

# Energy-Efficient Transceivers for Ultra-Highspeed Computer Board-to-Board Communication

Michael Jennings\*, Bernhard Klein\*, Ronny Hahnel\*, Dirk Plettemeier\*, David Fritsche†, Gregor Tretter†, Corrado Carta†, Frank Ellinger†, Tobias Nardmann¶, Michael Schroter¶, Krzysztof Nieweglowski‡, Karlheinz Bock‡, Johannes Israel||, Andreas Fischer||, Najeeb Ul Hassan§, Lukas Landau§, Meik Dörpinghaus§, Gerhard Fettweis§

\*Chair for RF Engineering, Faculty of Electrical and Computer Engineering, Technische Universität Dresden, {michael.jenning, bernhard.klein, dirk.plettemeier}@tu-dresden.de

†Chair for Circuit Design and Network Theory, Faculty of Electrical and Computer Engineering, Technische Universität Dresden, {david.fritsche, corrado.carta, frank.ellinger}@tu-dresden.de

¶Chair for Electron Devices and Integrated Circuits, Faculty of Electrical and Computer Engineering, Technische Universität Dresden, tobias.nardmann@tu-dresden.de, mschroter@ieee.org

‡Electronic Packaging Lab, Faculty of Electrical and Computer Engineering,

Technische Universität Dresden, {krzysztof.nieweglowski, karlheinz.bock}@tu-dresden.de

||Chair of Numerical Optimization, Department of Mathematics, Technische Universität Dresden, {johannes.israel, andreas.fischer}@tu-dresden.de

§Vodafone Chair Mobile Communications Systems, Faculty of Electrical and Computer Engineering, Technische Universität Dresden,

{najeeb\_ul.hassan, lukas.landau, meik.doerpinghaus}@tu-dresden.de, gerhard.fettweis@vodafone-chair.com

**Abstract**—Enabling the vast computational and throughput requirements of future high performance computer systems and data centers requires innovative approaches. In this paper, we will focus on the communication between computer boards. One alternative to the bottleneck presented by copper wire based cable-bound communication is the deployment of wireless links between nodes consisting of processors and memory on different boards in a system. In this paper, we present an interdisciplinary approach that targets an integrated wireless transceiver for short-range ultra-high speed computer board-to-board communication. Based on our achieved results and current developments, we will also estimate energy consumption of such a transceiver.

## I. INTRODUCTION

Today's global connectivity through the Internet with its numerous applications and services require enormous computing and networking resources. Without any doubt, those requirements are increasing [1]. To cope with this trend, both computational and networking capabilities need to be addressed. On the computational front, the two most common approaches are increasing the performance of individual cores and parallelism. The latter is a multi-level approach. Nowadays, Central Processing Units (CPUs) have typically two or four cores. In server computers, up to 16 cores per single CPU are available. The next level is the use of multiple CPUs per node, employing main boards with up to four CPU sockets. Combining multiple of such nodes in one or more racks is the next level of parallelism, resulting in the construction of large data centers.

Efficiently utilizing the available computational resources requires a tremendous amount of communication within each

level and across levels. For the communication within one rack or between multiple racks, cable based technologies like Ethernet, InfiniBand, or optical, e. g., optical Ethernet, are employed or introduced. They all require switching devices and running cables to enable communication between all nodes.

Board-to-board communication is a major issue addressed in the Deutsche Forschungsgemeinschaft funded Collaborative Research Center 912 *Highly Adaptive Energy-Efficient Computing* [2] that was initiated at TU Dresden in 2011. One of the goals of the project is the use of wireless links with data rates of up to 100 Gbit/s between two computer boards in a High Performance Computing (HPC) system. This scenario is depicted in Fig. 1. To achieve this goal, wire based interconnects between nodes on different boards are replaced by wireless links. Carrier frequencies will be above 100 GHz, to keep the relative bandwidth at a manageable level. The following section will briefly describe the proposed wireless system HAEC is targeting. Afterwards, a prediction of the energy consumption based on our previous research and future developments within HAEC will be given. Finally, conclusions based on current results are presented.

