Contactless Heterogeneous 3-D ICs for Smart Sensing Systems

DOI:
10.1016/j.vlsi.2018.04.001

Document Version
Accepted author manuscript

Link to publication record in Manchester Research Explorer

Citation for published version (APA):

Published in:
Integration, the VLSI Journal

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
Contactless Heterogeneous 3-D ICs for Smart Sensing Systems

Ioannis A. Papistas*, Vasilis F. Pavlidis*

Advanced Processor Technologies Group
School of Computer Science, University of Manchester, M13 9PL, Manchester, United Kingdom

Abstract
A heterogeneous contactless transceiver circuit is designed to provide inter-tier signalling for a 3-D system considering specific bonding constraints. The system is composed of two tiers, a 65 nm processing tier and a 0.35 μm sensing tier. Face-to-back integration is chosen to support fluidic sensing. Half duplex communication between the tiers is provided through inductive links. Each tier is considered to be fabricated in a different technology to enable low manufacturing cost and benefit from the advantages each technology offers. Both the uplink and downlink transceivers achieve data rates that reach 1 Gbps with non-return-to-zero data encoding. Energy efficiency is the predominant objective, with the uplink dissipating 5.28 pJ/b and 8.67 pJ/b for the downlink. A 6.8× power reduction is demonstrated when using heterogeneous technologies, compared to a state-of-the-art 0.35 μm transceiver, while the dissipated energy is decreased by 37.5% as compared to a state-of-the-art 65 nm transceiver. Process variation analysis is also performed to ensure the proposed circuit operates correctly across several process corners, covering a broad design space. To improve system robustness, an overhead of 2.3% on the peak power and < 1% on the average power is shown, respectively.

Keywords: 3-D Integration, Heterogeneous Systems, Contactless Communication, Heterogeneous Inductive Links.

1. Introduction

The introduction of smart devices, capable of sensing, processing, and communicating with the physical world leads to the Internet-of-Things (IoT); a vast network of interconnected devices via the internet [1]. The continuous drive for smaller, low power, and high performance electronics has enabled, among others, the development of technologies such as the IoT, by embedding low cost electronics into “naïve” objects. The physical properties of the naïve objects are augmented with smart properties, allowing operation of those objects as digital sources of data and communication with other smart devices without human supervision.

The abstraction layers of an IoT system architecture are depicted in Figure 1 [2]. The perception layer contains the IoT edge or, alternatively, the interface to the physical world. The amount of data collected by a network of edge devices can rapidly reach the terabyte level, improving the resolution of the observed phenomenon and broadening the perception of the observation [3]. To avoid an increase in the network traffic, the functionality of the edge devices can be extended to include real time computing, utilising IoT based operating systems [3]. For example, real time decision making in an industrial environment is improved with the IoT [3]. Nevertheless, sensors or networks of sensors can be deployed where access to a power source or the sensors themselves is limited. Therefore, high autonomy and energy efficiency are key requirements for the IoT. Furthermore, form factor is another major limitation, since the footprint of the device and the accompanying battery should be minimised.

Considering these important issues, heterogeneous 3-D integration provides a promising platform for edge devices of the Internet-of-Things comprising multi-functional, small form factor, and low power systems [3]. A heterogeneous 3-D system is composed of stacked integrated circuits, manufactured in disparate technologies. Digital, analog, and sensors circuits can thus coexist in a single 3-D package reducing the form factor of the system [7]. This approach is suitable for IoT appli-
cations where multi-functionality and power consumption are paramount.

Nevertheless, several challenges exist for developing heterogeneous 3-D systems including heat dissipation, testing, and inter-tier communication [7]. Considering inter-tier interconnects, through silicon vias (TSVs) provide high density signalling with low latency and power [8]. However, TSVs entail an overhead in cost due to the related manufacturing complexity and possibly low yield [5, 6, 17]. For example, to alleviate the impact of copper pumping due to the TSVs, an additional high thermal annealing process is required, increasing the manufacturing cost [6]. Several other reliability issues need to be considered, such as copper diffusion from the TSV to the substrate, mechanical stresses, and electromigration, each requiring additional processing steps. Furthermore, excessive substrate thinning is imperative for state-of-the-art TSV integration, where the TSV has a diameter of 5 μm or smaller. Consequently, a significant processing cost is incurred due to the handling of the thin wafers. Alternatively, contactless solutions based on AC coupling have been proposed [11, 17, 18, 19, 20].

Both inductive and capacitive contactless communication are compatible with conventional CMOS processes and therefore no additional manufacturing steps are required for underpinning inter-tier communication [12, 11, 18]. Wireless interconnects offer several advantages to heterogeneous 3-D integration including die detachability [12] and inter-tier communication without the requirement of level shifters [16]. However, capacitive coupling is limited to face-to-face bonding, significantly narrowing the candidate applications for this communication mechanism. For example, face-to-face integration is not a feasible option for sensing applications (e.g. fluidic sensing), where the sensing tier should be integrated face up. Due to these limitations, inductive links are investigated in this work.

A transceiver circuit is proposed for contactless 3-D systems characterised by both functional and technological heterogeneity. This twofold heterogeneity originates from the nature of the target applications and related cost issues. Lab-on-chip applications including fluidic sensing and characterisation are excellent candidates for these heterogeneous systems. The target system is a two tier system consisting of a processing tier and a sensing tier implemented, respectively, in a 65 nm [17] and a 0.35 μm technology [18]. The choice of the 0.35 μm technology is due to the typical use of this technology for sensing circuits [18, 19]. Older and therefore lower cost technologies are preferred for sensing applications since speed is not a critical requirement. Consequently, significantly different technologies are utilised for the sensor circuits and the digital circuits that control the sensor and process the sensory data.

This work considerably extends the circuit presented in [17]. Emphasis is placed on the contactless inter-tier communication and the issues relating to the dissimilar manufacturing technologies of the tiers. Process corners, mismatch, and voltage variations are considered through back-annotated simulations to verify the accuracy of the presented results. However, the primary contribution of this work is a design methodology for low power inductive links by carefully considering the salient features of each manufacturing technology. Tradeoffs related to the design of the link including the coupling between the inductors, the area and power consumed by the inductors are also explored.

