A 1.25pJ/b 73Gb/s/ch 210GHz Transceiver Front-End for Wireless 3D IC Thru-Silicon Interface

Chen Jiang, Xiaohan Shen, Xiaodi Feng, Chixiao Chen and Qi Liu
State Key Lab of Integrated Chips and Systems, Frontier Institute of Chip and System, Fudan University
Shanghai 200433, China
Email: cjiang@fudan.edu.cn

***Abstract*—The 3D integration technique brings significant density and efficiency gains for high-performance computing by stacking memory and logic dies. However, the TSV-based solution suffers from high complexity and cost, yield limitations, as well as thermal and reliability challenges. This paper introduces a novel THz through-silicon interface as an alternative solution for 3D integration. Implemented in 28nm CMOS, two transceiver designs are presented: a QPSK transceiver achieving 73Gb/s with 1.25pJ/b efficiency, and a PAM-4 transceiver achieving 71.4Gb/s with 1.22pJ/b efficiency, demonstrating the potential of THz wireless interfaces for high-bandwidth, energy-efficient 3D integration.**

I. Introduction

As the pace of Moore's Law slows, chiplet technology has emerged as a crucial and transformative approach for driving continued innovation in the semiconductor industry By stacking of memory and logic dies, the 3D integration significantly enhances the scale, density and interconnect efficiency of the HPC and AI processors. The TSV-based interface enables vertical data links among dies, however it suffers from costly and complicated fabrication process, high yield loss, as well as reliability and thermal issues. Therefore, it's only economically justifiable for premium products. Wireless schemes relying on electro-magnetic coupling is promising to eliminate the expensive TSVs. Prior works have demonstrated vertical links based on inductive coupling [1,2], however this scheme cannot support consecutive identical digits due to the inherent high-pass filtering characteristic of inductive channels. Furthermore, these links are sensitive to environmental noise and interferences. To mitigate these drawbacks, the modulation scheme is an option [3]. To achieve higher data rates and interconnect density, pushing the carrier frequency higher is a natural trend. This paper presents a high-speed low-cost wireless 3D IC thru-silicon interface using THz coupling, which provides abundant bandwidth and short wavelength to achieve competitive energy and area efficiency. This scheme significantly simplifies the fabrication and relaxes the alignment requirement by 10~100x, which helps to maintain high yield and low cost. It is also highly scalable, which is critical to provide a Tbps-level total bandwidth for cutting-edge HPC chiplet systems. Moreover, it provides greater flexibility for reconfigurable interconnect topologies, and supports a wider range of thermal management solutions, paving the way for large scale 3D integration applications.

To balance available bandwidth and energy efficiency, a carrier frequency of 210 GHz is selected. To demonstrate the feasibility, QPSK and PAM-4 transceivers are designed in a standard 28nm CMOS technology. The QPSK transceiver can support up to a 73Gb/s data rate with a 1.25pJ/b energy efficiency, while the PAM-4 transceiver can provide up to a 71.4Gb/s data rate with a 1.22pJ/b energy efficiency.

This work is supported by the National Natural Science Foundation of China under Grant 92473101 and 62341409.

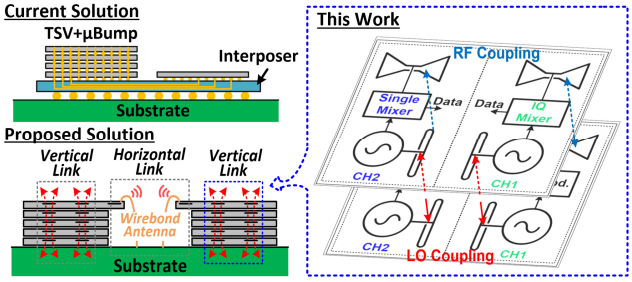

Fig. 1. Concept of the THz wireless 3D IC interface, as well as the proposed RF and LO coupling scheme.

