2024 7th International Balkan Conference on Communications and Networking (BalkanCom) | 979-8-3503-6595-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/BalkanCom61808.2024.10557192

Waveform Design for Over-the-Air Computing under Sampling Error

Nikos G. Evgenidis[1], Nikos A. Mitsiou[1], Sotiris A. Tegos[1,2], Panagiotis D. Diamantoulakis[1], Panagiotis Sarigiannidis[2] and George K. Karagiannidis[1,3]
[1]Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece
[2]Department of Informatics and Telecommunication Engineering, University of Western Macedonia, 50100 Kozani, Greece
[3]Artificial Intelligence & Cyber Systems Research Center, Lebanese American University (LAU), Lebanon
nevgenid@ece.auth.gr, nmitsiou@auth.gr, sotiristegos@ieee.org
padiaman@auth.gr, psarigiannidis@uowm.gr, geokarag@auth.gr

***Abstract*—To accommodate the large number of devices expected to operate in next-generation networks, a paradigm shift toward over-the-air (OTA) computing has been proposed, which takes advantage of the superposition principle of multiple access channels aiming to achieve better resource management as it supports simultaneous transmission in time and frequency. However, related studies have focused on analog transmission schemes without considering the components of modern transceivers. Therefore, to facilitate the use of OTA computing in modern systems, we investigate the impact of different waveforms transmission in OTA computing by taking into account the sampling errors that occur at the receiver side due to synchronization problems. To this end, the average minimum square error (MSE), under time sampling error for any utilized waveform, is investigated. Then, the MSE minimization problem is formulated and solved using alternating optimization to extract an efficient power allocation scheme. Simulation results for the raised cosine (RC) and the better-than-raised-cosine (BTRC) waveforms validate the theoretical part of our work and illustrate the efficiency of the extracted power allocation scheme while also providing a fair comparison between the RC and the BTRC waveforms.**

***Index Terms*—over-the-air computing, synchronization, time sampling error, resource allocation**

I. Introduction

Beyond 5G wireless networks are expected to support many new applications for devices, of which computing is one of the most important due to its necessity in many real-world scenarios, such as autonomous driving, etc. Therefore, more suitable techniques to better handle resource availability are needed for such goals [1]. For the task of computing, OTA computing is considered an interesting option since it utilizes the multiple access channel (MAC) superposition principle, which allows the transmission of multiple devices simultaneously, enabling wireless data aggregation [2] with reduced computational complexity due to the distributed nature of the method. However, OTA computing has not been studied in combination with fundamental elements of modern communication systems such as digital waveforms and filters, which limits its applicability and integration into current systems.

The ability of OTA computing to approximate target functions that can be written in nomographic form was facilitated in the seminal works [3], [4], while in [5] it was shown that uncoded analog transmission is the optimal transmission method for OTA computing, thus distinguishing it from modern systems that mostly rely on the use of digital waveforms for transmission [6]. With this in mind, optimal power allocation techniques for OTA computing have been studied in [7], [8], while optimal power allocation schemes and techniques to improve accuracy performance were proposed in [9].

OTA computing has also been studied in combination with many emerging technologies to improve its performance. For example, in [10], [11], reconfigurable intelligent surfaces (RISs) were considered as a means to increase accuracy through optimal handling of RISs, resulting in improved channel conditions. Similarly, MIMO systems have been proposed as a possible way to improve the performance of OTA computing. As such, many different directions have been explored in [12], [13], including joint hybrid beamforming and multiple target function computation under mean squared error (MSE) threshold and outage probability constraints. Federated learning (FL) has also been considered as an application to be enabled by OTA computing, since the distributed structure of the former fits the OTA computing model. In this context, OTA computing has been studied as an effective way to provide the update from the distributed devices to the fusion center (FC) [14], [15]. The convergence of FL even if not all devices participate with updates for each training round has also been studied in [16]. Although OTA computing has been studied extensively, most works are based on the assumption of analog transmission without considering the existing components of current communication systems. In [17], a DNN framework was proposed to enable OTA computing for digital systems without discussing the fundamental principles related to the transmitted waveforms and the impact of practical issues such as sampling error. In order to enable OTA computing in modern devices, it is crucial to investigate the performance of OTA computing under such scenarios.

