_N\) for variance.

2.2.6 Number of transmissions

The number of transmissions \(N_{tx}\) refers to the number of times, on average, a message has to be sent to guarantee it is successfully received. If we consider the model of selective repetition quest, the mathematical expression is directly tied to the WER or \(\rho\):

\[
N_{tx} = \frac{1}{1 - \rho}
\]

As explained in 2.2.3, the value of \(\rho\) is different in the cases of coded and uncoded data. Therefore, we can obtain the following two expressions:

\[
N_{tx\, UD} = \frac{1}{1 - \rho_{UD}}
\]

(2.16)
\[ N_{tx\,CD} = \frac{1}{1 - \rho_{CD}} \] (2.17)

### 2.3 Analysis

With the expressions obtained in the previous section, it is now possible to test the theoretical performance of coded and uncoded data in order to determine the SNR values at which transmitting without encoding the messages starts outperforming the coded method.

However, it is also important to remember that uncoded data is based on messages with length \( n \), while coded data uses messages with length \( k \). If we let \( L \) be the total length in bits of the transmission, this can be expressed as follows:

\[
\frac{L}{n} N_{tx\,UD} \leq \frac{L}{k} N_{tx\,CD}
\] (2.18)

\[
\frac{N_{tx\,UD}}{N_{tx\,CD}} \leq \frac{n}{k}
\] (2.19)

The moment the requirement above is met, it becomes more efficient to transmit without encoding the messages.

In this section we evaluate the expressions above in the cases of same SNR values and different SNR values as explained in subsection 2.2.4.

An error-correction capability of 1 and 2 is considered in both cases.

#### 2.3.1 Same SNR values

In this section, we assume the SNR values are the same for both coded and uncoded data. This also means the Bit Error Rate \( \varepsilon \) is the same in both cases.

\[
SNR_{CD} = SNR_{UD} = SNR
\]

\[
\varepsilon_{UD} = \varepsilon_{CD} = \varepsilon
\]

The full computations can be found in the attached annex.
2.3.1.1 Error-correction capability \( t = 1 \)

In this subsection we assume BCH’s error-correction capability is 1, noted \( BCH(n,k,1) \) and the values of \( k \) and \( n \) are appropriate (refer to Table 2.1).

\[
\frac{L}{n} N_{tx UD} \leq \frac{L}{k} N_{tx CD}
\]

\[
k(1 - \rho_{CD}) \leq n(1 - \rho_{UD})
\]

\[
\varepsilon \leq \frac{n - k}{n(1 + k) - k} \quad (2.20)
\]

The above expression can accurately be approximated to:

\[
\varepsilon \leq \frac{1}{k} \left( 1 - \frac{1}{n} \right)
\]

2.3.1.2 Error-correction capability \( t = 2 \)

In this subsection we assume BCH’s error-correction capability is 2, noted \( BCH(n,k,2) \) and the values of \( k \) and \( n \) are appropriate (refer to Table 2.1).

\[
\frac{L}{n} N_{tx UD} \leq \frac{L}{k} N_{tx CD}
\]

\[
k(1 - \rho_{CD}) \leq n(1 - \rho_{UD})
\]

\[
\varepsilon^2 \left( n - k + nk - \frac{nk(n-1)}{2} \right) + \varepsilon(-2n + 2k - nk) + n - k \geq 0 \quad (2.21)
\]

2.3.2 Different SNR values

In this section we assume the SNR values are not the same for coded and uncoded data, as explained in section 2.2.4. This also means the Bit Error Rate \( \varepsilon \) is not the same in both cases.

\[
SNR_{CD} = \frac{k}{n} SNR_{UD} \quad (2.22)
\]
\[ \varepsilon_{UD} = Q\left(\sqrt{\text{SNR}_{UD}}\right) \quad (2.23) \]
\[ \varepsilon_{CD} = Q\left(\sqrt{\text{SNR}_{CD}}\right) \quad (2.24) \]

The full computations can also be found in the attached annex.

### 2.3.2.1 Error-correction capability \( t = 1 \)

In this subsection we assume BCH’s error-correction capability is 1, noted \( \text{BCH}(n,k,1) \) and the values of \( k \) and \( n \) are appropriate (refer to table 2.1).