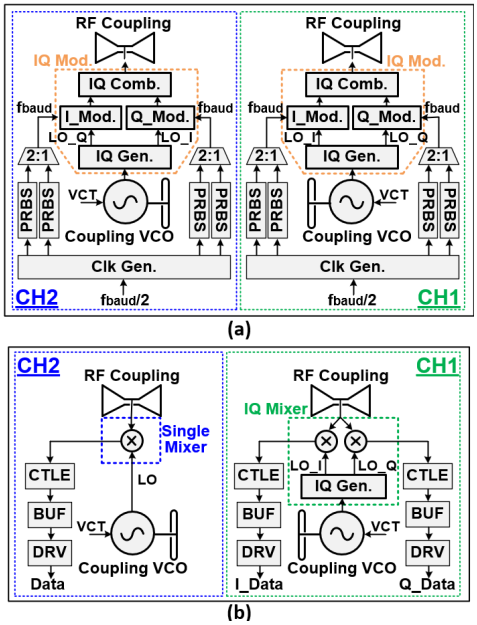

Fig. 2. Structures of the prototype: (a) the transmitter front end, and (b) the receiver front end.

II. Architecture of the Terahertz Thru-Silicon Interface Transceiver

Concept of the THz wireless 3D IC interface, as well as the coupling scheme between the dies is shown in Fig. 1. Detailed structures for the transceiver front-ends are shown in Fig. 2. In this design, two independent channels are implemented. In the TX, a VCO is used to generate the LO signal, which is sent to the RF coupling antenna after a QPSK modulation. On-chip PRBS circuits are implemented for convenience of validation. For the RX, two topologies are implemented and compared: quadrature mixer (CH1) and single mixer (CH2), both incorporate CTLE, buffer, and extra driver stages for the 50Ω oscilloscope load. Energy efficiency and interconnect density are crucial in this application, therefore, instead of relying on power and area consuming frequency synthesizers, LO synchronization is obtained with wirelessly coupled VCOs. As shown in Fig. 1,

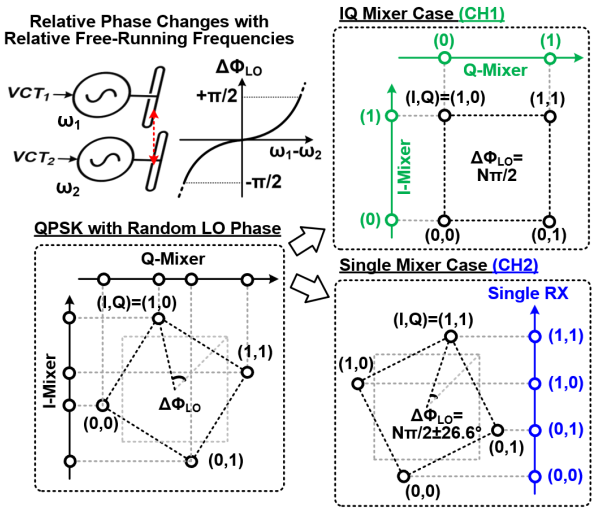

Fig. 3. The relative phase of the TX and RX LO is controlled by adjusting the free-running frequencies of the VCOs. Specific phase relationships are established to optimize the links' implementation.

the LO and RF are coupled with orthogonally-placed linear-polarized antennas to avoid unwanted interference. As shown in Fig. 3, under mutual injection, the relative phase between the TX and RX VCOs is tunable by changing their relative free-running frequency [4]. The theoretical tuning range exceeds ±90°. In CH1, for the quadrature RX, the relative LO phase is tuned (e.g., TX and RX LO in phase) so that the I and Q subchannels can receive independent NRZ signals with minimized crosstalk. Therefore, simple data slicers, instead of power-consuming high-speed ADCs, can digitize the signal. In CH2, for the single-mixer RX, the relative phase is tuned (e.g., RX LO lags TX by 26.6°) so that the rotated QPSK constellation points will be equally-spaced after projected to the single mixer output, forming a PAM-4-like waveform. This scheme achieves the same data rate while eliminating the quadrature LO generation and RF power splitting. This allows for increased LO and RF power at the mixer, boosting SNR for PAM-4 signaling. The simpler design also reduces power and area.