Therefore, our work aims to investigate the performance of OTA computing when used in modern communication systems, along with the practical issues that arise in the latter. To this end, an optimization problem is formulated based on the transmitted waveforms, and sampling errors are considered as part of it to extract an optimal power allocation scheme for such a scenario. Simulation results illustrate a performance

gain by using the proposed power allocation scheme over the current optimal scheme proposed in [7].

II. System Model

A. OTA Computing Preliminaries

In this work, we consider an OTA computing system consisting of numerous transmitting devices and a receiver, that also acts as an FC. Let K be the number of transmitting devices participating in the system, the measurements of which are all independent to one another. The main objective of OTA computing is the calculation of a function $f : \mathbb{R}^K \to \mathbb{R}$ of all transmitted data, denoted as $f(x_1, x_2, \cdots, x_K)$. When f is a nomographic function, it is known that there exists an appropriate function decomposition consisting of a pre-processing function $\varphi_k : \mathbb{R} \to \mathbb{R}, \forall k \in \{1, \cdots, K\}$ and a post-processing function $\psi : \mathbb{R} \to \mathbb{R}$, such that the target function f can be expressed as

$$f(x_1, x_2, \cdots, x_K) = \psi\left(\sum_{k=1}^{K} \varphi_k(x_k)\right), \tag{1}$$

where x_k denotes the data measurement of the k-th device. Due to the stochastic nature of the wireless medium, all transmitted data are subject to channel fading and noise at the receiver, resulting in

$$\hat{f} = \psi\left(\sum_{k=1}^{K} h_k \varphi_k(x_{k,t}) + n\right), \tag{2}$$

where h_k denotes the channel fading of the k-th device, which is assumed to follow a Rayleigh distribution, and n denotes the additive white Gaussian noise (AWGN) with $\mathbb{E}_n[n] = 0$ and $\mathbb{E}_n[n^2] = \sigma^2$, where σ^2 is the noise power and $\mathbb{E}_X[\cdot]$ denotes the expectation with respect to (w.r.t.) to the random variable X. We define the set of all devices as $\mathcal{K} = \{1, \cdots, K\}$, where the devices are ordered in ascending order of their channel gains. For the transmitted data of each device it is assumed that $\mathbb{E}_{x_k}[x_k] = 0,\ \forall t$ and $\mathbb{E}_{x_k}[x_k^2] = 1$. In the context of this paper, and without loss of generality, we assume that the receiver and all transmitting devices are equipped with a single antenna. We assume that perfect channel state information (CSI) is available at both the transmitter and the receiver.

B. Basic Waveforms

Most modern communication systems are based on the use of suitable waveforms to tackle the effects of phenomena such as intersymbol interference (ISI). In general, ISI occurs when the transmitted symbols are interfered with past and future symbols, resulting in incorrect reconstruction of the original symbol at the base station (BS). Let T be the symbol period for all devices and $z_k(t)$ be the waveform associated with the k-th device and its data x_k. Because of its ability to counter ISI, one of the most commonly utilized waveforms for current digital systems is the raised cosine (RC) waveform, expressed in the time domain as

$$z_{\mathrm{RC}}(t) = \frac{1}{T}\mathrm{sinc}\left(\frac{t}{T}\right)\frac{\cos\left(\pi\alpha\frac{t}{T}\right)}{1-\left(2\alpha\frac{t}{T}\right)^2}, \tag{3}$$

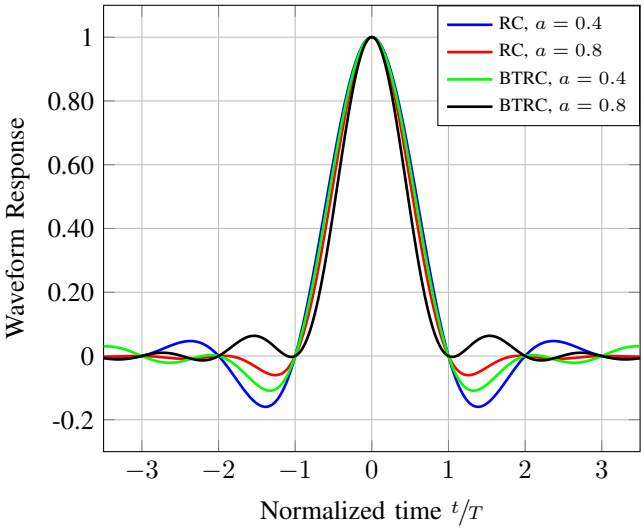

Fig. 1. RC and BTRC waveforms for roll-off factors $\alpha = 0.4$ and $\alpha = 0.8$.

