Design Considerations of Time-interleaved Discrete-time Beamformers Toward Wideband Communications

Chung-Ching Lin, *Member, IEEE,* Qiuyan Xu, *Graduate Student Member, IEEE,* Huan Hu, *Member, IEEE,* and Subhanshu Gupta, *Senior Member, IEEE,*

***Abstract*—Efficient exploitation of wide bandwidth data communication requires antenna array providing high gain across all frequency components for both transmit and receive equipment. In contrast to the frequency-dependent phased array, true-time-delay (TTD) arrays are appealing yet insufficiently investigated alternative for both fast initial access (IA) process and wideband directional data communications. In this brief, the mathematical relationship of delay before and after the frequency conversion is first discussed, which lays a foundation of the TTD beamformed system followed by a step-by- step design procedure and its tradeoff. Finally, system-level analysis is presented to estimate the minimum interleaving factor and silicon area for both IA and data communications associating the circuit design tradeoffs. The three-part analysis aims to provide a quick starter guide to design a TTD array.**

***Index Terms*—Wideband beamforming, true-time-delay, time interleaving, discrete-time signal processing.**

I. Introduction

Wideband beamformed yet energy-efficient solutions are highly desirable for next generation communication system to enable high speed data processing. State-of-the-art beamformed systems are, however, limited in the adoption of phase shifter [1], [2]. This is because that the frequency-dependent response sets a limit on their operational fractional bandwidth. TTD technique has been widely used to replace the phase shifter, striving to achieve beam-squint free data receiving, targeting for wide fractional bandwidth operations. On the other hand, TTD technique can also be applied for beamtraining to achieve low-latency initial access process in the same beamformed system [3]. Compared to most of the existing TTD approaches using passive elements [4], all pass filter [5], and digital approach [6], baseband (BB) TTD using sample-and-hold-circuit [3], [7]–[9] has attracted attention owing to its compact and digital friendly nature. Though operation mechanism and measurement results have been demonstrated to prove the efficacy of large delay range and fine resolution in TTD arrays [3], [10], [11], the circuit design considerations and trade-off analysis are insufficiently addressed. Additionally, any circuit functionalities should fit to a specific system needs and always make compromise between performance, and form-factor. However, system-level considerations has merely being quantified and discussed. This paper aims to address these gaps from circuit/component parameter selection for a scalable architecture.

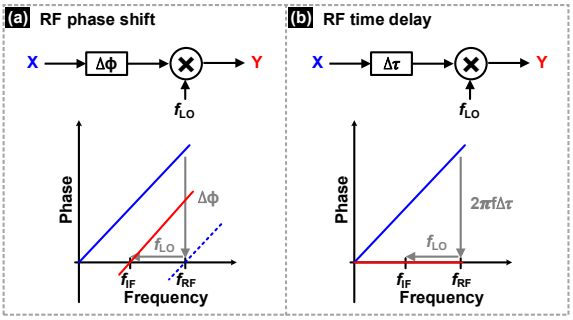

Fig. 1. Signal compensation at RF using (a) phase shifter; and (b) time delay unit.

The rest of the paper is organized as follows. Section II summarizes the mathematical derivation for the delay effect at the presence of frequency mixing. Section III discusses and analyzes the design procedures and tradeoffs of the BB TTD approach for beamtraining and beamforming. The scalability discussion is presented in Section IV, followed by conclusions and future work in Section V.

II. BB TTD and Equivalency to RF TTD

Considering a single channel of a beamformed system, the phase shifter compensation scheme is shown in Fig. 1(a) causing beam squint on the band edges [7]. Delay compensation poses several design challenges when performed at RF frequencies to solve this beam squint. First, scaling the carrier frequency from sub-6-GHz to mmW bands limits the achievable delay range that a unit delay cell can provide. Second, the fine delay tuning (resolution) is hard to control at such a high frequency. Third, as one move from one frequency to another, the delay cell has to be redesigned to meet linearity, noise, and power consumption.

