Contactless Inter-Tier Communication for Heterogeneous 3-D ICs

Ioannis A. Papistas and Vasilis F. Pavlidis
Advanced Processor Technologies Group
School of Computer Science, The University of Manchester
{papistai, pavlidis}@cs.man.ac.uk

Abstract—A heterogeneous contactless transceiver circuit is designed to provide half duplex communication for a 3-D system considering specific bonding constraints. The system is composed of two tiers and is integrated face-to-back to support fluidic sensing. Communication between the tiers is achieved through inductive links. Each tier is considered to be fabricated in a different technology node 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 primary objective, with the uplink dissipating 4.93 mW and the downlink 10.53 mW. A 3.2× power reduction is achieved when using heterogeneous technologies, compared to a state-of-the-art 0.35 µm transceiver, while the dissipated energy is decreased by 34% as compared to a state-of-the-art 65 nm transceiver.

I. INTRODUCTION

Heterogeneous 3-D integration provides a promising platform for edge devices of the Internet-of-Things (IoT) [1] comprising multi-functional, small form factor, and low power systems [2]. 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 [3]. This approach is suitable for Lab-on-Chip (LoC) [4] or the IoT applications where multi-functionality and power are important.

Nevertheless, several challenges exist for developing heterogeneous 3-D systems, such as heat dissipation, testing, and inter-tier communication [3]. Considering inter-tier interconnects, through silicon vias (TSVs) provide high density signaling with low latency and power [5]. However, TSVs are an expensive means due to the related manufacturing complexity and possibly low yield [6]. Alternatively, contactless solutions based on ac coupling have been proposed [7]–[10].

Both inductive and capacitive contactless communication are compatible with conventional CMOS process and therefore no additional manufacturing steps are required for underpinning inter-tier communication [8], [11], [12]. Wireless interconnects offer several advantages to heterogeneous 3-D integration including die detachability [8] and inter-tier communication without the requirement of level shifters [13]. However, capacitive coupling is limited to face-to-face bonding, supporting primarily two tier systems and therefore significantly narrowing the candidate applications for this communication mechanism. For example, face-to-face integration is not a feasible option for LoC applications where the sensing tier should be integrated face up. Due to these limitations, inductive links are investigated in this work.

Systems with functional heterogeneity have recently been explored, [14], where 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 [15] 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 node. In [16], 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 applications requires a different design approach since power and cost are often more important than speed [17]. Thus, sensing circuits are usually implemented in older and therefore low cost technologies. Consequently, there is a significant technology gap between the sensor circuits and the digital circuits that control the sensor and process the sensory data.

In this paper, a transceiver circuit is proposed for a 3-D system characterised by both functional and technological heterogeneity. The target system is a two tier system consisting of a processing tier and a sensing tier implemented, respectively, in a 65 nm [18] and a 0.35 µm technology [19]. Homogeneous inductive links have been demonstrated in both of these technologies, where the most advanced manufacturing process for a prototype inductive link is a 65 nm technology [18]. The choice of the 0.35 µm technology is due to the typical use of this technology for sensing circuits [19], [20]. In addition, these homogeneous links serve as a baseline for comparison.

The remainder of this paper is organised as follows. In Section II, the salient features of the envisioned two tier system are introduced. The design of the inductive link transceiver that provides communication between the tiers is presented in Section III, where related design tradeoffs are also explored. Finally, some conclusions are summarised in Section V.

This work was supported in part by EPSRC UK under Grant EP/M009238/1 and scholarships from the School of Computer Science and the Doctoral Training Centre at the University of Manchester.
II. HETEROGENEOUS CONTACTLESS 3-D SYSTEM

An overview of the heterogeneous 3-D system is presented in this section. Each tier is assumed to be manufactured at 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 Fig. 1(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 cost, where similar communication distances are achieved. Moreover, there is no need for level conversion, even though the supply voltage of each technology is different. Off-chip communication between the tiers via wire bonds is also avoided to reduce the effect of parasitic impedance and achieve a higher performance and lower power as compared to an only wire bonded SiP System.