where α is the roll-off factor. A slightly different waveform that is known to achieve better BER performance in conventional communication systems is the better-than-raised-cosine (BTRC) or flipped-exponential waveform. The BTRC waveform also satisfies the Nyquist criterion and is given in the time domain as

$$z_{\mathrm{BTRC}}(t) = \frac{1}{T}\mathrm{sinc}\left(\frac{t}{T}\right)\frac{4\beta\pi t\sin\left(\frac{\pi\alpha t}{T}\right)+2\beta^2\cos\left(\frac{\pi\alpha t}{T}\right)-\beta^2}{(2\pi t)^2+\beta^2}, \tag{4}$$

where $\beta = {}^{2T\ln 2}/_{\alpha}$. The impact of the roll-off factor on the waveform response can be seen in Fig. 1, where RC and BTRC waveforms are plotted for two values of α.

The bandwidth of both waveforms is dependent on the selected roll-off factor of the system, as follows

$$W_{SC} = (1+\alpha)W. \tag{5}$$

In general, the roll-off factor controls both the bandwidth excess and the amount of ISI in the system. Larger values indicate a system that is more resilient to ISI due to the smaller amplitude of the sidelobes as illustrated in Fig. 1, but with a larger spectrum allocation, while the opposite holds for small values of it, thus creating a critical trade-off between communication performance and resource efficiency.

It is important to note that most modern communication systems rely on a two-part split of the used waveform, i.e., for practical reasons, both the transmitting device and the BS are equipped with the square-root filter of the waveform, which leads to an optimal SNR at the receiver side [18] when sampling takes place. This technique does not affect the actual received signal in any way, but is preferred for practicality and sampling reasons, while also ensuring a fair comparison amongst different waveforms since the energy of both the square-root RC and the BTRC will be equal to $E = 1$.

C. OTA Transmission under Sampling Error

Without loss of generality, we assume that the target function is the arithmetic mean of all transmitted data, in which case, pre- and post-processing functions are not required to be utilized on either side of the transceiver, except for appropriate power allocation of the participating devices. Let $b_k \in \mathbb{C}$ be the transmit factor at the k-th device, where $|b_k|$ symbolizes the transmit power and $\angle b_k$ symbolizes the phase of the transmit signal. Similarly, $a \in \mathbb{C}^*$ symbolizes the receiver gain factor. All transmitting devices are assumed to have a common maximum power P, so that $|b_k|^2 \leq P$ for all $k \in \mathcal{K}$. Due to the perfect CSI availability, the phase of b_k can always be chosen such that the phase shift introduced by fading is always eliminated. Thus, we consider that the receiver gain, the transmit power, and the channel coefficients are all real numbers, i.e., $a, b_k, h_k \in \mathbb{R},\ \forall k \in \mathcal{K}$.

For our case of study, time sampling error, ϵ, occurs and we assume that it follows a Gaussian distribution centered around the ideal sampling time [19] and it has variance σ_ϵ^2, hence $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. If the sampling error is the only imperfection at the receiver, the received signal can be described as

$$\hat{y} = a\left(\sum_{k=1}^{K} x_k z_k(\epsilon) b_k h_k + n\right). \tag{6}$$

It should be highlighted that (6) corresponds to the case of flat-frequency channel fading, where no ISI is present, and the ideally received signal is given by

$$r = \sum_{k=1}^{K} x_k. \tag{7}$$

III. Optimization Problem

In this section, we formulate an optimization problem aiming to minimize the MSE of OTA computing between the ideal and actual received signal for both RC and BTRC. To this end, sampling error is considered to be present, while the optimal power allocation for any waveform during the OTA transmission is also extracted.