Alternatively, compensation can be performed after the down-conversion as shown in Fig. 2 (a)-(c) [12]. Similar to RF phase shift, IF phase shift also experiences beam squint due to non-constant phase response at the output Y [12]. If we

This work is support from Semiconductor Research Corporation (SRC) through the Center for Design of Analog Digital Integrated Circuits (CDADIC) under Task. 2550.067 and National science foundation (NSF) Career Award (1944688). The first two authors contributed equally to this work.

C. Lin, Q. Xu, and S Gupta are with the School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA 99164 USA (e-mail: subhanshu.gupta@wsu.edu).

H. Hu was with the School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA 99164 USA, He is now with the School of Microelectronics, Southern University of Science and Technology, Shenzhen 518055, China (e-mail: huh3@sustech.edu.cn).

This article has been accepted for publication in IEEE Transactions on Circuits and Systems--II: Express Briefs. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCSII.2023.3282628

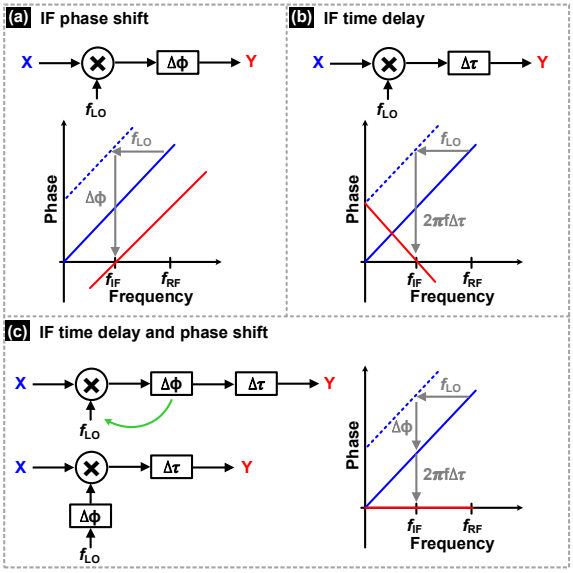

Fig. 2. Signal compensation at IF with (a) phase shift; (b) time delay; and (c) time delay with phase shift.

compensate delay at IF, the received signal Y can be expressed as:

$$\begin{aligned} Y &= X \times e^{-j2\pi f_{\mathrm{LO}} t} \times e^{j2\pi(f-f_{\mathrm{LO}})\Delta\tau} \\ &= e^{j2\pi(f-f_{\mathrm{LO}})t} \times e^{j2\pi(f\tau_n + f\Delta\tau - f_{\mathrm{LO}}\Delta\tau)} \end{aligned} \tag{1}$$

where $X = e^{j2\pi f(t+\tau_n)}$ denotes input signal, and τ_n is the delay of the n^{th} channel. Referring to (1), zero phase condition only happens at a single frequency f_{O} with $\Delta\tau$ as:

$$\Delta\tau = -f_{\mathrm{O}}/(f_{\mathrm{LO}} + f_{\mathrm{O}}) \times \tau_n \tag{2}$$

This result draws a conclusion that applying delay at IF only is insufficient and introduces beam squint eventually. To solve the beam squint issue, an additional phase shift is required as shown in Fig. 2(c). The received signal with both phase shift and time delay at IF is expressed as [12]:

$$Y = X \times e^{-j2\pi f_{\mathrm{LO}} t} \times e^{j2\pi(f-f_{\mathrm{LO}}\Delta\tau)} \times e^{-j\Delta\phi} \tag{3}$$

A frequency-independent output is obtained by combining both IF delay ($\Delta\tau = -\tau_n$) and phase shift ($\Delta\phi = 2\pi(f - f_{\mathrm{LO}}) \times \tau_n$) [12]. Furthermore, the phase shift can be moved to the LO side, further simplifying the design of the signal chain. The aforementioned explanation indicates the IF time delay with a corresponding phase shift has an ability to align the phase not only at the single frequency but also at the band edges of the received signal. This key conclusion opens new opportunities to implement large delay range-to-resolution ratio in the baseband. In Section-III, we will analyze and discuss a beamforming architecture that uses the baseband delay in the implementation.