The remainder of this paper is organised as follows. In Section 1, a literature review of state-of-the-art inductive links is provided. In Section 2, the primary characteristics of the envisioned two tier system are introduced. The design of the on-chip inductors and transceiver circuits that provide communication between the tiers are described in Section 3. Simulation results for the corner analysis of the circuits and related design tradeoffs are presented in Section 4. Finally, some conclusions are drawn in Section 5.

2. Previous Work

A review of existing inductive links is presented in this section. Circuits, low power techniques, and typical fabrication processes are presented. Moreover, a scaling scenario of the resulting performance and power dissipation is presented. Several inductive links have been proposed [21, 22, 18, 12, 15], where the majority of these links support data transfer between a processor and a memory module. Homogeneous inductive links (i.e. the same manufacturing technology is utilised for both the transmitter and receiver circuits) have been fabricated in many processes, where the most advanced manufacturing process for a prototype inductive link is a 65 nm technology [17].

The energy per bit and performance of homogeneous inductive links are illustrated in Figure 1. As the supply voltage decreases with technology scaling, the power decreases. This reduction is however bounded by the communication distance X, between the paired inductors, which does not scale at the same rate as the supply voltage. Aggressive thinning of the substrate (e.g. < 50 μm) improves coupling but increases the processing cost. Moreover, the yield is potentially reduced as thinned dies are brittle. For technologies down to 0.18 μm [23], the link performance is fixed at 1 Gbps, since increasing the performance results in an excessive overhead in power consumption. The scaling of the supply voltage has enabled the development of high performance inductive links with low power and data rates reaching 8 Gbps [22].

Low power techniques have been utilised to further reduce power consumption. In [24], assuming multiple inductive link channels, a daisy chain topology allows power reduction by reusing the current in each transmitter. A decrease of 35% is observed by implementing a 4-stage daisy chain transmitter. Additionally, a dual coil scheme is used in [25], where PMOS devices are not required for the transmission of data. Thus, the number of transistors is reduced and, without PMOS devices, low voltage operation is achieved with a Vdd = 0.55 V, achieving a 10 fJ/b energy consumption. These techniques however include more than one on-chip inductor which incurs significant area overhead, being at odds with the small form factor requirement. Note, however, that if the area requirements are relaxed, these techniques can complement the proposed transceivers. Moreover, using additional inductive links is challenging, since the presence of a sensor circuit poses stricter
area constraints compared to contactless systems that aim at memory-processor communication.

Furthermore, systems with functional heterogeneity have recently been explored [21], where homogeneous inductive links within a 3-D network-on-chip interconnect a multicore processor with accelerators on several tiers. Alternatively, an interface for memory control manufactured in a 65 nm technology is proposed in [22] where the channel of the transistors in the memory modules is elongated to emulate a 100 nm node. However, this circuit is inherently homogeneous, utilising only one technology. In [26], despite the one technology node difference between the tiers, only digital systems are interconnected. Additionally, these systems target memory-processor communication where speed is of high importance. However, the nature of sensing and more broadly IoT applications requires a different design approach since power and cost are often more important than speed [27].

Alternative approaches to CMOS sensors are also investigated in [28, 29, 30], where large area electronics (LAE) are utilised. LAE sensors are combined with CMOS to exploit advantages from both fields, while communication between them is achieved via inductive links. A communication distance of 7.5 m is achieved, while power efficiency is in the range of picojoule per bit. Nevertheless, the data rates achieved are significantly lower compared to integrated inductive links, a difference of three orders of magnitude. Moreover, the integration of on-chip inductors, which are up to 100× larger than the inductors used in the envisioned contactless 3-D circuits require excessive silicon area and specialised CMOS processes [31, 32] and, consequently, are not well suited for these systems. These requirements greatly increase the cost of integration. Additionally, the design approaches and methodologies followed for larger implementations are not directly applicable to circuits targeting deep sub-micrometer scale technologies. 3-D printed passive components suffer from similar drawbacks, as shown in [33], and are not, therefore, a viable candidate for co-integration within a heterogeneous 3-D IC.

### 3. Heterogeneous Contactless 3-D System

An overview of the heterogeneous 3-D system is presented in this section. The design space of the proposed system is explored, considering the inherent system constraints and technology characteristics.

Each tier is assumed to be manufactured in a different technology for improved yield and therefore lower cost. There exist several ways to integrate the two tiers of the system, which communicate wirelessly, including a 3-D SiP with contactless links as shown in Figure 3(a), a hybrid 3-D stack with TSVs and contactless links, and a purely contactless system. A wireless approach is preferred over TSVs not to increase manufacturing complexity, where similar communication distances are achieved. Moreover, there is no need for voltage conversion, although the supply voltage of each technology is different. Communication between the tiers via wire bonds and off-chip interconnects is also avoided eliminating the parasitic impedance of the wires, thereby offering an improved performance and lower power as compared to an only wire-bonded SiP System.

The two tiers are stacked face-up, to enable fluidic sensing. For the bottom tier, a face down approach is feasible, however increasing the communication distance between the coupled inductors. Consequently, a larger on-chip inductor would be required to achieve the same communication data rate, resulting in higher cost or increased power. Since both tiers are stacked face-up, microbump technology cannot be implemented and, therefore, wire-bonds are preferred to provide power and ground. To reduce the signal attenuation through the substrate, a high resistivity substrate is chosen for both technologies. A high-k, ten metal layer process is chosen for the 65 nm technology, while four metal layers are used in the 0.35 µm technology. No thick metals are implemented, as the increased resistance due to the slightly thinner metal layers effectively dampens the oscillations in the receiver tier, without excessively increasing power consumption. There is no requirement to thin the bottom tier substrate, alternatively, the top tier substrate is reasonably thinned depending upon the chosen outer diameter of the inductor to reduce power and facilitate inter-tier communication.