## II. WIRELESS PLATFORM

### A. Overview

The wireless platform is jointly addressed on all major research areas: antennas, analog frontends, digital domain and packaging. This is necessary to cope with the challenges of wireless links above 100 GHz and its interfacing and interconnecting of the circuits to support the massive I/O

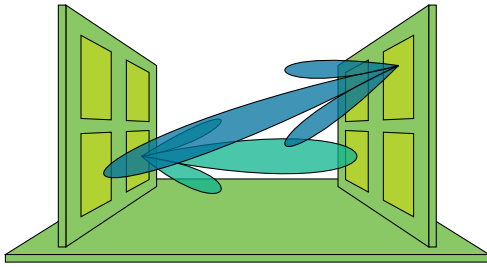


Figure 1. Illustration of a wireless board to board communication scenario in an HPC box. For simplicity, only two boards with four chips each and two possible links are shown. Full flexibility of the links enabled by beam switching is targeted.

demands. The demand for data rates above 100 Gbit/s and, hence, bandwidth either require sophisticated modulation techniques or higher carrier frequencies where bandwidth scales accordingly. Higher order modulation schemes require high SNR levels and fast Analog-to-Digital Converters (ADCs) with a high resolution. Both requirements are even more challenging when energy efficiency is essential. Large bandwidth on the other hand requires both frontend components, like the Low Noise Amplifiers (LNAs), as well as antennas complying with this demand. To keep losses small, it is practical to integrate all components as close as possible. Recent advances in 3D packaging enable the vertical integration of different semiconductor and RF substrates in one stack. Fig. 2 illustrates the idea of a 3D chip stack that is being pursued in HAEC. The benefit of such a 3D integration is an increased packaging density. The components of this stack will be described in more detail throughout this section.

### B. Antennas and Feeding Network

Within HAEC, several on-chip antenna elements have been developed and fabricated [3]–[5]. Microscopic photographs of the small elements that meet the bandwidth requirements of 30 GHz are shown in Fig. 3. All of these antenna elements

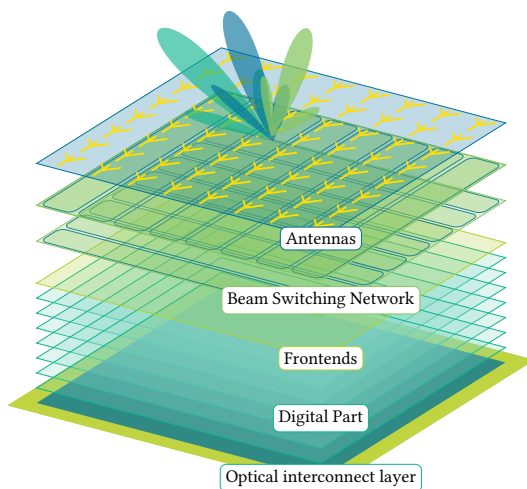


Figure 2. 3D chip stack with frontend and antenna array.

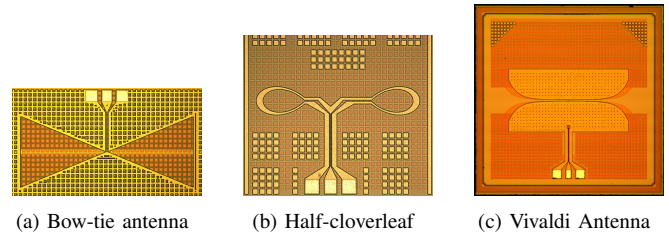


Figure 3. Microscopic photographs of on-chip antennas.

were designed for a SiGe semiconductor integrated technology. To achieve notable gain values, it was necessary to have the wafers thinned down to a thickness of 100  $\mu\text{m}$  and to utilize a metallic reflector. Thinning to 100  $\mu\text{m}$  was available as an additional process step from the fab. At 180 GHz and on a SiGe substrate, 100  $\mu\text{m}$  is close to a quarter of the wave length of  $\approx 120 \mu\text{m}$ . During measurements, the wafer chuck acts as the metallic reflector. For integration within the chip stack, a backside metallization step will be required.

To improve antenna performance, it will be necessary to use a different technology. Semiconductor materials have a relatively high permittivity and conductivity, both limiting bandwidth and efficiency. RF substrates or other materials used in stacked buildups, like Benzocyclobutene (BCB) or Polyimide, have better characteristics, which will result in higher gain, larger bandwidth and increased efficiency. The disadvantage is that additional process steps are required. This issue will be further addressed in section II-F.

The usage scenario of the wireless interface within HAEC is the communication between computer boards in HPC computers. It is therefore safe to assume that the configuration of the boards is fixed over longer periods. Additionally, it is possible to have the location of the chips at previously optimized locations. This reduces the complexity of the antenna feeding networks and allows using beam switching instead of a fully capable beam steering approach. The beam switching approach will be realized by a Butler matrix network [6], [7]. Using such a network, it can be shown that there are optimum locations on a board for the placement of chips [8]. The Butler matrix presented in [6] shows losses in the range of 1 dB to 2 dB. A microscopic photograph of the Butler matrix realized in a BCB sequential build-up process is shown in Fig. 4.