III. Analysis and Implementation of Interleaved Switched Capacitor TTD Array

Discrete-time delay unit has been widely used for its wide achievable delay range capability and its possible implementations are illustrated in Fig. 3(a)-(d) using switched capacitor circuits. Shown in Fig. 3(a), the key concept to create delay is to sample a continuous time signal (under Nyquist theorem) onto a capacitor (i.e., C_{S}). Then, the release time of the second switch (right) is controlled to transfer the charge to the

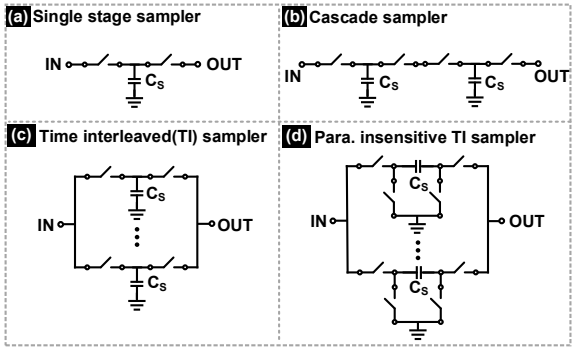

Fig. 3. Various approaches of time delay unit design using switched capacitor circuits

subsequent stages. A larger delay can be achieve by cascading multiple switched-capacitor circuits in Fig. 3(a) and constructs as Fig. 3(b). However, additional loss is introduced due to several series sampler [13]. The interleaved sampler [7] is shown in Fig. 3(c). The effective delay is determined by the number of interleaving levels. Nevertheless, the parasitic capacitance (i.e., contribute by the two switches) at the top plate of C_{S} degrades the signal-to-noise-and-distortion ratio (SNDR). An improved sampler with additional two switches are shown in Fig. 3(d), which mitigates the parasitic effect with improved SNDR performance [3].

Figure 4 shows a generic architecture for implementing N-element basedband TTD spatial signal processor (SSP) based on the time-interleaved (TI) parasitic insensitive sampler. Its design tradeoffs are also presented in the figure. In the subsequent section, the inclusion of the input buffer, summer design requirement, and the stray capacitance effect at the summer virtual node will be covered.

The input buffer is essential in implementing TTD circuits for two reasons. First, since the TTD implementation is based on switches and capacitor, the sampler input impedance, switch on-resistance, and sampling capacitor determines the input network time constant. Noted that the input impedance and the switch on-resistance are connected in series, their magnitude need to be minimized. Second, similar to any sample-and-hold system, the switching behavior introduce residual charge back towards the input. This undesired phenomenon also known as "kickback" should be minimized. A source follower is usually adopted for this because of its high input impedance and low output impedance which is inversely proportional to the source follower transconductance thus posing a tradeoff with power consumption at the first order. Though advanced input buffer designs have been investigated and applied in TI-ADCs [14]–[16], the tradeoff still presents in terms of spurious-free dynamic range (SFDR), input common mode range, common mode voltages, and silicon area. Note that the large current and high power supply are inevitable in these cases which take significant portion of the overall power consumption.

The noise performance of the switched-capacitor delay cell is desired to be dominated by the thermal noise from the sampling capacitor. Considering the signal-to-noise ratio (SNR) with a given signal power, the SNR is then determined by equating both the thermal and quantization noise. The

This article has been accepted for publication in IEEE Transactions on Circuits and Systems--II: Express Briefs. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCSII.2023.3282628