The bottom tier of the system controls the sensor module and post-processes the received data. The sensor tier is partitioned into the blocks that sense, transmit the sensed data, and receive the control data. Throughout this paper, the term uplink describes a transceiver with the transmitter module designed in the 65 nm technology and the receiver circuit in the 0.35 µm technology, respectively. The uplink communicates the control data to the sensor tier.

Conversely, for the downlink transceiver, the technologies for the transmitter and receiver circuits are 0.35 µm and 65 nm, respectively. The downlink transmits the sensed data to the processor tier. The block diagram of the proposed transceivers is illustrated in Figure 3(b). The uplink and downlink transceivers share the on-chip inductors and therefore only half duplex communication is supported. To extend the functionality of the inter-tier interconnects to full duplex communication, another inductive link should be employed which entails an additional area overhead. Alternatively, another approach is to shift
transmission phase of each transmitter circuit and allow data transmission from both tiers.

The efficacy of the communication between the coupled inductors depends upon the achieved coupling given the separation distance \( X \) and the outer diameter \( d_{\text{out}} \) of the on-chip inductors, as depicted in Figure 4(a). The coupling between the paired inductors is given by

\[
k = \frac{M}{\sqrt{L_T L_R}},
\]

where \( M \) is the mutual inductance between the inductors and \( L_T \) and \( L_R \) are the self-inductance of the transmitter and receiver inductors, respectively. The mutual inductance and the transimpedance \( V_{R,i}/I_{T,i} \), since \( V_{R,i} = \frac{M I_{T,i}}{L_R} \), of the paired inductors depend upon the geometry of the inductors. Consequently, a higher \( k \) increases the amplitude of the received signal \( V_{R,i} \). For the chosen technologies, a minimum communication distance of 80 \( \mu m \) is feasible without a significant cost overhead [23]. This distance allows a reasonably small size for the on-chip inductor. Note that more aggressive thinning as assumed, for example, in [24] results in excessive cost.

A minimum coupling of 0.1 is required to support inter-tier communication. Below 0.1, multiple amplification stages are needed to efficiently recover the transmitted signal [25], which drastically increase power. As the objective is to maintain low power, these techniques are rather inapplicable and are therefore not considered.

Another factor that affects the coupling is the spatial misalignment between the coupled inductors. The impact of misalignment on the coupling is depicted in Figure 4(b) up to a distance of \( d_{\text{out}}/2 \). For this simulation, inductors with outer diameter of \( d_{\text{out}} = 300 \mu m, 250 \mu m, \) and \( 200 \mu m \) are considered to demonstrate the change in coupling for an increasing misalignment distance. A misalignment of up to \( d_{\text{out}}/4 \) can be tolerated for all inductors, since coupling is above the limit of 0.1. Moreover, depending upon the power margins, the tolerance to misalignment can be controlled by tuning the sensitivity of the receiver circuits. In a system, where multiple inductive links are used, the interference from the neighbouring inductive links should also be considered in the design process to evaluate the overall effect of misalignment on the transceiver performance.

The total power consumed by the inductive link is,

\[
P_{\text{tot}} = P_{T,65} + P_{R,350} + P_{T,350} + P_{R,65},
\]

where \( P_{T,65} \) is the power consumed by the 65 \( nm \) transmitter and \( P_{T,350} \) by the 0.35 \( \mu m \) transmitter, respectively. Equivalently, \( P_{R,65} \) and \( P_{R,350} \) are the power consumption of the receivers on each tier. The objective is to minimise the total power of the inductive link, \( P_{\text{tot}} \). Since the two tiers are fabricated in different technologies, the supply voltage of the target technologies is significantly different, greatly affecting the power consumed on each tier. The nominal supply voltage is \( V_{dd} = 1.2 \) V and \( V_{dd} = 3.3 \) V for the 65 \( nm \) and 0.35 \( \mu m \) technology, respectively. Given these voltages, greater power savings result...
by reducing the current drawn by the transceiver circuits in the sensor tier.

4. Heterogeneous Transceiver Design

The design method for the heterogeneous inductive link is presented in this section. In subsection 4.1, the coupled inductors are presented, while in subsection 4.2, the proposed circuit is described. Tradeoffs related to the design methodology are explored in subsection 4.3.

4.1. On-Chip Coupled Inductors

The design of the coupled inductors is discussed in this subsection. An \( RLC \) model of the inductors is described along with the utilised signalling scheme.

In Figure 4(a), the coupling coefficient is illustrated for a communication distance of \( X = 80 \) \( \mu \)m and increasing outer diameter \( d_{out} \). Ansys Maxwell [55] is used to accurately evaluate the coupling level. To achieve the minimum coupling of \( k = 0.1 \), an inductor with an outer diameter of \( d_{out} = 150 \) \( \mu \)m suffices. However, since the focus is on reducing the power dissipation, a larger inductor is employed to increase the coupling level allowing for lower power. A coupling of \( k = 0.22 \) is achieved using inductors with four turns, an outer diameter of \( d_{out} = 300 \) \( \mu \)m, and a wire width of \( w = 5 \) \( \mu \)m. To improve further the level of coupling, \( d_{out} \) increases exponentially resulting in a considerable area overhead, without the equivalent increase in coupling, according to \([13]\)

\[
k = \left( \frac{0.25}{(X/d_{out}) + 0.25} \right)^{1.5}.
\]

A schematic of the coupled inductors and the corresponding \( RLC \) characteristics are depicted in Figure 4(a). Despite the 4-node difference between the technologies of the tiers, the top interconnect layers used for the inductors exhibit similar physical characteristics (e.g. thickness). The illustrated \( RLC \) characteristics for the paired inductors and the coupling coefficient \( k \) are extracted from Ansys Maxwell [55] simulations. The inductor pair behaves as a resonant band pass filter, with a peak resonance frequency of \( f_{res} = 3.2 \text{ GHz} \), illustrated in Figure 5(a). Since the induced voltage in the receiver is

\[
V_{Rx} = j\omega M_{Tx},
\]

the bandwidth of the inductive link is bounded by the resonant frequency of the band pass filter.