By developing the original equation, it is possible to find the bit error probability corresponding to the point where the ratio goes below the threshold:

\[
\frac{L}{n} N_{tx\,UD} \leq \frac{L}{k} N_{tx\,CD} \\
k(1 - \rho_{CD}) \leq n(1 - \rho_{UD}) \\
\varepsilon_{UD} \leq 1 - (1 - \varepsilon_{CD}) \frac{n}{k} + \frac{k \varepsilon_{CD}}{1 - \varepsilon_{CD}} \quad (2.25)
\]

This expression shows how, for any value of \( \varepsilon_{CD} \) within the probability limits \([0,0.5]\), a value of \( \varepsilon_{UD} \) exists which satisfies equation \((2.18)\).

### 2.3.2.2 Error-correction capability \( t = 2 \)

In this section we assume BCH’s error-correction capability is 2, noted \( \text{BCH}(n,k,2) \) and the values of \( k \) and \( n \) are appropriate (refer to table 2.1).

\[
\frac{L}{n} N_{tx\,UD} \leq \frac{L}{k} N_{tx\,CD} \\
k(1 - \rho_{CD}) \leq n(1 - \rho_{UD}) \\
\varepsilon_{UD} \leq 1 - (1 - \varepsilon_{CD}) \sqrt{n \left(1 + \frac{n(n - 3)}{2}\right)} + \varepsilon_{CD}(n - 2) + 1 \quad (2.26)
\]
This expression shows how, for any value of $\epsilon_{CD}$ within the probability limits [0,0.5], a value of $\epsilon_{UD}$ exists which satisfies equation (2.18).
3 Access Scheme

Predictions of future technological tendencies indicate that FETs (Field-Effect Transistors) will gradually experience a significant reduction in size in the next fifteen years. As chips get smaller, interferences on the bus lines – and therefore, the probability of faulty transmissions – increase. This predicted evolution is so quick that material innovation with traditional scaling will not be able to keep up with future requirements of NoCs infrastructures, making the need for an alternative to the classic wired NoC increasingly apparent.

In order to tackle the foreseen throughput drawbacks on NoCs, a tentative solution is to use a Wireless Network-on-Chip (WNoC). This way, communications between distant cores are carried out in a single hop. Therefore, multi-hop communications are replaced by single hops of large bandwidth. On the other hand, wireless networks-on-chip introduce the inconveniences of interferences and needing to have a good enough quality of the communications channel.

Other solutions to support high throughputs are also possible. For instance, 3D implementations of NoCs could help reduce latency. However, the high level of integration will increase the heat generated in these types of networks. Optical based networks-on-chip are also possible, but integrating the semiconductor domain and optical fiber is not an easy task.

Fig. 3.1 Wireless NoC structure
In WNoCs, some difficulties in implementation arise, such as collisions among functional units attempting to transmit simultaneously. In this aspect, it is possible to increase the throughput by means of using slotted protocols and efficient design of MAC protocols. In particular, this work proposes to use Slotted-ALOHA protocols that use time-diversity to increase throughput. In this approach, the cores act as nodes of a wireless micro-network. The wireless micro-network could be also understood as a Machine to Machine (M2M) communication since cores (machines) communicate without any human support, transmit small amounts of data and require high throughputs.

WNoCs allow one-hop data transfers between distant nodes, resulting in performance benefits: a reduction in interconnect delay and energy dissipation, which improves the scalability of the system (number of cores).

However, even though WNoCs solve the crosstalk problem (interferences between bus lines), it comes at the price of a reduction in reliability in scenarios with a high number of cores. In addition, this alternative requires the use of miniaturized on-chip antennas as an enabling technology.

It is possible to mitigate the reduction in reliability by using the appropriate access scheme and optimizing it with new techniques such as diversity and interference cancellers in order to obtain a system that is significantly better than the original wired NoCs.

One important advantage that WNoCs have over other wireless systems is that the precise typology and the number of transmitting devices (i.e., nodes) are known and therefore it becomes significantly easier to optimize the network. This is especially relevant in the application of a diversity scheme, as explained in section 3.3.

The following sections explain why ALOHA was chosen as the access scheme for WNoCs and present a series of improvements that can be implemented in order to greatly increase its performance in networks-on-chips.

### 3.1 ALOHA

ALOHA was one of the first protocols designed for wireless packet data networks and remains today one of the simplest. Despite its simplicity, it can be greatly effective if combined with the appropriate improvements, making it one of the most effective protocols for use in WNoCs.