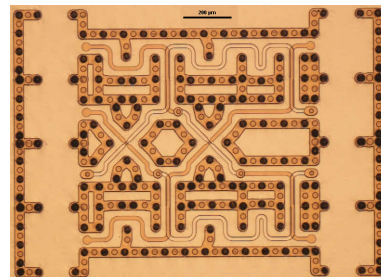


Figure 4. Photograph of the Butler matrix.

As shown in [9], a 2D antenna array can be fed by cascading two Butler matrix stages. It is therefore planned to apply this cascading principle to achieve spatial beam switching capabilities in two dimensions.

### C. Frontend

During the first phase of HAEC, the research focus was on the receiver side of the frontend. A very efficient LNA was presented in [10]. The achieved characteristics are a bandwidth of 44 GHz, a maximum power gain of 16.9 dBm and  $-3.5$  dBm output power at 1 dB compression. This is achieved at a DC power consumption of only 18 mW.

For the down conversion of the 200 GHz signal to baseband, a mixer was designed [11]. It achieves a 3 dB RF bandwidth of 32.5 GHz (180 GHz to 212.5 GHz) and exhibits a voltage conversion gain of 5.5 dB at a DC power consumption of 17 mW.

Microscopic photographs of both circuits are shown in Fig. 5. Both circuits were designed and fabricated for the same  $0.13 \mu\text{m}$  BiCMOS technology. The semiconductor fabrication process was carried out by IHP [12]. The fundamental models provided by IHP were replaced by optimized geometry scalable models also derived within HAEC [13].

Fundamentally, the electrical characteristics (such as output power for a given power consumption and speed) of the analog high-frequency frontend circuitry and, hence, of the overall system are determined by the selected semiconductor process technology and their specific generations, which all differ widely in cost and performance. For example, using a faster

technology often allows saving energy at the same system performance, as it is illustrated in Fig. 6. Therefore, it is planned to explore and compare the impact of other process technologies like InP on circuit and system performance. To achieve this, geometry scalable models were developed and extended, already showing very good agreement with actual measurements [14]. Based on those works, existing models will be extended. This is especially necessary for InP heterojunction bipolar transistors. Furthermore, the fabrication of test chips is planned to allow verification and validation of the models up to 325 GHz.

### D. Analog-to-Digital Converter (ADC) and Modulation

The analog-to-digital conversion is realized as a single core 3 bit ADC in a digital CMOS technology [15]. The ADC achieves a rate of 24 GSa/s with 2.2 equivalent number of bits. Conversion of a 800 mV differential input signal is carried out with a DC power consumption of 406 mW from 1.4 V and 1.75 V supplies. Assuming a two-cycle conversion, a latency of 83 ps is achieved. The fabricated ADC is shown in Fig. 7.

The energy consumption normalized to the data rate is approximately 5.6 pJ/bit. To achieve 100 Gbit/s, interleaving four ADCs will be necessary. This results in a DC power consumption of more than 1.6 W. This number was a motivation to investigate other approaches, to further maintain the ambitious energy efficiency goals. One alternative approach is the utilization of 1 bit quantization and oversampling. Promising theoretical results have been published in [16]–[18].

Regarding modulation it has been shown that 3.5 bit per channel use can be achieved by applying faster than Nyquist signaling when considering root raised cosine matched filtering with i.i.d. QPSK input symbols and a symbol-by-symbol detection. In this case, the roll-off factor was 1 and the sampling and signaling rate was increased by a factor of 1.75 with respect to the Nyquist rate.

On the hardware implementation side, this requires a different concept. Simply increasing the sampling frequency up to 100 GHz to achieve the oversampling will result in massive demultiplexing to convert the very fast serial digital signal to parallel streams that can be processed with digital hardware. This additional process step will result in a comparable DC power consumption. Given the nature of the signal,

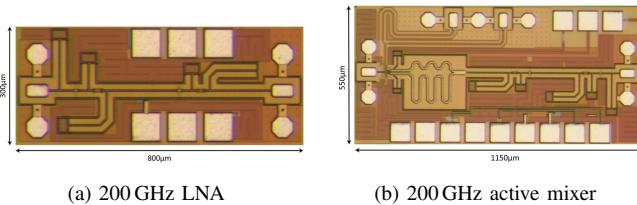


Figure 5. Chip photographs of integrated circuits.