Additionally, to reduce the complexity and, therefore, the power of the receiver circuit, non-return-to-zero encoding is used for transmitting data as depicted in Figure 6(a). A high frequency positive voltage pulse is generated on the transmitter inductor for a transition from logic zero to logic one, or a negative pulse for the opposite transition. The transmitted pulse is sensed by the inductor of the receiver and is transformed into a digital signal through the receiver circuit. A simulation of the transmitted data \( V_{Tx} \) and the received signal \( V_{Rx} \) is plotted, respectively, in Figures 6(b) and 7(c). An oscillation is observed at the receiver signal due to the inductor parasitic impedance. Nevertheless, the oscillation is dampened sufficiently fast to prevent inter-symbol interference and there is no need to increase the series resistance of the inductor pair.

The quality factor for both on-chip inductors is illustrated in Figure 7(a) against an increasing trace width for the target frequency of 1 GHz. As expected, the quality of the inductor increases monotonically to wider traces. For the selected width \( w = 5 \) \( \mu \)m, the on-chip inductors exhibit a self-inductance of \( L_{350} = 11.4 \) nH and \( L_{65} = 11.5 \) nH for the 0.35 \( \mu \)m and 65 nm tiers, respectively, while both inductors present a quality factor of \( Q = 4 \).

While the quality factor is an important figure of merit in the majority of applications comprising on-chip inductors (e.g. a voltage controlled oscillator), in near field inductive links, the parasitic resistance of the on-chip inductor is exploited to efficiently damp the oscillations caused by the coupled inductors. Moreover, the quality factor is inversely proportional to the bandwidth of the inductors and, effectively, limits the performance of the inductive link. Therefore, the inductor is not necessarily optimised for the highest quality factor. Alternatively, implementing high-\( Q \) inductors improves the power efficiency of the link. In Figure 7(b), the per cent loss through the coupled inductors is depicted, as the trace width (and therefore the quality factor) increases. The dashed and dashdotted lines denote the peak voltage of the transmitted and received signal, respectively. For this example, a 65 nm H-bridge circuit drives
the minimum amplitude of $V$ of the transmitter and the sensitivity of the receiver circuit, i.e. $M$. The distance between the tiers on the receiver side approaches a maximum that for a width of 5 $\mu$m and exhibiting high losses through the on-chip inductors. Note resulting in overdesign of the active components of the transceiver and achieving the target performance, without re-designing the active components.

In each clock cycle, during the positive half-period the cross-coupled transistors are discharged at different signal levels. The precharge state is modulated by the size of the clock driven PMOS devices, $M6$. The NMOS devices controlled by the clock signal ($M2$, $M6$). The minimum device width that rectifies the signal and supports the desired data rate is utilised.

$\frac{V_{\text{peak,rx}}}{V_{\text{peak,tx}}}$

Due to the limited coupling between the two inductors, a synchronous sensing scheme is preferred. Synchronous receivers sample the received pulse within a specified time interval and therefore the effect of crosstalk noise is reduced and accidental switching due to glitches is prevented.

$\text{Peak Voltage [V]}$

$\text{Communication Loss [%]}$

Figure 6: Transient SPICE waveforms for, respectively, (a) the signal of the transmitted data $T_{\text{data}}$, (b) the signal at the transmitter inductor $V_{T_x}$, and (c) the signal at the receiver inductor $V_{R_x}$. The observed damped oscillation of the signal $V_{R_x}$ is due to the parasitic impedance of the inductor.

$\text{Figure 7: (a) The quality factor of both on-chip inductors and (b) the per cent loss through the inductive link versus an increasing trace width.}$

$\text{loss} = 100 \times \frac{V_{\text{peak,tx}} - V_{\text{peak,rx}}}{V_{\text{peak,tx}}}, \quad (5)$

where $V_{\text{peak,tx}}$ and $V_{\text{peak,rx}}$ are the peak voltages of the transmitted and received signal, respectively. The losses through the link significantly decrease as the quality factor increases, resulting in a power efficient design.

$\text{Communication Loss [%]}$

Therefore, a balance is necessary between the self inductance and parasitic resistance to effectively damp the undesirable oscillations and achieve the target performance, without resulting in overdesign of the active components of the transceiver and exhibiting high losses through the on-chip inductors. Note that for a width of 5 $\mu$m or greater, the peak voltage on the receiver side approaches a maximum 0.4 V, while the transmitter operates relatively close to the optimal. Nevertheless, for traces wider than 5 $\mu$m the damping time approximates 1 ns, which is the target data rate of the inductive link. Consequently, traces wider than 5 $\mu$m deteriorate the performance and robustness of the inductive link, limiting the appropriate width for the specific circuit to $w = 5 \mu$m.

4.2. Transceiver Circuit

The transceiver design for the communication between the processing and sensing tiers is described in this subsection. Commercial libraries are used for both the 65 nm technology [12] and the 0.35 $\mu$m technology [18].

The transceiver consists of an H-Bridge circuit driving the inductor, as shown in Figure 8(a) [14]. For a specific physical distance between the tiers $X$ and coupling level $k$, the size of the transmitter and the sensitivity of the receiver circuit, i.e. the minimum amplitude of $V_{R_x}$ that can be successfully sensed, are interdependent. In addition, the devices $M0$ and $M1$ shown in Figure 8(a) must be appropriately sized to tolerate any mismatch of the voltage that differentially drives the inductor and satisfy the performance requirements. The delay buffer implemented by the three inverters shown in Figure 8(a) determines the width of the transmitted signal.

Alternatively, the receiver circuit is a synchronous sense amplifier driven by a differential pair, as illustrated in Figure 8(b) [14]. The solid line is the output $D$ of the cross-coupled pair, while the dashed line denotes the output $\overline{D}$. The dashdotted line is the common ground $\text{com}$ of the differential pair. The cross-coupled inverters are discharged at different
The voltage \( V \) of 65 nm is required. Hence, power decreases. Equivalently, for the amplified and consequently the width of the following devices input gain. Due to the increased input gain, the signal is pre-controlled, while regulating the voltage \( V_{RX} \) sensed by the receiver inductor. For the 0.35 \( \mu m \) receiver, a bias voltage of \( V_{BIAS} \) is chosen for this circuit to produce a high input gain. Due to the increased input gain, the signal is pre-amplified and consequently the width of the following devices in the circuit of the receiver is smaller, since less amplification is required. Hence, power decreases. Equivalently, for the 65 nm receiver the same approach is applied with a bias voltage of \( V_{BIAS} = 0.6 \) V and \( R_{BIAS} = 3 \) k\( \Omega \).