The first version of the protocol, pure ALOHA, established the following process should take place when a packet is transmitted:

- Send data as soon as there is data to send without checking if the channel is busy (there is no “listening” step).
• If a collision occurred, wait a random time (backoff) and try sending the message again.

The throughput in pure ALOHA is:

\[ S_{\text{pure}} = Ge^{-2G} \]  

(3.1)

Where G corresponds to the load and is defined as:

\[ G = \frac{\text{stations}}{\text{slots}} = \frac{m}{n} \]  

(3.2)

![Fig. 3.2 Throughput in Pure ALOHA](image)

The maximum throughput obtainable occurs when \( G = 0.5 \) and is 0.184, which means only 18.4% of the time is used for successful transmissions.

### 3.2 Slotted-ALOHA

Slotted-ALOHA improves the original pure ALOHA by, as the name implies, introducing discrete timeslots. This means that time gets split in a sequence of
Access Scheme

intervals, all equal in duration, called slots. In Slotted-ALOHA, a station is only allowed to transmit at the beginning of a slot.

This results in the elimination of partially overlapping packets and thus permits an improvement in throughput and better synchronization. Specifically, the maximum throughput is doubled.

The new throughput in Slotted-ALOHA is:

\[ S_{\text{slotted}} = Ge^{-G} \quad (3.3) \]

Resulting in a maximum throughput of 0.368, reached when \( G = 1 \), meaning 36.8% of the time is used for successful transmissions.

![Fig. 3.4 Throughput in Slotted-ALOHA](image)

### 3.3 Diversity Slotted-ALOHA

Diversity Slotted-ALOHA is an enhancement of Slotted-ALOHA. Time-diversity refers to the concept of a station sending copies of the original message within the same frame in order to increase the probability that it will be received successfully.
The number of the message’s copies sent in addition to the original is denoted as \( k \) and the new value of the load \( G \) is:

\[
G = \frac{m}{n} (k + 1)
\]

\[(3.4)\]

**Fig. 3.5** Example of Slotted-ALOHA with diversity

**Figure 3.5** features an example of ALOHA with diversity in action. There are four devices attempting to transmit data over the channel. It’s important to note how every device sends four copies of their respective original messages, i.e. instead of sending just one message, three identical extras are also transmitted. For instance, Device 1 sends messages in slots 2, 3, 4 and 6. This means diversity is being used and the value of \( k = 3 \), since one message plus three extra copies are being sent.

In Frame 1, all devices present collisions in some of the slots. However, Device 4 is the only one to transmit in slot 1, and Device 1 is the only one to transmit in slot 4. This means those two messages did not collide with any other and therefore were received successfully.

Only Device 2 and Device 3 remain, which leads them to try retransmitting their messages in Frame 2. Unfortunately, both devices end up transmitting in the exact same slots, leading to another failure.

Finally, in Frame 3, Device 2 manages to avoid a collision in slots 1 and 3, while Device 3 does the same in slots 3 and 4. Both messages are therefore successfully received.

In Diversity Slotted-ALOHA, the throughput is defined as:

\[
T = \frac{L}{N_T}
\]

\[(3.5)\]
Where \( L \) is the number of transmitted bits and \( N_T \) is the number of frames required in order to successfully transmit \( L \) bits. More specifically, if there are \( n \) devices that transmit at the same time, the number of transmitted frames is:

\[
n = \sum_{q=1}^{N_T} n_{dq}
\]  

(3.6)

Where \( n_{dq} \) is the number of detected contending devices in each of the \( N_T \) transmissions and is defined as:

\[
n_{dj} = \left( n - \sum_{p=1}^{j-1} n_{dp} \right) p_{dj}
\]  

(3.7)

Where \( p_{dj} \) is the detection probability in the \( j \)-th retransmission. This probability represents the number of devices that are different detected at the destination and is defined as:

\[
p_{dj} = \sum_{r=0}^{n_j} \sum_{s=q}^{n-1} p(N_{su} = r | N_s = s) p(N_s = s)
\]  

(3.8)

Where \( p(N_{su} = r | N_s = s) \) is the probability the number of successful transmitting cores is equal to \( r \) conditioned to have \( s \) successful slots (we are transmitting in a slotted protocol). Finally, \( p(N_s = s) \) is the probability of having \( j \) successful slots and has been computed according to the next binomial distribution:

\[
p(N_s = s) = \binom{m}{s} p_s^s (1 - p_s)^{m-s}
\]  