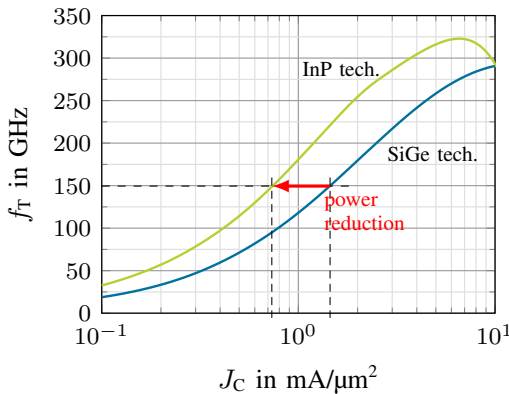


Figure 6. Illustration of energy saving by means of a faster semiconductor technology; comparison of 350 GHz InP and 300 GHz SiGe technology. Assuming 150 GHz  $f_T$  requirement translates to 50% energy saving.

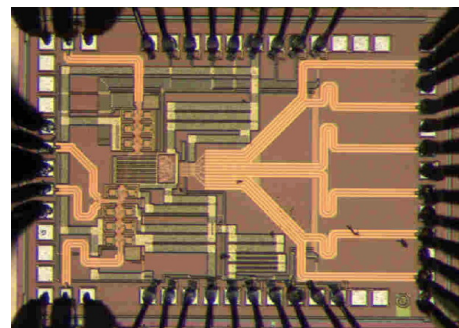


Figure 7. Fabricated 24 GSa/s ADC.

which carries information in the transition time of the binary signal, alternative architectures will be investigated, borrowing concepts from the time-to-digital converters, which are typically employed in digital phase locked loops.

### E. Coding

To compete with the state-of-the-art memory access delays, the delay requirement of  $<100$  ns of the wireless communications links is critical. The main contribution to the link latency is due to the channel coding. In this regard, research focus was put on the analysis of the decoding algorithms for Low-Density Parity-Check Convolutional (LDPC) codes with stringent constraints on latency. In [19], [20] it was demonstrated that using a convolutional version of the LDPC code, compared to LDPC block codes a gain in terms of Bit Error Rate (BER) can be achieved. This was even achieved under a sub-optimum message passing decoding at the structural latency of less than 500 bit. Although convolutional codes have traditionally been considered to be suitable for latency constrained applications, results presented in [19] show that LDPC codes outperform the convolutional codes for equal latency. Fig. 8 shows the required  $E_b/N_o$  to achieve an error rate of  $10^{-5}$  against various values of latency, calculated in terms of information bits. The results are obtained by using finite length codes and Monte Carlo simulations. LDPC codes outperform the convolutional and block codes for the latency value of our interest, i.e., 500 bit. Considering a transmission rate of 100 Gbit/s, a 500 bit latency corresponds to a structural latency of 5 ns. In order to meet the 100 ns latency requirement on the wireless links, a window length of 1000 bit corresponding to 10 ns is considered. The remaining 90 ns are reserved for the processing.

### F. Packaging

Integrating the different transceiver components that are fabricated in different technologies is a challenging task. A possible solution is the integration of mm-wave components on or in one silicon wafer as shown in [21]. This approach for heterogeneous integration offers the flexibility of combining various technologies on one common substrate that can be processed in consecutive steps. Considering the ambitious goals

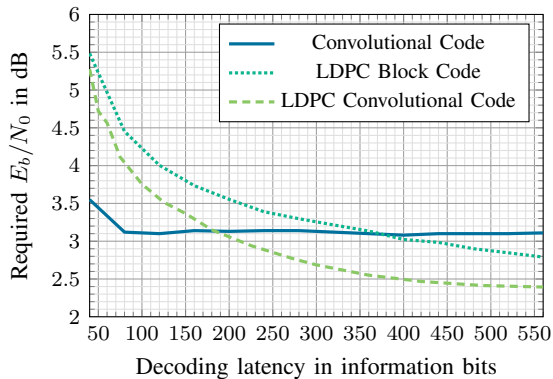


Figure 8. Comparison of Codes with respect to delay vs. required energy per bit.

of HAEC, a horizontal integration is likely to be insufficient. Fig. 9 depicts the structure of transmitter components within the chip stack. The antennas and the Butler matrix are likely to be implemented in a polymer based technology for cost and performance optimization, since they can be fully designed using passive elements. Amplifiers are going to be realized in a BiCMOS technology, and the digital domain in a CMOS one. The power divider can either be done in a polymer based technology or a semiconductor one. The interconnection of those layers will have to be done by vertical vias that are optimized for energy efficient RF performance.

