INTEGRATED DISTRIBUTED WIRELESS SENSING WITH OVER-THE-AIR FEDERATED LEARNING

Shijian Gao

Samsung Electronics
SoC R&D Lab, San Diego

Jia Yan, Georgios B. Giannakis

University of Minnesota
Department Electrical and Computer Engineering

ABSTRACT

Over-the-air federated learning (OTA-FL) is a communication-effective approach for achieving distributed learning tasks. In this paper, we aim to enhance OTA-FL by seamlessly combining sensing into the communication-computation integrated system. Our research reveals that the wireless waveform used to convey OTA-FL parameters possesses inherent properties that make it well-suited for sensing, thanks to its remarkable auto-correlation characteristics. By leveraging the OTA-FL learning statistics, i.e., means and variances of local gradients in each training round, the sensing results can be embedded therein without the need for additional time or frequency resources. Finally, by considering the imperfections of learning statistics that are neglected in the prior works, we end up with an optimized the transceiver design to maximize the OTA-FL performance. Simulations validate that the proposed method not only achieves outstanding sensing performance but also significantly lowers the learning error bound.

Index Terms— over-the-air, federated learning, gradient statistics, wireless sensing, resource allocation

1. INTRODUCTION

Federated learning (FL) is a distributed learning paradigm that allows multiple users to collaboratively learn a shared model under the coordination of a central server and without exchanging data directly [1]. While FL is data-efficient, it poses challenges to spectrum resources when each device requires a dedicated frequency band to upload its local model parameters for global model aggregation. In response to this challenge, over-the-air aggregation [2] has emerged as a novel solution, leveraging the signal-superposition property of multiple-access channels. By allowing simultaneous transmissions of all device-end updates through a shared spectrum, over-the-air aggregation enables integrated communication and computation to accomplish a learning task.

The concept of over-the-air federated learning (OTA-FL) has garnered significant research interest for its potential of substantial resource savings while maintaining comparable performance to conventional FL with orthogonal transmission. Existing OTA-FL studies, despite dealing with wireless fading, simplify assume error-free learning statistics, i.e., means and variances of local gradients in each training round. Although the statistics occupy negligible frequency resources, the resultant processing leads to a severe sub-optimality in practical scenarios. Moreover, current OTA-FL research overlooks the inherent freedom introduced by OTA transmission, which can be leveraged to provide additional sensing functionality to enhance security and privacy measures.

This work aims to enhance OTA-FL by addressing these deficiencies through two key contributions. Firstly, we recognize the inherent suitability of OTA-FL for wireless sensing, leveraging the high-dimensional nature and desirable auto-correlation of local model vectors, as well as the potential for improved sensing accuracy through device diversity. Motivated by these observations, we delve into the seamless integration of wireless sensing and OTA-FL by using learning statistics as an off-the-shelf gateway to deliver sensed results without introducing additional overhead. This integrated design enables the coherent fusion of sensing, communication, and computation in a wireless manner, which advances the concept of integrated communication and computation [2] and integrated sensing and communication [3]. Furthermore, we emphasize the crucial role of learning statistics, which not only enhances sensing capabilities but also has a substantial impact on learning performance. Strikingly, previous studies have largely overlooked the influence of learning statistics, assuming error-free transmission and negligible communication resources. By addressing these concerns, we establish a sensing-enabled OTA-FL framework with a more practical problem formulation, enhanced learning performance, and efficient resource utilization. We have validated these advantages through comprehensive simulations.

2. PROPOSED METHODOLOGY

2.1. Signal Model

Consider a general FL system with K single-antenna wireless devices, and a single multi-antenna edge server. At the t-

This work is supported by NSF grants 2126052, 2103256, 2102312, 2128593, 1901134, and 2220292

The corresponding author is Jia Yan (yanj@umn.edu).