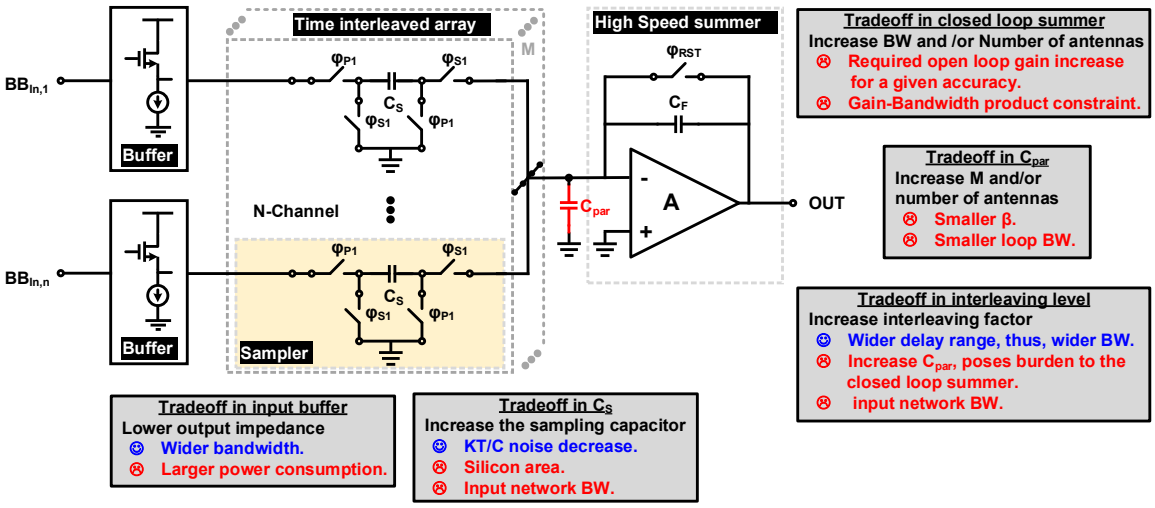

Fig. 4. Generic TTD SSP architecture and its design tradeoff

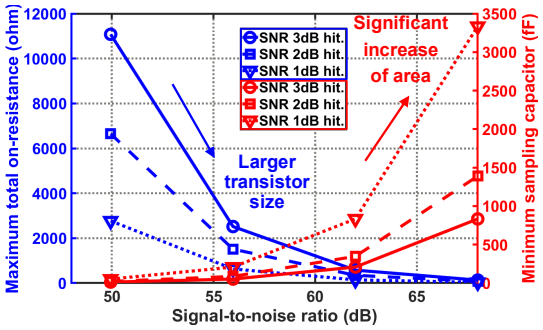

Fig. 5. Sampling capacitor and on-resistance design considerations at a given signal-to-noise ratio (SNR)

minimum sampling capacitor.$C_{S,min}$, can be expressed as:

$$C_{S,min} = m \times K \times T \times \frac{1}{(V_{LSB})^2} \tag{4}$$

where K is the Boltzmann constant, T is the temperature, and V_{LSB} is the smallest voltage level at a given resolution. The factor m represents SNR penalty after inclusion of the thermal noise, m can be computed as 12, 20, and 48 for 1, 2, and 3 dB SNR penalties respectively. The maximum allowable on-resistance, $R_{\mathrm{S,\,max}}$, is then decided by the settling requirement and can be expressed as (assuming the voltage is settled within 0.25 V_{LSB} accuracy):

$$R_{S,max} = \frac{1}{f_{in} \times ln(2^B \times 4) \times C_{S,min}} \tag{5}$$

where B is the resolution in bit, and f_{in} is the maximum input frequency under Nyquist criterion. Here, the design procedure can be summarized as follows:

1) Choose the sampling capacitor such that its thermal noise power level is less than quantization noise for a given resolution.
2) Size the switch on-resistance based on RC time constant calculation such that the settling error (ϵ) is acceptable. For instance, for a 10-bit design with an ϵ less than 0.25LSB, the required time constant is required to be 8.3 times larger than the inverse of the bandwidth-of-interest.
3) Plug the initial value in the simulator to perform transistor-level simulation.

Figure 5 further illustrates this tradeoff, considering a full scale input of 0.5V. As suggested, a larger sampling capacitor introduces less thermal noise to achieve a better SNR, but suffers from a reduced bandwidth if the on-resistance is not scaled proportionally. One may wonder if it's feasible to enlarge the transistor size to scale down the on-resistance. However, the associated junction capacitance starts to attribute nonlinear component, causing undesired distortion. Additionally, as mentioned earlier, the output impedance of the buffer is connected in series with the on-resistance further exacerbate the the on-resistance design selection. Thus, the noise-bandwidth relationship indicates another critical tradeoff here.