III. Circuit Design Details of the Prototype

A. The 210-GHz Wireless Coupling VCO

Structure of the 210-GHz wireless coupling VCO is shown in Fig. 4. The VCO oscillates at a fundamental frequency of 105GHz, and the second harmonic is extracted and used as the LO. To optimize the harmonic power generation, the return-path gap coupler based self-feeding structure [5] is adopted. The structure is modified to allow vertical coupling with the VCO on the other die. At the fundamental frequency, standing waves form in the top λ/4 slot, which gives rise to the radiation and mutual injection of the VCOs. The bottom λ/4 slot is split and folded to save area and cancel the radiation effect [5]. The VCO frequency is tuned by changing its gate bias [6]. An output buffer is added to isolate the VCO from the modulator or mixer, so that the oscillation will not be impacted. According to simulation, under 38mW power consumption, the LO power at the buffer output is -2.6dBm.

B. The Proposed Low-Loss QPSK Modulator

Due to the limitations of silicon transistors, conventional mixers present high loss at this frequency. To minimize loss

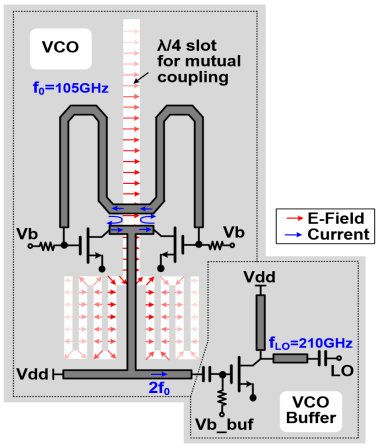

Fig. 4. The 210-GHz wireless coupling VCO structure.

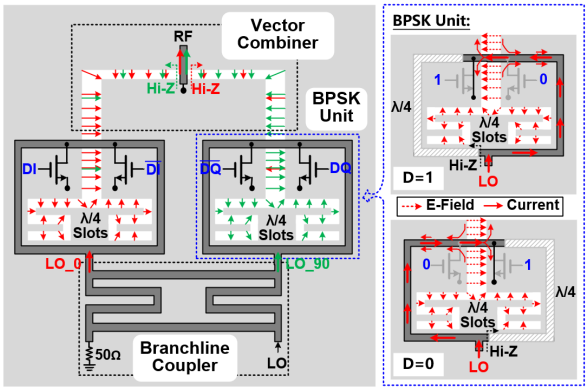

Fig. 5. The proposed low-loss QPSK modulator structure, and operation principle of the BPSK modulator unit.

and enhance spectrum utilization, as shown in Fig. 5, a new quadrature modulator topology is proposed, which is composed of a branchline coupler for quadrature LO generation, two BPSK modulator units for both I- and Q-path, as well as a vector combiner for IQ summation. The BPSK modulator unit is composed of a transmission line (TL) to slot mode conversion structure and two parallel NMOS switches. When D=1, the left λ/4 TL transforms the small switch ON impedance into a high impedance, directing most of input power through the right TL, which then induces a slot mode. When D=0, the same mechanism induces a slot mode with the opposite electrical fields. Therefore, relying on the parallel switches and the highly symmetrical structure, a low-loss high-quality BPSK modulation is achieved. The slot lines at the bottom of the BPSK units are λ/4 long to force the wave to propagate along the top slots. In the vector combiner, the slot line on the left half is designed to transform the I-channel BPSK modulator output into a high impedance, so that most of the Q-channel signal power will be converted back to TL mode at the IQ joining point instead of entering the I-channel modulator structure. With the same mechanism for I-channel signal and law of superposition, a vector summation of the IQ signals is achieved. Therefore, a QPSK modulation is obtained, which doubles the achievable total data rate. The simulated modulation loss including the branchline coupler is 7.4dB at 210GHz without consuming any dc power.