### III. ESTIMATED ENERGY CONSUMPTION

Since several components of the wireless links are still under development, it is very challenging to give an accurate estimate of the energy consumption of the wireless links per transmitted bit. Nevertheless, in the following we will estimate the energy consumption based on measurement results acquired for the components developed so far and based on predictions for components which have not been designed yet. Our aim is to achieve a data rate of 100 Gbit/s while fulfilling the latency requirement of 100 ns.

For the calculation of the link budget we consider the parameters given in Table I. Here, we will exemplarily assume the use of 2D antenna arrays with  $16 \times 16$  antennas elements which are fed by a Butler matrix switching network with a predicted loss of 20 dB. For the energy estimation we consider as a worst case the longest link between two printed circuit boards, for which we assume a distance of 14 cm. Moreover, we consider a channel bandwidth of 30 GHz and receive SNR of 15 dB at the ADC for a data rate of 50 Gbit/s (100 Gbit/s will be achieved by two orthogonal polarizations). With the parameters in Table I we thus have to use a transmit power of 6 dBm. Under these assumptions the power consumption of the analog part of the transmitter without the Digital-to-Analog Converter (DAC) is estimated to be 240 mW. For the analog part of the receiver without ADC the power consumption is estimated to be 110 mW. In addition, complex modulation and a channel coding rate of  $1/2$  is considered. This corresponds to two real data streams with a coded bit rate of 50 Gbit/s each, which can be achieved by employing 1 bit quantization with oversampling. For the new concept for 1 bit ADC as previously proposed, the estimated power consumption is 95 mW for 4-



Figure 9. Envisioned structure of the transmitter components in a chip stack. Different background colors indicate different technologies.

Table I  
LINK BUDGET PARAMETERS FOR BOARD-TO-BOARD COMMUNICATIONS  
USING 16×16 ANTENNA ARRAYS

	Unit	Value
RX noise figure	dB	15
Path loss exponent	–	2
Path loss for shortest link 0.025 m (232.5 GHz)	dB	47.7
Path loss for largest link 0.143 m (232.5 GHz)	dB	62.9
Array gain	dB	24
Butler matrix inaccuracy	dB	20
Polarization mismatch	dB	3
Implementation loss	dB	5
RX temperature	K	323

times oversampling for I- and Q-component each (totaling to 190 mW). The corresponding DAC is estimated to use the same power. This yields an energy consumption of 14.6 pJ/bit. The estimated energy for channel decoding is 13.35 pJ/bit using 5 decoding iterations. Finally, some margin for a demapping and demultiplexing process is added. Their complexity corresponds at maximum to one channel decoding iteration. Thus, the overall estimated energy consumption is 30.62 pJ/bit.

#### IV. CONCLUSION

In this paper, we presented a wireless transceiver structure for computer board to board communication that is currently under development. This architecture successfully addresses the technical challenges of realizing energy efficient wireless links with data rates up to 100 Gbit/s. Based on actual measurements of successfully developed components and the inclusion of alternative approaches to further improve energy efficiency, we were able to estimate the energy consumption of such a link. Although the energy consumption is higher than in comparable waveguide-based optical links, the wireless links are superior in terms of flexibility as they can be dynamically established between nodes whenever needed.

#### ACKNOWLEDGMENT

This work is supported by the German Research Foundation (DFG) in the Collaborative Research Center 912 "Highly Adaptive Energy-Efficient Computing" [2].

#### REFERENCES

- [1] "The Zettabyte Era – Trends and Analysis," Cisco Systems, Inc., San Jose, CA, White Paper, 2014. [Online]. Available: [http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/VNI\\_Hyperconnectivity\\_WP.pdf](http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/VNI_Hyperconnectivity_WP.pdf)
- [2] SFB 912, "Highly Adaptive Energy-Efficient Computing." [Online]. Available: <http://tu-dresden.de/sfb912>
- [3] M. Jenning and D. Plettemeier, "180 GHz on-chip integrated bow-tie antenna," in *IEEE International Symposium on Antennas and Propagation*, July 2014.
- [4] B. Klein, M. Jenning, P. Seiler, and D. Plettemeier, "Wideband half-cloverleaf shaped on-chip antenna for 160 GHz – 200 GHz applications," in *IEEE Antennas and Propagation Society International Symposium*, July 2014, pp. 364–365.