After signal at each of the channel gets delayed properly, an amplifier with capacitive feedback as summer transfers charges to the output stages. Assuming total of N channels and a feedback capacitor of C_F, feedback factor is calculated as $C_F/(C_F+NC_S)$. With an open loop bandwidth of BW_{O}, the closed-loop summer bandwidth will be its BW_{O} divided by its feedback factor. This indicates that for a larger array size, the bandwidth requirement increases proportionally. Additionally, amplifier design in general falls into gain-bandwidth product constraint. To break this constraint, advanced techniques with extra power consumption and silicon area are required.

It is also important to minimize the parasitic capacitance C_{par} at the virtual ground node of the summation amplifier. The C_{par} is contributed by the summer input capacitance, switch junction capacitance, and routing capacitance. The latter is more substantial when a large number of interleaving levels M (i.e., a longer achievable delay range) are implemented. A large C_{par} deteriorates the closed-loop bandwidth, posing challenges to the summer design. For example, C_{par} at the non-inverting input is mostly routing capacitance and transistor parasitic, while the output node capacitance is contributed from metal routing, transistor parasitic and feedback capacitor bottom plate capacitance. The total capacitance located at the inverting input node has extra C_S and feedback capacitor top plate parasitic capacitance, and is considered to be the largest among all three nodes. In general, C_{par} at the inverting input node has more pronounced effect on the summation amplifier

This article has been accepted for publication in IEEE Transactions on Circuits and Systems--II: Express Briefs. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCSII.2023.3282628

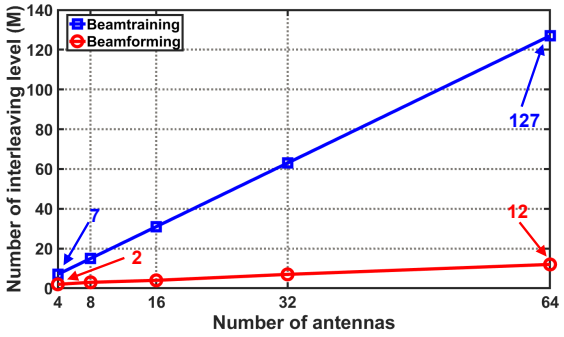

Fig. 6. Interleaving levels scale with number of antennas assuming BW/f_c = 0.2 and $\theta = \pm 60°$.

performance. Considering C_{par} within a closed-loop system, the feedback factor β becomes $C_F/(C_F{+}NC_S + C_{\text{par}})$. With C_{par} at the inverting input virtual ground node of the amplifier. As the feedback factor, β, decreases, gain and bandwidth of the amplifier decreases accordingly. Therefore it is important to consider parasitic caps in the amplifier design procedure.

The concept of discrete time TTD operation is inspired from the TI-ADC scheme, the calibration shall be included to minimize the inter-channel mismatches and DC offset. Here, few calibration examples can be done in either analog way [17] or digital domain [18] to minimize the mismatches and prevent the SNDR degradation.

IV. System and Application Level Tradeoff Consideration and Investigation

The critical steps in designing the discrete time beamformer are deciding the number of interleaving factor M always being driven by applications, fractional bandwidth, and available silicon area. This section aims to have a brief discussion about it and provide the reader with a quick guide in determining the factor M from system-level perspectives.

Based on different delay range requirement for specific functionality in the SSP, the interleaving factor considering beamtraining (i.e., initial access) and beamforming (i.e., data communication) can be determined as follows [3], [7]:

$$M_{\text{BT}} \geq 1 + 2(N-1) \tag{6}$$

$$M_{\text{BF}} \geq 1 + (N-1)sin(\theta)(BW/f_{\text{C}}) \tag{7}$$

where N is the array size, θ represents the angle-of-incidence, and BW/f_c is the fractional bandwidth. Now, let us explain the impact of M with respect to each parameters. Note, the discussion of beamtraining is focused on the one-dimensional linear array, interested reader can refer to [19], [20] for two-dimensional array discussion.