By definition, ideal sampling of any waveform occurs exactly at intervals of symbol period T. Since we assume that ideal sampling occurs at $t = 0$, it is clear that introducing the sampling error ϵ will lead to (6) and (7). Consequently, the MSE in this scenario can be expressed as

$$\mathrm{MSE}(a, \boldsymbol{b}) = \mathbb{E}_\epsilon\left[\sum_{k=1}^{K}\left(a z_k(\epsilon) b_k h_k - 1\right)^2 + \sigma^2 a^2\right], \tag{8}$$

where the expectation w.r.t. $x_k, \forall k \in \mathcal{K}$, and n have been calculated and $\boldsymbol{b} = [b_1, \cdots, b_K]$ is the transmission power vector of all users. Knowing the sampling error distribution allows the numerical calculation of the expected values $\bar{\epsilon}_1 = \mathbb{E}_\epsilon[z_k(\epsilon)]$ and $\bar{\epsilon}_2 = \mathbb{E}_\epsilon[z_k^2(\epsilon)]$ and it is straightforward to prove that the MSE under time sampling error is now described by

$$\mathrm{MSE}(a, \boldsymbol{b}) = \sum_{k=1}^{K}\left(\left(a b_k h_k\right)^2 \bar{\epsilon}_2 + 1\right) - 2\sum_{k=1}^{K} a b_k h_k \bar{\epsilon}_1 + \sigma^2 a^2. \tag{9}$$

Then, we can formulate the following optimization problem w.r.t. the power allocation factors at the transmitting devices and receiver sides

$$\begin{aligned} \min_{a,\boldsymbol{b}} \quad & \textstyle\sum_{k=1}^{K}\left(\left(a b_k h_k\right)^2 \bar{\epsilon}_2 + 1\right) \\ & -2\textstyle\sum_{k=1}^{K} a b_k h_k \bar{\epsilon}_1 + \sigma^2 a^2, \\ \textbf{s.t.} \quad & \mathrm{C}_1 : b_k \leq P,\ \forall k \in \mathcal{K}. \end{aligned} \tag{P1}$$

It is noted that (**P1**) is non-convex w.r.t. $a, \boldsymbol{b}$ jointly, but it is straightforward to prove that it is convex w.r.t. a and $\boldsymbol{b}$ separately. For this reason, we will use alternating optimization to tackle (**P1**).

For constant a, each device aims to independently minimize its MSE term given as $S_k = (a^2 h_k^2 \bar{\epsilon}_2) b_k^2 - 2a h_k \bar{\epsilon}_1 b_k + 1$. By fundamental properties of quadratic functions, S_k obtains its minimum at $b_k = {\bar{\epsilon}_1}/{(a h_k \bar{\epsilon}_2)}$, but due to the power constraint, this power can be selected only if $\sqrt{P} > {\bar{\epsilon}_1}/{(a h_k \bar{\epsilon}_2)}$, otherwise $b_k = \sqrt{P}$ must hold. Thus, the optimal power policy for a given a is given by

$$b_k = \min\left\{\sqrt{P}, \frac{\bar{\epsilon}_1}{a h_k \bar{\epsilon}_2}\right\}, \quad \forall k \in \mathcal{K}. \tag{10}$$

We observe that if there exists a device $i \in \mathcal{K}$ that can select $b_k = {\bar{\epsilon}_1}/{(a h_k \bar{\epsilon}_2)}$, any device $j \in \{\mathcal{K} | j > i\}$ can also select the same inverse channel-like transmit power due to the ascending channel order. Taking this into account, we decompose the original problem into K subproblems considering the different transmit power levels that can be selected. Then, for the i-th subproblem, the aim is to minimize the MSE, which can be written as

$$\begin{aligned} \mathrm{MSE}_i(a) = a^2\left(\sum_{k=1}^{i} P h_k^2 \bar{\epsilon}_2 + \sigma^2\right) - 2a\sum_{k=1}^{i}\sqrt{P} h_k \bar{\epsilon}_1 \\ + \sum_{k=i+1}^{K}\left(1 - \frac{\bar{\epsilon}_1^2}{\bar{\epsilon}_2}\right) + i. \end{aligned} \tag{11}$$

It should be noted that for the power allocation of the i-th device in (10) to be feasible, the following must hold for the receiver gain factor

$$\frac{\bar{\epsilon}_1}{\sqrt{P} h_{i+1} \bar{\epsilon}_2} < a. \tag{12}$$

If (12) is satisfied, the minimum value of (11) is reached when

$$a = a_i = \frac{\sqrt{P}\bar{\epsilon}_1 \sum_{k=1}^{i} h_k}{P\bar{\epsilon}_2 \sum_{k=1}^{i} h_k^2 + \sigma^2}, \tag{13}$$

and thus the global minimum of (11) is equal to

$$\mathrm{MSE}_i\left(\max\left\{\frac{\bar{\epsilon}_1}{\sqrt{P} h_{i+1} \bar{\epsilon}_2}, a_i\right\}\right). \tag{14}$$