4.3. Design Tradeoffs

Tradeoffs related to the design of heterogeneous inductive link transceivers are explored in this subsection. The key idea is to exploit the heterogeneity of the system to decrease power as compared to existing homogeneous links.

Both transceivers are simulated with Cadence Spectre [5], exhibiting a data rate of 1 Gbps. The transmitted data \( T_x \) and the received data \( R_x \) are illustrated, respectively, in Figures 3 and 4 for the uplink and downlink transceiver. A full swing signal at the nominal voltage of each tier is produced, without the usage of level shifters. For a specific physical distance \( X \), coupling level \( k \), and data rate, the power can be lowered by carefully sizing the devices in both the \( T_x \) and \( R_x \) circuits, considering the different sensitivity of these circuits when implemented with dissimilar technologies.

For the uplink receiver, the size of the devices \( M2, M3, M6 \) that minimises the power, yields a sensitivity threshold of 300 mV. Further lowering this sensitivity threshold, requires a higher input gain. A higher input gain is, in turn, achieved by increasing the size of the amplification stage \( M3 \) and the clock controlled devices \( M2, M6 \), increasing the current flowing through the sense amplifier, given by [58]

\[
I_{com} \approx 2g_m(V_{IN} - V_{TH})^2\left(1 - \frac{0.75}{1 + \frac{V_{IN} - V_{TH}}{V_{TH}}}\right),
\]

where \( g_m \) is the transconductance of the differential pair \( M3 \) in Figure 3(a), \( V_{IN} \) is the input signal of the differential pair and \( V_{TH} \) is the threshold voltage of the differential pair devices, respectively. This increase in the common current is translated into an increase in the current difference between the drain terminals of the differential pair as [58]

\[
\Delta I_{diff} = \sqrt{2g_m I_{com} \Delta V_{diff}}.
\]

This approach, nevertheless, results in larger devices and higher currents in the sensing tier, where the supply voltage is 3.3 V as the power is proportional to

\[
P \propto I_{com} V_{dd}.
\]

Thus, an alternative design methodology is proposed. By increasing the size of \( M0 \) and \( M1 \) in the 65 nm transmitter to drive a higher current and using the minimum-power sensitivity (i.e. 300 mV) for the 0.35 \( \mu m \) receiver, the communication specifications are satisfied, without significantly increasing power.

Alternatively, the devices of the 65 nm receiver are not sized for minimum power but rather for highest sensitivity. Thus, the 65 nm receiver is designed to sense signals with amplitude as low as 75 mV. The resulting increase in the power of the receiver, however is compensated by the power savings due to the smaller sizes of the devices \( M0 \) and \( M1 \) in the 0.35 \( \mu m \)
transmitter. Consequently, the devices of the 65 nm tier are slightly larger than the devices of the 0.35 μm tier, which are sized for minimum power (see Table 2). Therefore, the downlink transceiver benefits from the higher sensitivity of the 65 nm receiver which allows for a 70% decrease in the size of the devices of the 0.35 μm transmitter.

5. Simulation Results

Simulation results for the verification of the proposed design are presented in this section. As power is the primary objective of the proposed inductive links, a power comparison is performed with prior designs at 65 nm [22] and 0.35 μm [13] technologies, which serve as a baseline. The effect of process variations on the proposed design is investigated in subsection 5.1. The effect of device mismatch is explored in subsection 5.2, while the effect of misalignment on the power of the system is presented in subsection 5.3. A discussion on the results is presented in subsection 5.4. Back-annotated results including the parasitic effects of the layout of the transceiver are given in subsection 5.5.

At the data rate of 1 Gbps, the 65 nm transmitter consumes $P_{T, 65} = 2.5$ mW and the 0.35 μm receiver $P_{R, 350} = 2.78$ mW, tallying a peak power of $P_{aplink} = 5.28$ mW for the uplink circuit, where $P_{aplink} = P_{T, 65} + P_{R, 350}$. The average power consumption for the uplink transceiver is $P_{avg, T, 65} = 2.03$ mW for the transmitter and $P_{avg, R, 350} = 498.4$ μW for the receiver. Alternatively for the downlink transceiver, $P_{T, 350} = 6.5$ mW are consumed by the 0.35 μm transmitter and $P_{R, 65} = 2.37$ mW by the 65 nm receiver, respectively, for a total peak power of $P_{downlink} = 8.67$ mW, where $P_{downlink} = P_{T, 350} + P_{R, 65}$. The average power consumption for the downlink transceiver is $P_{avg, T, 350} = 2.37$ mW for the transmitter and $P_{avg, R, 65} = 24.18$ μW for the receiver.

![Figure 10: Transient simulation of the transmitted signal $T_{xData}$ and the rectified signal $Rx_{Data}$ for (a) the uplink transceiver and (b) the downlink transceiver.](image)

The performance, power, and inductor parameters of this work compared to the state-of-the-art inductive links in 0.35 μm and 65 nm are listed in Table 1. Due to the variety of design choices in communication distance, outer diameter of the implemented on-chip inductors, and transceiver circuits, a direct comparison is not feasible. Both $P_{aplink}$ and $P_{downlink}$ of the proposed transceivers exhibit a reduction of 87.2% and 80.9%, respectively, compared to the transceiver in [14], mainly due to the difference in the communication distance and secondarily due to the tradeoff explored between the power and sensitivity in the proposed design methodology for heterogeneous 3-D ICs.