Assuming a fractional BW of 0.2, and θ of $\pm 60°$, Fig. 6 shows the minimum required interleaving levels for beamforming and beamtraining as the array size scales from 4 to 64. As expected, as the number of antenna increases, the total required delay range increases, so does the interleaving levels. Beamtraining requires larger delay and thus more interleaving levels compared to beamforming in a system with certain number of antennas. That is to say, if the designed hardware is to provide both functionalities to support both the beamtraining and beamforming modes, the interleaving level is determined by the beamtraining requirement.

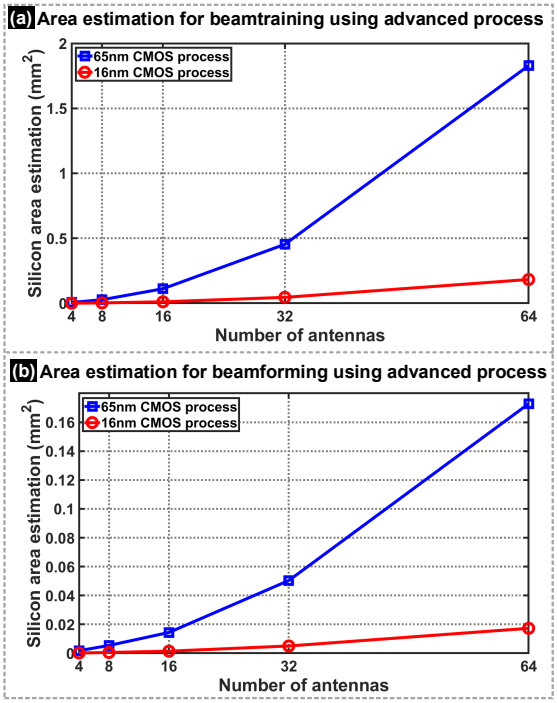

Fig. 7. Area estimation for both beamtraining and beamforming using proposed technique.

We further compare the area overhead with process scaling. Power consumption of analog/hybrid/digital TTD arrays have been compared in existing works indicating analog/hybrid arrays providing better energy savings [7], [10]. For the area analysis, we will considering the number of sampler increases proportional to the number of antenna elements and using a parasitic-insensitive switched-capacitor summer with four switches and one metal-insulator-metal (MIM) capacitor. For the sake of brevity, we assume the area is dominated by the size of the sampling capacitor and the switches are laid out under the MIM capacitor. The area for beamtraining and beamforming modes with a capacitor density of D_{MIM} (fF/μm^2) can thus be expressed as:

$$A_{\text{BT}} \geq (1 + 2(N-1))ND_{\text{MIM}} \tag{8}$$

$$A_{\text{BF}} \geq (1 + (N-1)sin(\theta)(BW/f_{\text{C}}))ND_{\text{MIM}} \tag{9}$$

The area estimation is further quantified using available MIM capacitors in 65nm CMOS and 16nm CMOS with a capacitor density of 2fF/μm^2 and >20fF/μm^2 [21] respectively as examples. Figure 7 shows the area estimation versus number of antennas. We assume a sampling capacitor of 50fF with 9 times of its area to consider the required dummy cells for better matching. As expected, the area requirement for beamtraining dominates the overall area. Fortunately, the area can be smaller as we move the process to a finer technology node with lower routing complexity. Additionally, the power consumption of phase shifter can also be reduced as the process moves to a finer technology node. Admittedly, this first-order estimation does not consider other circuitry and IO pads with ESD. However, the discussion shows the feasibility to deploy in a large-scale array. The complex routing does affect the chip design which shares similar challenges as other kinds of large-scale array.

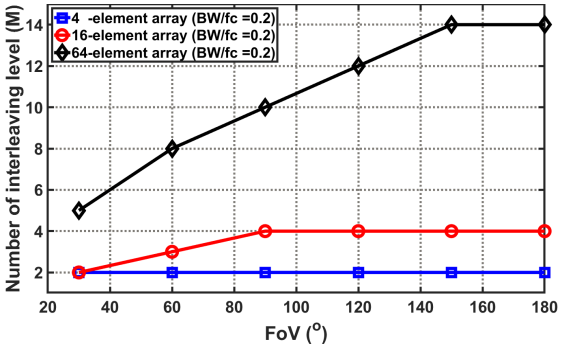

Fig. 8. Minimum number of interleaving level in respect to the FoV and the number of antennas under the assumption of a fractional bandwidth of 0.2.