In this paper, emphasis is placed on the contactless inter-tier communication and the issues relating to the dissimilar manufacturing technologies of the tiers. Consequently, trade-offs related to the design of the link including the coupling between the inductors, the area and power consumed by the inductors are explored.

The bottom tier of the system controls the sensor module and post processes the received data. The sensor tier is partitioned to the blocks that sense, transmit the sensed data, and receive the control data. Throughout this paper, uplink is the transceiver circuit 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 transceiver of the downlink, 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 efficacy of the communication between the coupled inductors depends upon the achieved coupling given the separation distance X and the outer diameter d_out of the on-chip inductors, as depicted in Fig. 1(b). Thinning the substrate of the upper tier improves the coupling coefficient, without excessively increasing the size of the on-chip inductor. For the chosen technologies, a minimum communication distance of 80 μm is feasible without a significant cost overhead [19].

A minimum coupling of 0.1 is required to support inter-tier communication. Below 0.1, multiple amplification stages are required to receive the transmitted signal [13], which drastically increase the power of the transceiver. As the objective is to maintain low power, these techniques are rather unsuitable and are therefore not considered.

The total power consumed by the inductive link is the sum of the power dissipated by both the uplink and downlink transceiver circuits. The goal is to minimise the power of both links, given the communication distance and outer diameter of the inductor pair. Since the two tiers are fabricated in different technologies, the power consumed by each tier is significantly affected by the supply voltage. The nominal supply voltage is respectively V_dd = 1.2 V and V_dd = 3.3 V for the 65 nm and 0.35 μm technology. Given these voltages, greater power savings for the inductive link result by reducing the current drawn by the transceiver circuits of the sensor tier. Alternatively, the processing tier can be designed to satisfy the bandwidth requirements.

III. HALF DUPLEX TRANSCEIVER DESIGN

The proposed transceiver circuits for inter-tier communication are presented in this Section. In subsection III-A, the chosen coupled inductors are presented. In subsection III-B, the design of the proposed transceivers is described for both the uplink and downlink communication.

A. Coupled Inductors and Signalling

The design of the coupled inductors is presented in this subsection. An RLC model of the inductors is described along with the utilised signalling approach.

The coupling coefficient for a communication distance of X = 80 μm and increasing outer diameter d_out is shown in Fig. 1(b). Ansys Maxwell [21] is used to accurately evaluate the coupling level. To achieve the minimum coupling an inductor with an outer diameter of d_out = 150 μm is required. However, since the focus is on reducing the power dissipation, a larger inductor can be employed to increase the coupling level and facilitate the inter-tier communication. A coupling of k = 0.22 is achieved using inductors with four turns, an outer diameter of d_out = 300 μm, and a wire width of w = 5 μm. To further increase the coupling, d_out increases exponentially resulting in excessive area consumption. Additionally, to reduce the complexity and therefore power demand of the receiver circuit non-return-to-zero encoding is used for transmitting data.

A schematic of the coupled inductors is depicted in Fig. 2, where the corresponding RLC characteristics are also shown.
Due to the limited coupling between the two inductors, as shown in Fig. 3(a) [23], the inductor and to satisfy the performance requirements. The voltage that differentially drives the inductor, as shown in Fig. 3(a) [23]. For a specific physical distance between the tiers \( k \), and coupling level \( k \), the minimum amplitude of \( V_{Rx} \) that can be successfully sensed, are interdependent. In addition, the \( M0 \) and \( M1 \) devices shown in Fig. 3(a) must be appropriately sized to avoid any mismatch of the voltage that differentially drives the inductor and to satisfy the performance requirements. The delay buffer implemented by the three inverters shown in Fig. 3(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 Fig. 3(b) [23]. 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.

The differential pair (\( M3 \)) provides the received signal to the synchronous sense amplifier. The signal received by the sense amplifier is sampled according to the \( T_{sample} \) clock signal. The peak performance of the inductive link is determined by the duty cycle of the sampling clock. Alternatively, the sensitivity of the receiver is determined by the driving strength of the devices controlled by the clock signal (\( M2, M6 \)). The minimum device width that rectifies the signal and supports the desired data rate is utilised.