Alternatively, $P_{aplink}$ and $P_{downlink}$ exhibit a reduction of 47.7% and an increase of 10%, equivalently, as compared to the 7.8 mW in [22]. The decreased power manifested by the uplink transceiver is credited to the low power consumption of the 65 nm transmitter and the specific performance constraints of the sensing module. Note that in a homogeneous 65 nm circuit, the performance of the link would not be upper bound by the slower 0.35 μm process while the power would be further reduced by the 65 nm receiver. Alternatively, the increased consumption of the downlink transceiver is due to the power hungry 0.35 μm transmitter.

The performance of the transceiver designed in this work can be improved up to 2 Gbps without significant increase in power. However, further improvement would result in considerable increase in power due to the performance of the sensing tier. Additionally, by thinning the substrate below 80 μm (as in [24]) a proportional decrease in power can be observed. Nevertheless, excessive thinning is avoided as it can lead to lower yield due to the handling of thin and, therefore, brittle wafers.

5.1. Effect of Process and Voltage Variations

An analysis of the impact of process variations on the proposed circuits is presented in this subsection. Note that in neither [13] nor [22], the impact of process variations is considered. For both the uplink and downlink transceivers, variations in the characteristics of the devices for each technology are considered. Moreover, variations on the supply voltage, $V_{dd}$, and the voltage bias, $V_{bias}$, of the differential pair are investigated.

For both tiers, five process corners are considered for the performance of the devices, namely the typical (TT), fast-fast (FF), slow-slow (SS), fast nmos-slow pmos (FNS P), and slow nmos-fast pmos (SNFP) corners. Since disparate technologies are used on each tier and a two tier system is considered, there exists a total of 25 permutations to cover the entire design space. The sensitivity of the design to different corners depends upon the technology node, the nature of each circuit (i.e. either a transmitter or a receiver), and the allowed variation of the sensed signal.

For the transmitter circuit, simulations are performed for the five process corners, while the receiver is in the typical process corner. In these simulations, the overall effect of variations on the operation of the circuits is not significant. The effect on power is illustrated in Figure 11(a), where the average power of the 65 nm tier and the 0.35 μm tier are, respectively, shown by the white and black bars. The effect on power for each of the
different process corners on the 0.35 μm transmitter is substantial, while for the 65 nm technology, the corresponding effect is lower. For the remaining corners of the 25 permutations, the operation of the transceiver is negligibly affected, due to variations effects not degrading the quality of the propagated signal through the inductive link.

In addition to the process corners, a ±10% margin is assumed for the nominal supply voltage, \( V_{dd} \) of the transmitter in each technology, while the receiver is supplied by the nominal voltage. For the transmitter of the 65 nm tier, the fluctuations of the power supply \( V_{dd} \) do not corrupt the data transfer to the receiver tier. Differently, the 0.35 μm transmitter operates normally for a supply voltage reduced by up to 10%. However, when the supply voltage increases over 3.45 V faulty operation is observed and bit errors occur. This higher voltage increases the amplitude of the signal \( V_{in} \) at the receiver, while the low impedance of the inductor does not dampen the oscillation adequately fast. A larger series resistance can be used for the coupled inductors to prevent the effect of overdrive, which, however, results in increasing power. Alternatively, normal operation of the transceiver is restored by reducing the duty cycle of sampling through adjustment of the clock signal \( T_s \).

The receiver circuit is more sensitive to both voltage and process variations. By maintaining the transmitters at the typical process corner, simulations are performed on all process corners for the receiver circuits. For the FF and FNSP corners, upsizing of the 0.35 μm receiver circuit is required to ensure the correct operation of the transceiver. Compared to a 0.35 μm receiver sized to operate at the lowest power, M3 is increased by 2 μm and M6 by 0.8 μm. The upsized widths are listed for the robust receivers in Table 1. Increasing the device width is necessary to further amplify the received signal and generate a larger voltage difference at nodes \( D \) and \( \overline{D} \) shown in Figure 11(a). Therefore, the cross-coupled inverters operate correctly, without bit errors. However, the peak power is increased by 0.62 mW reaching 2.78 mW. The biasing stage is accordingly adapted considering the transistor sizes.

Alternatively, for the SNFP and SS corners, the sampling time \( T_s \) is increased to 550 ps and 750 ps, respectively, to properly sense the received data. Since the sampling time is not increased above 1 ns, the performance of the circuit is not degraded. For the 65 nm receiver, no overdesign is required for the FF and FNSP corners. For the SNFP and SS corners, the sampling time is adjusted to 300 ps. The average power consumption of the receiver circuits is illustrated in Figure 11(b), where the black bars denote the 0.35 μm circuit and the white bars denote the 65 nm circuit, respectively.

Power supply variation is also explored for the receiver circuits. However, the functionality of neither the 65 nm nor the 0.35 μm receiver is affected by a ±10% fluctuation in the power supply. Alternatively, the operation of the receiver circuit can be hindered by the voltage variation on the biasing voltage \( V_{bias} \) due to the differential pair. A large fluctuation on the biasing voltage results in incorrect biasing of the differential pair and, consequently, in erroneous operation of the receiver. The sensitivity on \( V_{bias} \) Variations for both the 0.35 μm and 65 nm receiver circuits is tested, by sweeping the typical value of \( V_{bias} \) for each circuit.

For the 65 nm receiver, a \( V_{bias} = 0.6 \) V is utilised. The differential pair, however, does not operate for biasing voltages be-

<table>
<thead>
<tr>
<th>Table 1: Comparison to state-of-the-art.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metrics</td>
</tr>
<tr>
<td>Data rate [Gbps]</td>
</tr>
<tr>
<td>Power [mW]</td>
</tr>
<tr>
<td>Communication Distance (X) [μm]</td>
</tr>
<tr>
<td>Outer Diameter ( d_{out} ) [μm]</td>
</tr>
<tr>
<td>Prototyping Cost [-]</td>
</tr>
</tbody>
</table>

Figure 11: Average power consumption at different process corners for the transmitter circuits and the receiver circuits, respectively.
low 0.58 V and consequently a −10% constraint is not fulfilled. To ensure a −10% margin, the $V_{bias}$ is increased to 0.63 V. Simulations indicate that the 10% margin is satisfied. By increasing the biasing voltage to 0.63 V, the average power consumption is increased by 1.12 $\mu W$, which is negligible compared to the overall power consumed by the transceivers. Alternatively, for the 0.35 $\mu m$ receiver, $V_{bias} = 1.5$ V and a range between 1.35 V and 1.65 V is assumed. Erroneous operation is observed for $V_{bias} < 1.39$ V and $V_{bias} > 1.61$ V. Therefore, for a range of ±7.3% $V_{bias}$ the differential pair operates correctly. To further increase the margin for $V_{bias}$ significantly increases the power of the circuit.