By comparing the values of the sequence $\mathrm{MSE}_i,\ \forall i \in \mathcal{K}$ described by (14), we can identify the number of devices i^* that must transmit with maximum power, which is equal to

$$i^* = \underset{1 \leq i \leq K}{\operatorname{argmin}}\ \mathrm{MSE}_i. \tag{15}$$

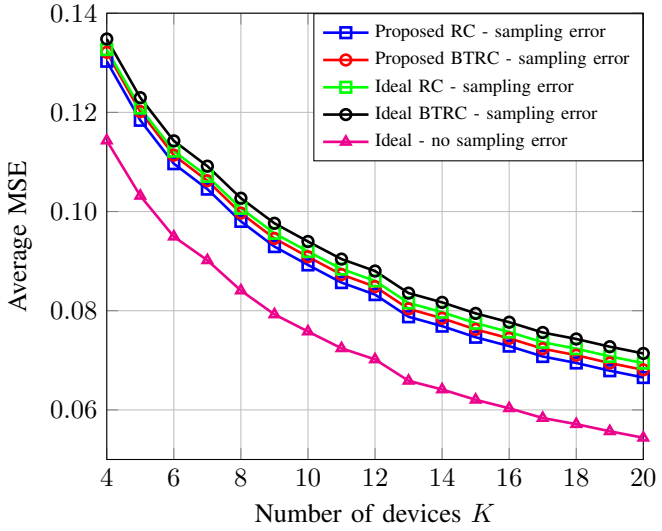

Fig. 2. Waveform MSE performance for varying number of devices K and sampling error variance $\sigma_\epsilon = 0.1$.

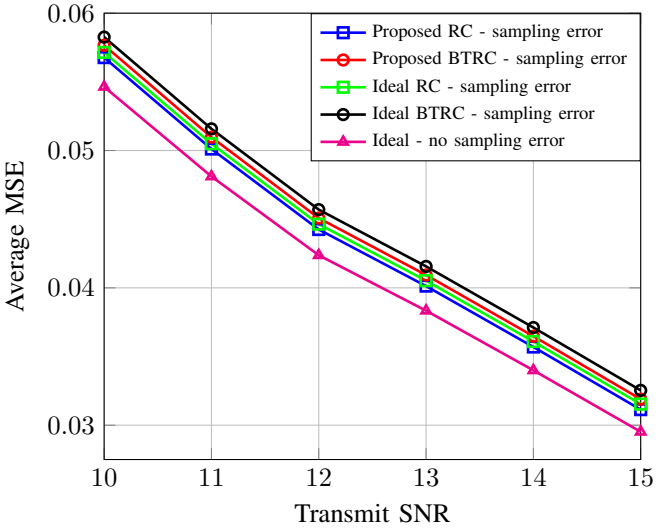

Fig. 3. Waveform MSE performance for varying transmit SNR P with roll-off factor $\alpha = 0.8$ and sampling error variance $\sigma_\epsilon = 0.1$.

Then, the optimal power allocation at the devices and the receiver can be calculated by combining (15), (13) and (10) in this specific order.

It is important to emphasize that, as observed by the previous analysis, the optimal power allocation policy of OTA computing is affected by the transmitted waveform, and specifically by the first two moments of the waveform amplitude, $\bar{\epsilon}_1$ and $\bar{\epsilon}_2$, when sampling error occurs. Therefore, proper selection of the transmitted waveform can lead to improved MSE performance.

IV. Simulation Results

In this section, the simulation results are presented. For all simulations, we assume that channel fading follows the circularly symmetric complex Gaussian distribution, i.e., $h_k \sim \mathcal{CN}(0,1), \forall k \in \mathcal{K}$ and without loss of generality, the sampling error follows a Gaussian distribution with mean value $\mu_\epsilon = 0$ and variance $\sigma_\epsilon^2 = 0.1$, where time is considered normalized w.r.t. the symbol time period T. Unless otherwise stated, we assume that there are $K = 20$ devices participating in the OTA transmission scheme, each utiling 40% additional bandwidth, i.e., $\alpha = 0.4$, and that the maximum transmit power at each device is such that the transmit SNR is $P_{\max}/\sigma^2 = 10$ dB. In order to evaluate the simulated results the average MSE is used as performance metric, which is defined as $\mathbb{E}[\text{MSE}]/K$.