C. The Proposed Low-Noise Receiver Mixer

Enhancing the conversion gain and noise figure is the key for mixer design. Since the gate of the transistor presents a higher impedance than the source, under the same

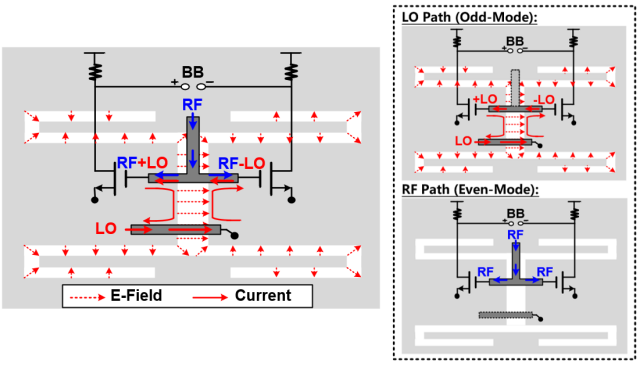

Fig. 6. The proposed receiver mixer structure, and concept of the LO and RF signal distribution under different modes.

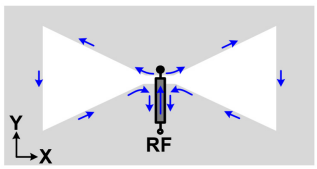

Fig. 7. The coupling antenna structure.

power level, compared to injecting LO and RF separately from the gate and source, it is more efficient to inject both from the gate to produce a higher swing, thus a larger nonlinearity to create a better conversion gain. Simulation suggests ~3.5dB NF improvement and ~4dB less LO power requirement using this scheme. Since the RF and LO signals are uncorrelated, using traditional power combiners will cause 3-dB power loss. To avoid this loss, a new mixer topology is proposed, as shown in Fig. 6, which is based on a slot balun structure [7,8]. Injected through the balun, the LO signal will excite odd mode at the gate of the transistors. The RF signal is injected from the center tap, which will excite even mode at the gate. Since the center tap is a virtual ground under the odd mode, it will not impact the LO path. Relying on superposition of the modes, the RF and LO signals are combined with minimized loss and fed into the mixer gates. Following the mixer, a CTLE and TIA are designed. To drive the 50Ω load of the oscilloscope, additional output buffers are inserted. Under a 1.7mW power consumption for each mixer cell, the simulated NF of the whole quadrature and single RX chain is 21.6dB and 16.7dB, respectively.

D. The RF Coupling Antenna

A bow-tie shaped slot antenna is used to achieve a larger bandwidth. The simulated 3dB bandwidth of the coupling is more than 60GHz, which is large enough to support a 100Gb/s level data rate. The simulated coupling loss is 4.4dB and 9.8dB for die thickness of 50µm and 100µm, respectively. The simulated tolerance for misalignment is ~100µm, providing robustness against fabrication variations.

IV. Experimental Results

The prototype is implemented in a standard 28nm CMOS technology, conforming to all metal density rules. Manually inserted metal dummies in the THz structures have a minor impact on the link performance. The die photos are shown in Fig. 8, and the assembled sample is shown in Fig. 9. For probing convenience, the RX die is placed on top of the TX die in the sample. A dummy die is placed on top to maintain a similar EM environment for the THz circuits. Limited by

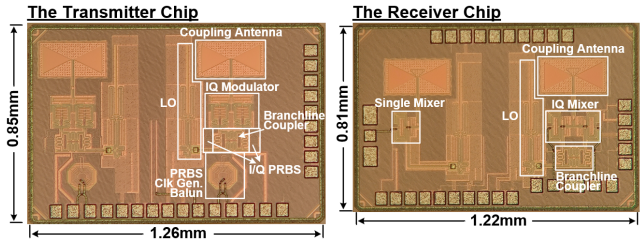

Fig. 8. Microphotograph of the transmitter and receiver chips.