5.2. Monte Carlo Mismatch Analysis

The results of Monte Carlo simulations are presented in this subsection to evaluate the sensitivity of the differential pairs (M3) to device mismatch [35]. Regardless of the process corners, random mismatch between the NMOS devices of the differential pair affects voltage sensing. Mismatch can result from process variations either on the $V_t$ or the aspect ratio $W/L$ of the transistors [33]. To reduce the effect of mismatch the length $L$ of the related devices is increased, maintaining, however, the ratio of $W/L$ constant. Thus, the effect of mismatch on the transistors is reduced.

Initially, the minimum length $L$ is utilised for all devices in both technologies. However, bit error rates occur in a mismatch analysis with Monte Carlo simulations of 100 points and 3σ process variations for the devices M3 of the differential pair. To reduce the effect of mismatch, the length of the devices in the differential pair is increased. For the 0.35 $\mu m$ technology node, the length of M3 is adjusted to 500 nm, while in the 65 nm technology node, the length is doubled to 120 nm. As expected, the effect of random process variations on the smaller technology node is 22.8% greater compared to the 0.35 $\mu m$ technology.

Based on this mismatch analysis, the distribution of the average power consumption for the 0.35 $\mu m$ and the 65 nm receivers is depicted in Figures 12(a) and 12(b), respectively. The distribution of power for the 0.35 $\mu m$ receiver ranges from 495 $\mu W$ to 515 $\mu W$, with $\sigma_{350} = 3.74$ $\mu W$ and $\mu_{350} = 501$ $\mu W$. Alternatively, the power on the 65 nm tier is much closer to the mean value of $\mu_{65} = 23.1$ $\mu W$, with $\sigma_{65} = 82.9$ nW. Since the power is proportional to

$$ P \propto CV_d^2, \quad (9) $$

increasing the device length and keeping the ratio $W/L$ fixed, a small increase in power is incurred. Indeed, the average power of the 65 nm receiver increases from 16.06 $\mu W$ ($L_{65} = 65$ nm) to 23.1 $\mu W$ ($L_{65} = 120$ nm). The average power for the 0.35 $\mu m$ receiver increases from 467.5 $\mu W$ ($L_{350} = 350$ nm) to 498.4 $\mu W$ ($L_{350} = 500$ nm).

5.3. Misalignment Analysis

In this subsection, the impact of misalignment on the performance of the transceiver circuits is presented. Apart from process and voltage variations, a unique source of process variation for inductive links is the misalignment between the coupled inductors. The coupling coefficient, as described in Section 8, is a function of the communication distance $X$ and outer diameter $d_{out}$ and depends upon the spatial alignment of the inductors. For the implemented inductors with an outer diameter of 300 $\mu m$, a nominal coupling of 0.22 is achieved. Based on Figure 13, the coupling level is higher than $k = 0.1$ for any misalignment of the inductors smaller than 102 $\mu m$. A decrease in the coupling coefficient, results in a decreased amplitude for the received signal $V_{Rx}$ and, consequently, erroneous operation of the receiver circuit.

The required power for the transceivers to correctly operate while being tolerant to misalignment is illustrated in Figure 13. The solid line is the average power of the uplink transceiver, while the dashed line is the average power of the downlink transceiver, respectively. As the design methodology focuses on power efficiency exploiting the power versus sensitivity trade-off, the 0.35 $\mu m$ circuits are designed for lowest power, while the 65 nm circuits for highest sensitivity. Consequently, for the uplink transceiver to tolerate misalignment, the sensitivity of the 0.35 $\mu m$ receiver can be enhanced resulting to an increase in power. The increase in power is appropriately split between the transmitter and the receiver by upsizing the receiver circuit to enhance sensitivity and by increasing the driving strength of the transmitter. Therefore, a small overhead on peak power consumption is ensured. Alternatively, by unilaterally increasing either the strength of the transmitter or the sensitivity of the receiver, a worse solution is produced. Due to the high power requirements of the H-Bridge transmitter, increasing the
misalignment between the coupled inductors and solely adjusting the transmission power results in a significant increase in power. Equivalently, exclusively tuning the sensitivity of the 0.35 µm receiver to accommodate for the additional misalignment, a considerable power surge is incurred due to the higher core voltage of the older manufacturing process.

For the downlink transceiver the sensitivity of the 65 nm receiver cannot be further increased. Consequently, the amplitude of the transmitted signal is required to be significantly higher for the receiver to operate normally, resulting in a significant overhead on power for the 0.35 µm transmitter. Alleviating the mismatch between the processing and sensing tiers, ensures system robustness and low power operation.

5.4. Transceiver Design for Robustness

The primary issues for the robustness of the heterogeneous transceivers are discussed in this subsection. The modified size of the devices for the transmitter and receiver circuits, which guarantee operation under all of the aforementioned sources of variations, are listed in Table 2. The nominal transistor length of each technology is utilised, apart from the device denoted with an asterisk. For device M3 the length is increased to reduce the impact of mismatch.

In Table 2, area, power, and performance are reported for both the uplink and downlink transceivers. The table is divided in three parts. In the first four rows, the inductor area, data rate, and energy efficiency are given. The peak and average power, where no variations are considered, are listed in rows five and eight. The increase in peak and average power that guarantees robustness under several types of variations are notated as $\Delta P_{\text{peak}}$ and $\Delta P_{\text{avg}}$, respectively. The overhead $\Delta P_{\text{peak}}$ due to process variations and misalignment is, respectively, reported in rows six and seven. The overhead $\Delta P_{\text{avg}}$ due to variations in $V_{\text{bias}}$ and device mismatch is, respectively, reported in rows nine and ten.