(3.9)

Where \( p_s \) is the probability a transmission is successful. This probability has been considered as a Poisson Distribution since it is assumed that the number of cores that transmit at the same time could be high enough. This assumption can be reasonable in systems such as data centers or high performance computers since the number of cores is high (>>100). According to this approach the value of \( p_s \) is:

\[
p_s = Ge^{-G}
\]  

(3.10)
3.4 Diversity Slotted-ALOHA with SIC

Serial Interference Cancellation is a tool that makes it possible, in certain situations, to decipher messages even when an overlap occurs. In essence, it turns some collisions into successful transmissions.

For instance, in a scenario where the messages from two users overlap in the same timeslot but one of the packets is obtainable from a different slot, a simple XOR operation suffices to obtain the second message.

![Diagram](image_url)

**Fig. 3.6 Example of Slotted-ALOHA with diversity and SIC**

**Figure 3.6** features an example of ALOHA with diversity and SIC in action. In the first frame, Device 1 and Device 5 transmit in the same slot, resulting in a collision. The receiver is only able to see the sum of the transmissions, i.e., the result of the exclusive OR (XOR) operation between both original messages. However, the receiver possesses the information of which slots each device transmitted in. Moreover, it can retrieve the contents of message 1 from slot 4, since Device 1 is the only one to transmit in that slot.

This is where Serial Interference Cancellation comes into play. By applying the XOR operation between the bits received in slot 1 and message 1, the receiver is now also capable of resolving message 5, since the end result is as if message 1 had never been in that slot.

In Frame 2, the same process takes place. Device 4 is the only one to transmit in slot 6 and therefore it is successfully received. Moreover, the receiver realizes Device 4 and Device 6 are the only ones that transmitted in slot 3, and
since it now “knows” the contents of message 4, SIC comes into play to resolve message 6.

Finally, in Frame 3, there is no need to use SIC as Device 2 and Device 3 are able to send their messages without collisions occurring in slots 4 and 5 respectively.

As for throughput, after receiving a set of samples the serial interference canceller detects the cores that have transmitted in an iterative way. In particular in each iteration of the interference canceller one core is detected. So, it means that each iteration of the interference canceller can be understood as a retransmission with diversity in which only a single core is decoded. The process stops until the detection probability is lower than the collision probability of the iterative algorithm. In particular, the collision probability for a Poisson distribution is:

\[ p_c = 1 - e^{-G}(1 + G) \]  \hspace{1cm} (3.11)

Where G, in this case, corresponds to the load of a slotted-aloha system using diversity is slot.
4 Results

This section aims to provide results to the objectives stated in sections 2 and 3.

The first subsection presents the results of the simulations carried in order to compare the performance of coded and uncoded data and to determine the SNR values at which transmitting without encoding the messages starts outperforming the coded method. As explained in previous sections, this includes the cases where SNR values either differ or are the same for coded and uncoded data. Error-correction capabilities of 1 and 2 are considered.

The second subsection focuses on the performance of the ALOHA access scheme and aims to present proof of the improvements the different implementations discussed in section 3 provide.

Note: the near-100% error probability caused by very low SNRs (especially SNR values below -5dB), causes the results in that interval to be more unreliable.

4.1 Analyzing NoC transmissions at different SNRs

It is possible to simulate the initial assumption using theoretical values in order to find the theoretical point we are looking for.

To accomplish this, it suffices to represent equation (2.19) in a given SNR interval.

The following subsections provide first the results of a test carried out with theoretical values before presenting the simulations’ results of real transmissions.

The SNR interval considered is [-10dB, 10dB] and the values of \( N_{tx UD} \) and \( N_{tx CD} \) (and subsequently \( \rho_{CD} \) (2.5) and \( \rho_{UD} \) (2.6)) used are (2.16) and (2.17) respectively.

The following equations will also be used:

\[
(ratio) \ r = \frac{N_{tx UD}}{N_{tx CD}} \tag{4.1}
\]

\[
(threshold) \ th = \frac{n}{k} \tag{4.2}
\]
4.1.1 Same SNR

As explained in sections 2.2.4 and 2.3.1, the results provided in this section assume the SNR value is the same in the cases of coded and uncoded data.