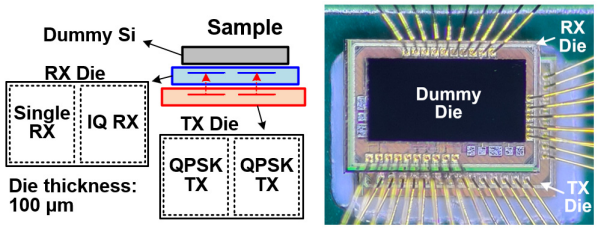

Fig. 9. Overview and photo of the prototype.

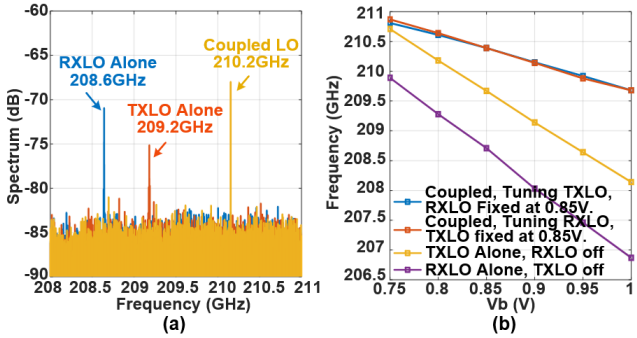

Fig. 10. The measured LO spectrums and frequency tuning ranges.

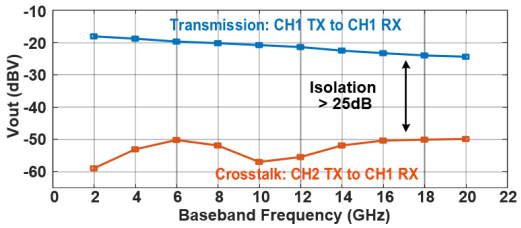

Fig. 11. The measured isolation between the two channels.

our packaging facility capability, all dies are only thinned to 100µm. According to simulation, further thinning to 50µm could improve the coupling loss by ~5dB. Adding testing structure to directly measure the THz signal would impact the performance. Instead, the measurement is done by capturing the leaked radiation with a horn antenna above the sample. A harmonic mixer then down-converts the signal to below 10GHz for spectrum analyzer measurement. Fig. 10 shows the measured spectrum. With only TX or RX on, the VCO free-running frequencies are 209.2GHz and 208.6GHz, respectively. When locked, the frequency is 210.2GHz. Fig. 10 also shows the measured frequency tuning ranges, which shows the VCOs are frequency locked over the whole range. The measured isolation between the two links is 25~35dB up to a 20GHz baseband frequency (shown in Fig. 11), ensuring low enough crosstalk for both NRZ and PAM-4 signals. To facilitate the time-domain measurements, high-speed PRBS circuits are implemented on the TX die. A DC-50GHz RF probe is used to acquire the RX outputs. A real-time oscilloscope is used for analyzing the link quality. The measured eye diagrams and BERs under different conditions

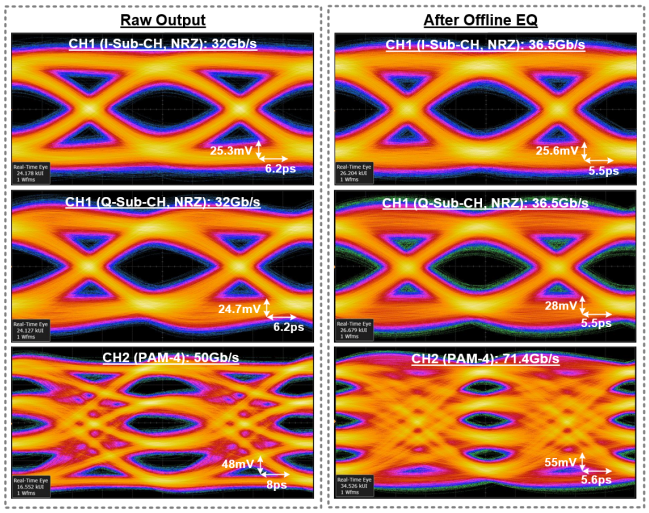

Fig. 12. The measured eye diagrams.