To accommodate all process corners, the 0.35 µm receiver peak power is increased by 0.62 mW due to upsizing of the devices. Hence, correct operation is satisfied for all process corners in both receiver circuits. The peak power is also affected by misalignment between the inductors. For example, reducing the coupling to $k = 0.2$, the peak power is increased by 1 mW and 2.47 mW for the uplink and the downlink transceivers, respectively. Alternatively, to establish a ±10% margin for the $V_{\text{bias}}$ of the 65 nm receiver, the average power consumption is increased by 1.12 µW. The peak power is not affected. To reduce the impact of device mismatch on the differential pair, the length of the differential pair devices is increased to 500 nm and 120 nm for the 0.35 µm and the 65 nm technology, respectively. A power overhead of 7 µW is observed for the 65 nm receiver, while for the 0.35 µm receiver an overhead of 30.9 µW is demonstrated.

The reported power overhead for all process variations except for the mismatch is negligible, less than 3%. The misalignment, however, incurs a significant power overhead that cannot be effectively addressed by careful redesign of the transceiver circuit. Rather, the manufacturing process should guarantee sufficient alignment between the tiers, such that low power operation of the inductive link is ensured.

5.5. Back-Annotated Verification

For the verification of the proposed transceiver, the back-annotated results of the layout extracted view are given in this subsection. The simulated circuits contain inductors which comply with design rule check (DRC) and design for manufacturability (DFM) including all the parasitic impedances by utilising Helic Veloce RF and Helic Veloce Raptor X. These tools ensure high accuracy comparable to silicon measurements.

The layout of the transceiver circuit for the 65 nm and the 0.35 µm tiers is illustrated in Figures 14(a) and 14(b), respectively. For clarity, the layout of the on-chip inductors are illustrated separately in Figures 15(a) and 15(b), respectively, for the 65 nm and 0.35 µm tier. For the receiver circuits, symmetry is crucial to reduce the mismatch between the devices that form the differential pair. Even ostensibly insignificant discrepancies between the NMOS differential pair can potentially impair the functionality of the differential pair, and consequently the functionality of the transceiver. Therefore, special care is taken to symmetrically place the differential pair and reduce parasitic effects by folding and fusing the devices to avoid metal connections. Furthermore, by folding the differential pair into a square topology the sensitivity to process mismatch is mitigated.

![Figure 13: Power consumption of the proposed transceiver circuits, with misaligned inductors for a maximum misalignment of 102 µm and $k \geq 0.1$.](image-url)
Table 3: Key metrics for the uplink and downlink transceiver.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Uplink</th>
<th>Downlink</th>
<th>Power Overhead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inductor area</td>
<td>$[\mu m^2]$</td>
<td>$300 \times 300$</td>
<td>-</td>
</tr>
<tr>
<td>Data rate</td>
<td>$[Gbps]$</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>$P_{tot}$</td>
<td>$[mW]$</td>
<td>13.95</td>
<td>2.34%</td>
</tr>
<tr>
<td>Energy efficiency</td>
<td>$[\mu J/b]$</td>
<td>5.28</td>
<td></td>
</tr>
<tr>
<td>Peak power $P_{peak}$</td>
<td>$[mW]$</td>
<td>2.5</td>
<td>8.67</td>
</tr>
<tr>
<td>$\Delta P_{peak}$ due to process corners</td>
<td>$[mW]$</td>
<td>-</td>
<td>0.62</td>
</tr>
<tr>
<td>$\Delta P_{peak}$ due to misalignment ($k = 0.2$)</td>
<td>$[mW]$</td>
<td>1</td>
<td>2.47</td>
</tr>
<tr>
<td>Average power $P_{avg}$</td>
<td>$[mW]$</td>
<td>2.03 mW</td>
<td></td>
</tr>
<tr>
<td>$\Delta P_{avg}$ due to $V_{bias}$ variation</td>
<td>$[\mu W]$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$\Delta P_{avg}$ due to device mismatch</td>
<td>$[\mu W]$</td>
<td>-</td>
<td>30.9</td>
</tr>
</tbody>
</table>

6. Conclusion

A heterogeneous transceiver for contactless links is presented, providing inter-tier communication between a processing and a sensing tier. By exploiting the sensitivity versus power tradeoff enabled by the use of heterogeneous technologies, a significant decrease in power is achieved for inter-tier communication with the inductive links. In addition, although heterogeneous technologies are considered, communication is achieved between the modules without the need of level shifters.
Figure 15: Layout of (a) the 65 nm on-chip inductor and (b) the 0.35 µm on-chip inductor.

Figure 16: Back-annotated power waveform for the uplink transceiver where (a) the power trace for transmitting data with the 65 nm transmitter, and (b) the power trace for sensing data with the 0.35 µm receiver.

Figure 17: Back-annotated power waveform for the downlink transceiver where (a) the power trace for transmitting data using the 0.35 µm transmitter, and (b) the power trace for sensing data with the 65 nm receiver.

EDITING FROM TECHNOLOGY HETEROGENEITY, THE ENERGY EFFICIENCY OF THE TRANSCIEVERS IS 5.28 pJ/b FOR THE UPLINK AND TO 8.67 pJ/b FOR THE DOWNLINK. COMPARED TO PROTOTYPE 0.35 µM AND 65 NM TRANSCIEVER CIRCUITS, RESPECTIVELY, A 6.8X AND A 37.5% REDUCTION IN POWER IS OBSERVED AT A BANDWIDTH OF 1 Gbps.

The transceivers are designed considering process and voltage variations to guarantee system robustness. Moreover, a mismatch analysis is provided since differential sense amplifiers are affected by random process variability. To improve system robustness against process variations and mismatch, a low overhead of < 1% in the average power consumption of the inductive link is required. For the peak power an overhead of 2.34% is exhibited.

References