4.1.1.1 Error-correction capability $t = 1$

Using theoretical values, it is possible to obtain the graphical representation of equation (2.19):

![Graph showing minimum SNR for uncoded-coded data when $t = 1$](image)

**Fig. 4.1** Minimum SNR for uncoded-coded data when $t = 1$

Where the red curve corresponds to the ratio (4.1) and the blue line corresponds to the threshold (4.2).

The results would indicate the ratio $r$ has a value below the threshold when the SNR value is $2.9\, dB$ or higher.

This means it theoretically becomes more efficient to transmit uncoded messages than coded messages when the SNR in the channel is of value $2.9\, dB$ or higher.
Once the theoretical analysis has been completed, it is time to check if the simulations' results match.

As figure 4.2 shows, the results of the simulation match the theoretical ones within a margin of error.

4.1.1.2 Error-correction capability $t = 2$

Using theoretical values, it is possible to obtain the graphical representation of equation (2.19):
Fig. 4.3 Minimum SNR for uncoded>coded data when $t = 2$

Where the red curve corresponds to the ratio (4.1) and the blue line corresponds to the threshold (4.2).

The results would indicate the ratio $r$ has a value below the threshold when the SNR value is $1.2dB$ or higher.

This means it theoretically becomes more efficient to transmit uncoded messages than coded messages when the SNR in the channel is of value $1.2dB$ or higher.

Once the theoretical analysis has been completed, it is time to check if the simulations' results match.
As figure 4.4 shows, the results of the simulation match the theoretical ones within a margin of error.

### 4.1.2 Different SNR

As explained in sections 2.2.4 and 2.3.2, the results provided in this section assume the SNR value is different in the cases of coded and uncoded data.

#### 4.1.2.1 Error-correction capability \( t = 1 \)

Using theoretical values, it is possible to obtain the graphical representation of equation (2.19):
Fig. 4.5 Minimum SNR for uncoded-coded data when $t = 1$

Where the red curve corresponds to the ratio (4.1) and the blue line corresponds to the threshold (4.2).

The results would indicate the ratio $r$ has a value below the threshold when the SNR value is $1.8dB$ or higher.

This means it theoretically becomes more efficient to transmit uncoded messages than coded messages when the SNR in the channel is of value $1.8dB$ or higher.

It is also possible to represent equation (2.25) in a [-10dB, 10dB] SNR interval:
**Fig. 4.6** Graphical representation of the equation when $t = 1$

In *figure 4.6*, the red curve corresponds to $\varepsilon_{UD}$, while the blue curve corresponds to the second half of the equation, i.e.:

$$1 - (1 - \varepsilon_{CD}) \sqrt[n]{\frac{k}{n} + \frac{k\varepsilon_{CD}}{1 - \varepsilon_{CD}}}$$

The figure indicates the threshold is at $1.8\,dB$ and therefore is consistent with the result obtained before.

Once the theoretical analysis has been completed, it is time to check if the simulations’ results match.
As figure 4.7 shows, the results of the simulation match the theoretical ones within a margin of error.

4.1.2.2 Error-correction capability $t = 2$

Using theoretical values, it is possible to obtain the graphical representation of equation (2.19):
Where the red curve corresponds to the ratio (4.1) and the blue line corresponds to the threshold (4.2).

The results would indicate the ratio $r$ has a value below the threshold when the SNR value is 1.8$dB$ or higher.

This means it theoretically becomes more efficient to transmit uncoded messages than coded messages when the SNR in the channel is of value 1.8$dB$ or higher.

Once the theoretical analysis has been completed, it is time to check if the simulations’ results match.
As figure 4.9 shows, the results of the simulation match the theoretical ones within a margin of error.

4.2 Analyzing and optimizing WNoC using ALOHA

Error-correction capabilities of 1 and 2 are considered. SNR values chosen for high and low of 10dB and 0dB respectively.

4.2.1 Slotted-ALOHA
Figure 4.10 features transmissions using the Slotted-ALOHA access scheme, while not using any of the improvements (diversity or SIC), comparing the use of coded and uncoded messages in low and high SNR contexts.

As the number of contending devices increases, collisions become an increasingly bigger problem and detection probability suffers.