With regards to the simulated policies, by *"Proposed"* we refer to the power allocation scheme extracted in Section III while by *"Ideal"* we refer to the power allocation scheme which was extracted in [7] where the ideal case of no sampling error was considered.

Fig. 2 illustrates the average MSE performance for different numbers of transmitting devices. As shown, the presence of sampling error, even for relatively small values, has a significant impact of about 15% on the MSE of OTA computing. This indicates the importance of studying its effect and how to mitigate it. It can be observed that for both considered waveforms, the proposed power allocation policy achieves about 5% better performance than the policy that considers the ideal scenario without considering the sampling error. However, it is also important to emphasize that although the sampling error degrades the MSE performance, the latter still has a decreasing behavior with increasing number of devices, which is extremely important for applications such as FL to ensure the convergence of distributed training.

In Fig. 3, the average MSE performance is plotted against the transmit SNR for a roll-off factor $\alpha = 0.8$. As expected, the performance of the system improves as the value of SNR increases, regardless of the waveform used for transmission. It is also important to note that the proposed power allocation policy outperforms the ideal scenario where sampling error is not considered for the whole range of simulated SNR values, which together with the results of Fig. 2 proves the validity of the extracted power allocation policy in section III.

Fig. 4 shows the average MSE performance for different values of the roll-off factor. For the completely ideal scenario where no sampling error occurs, the performance is obviously not affected by changes in the values of α, as expected. Regarding the used waveform, it is important to note that RC always outperforms BTRC due to the absence of ISI. This behavior is similar to that of conventional communication systems, since BTRC can only achieve better performance compared to RC in the presence of ISI due to the smaller amplitudes of its sidelobes, which, however, do not affect the studied model. As observed, increasing the value of the roll-off factor leads to an increase in the average MSE, indicating that α should be chosen to be as small as possible. Apart from this, it is significant to observe that the proposed power allocation policy again outperforms the policy of the ideal scenario for the whole range of α. In fact, it is evident that even by selecting a non-optimal waveform for transmission

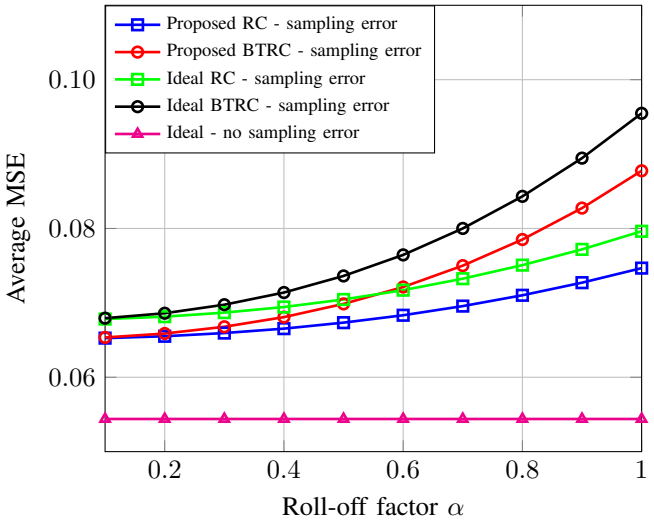

Fig. 4. Waveform MSE performance for varying roll-off factor α and sampling error variance $\sigma_\epsilon = 0.1$.

(BTRC in our simulation), it is still possible to outperform the performance of the optimal waveform (RC in our simulation) by using the proposed power allocation policy.

V. Conclusions

In the present work we modified the OTA computing model in order to expand its application from analog communications closer to that of modern communication systems. To this end, our extended model includes digital waveform transmission, since this is a key component of modern devices. More importantly, we formulated and solved an optimization problem to counter the effect of sampling error due to synchronization issues that arise in real-world scenarios. Our proposed policy showcases an improvement gain over the state-of-the-art optimal power allocation policy for varying number of devices and transmit SNR as well as different bandwidth utilization corresponding to different roll-off factors. Furthermore, our results focus on the most commonly utilized raised cosine and better-than raised cosine waveforms for both of which improvement is observed, thus highlighting the significance of the proposed policy.

Acknowledgment

This work was funded from the Smart Networks and Services Joint Undertaking (SNS JU) under European Union's Horizon Europe research and innovation programme (Grant Agreement No. 101096456 - NANCY).