- [5] R. Hahnel, B. Klein, and D. Plettemeier, "Integrated stacked vivaldi-shaped on-chip antenna for 180 GHz," in *IEEE International Symposium on Antennas and Propagation and North American Radio Science Meeting*, July 2015.
- [6] M. Jenning and D. Plettemeier, "Miniaturized integrated butler matrix for 180 GHz chip-to-chip communication," in *International Workshop on Antenna Technology*, 2014.
- [7] —, "1 x 4 antenna array for chip-to-chip communication at 180 GHz," in *International Conference on Electromagnetics in Advanced Applications*, August 2014.
- [8] J. Israel, J. Martinovic, A. Fischer, M. Jenning, and L. Landau, "Optimal antenna positioning for wireless board-to-board communication using a butler matrix beamforming network," in *International ITG Workshop on Smart Antennas*, 2013.
- [9] W.-Y. Chen, Y.-R. Hsieh, C.-C. Tsai, Y.-M. Chen, C.-C. Chang, and S.-F. Chang, "A compact two-dimensional phased array using grounded coplanar-waveguides butler matrices," in *European Radar Conference*, Oct 2012, pp. 421–424.
- [10] D. Fritsche, C. Carta, and F. Ellinger, "A broadband 200 GHz amplifier with 17 dB gain and 18 mW dc-power consumption in 0.13  $\mu$ m SiGe BiCMOS," *IEEE Microwave and Wireless Components Letters*, vol. 24, no. 11, pp. 790–792, Nov 2014.
- [11] D. Fritsche, J. D. Leufker, G. Tretter, C. Carta, and F. Ellinger, "A low-power broadband 200 GHz down-conversion mixer with integrated LO-driver in 0.13  $\mu$ m SiGe BiCMOS," *IEEE Microwave Wireless Components Letters*, 2015, accepted for publication.
- [12] "Low-volume & multi-project service," IHP GmbH – Innovations for High Performance Microelectronics/Leibniz-Institut für innovative Mikroelektronik, Tech. Rep., 2015. [Online]. Available: <http://www.ihp-microelectronics.com/en/services/mpw-prototyping/sigec-bicmos-technologies.html>
- [13] A. Pawlak and M. Schroter, "An improved transfer current model for RF and mm-wave SiGe(c) heterojunction bipolar transistors," *IEEE Transactions on Electron Devices*, vol. 61, no. 8, pp. 2612–2618, 2014.
- [14] T. Nardmann, P. Sakalas, F. Chen, T. Rosenbaum, and M. Schroter, "A geometry scalable approach to InP HBT compact modeling for mm-wave applications," in *IEEE Compound Semiconductor Integrated Circuit Symposium*, 2013, pp. 1–4.
- [15] G. Tretter, M. M. Khafaji, D. Fritsche, C. Carta, and F. Ellinger, "A 24 GS/s single-core flash ADC with 3 bit resolution in 28 nm low-power digital CMOS," in *RF Integrated Circuits Symposium*, 2015.
- [16] L. Landau and G. P. Fettweis, "Communications employing 1 bit quantization and oversampling at the receiver: Faster-than-nyquist signaling and sequence design," in *submitted to IEEE International Conference on Ubiquitous Wireless Broadband*, 2015.
- [17] L. Landau and G. Fettweis, "Information rates employing 1 bit quantization and oversampling at the receiver," in *International Workshop on Signal Processing Advances in Wireless Communications*, June 2014, pp. 219–223.
- [18] T. Halsig, L. Landau, and G. Fettweis, "Information rates for faster-than-nyquist signaling with 1 bit quantization and oversampling at the receiver," in *IEEE Vehicular Technology Conference*, May 2014, pp. 1–5.
- [19] N. ul Hassan, M. Lentmaier, and G. P. Fettweis, "Comparison of LDPC block and LDPC convolutional codes based on their decoding latency," in *International Symposium on Turbo Codes and Iterative Information Processing*, 2012, pp. 225–229.
- [20] G. P. Fettweis, N. ul Hassan, L. Landau, and E. Fischer, "Wireless interconnect for board and chip level," in *Design, Automation Test in Europe Conference Exhibition*, March 2013, pp. 958–963.
- [21] E. Topak, J.-Y. Choi, T. Merkle, S. Koch, S. Saito, C. Landesberger, R. Faul, and K. Bock, "Broadband interconnect design for silicon-based system-in-package applications up to 170 GHz," in *European Microwave Conference*, Oct 2013, pp. 116–119.