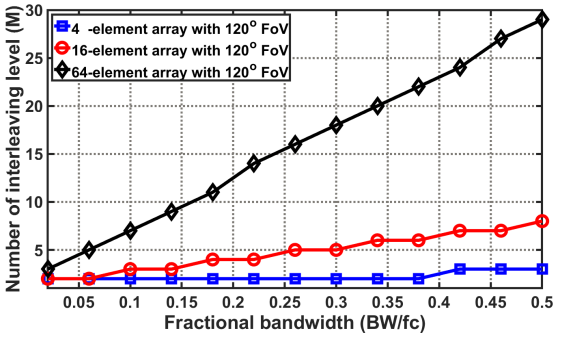

Fig. 9. Minimum number of interleaving level in respect to the fractional bandwidth and the number of antennas.

The effects of field-of-view (FoV), fractional bandwidth ($\mathrm{BW}/f_{\mathrm{C}}$) and array size in respect to the number of interleaving levels have also been investigated. Recall from (7), the required M is a function of the aforementioned three factors. FoV determines the maximum angle-of-arrival that a beamformed system can support As observed in Fig. 8, with a fixed fractional bandwidth of 0.2, the need of higher interleaving levels increases with an increased need of the FoV along with the array size. Similarly, the increased M is also observed with a large fractional bandwidth and array size as illustrated in Fig. 8. It is worth mentioning that the extreme scenario where BW/f_c is equal to 2 (i.e., 1000 MHz/500 MHz). Nevertheless, this case rarely happens at the presence of a RF front end operating at millimeter wave frequency band.

To summarize, in the beamforming mode, a larger array with wider fractional bandwidth requires a higher interleaving factor proportionally which implies a tradeoff between the number of antennas, FoV, fractional bandwidth, and hardware/power overhead. To leverage the interleaved architecture with a reasonable hardware/power overhead, a sub-array implementation within a hybrid architecture is thus feasible and preferred [10].

V. Conclusions

As the wider bandwidth available in the emerging applications, the TTD concept becomes more important in recent years. This brief analyzes the design considerations and tradeoffs in implementing the discrete time TTD interleaved beamtraining and beamforming system. The transistor-level design approach is first presented, describing design tradeoffs and component sizing selection procedures. Then, the system-level interleaving factor associated with hardware complexity has also been discussed. The two-part discussion in this brief suggests a quick design guide for the discrete time beamformed and beamtraining system towards wide bandwidth communication systems .

References

[1] N. Peng, P. Gu, X. You, and D. Zhao, "A ka-band cmos 4-beam phased-array receiver with symmetrical beam-distribution network," *IEEE Solid-State Circuits Lett.*, vol. 3, pp. 410–413, 2020.

[2] M. Yaghoobi, M. H. Kashani, M. Yavari, and S. Mirabbasi, "A 56-to-66 ghz cmos low-power phased-array receiver front-end with hybrid phase shifting scheme," *IEEE Trans. Cir. Syst. Regul. Pap.*, vol. 67, no. 11, pp. 4002–4014, 2020.

[3] C.-C. Lin, *et al.*, "Multi-mode spatial signal processor with rainbow-like fast beam training and wideband communications using true-time-delay arrays," *IEEE J. Solid-State Cir.*, vol. 57, no. 11, pp. 3348–3360, 2022.

[4] M. Li *et al.*, "An 800-ps origami true-time-delay-based cmos receiver front end for 6.5–9-ghz phased arrays," *IEEE Solid-State Circuits Lett.*, vol. 3, pp. 382–385, 2020.

[5] I. Mondal and N. Krishnapura, "A 2-ghz bandwidth, 0.25–1.7 ns true-time-delay element using a variable-order all-pass filter architecture in 0.13 μ m cmos," *IEEE J. Solid-State Cir.*, vol. 52, no. 8, pp. 2180–2193, 2017.