It’s important to note how at high SNR, whether coding is being used or not, as well as the error-protection level, do not matter, as transmission encoding has no effect on collided messages.

At low SNR, however, coding matters and the ability to correct errors allows for better detection probabilities.
Figure 4.11 features transmissions using the Slotted-ALOHA access scheme, while not using only SIC. The results are very similar to those of figure 4.10 but the addition of SIC causes an improvement of the detection probability across the board. This is due to the fact that SIC allows the retrieval of some of the messages that collided.

This time too, there is no difference between coded and uncoded data at high SNR as SIC improves both methods equally.

### 4.2.2 Diversity Slotted-ALOHA (+ SIC)

The diversity value used in this section is:

\[
k = \frac{\text{slots}}{\text{devices}}
\]  

(4.3)
Figure 4.12 features the addition of diversity to a plain Slotted-ALOHA scheme. Significant improvements can be noted here, such as the case with no code at low SNR, which even achieves a detection probability close to 100% when the number of contending devices is very low, whereas it could only reach about 30% without diversity.

Similar improvements can be seen in the rest of cases, where the biggest improvements in detection probability occur when the number of devices is low only to quickly dip into similar results to figure 4.10 when approaching a high number of devices.
Fig. 4.13 Slotted-ALOHA with diversity and with SIC

*Figure 4.13* features the addition of SIC to the previous scenario. Just like what happened when SIC was added in *figure 4.11*, the results present improvements across the board, but are even more significant here, especially when there are fewer errors in the messages (such as when the SNR is high or when code is used with an error-correction capability of 2).
5 Conclusions

This thesis attempted to analyze and test the effects of the channel’s noise on the Network-on-Chip communications and to propose circumstances under which it is appropriate to use coding techniques such as BCH on the data transmissions. Ultimately, the point of interest was finding the minimum signal-to-noise ratio that would allow us to avoid encoding the messages, which coincides with the moment the error-protection qualities of coding techniques are no longer worth the lower transmission rates that accompany them.

Table 5.1. Results summary

<table>
<thead>
<tr>
<th></th>
<th>Same SNR</th>
<th>Different SNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>t=1</td>
<td>2.9dB</td>
<td>1.7dB</td>
</tr>
<tr>
<td>t=2</td>
<td>1.2dB</td>
<td>-1dB</td>
</tr>
</tbody>
</table>

As table 5.1 shows, uncoded data starts outperforming coded data when the SNR is about 2.9dB (1.7dB when considering code rate). This means that, as long as the SNR in the channel can be kept at that level or higher, encoding messages would actually result in worse transmission rates.

It can also be noted that using higher error-correcting capabilities in a code such as BCH results in the SNR threshold becoming even lower. Thus, as the error-correction capability of the code is increased, the lower the SNR threshold that makes uncoded data superior becomes. For instance, transmitting without the use of a coding protocol is superior to transmitting encoded messages with an error-correction capability of 2 as long as the SNR is kept above 1.2dB (-1dB when considering code rate).

This thesis also worked on implementing and testing the effects of the ALOHA access scheme in a wireless Network-on-Chip scenario. The improvements proposed (slotted time-frame, time-diversity and serial interference cancellation) for the ALOHA scheme were explained and simulated. And as the results reflect, each implementation, even just independently, allowed for a higher detection probability and therefore better performance, most notably when the SNR in the channel is low (such as SNR=0).

Environmental study

The main objective of the implementations discussed in this work is to improve the efficiency of NoC communications.

The quest for the minimum SNR needed in order to transmit without encoding the messages allows us to minimize power consumption.
The WNoC alternative to the classic wired NoC, as long as the proposed implementation of the access scheme and its improvements, aim to provide a system where less transmissions are needed and thus requires lower power consumption.
6 Future lines of work

There exist several ways one could build up on the foundations this thesis has set up. A future line of work would be, for instance, to analyze the effects of Adaptive Diversity in Slotted-ALOHA schemes. Adaptive Diversity is yet another improvement that can be implemented to Diversity Slotted-Aloha in order to increase the efficiency of the access scheme. The basic idea is to assign a dynamic value to the diversity – that is, the number of copies of the original message sent – instead of a constant. This way, when fewer contending stations are left, the diversity value can be automatically adjusted to achieve greater success and efficiency. This method was not considered in this thesis as it goes beyond the scope of the initial approach and has already been studied in other papers.
7 References