References

[1] N. G. Evgenidis, N. A. Mitsiou, V. I. Koutsioumpa, S. A. Tegos, P. D. Diamantoulakis, and G. K. Karagiannidis, "Multiple Access in the Era of Distributed Computing and Edge Intelligence," 2024.

[2] G. Zhu, J. Xu, K. Huang, and S. Cui, "Over-the-Air Computing for Wireless Data Aggregation in Massive IoT," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 57–65, Aug. 2021.

[3] M. Goldenbaum and S. Stanczak, "Robust Analog Function Computation via Wireless Multiple-Access Channels," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3863–3877, Sep. 2013.

[4] M. Goldenbaum, H. Boche, and S. Stańczak, "Analyzing the space of functions analog-computable via wireless multiple-access channels," in *Proc. 8th Int. Symp. Wireless Commun. Syst.*, Nov 2011, pp. 779–783.

[5] M. Gastpar, "Uncoded Transmission Is Exactly Optimal for a Simple Gaussian Sensor Network," *IEEE Trans. Inf. Theory*, vol. 54, pp. 5247–5251, 2008.

[6] A. Assalini and A. Tonello, "Improved Nyquist pulses," *IEEE Commun. Lett.*, vol. 8, no. 2, pp. 87–89, 2004.

[7] W. Liu, X. Zang, Y. Li, and B. Vucetic, "Over-the-Air Computation Systems: Optimization, Analysis and Scaling Laws," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5488–5502, Aug. 2020.

[8] X. Cao, G. Zhu, J. Xu, Z. Wang, and S. Cui, "Optimized Power Control Design for Over-the-Air Federated Edge Learning," *IEEE J. Sel. Areas Commun*, vol. 40, no. 1, pp. 342–358, Jan. 2022.

[9] N. G. Evgenidis, V. K. Papanikolaou, P. D. Diamantoulakis, and G. K. Karagiannidis, "Over-the-Air Computing with Imperfect CSI: Design and Performance Optimization," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2023.

[10] P. S. Bouzinis, N. A. Mitsiou, P. D. Diamantoulakis, D. Tyrovolas, and G. K. Karagiannidis, "Intelligent Over-the-Air Computing Environment," *IEEE Wireless Commun. Lett.*, vol. 12, no. 1, pp. 134–137, 2023.

[11] W. Fang, Y. Jiang, Y. Shi, Y. Zhou, W. Chen, and K. B. Letaief, "Over-the-Air Computation via Reconfigurable Intelligent Surface," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8612–8626, Dec. 2021.

[12] X. Zhai, X. Chen, J. Xu, and D. W. Kwan Ng, "Hybrid Beamforming for Massive MIMO Over-the-Air Computation," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2737–2751, Apr. 2021.

[13] G. Zhu and K. Huang, "MIMO Over-the-Air Computation for High-Mobility Multimodal Sensing," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6089–6103, Aug. 2019.

[14] G. Zhu, Y. Du, D. Gündüz, and K. Huang, "One-Bit Over-the-Air Aggregation for Communication-Efficient Federated Edge Learning: Design and Convergence Analysis," *IEEE Transactions on Wireless Communications*, vol. 20, no. 3, pp. 2120–2135, March 2021.

[15] C. Xu, S. Liu, Z. Yang, Y. Huang, and K.-K. Wong, "Learning Rate Optimization for Federated Learning Exploiting Over-the-Air Computation," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 12, pp. 3742–3756, Dec. 2021.

[16] N. A. Mitsiou, P. S. Bouzinis, P. D. Diamantoulakis, R. Schober, and G. K. Karagiannidis, "Accelerating Distributed Optimization via Over-the-Air Computing," *IEEE Trans. Commun.*, vol. 71, no. 9, pp. 5565–5579, 2023.

[17] S. Razavikia, J. M. B. Da Silva, and C. Fischione, "ChannelComp: A General Method for Computation by Communications," *IEEE Trans. Commun.*, pp. 1–1, 2023.

[18] G. Turin, "An introduction to matched filters," *IRE Transactions on Information Theory*, vol. 6, no. 3, pp. 311–329, June 1960.

[19] A. Papoulis, "Error analysis in sampling theory," *Proceedings of the IEEE*, vol. 54, no. 7, pp. 947–955, July 1966.