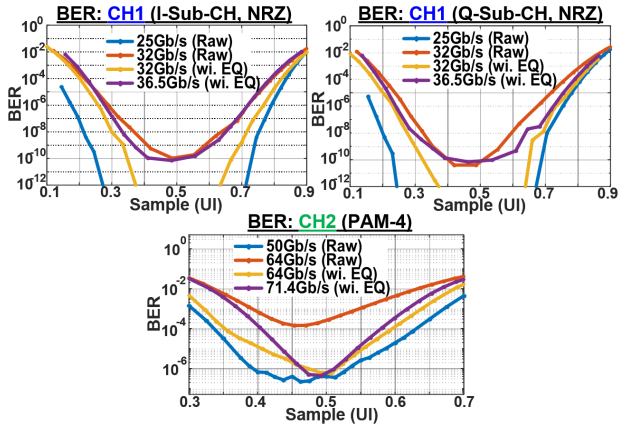

Fig. 13. The measured bit error rates.

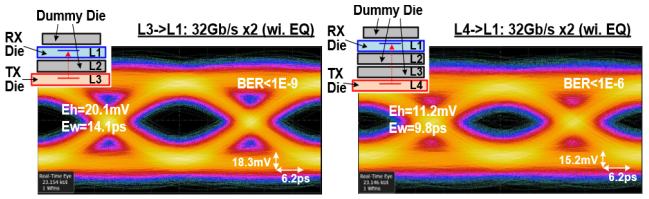

Fig. 14. The measured link performances under 3- and 4-die stack.

for both channels are shown in Fig. 12 and Fig. 13. Thanks to the large bandwidth achieved, the raw signal output can support a data rate up to 64Gb/s with a BER better than 10^{-10}. To further enhance the link performance, a light-weight 3-tap offline equalization FIR is applied. For CH1, a BER better than 10^{-10} is maintained up to a total data rate of 73 Gb/s at a 1.25pJ/b energy efficiency. For CH2, a BER better than 10^{-6} is maintained up to a data rate of 71.4Gb/s at a 1.22pJ/b energy efficiency. The prototype is further evaluated under more stack layers. As shown in Fig. 14, a good wireless link is achieved for 3-stack and 4-stack conditions. Under a 64Gb/s data rate, the measured BER is lower than 10^{-9} and 10^{-6}, respectively.

V. Conclusions

As demonstrated in TABLE I, thanks to the novel LO locking scheme and new circuit topologies proposed, this work achieves very competitive data rates, energy efficiency, and area efficiency compared to other state-of-the-art 3D integration solutions, including the TSV-based scheme. Furthermore, the wireless THz through-silicon interface offers significantly simplified fabrication, reduced cost, and enhanced reliability. To our best knowledge, this is the first demonstration of the THz wireless thru-silicon interface for 3D integration, showcasing its potential to revolutionize the large-scale 3D integration and enable innovative chiplet architectures.

TABLE I Performance Summary And Comparison.