[6] S. Jang, R. Lu, J. Jeong, and M. P. Flynn, "A 1-ghz 16-element four-beam true-time-delay digital beamformer," *IEEE J. Solid-State Cir.*, vol. 54, no. 5, pp. 1304–1314, 2019.

[7] E. Ghaderi, A. Sivadhasan Ramani, A. A. Rahimi, D. Heo, S. Shekhar, and S. Gupta, "An integrated discrete-time delay-compensating technique for large-array beamformers," *IEEE Trans. Cir. Syst. Regul. Pap.*, vol. 66, no. 9, pp. 3296–3306, 2019.

[8] E. Ghaderi and S. Gupta, "A four-element 500-mhz 40-mw 6-bit adc-enabled time-domain spatial signal processor," *IEEE J. Solid-State Cir.*, vol. 56, no. 6, pp. 1784–1794, 2021.

[9] A. Fikes, P. P. Khial, S. Nooshabadi, and A. Hajimiri, "Programmable active mirror: A scalable decentralized router," *IEEE Trans. Micro. Theo. Techn.*, vol. 69, no. 3, pp. 1860–1874, 2021.

[10] V. Boljanovic *et al.*, "Fast beam training with true-time-delay arrays in wideband millimeter-wave systems," *IEEE Trans. Circuits Syst. I: Regul. Pap.*, vol. 68, no. 4, pp. 1727–1739, 2021.

[11] H. Yan, V. Boljanovic, and D. Cabric, "Wideband millimeter-wave beam training with true-time-delay array architecture," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, 2019, pp. 1447–1452.

[12] M. Soer, "Switched-RC beamforming receivers in advanced CMOS : theory and design," Ph.D. dissertation, University of Twente, Nov. 2012,

[13] A. Nagulu *et al.*, "A full-duplex receiver with true-time-delay cancelers based on switched-capacitor-networks operating beyond the delay–bandwidth limit," *IEEE J. Solid-State Cir.*, vol. 56, no. 5, pp. 1398–1411, 2021.

[14] A. M. A. Ali *et al.*, "A 16-bit 250-ms/s if sampling pipelined adc with background calibration," *IEEE J. Solid-State Cir.*, vol. 45, no. 12, pp. 2602–2612, 2010.

[15] ——, "A 12-b 18-gs/s rf sampling adc with an integrated wideband track-and-hold amplifier and background calibration," *IEEE J. Solid-State Cir.*, vol. 55, no. 12, pp. 3210–3224, 2020.

[16] Z. Huang *et al.*, "A 6-ghz bandwidth input buffer based on ac-coupled flipped source follower for 12-bit 8-gs/s adc in 28-nm cmos," *IEEE Trans. Circuits Syst. II: Express Br.*, vol. 69, no. 10, pp. 4163–4167, 2022.

[17] M. Straayer *et al.*, "27.5 a 4gs/s time-interleaved rf adc in 65nm cmos with 4ghz input bandwidth," in *2016 IEEE International Solid-State Circuits Conference (ISSCC)*, 2016, pp. 464–465.

[18] S. Kundu *et al.*, "A 1.2 v 2.64 gs/s 8 bit 39 mw skew-tolerant time-interleaved sar adc in 40 nm digital lp cmos for 60 ghz wlan," *IEEE Trans. Cir. Syst. Regul. Pap.*, vol. 62, no. 8, pp. 1929–1939, 2015.

[19] Q. Xu *et al.*, "A switching-less true-time-delay-based beam probing approach for ultra-low latency wireless communications: System analysis and demonstration," *IEEE Trans. Circuits Syst. II: Express Br.*, vol. 69, no. 10, pp. 4113–4117, 2022.

[20] A. Wadaskar, V. Boljanovic, H. Yan, and D. Cabric, "3d rainbow beam design for fast beam training with true-time-delay arrays in wideband millimeter-wave systems," in *2021 55th Asilomar Conference on Signals, Systems, and Computers*, 2021, pp. 85–92.

[21] S. Wu *et al.*, "A 16nm FinFET CMOS technology for mobile SoC and computing applications," in *2013 IEEE International Electron Devices Meeting*, 2013, pp. 9.1.1–9.1.4.