In the receiver, the sensed signal is biased and amplified through \( V_{bias} \) and \( R_{bias} \). For the 0.35 \( \mu m \) receiver, a bias voltage of \( V_{bias} = 1.3 \) \( V \) ensures that the differential pair operates in the saturation region. An input resistance of \( R_{bias} = 1.5 \) \( k\Omega \) is chosen for this circuit to produce a high input gain. Therefore, the width of the other devices in the circuit of the receiver is smaller decreasing power. Equivalently, for the 65 \( nm \) receiver the same approach is applied with a bias voltage of \( V_{bias} = 0.6 \) \( V \) and \( R_{bias} = 2 \) \( k\Omega \).

IV. SIMULATION RESULTS

The operation of the heterogeneous links and power trade-offs are demonstrated in this section. As power is the primary objective of the proposed inductive links, a power comparison with prior designs at 65 \( nm \) [15] and 0.35 \( \mu m \) [23] technologies is performed. 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 [24], exhibiting a data rate of 1 \( Gbps \). The transmitted data \( Tx_{data} \) and the received data \( Rx_{data} \) are illustrated in Figs. 4(a) and 4(b), respectively, 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 \( Tx \) and \( Rx \) circuits, considering the different sensitivity of these circuits when implemented at 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 can be achieved by increasing the device size of the amplification stage (\( M3 \)) and the clock controlled devices (\( M2, M6 \)). This approach, nevertheless, results in larger devices and higher currents in the sensing tier, where the supply voltage is 3.3 \( V \). Thus, an alternative design methodology is followed. 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 the power.
the use of heterogeneous technologies, a significant decrease in
power consumption of the transceivers is achieved. By exploiting the sensitivity versus power tradeoff enabled by
heterogeneity, the power consumption of the transceivers is
reduced to 4.93 mW for the uplink and to 10.53 mW for the
downlink at 1 Gbps, compared to a 45.6 mW transceiver of a
prototype 0.35 μm system, while a decrease of 34% is
observed compared to a 65 nm implementation.

REFERENCES


Fig. 4. Transient waveforms of the transmitted signal $T_x$ and the recovered signal $R_x$ for (a) the uplink transceiver and (b) the downlink transceiver.

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 size of the devices $M0$ and $M1$ in the 0.35 μm transmitter. Consequently, as illustrated in the table of Fig. 3, 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. 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.

The 65 nm transmitter consumes $P_{Tx_{65}} = 2.77$ mW and the 0.35 μm receiver $P_{Rx_{350}} = 2.16$ mW, tallying $P_{uplink} = 4.93$ mW for the uplink circuit. Alternatively, for the downlink receiver, $P_{Rx_{65}} = 1.81$ mW are consumed by the 0.35 μm transmitter and $P_{Tx_{65}} = 2.35$ mW by the 65 nm receiver, respectively, for a total of $P_{downlink} = 10.53$ mW.

Both $P_{uplink}$ and $P_{downlink}$ of the proposed transceivers exhibit a reduction of 89.1% and 73.9%, respectively, compared to the transceiver in [23]. Considering the power within each tier, the power of the 65 nm tier is $P_{tot_{65}} = 5.12$ mW, as compared to 7.8 mW in [15], indicating a decrease of 34%. Note, however, that a direct comparison with [15] is not feasible, since different transmitter and receiver circuits are used. Alternatively, the circuits implemented in the 0.35 μm tier consume $P_{tot_{350}} = 13.11$ mW. In [23], the power of the transmitter is 43 mW, 5.2× larger compared to $P_{Tx_{350}} = 18.18$ mW. Concerning the receiver in [23], a reduction of 7.7% is observed for $P_{Rx_{350}}$. These results demonstrate that 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 inductive links.

V. CONCLUSION

A heterogeneous transceiver for contactless links is presented for inter-tier communication between a processing and a sensing tier. Although heterogeneous technologies are considered, communication is achieved between the modules without the need of level shifters. Benefiting from technology heterogeneity, the power consumption of the transceivers is reduced to 4.93 mW for the uplink and to 10.53 mW for the downlink at 1 Gbps, compared to a 45.6 mW transceiver of a prototype 0.35 μm system, while a decrease of 34% is observed compared to a 65 nm implementation.