	This Work		VLSI 2020 [1]	TCAS_I 2021 [2]	JSSC 2019 [3]	JSSC 2021 [9]	JSSC 2022 [10]
Interconnect Type	Wireless Modulated		Inductive Coupling	Inductive Coupling	Wireless Modulated(1)	TSV	TSV (Emulated)
Process	28nm CMOS		65nm CMOS	40nm CMOS	65nm CMOS	1y-nm DRAM	65nm CMOS
Carrier Freq. [GHz]	210		\	\	127	\	\
Modulation	QPSK (NRZx2)	PAM4	NRZ	NRZ	PAM4	NRZ	Duobinary
Data Rate [Gb/s]	36.5×2	71.4	1.27	3.6	20	5.0	5.25
Energy Efficiency [pJ/bit]	1.25(2)	1.22(2)	17.0	1.5	3.98	N/A	0.31(3)
Area Efficiency(4) [Gb/s/mm²]	471(2)	529(2)	20.3	2.5	33(1)	231	N/A
Die Thickness [μm]	100 (+10 glue)		70 (+10 glue)	8	N/A	N/A	20(3) (Emulated)
Integration Cost	Low		Low	Low	Low	Very High	Very High

1. For contactless interface between boards, the coupler on PCB is not counted in area efficiency.
2. Power consumption and area of the testing PRBS and RX output buffers is not counted.
3. On-chip emulated TSV. Equivalent die thickness is calculated with the size of the structure.
4. Area Efficiency = Data Rates / Area Overhead (includes TRX circuits, coupler, TSV, keep out regions, etc).

References

[1] B. J. Fletcher, T. Mak, and S. Das, "A 3D-stacked Cortex-M0 SoC with 20.3Gbps/mm² 7.1mW/mm² simultaneous wireless inter-tier data and power transfer," in IEEE Symp. on VLSI Circuits, Jun. 2020.

[2] K. Shiba, T. Omori, K. Ueyoshi et al., "A 96-MB 3D-stacked SRAM using inductive coupling with 0.4-V transmitter, termination scheme and 12:1 SerDes in 40-nm CMOS," IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 68, no. 2, pp. 692–703, Feb. 2021.

[3] Y. Kim, B. Hu, Y. Du et al., "A millimeter-wave CMOS transceiver with digitally pre-distorted PAM-4 modulation for contactless communications," IEEE J. Solid-State Circuits, vol. 54, no. 6, pp. 1600–1612, Jun. 2019.

[4] B. Razavi, "A study of injection locking and pulling in oscillators," IEEE J. Solid-State Circuits, vol. 39, no. 9, pp. 1415–1424, Sep. 2004.

[5] R. Han et al., "A SiGe terahertz heterodyne imaging transmitter with 3.3 mW radiated power and fully-integrated phase-locked loop," IEEE J. Solid-State Circuits, vol. 50, no. 12, pp. 2935–2947, Dec. 2015.

[6] C. Jiang, A. Cathelin, and E. Afshari, "An efficient 210GHz compact harmonic oscillator with 1.4dBm peak output power and 10.6% tuning range in 130nm BiCMOS," in IEEE Radio Freq. Integr. Circuits Symp., May 2016.

[7] C. Wang and R. Han, "Rapid and energy-efficient molecular sensing using dual mm-Wave combs in 65nm CMOS: A 220-to-320GHz spectrometer with 5.2mW radiated power and 14.6-to-19.5dB noise figure," in IEEE Int. Solid-State Circuits Conf., Feb. 2017.

[8] C. Jiang et al., "A 301.7-to-331.8GHz source with entirely on-chip feedback loop for frequency stabilization in 0.13μm BiCMOS," in IEEE Int. Solid-State Circuits Conf., Feb. 2018.

[9] K. C. Chun, Y. K. Kim, Y. Ryu et al., "A 16-GB 640-GB/s HBM2E DRAM with a data-bus window extension technique and a synergetic on-die ECC scheme," IEEE J. Solid-State Circuits, vol. 56, no. 1, pp. 199–211, Jan. 2021.

[10] J.-Y. Kim, J. Lee, K. Kim et al., "A 5 Gb/s time-interleaved voltage-mode duobinary encoding scheme for 3-D-stacked IC," IEEE J. Solid-State Circuits, vol. 57, no. 6, pp. 1913–1923, Jun. 2022.