th learning round, the local gradient generated by device k is a D-dimensional vector $\mathbf{g}_{k,t}$, whose entry-wise mean and standard deviation are $\bar{g}_{k,t}$ and $\nu_{k,t}$, respectively. With OTA aggregation, the transmit signal from device k is

$$\mathbf{x}_{k,t} \triangleq \{x_{k,t}[d] = p_{k,t}\frac{g_{k,t}[d] - \bar{g}_{k,t}}{\nu_{k,t}} : 1 \leq d \leq D\}, \quad (1)$$

with $p_{k,t}$ denoting the transmit equalization factor [4]. The local learning statistics, including the mean $\bar{g}_{k,t}$ and the standard deviation $\nu_{k,t}$, per device are transmitted individually following $\mathbf{x}_{k,t}$. A practical learning model may contain thousands or more parameters [5], implying that D can be extremely large. Meanwhile, $x_{k,t}[d]$ can be modeled as an independent zero-mean variable, leading to

$$\sum_{d} x_{k,t}[d]x_{k,t}[d+\tau] \approx \delta(\tau). \quad (2)$$

For this reason, $\mathbf{x}_{k,t}$ serves as a promising candidate for passive wireless sensing. This is analogous to the case with orthogonal frequency-division multiplexing (OFDM) waveform [6].

2.2. Sensing Embedding

To enable sensing functionality without calling for additional frequency resources, our proposed answer is that $\nu_{k,t}$ will behave as the gateway supporting wireless sensing. Two types of sensing tasks can be supported here: objection detection and object positioning.

Specifically for the former, by directly utilizing $\mathbf{x}_{k,t}$ for sensing, the binary decision obtained via a specific detector, e.g., the Neyman-Pearson detector [7], will be embedded into $\nu_{k,t}e^{j\theta}$ with $\theta = 0$ or π. As $\nu_{k,t}$ remains positive, it suffices for the server to recover the required standard deviation $\nu_{k,t}$, by taking the norm of $\nu_{k,t}e^{j\theta}$, and the sensed result θ via binary phase-shift keying (BPSK) demodulation.

The second task, object positioning, is more challenging as it further involves range estimation [8]. To improve the precision, the transmit signal $\mathbf{x}_{k,t}$ is wrapped with a transparent binary pseudo sequence $\mathbf{p}$ (e.g., M-sequence) for sensing [9]. The device m then unwraps the echo of $\mathbf{x}_{x,t} \odot \mathbf{p}$, and applies matched filtering for range estimation. Accordingly, the estimated distance is fed into an M-step quantizer, whose upperbound can be decided per the wireless cell size R, giving rise to the transmitted signal

$$\nu_{k,t}e^{j\frac{2\pi m}{M}},\ m \in \{0, 1, \ldots, M-1\}. \quad (3)$$

Through M-PSK demodulation, the edge serve obtains a series of estimated range as $\{\hat{R}_k\}_{k=1}^{K}$. Utilizing the known location of each wireless device, represented as $\{X_k, Y_k\}_{k=1}^{K}$, the position of the object can be determined by solving the following least-square estimation problem:

$$\begin{aligned}&(X^\star, Y^\star)\\ &= \arg\min_{X,Y} \sum_{k=1}^{K} \left|\sqrt{(X_k - X)^2 + (Y_k - Y)^2} - \hat{R}_k\right|^2. \end{aligned} \quad (4)$$

2.3. Transceiver Design

Once embedding the sensing functionality into OTA-FL, the remaining question is how to optimize the transceivers to yield the best learning performance. Let $\nabla F(\mathbf{w}_t)$ and $\mathbf{r}_t(\mathbf{f})$ stand for the gradient of the loss function $F(\cdot)$ at $\mathbf{w}_t$ and the received aggregated gradient using the combiner $\mathbf{f}$. The global model update obeys

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \eta(\nabla F(\mathbf{w}_t) - \mathbf{e}_t), \quad (5)$$

with η being the learning rate, and $\mathbf{e}_t := \nabla F(\mathbf{w}_t) - \mathbf{r}_t(\mathbf{f})$ incorporating the overall gradient error in OTA-FL [4], respectively. The error term $\mathbf{e}_t$ is first bounded by taking both the channel noise, as well as the imperfections of $\bar{g}_{k,t}$ and $\nu_{k,t}$ into account, in contrast to prior works that solely consider the first source. Based on the derived bound[1], convergence analysis can be conducted, to establish that

$$\begin{aligned}&\lim_{t\to\infty} \mathbb{E}\{F(\mathbf{w}_t) - F(\mathbf{w}^*)\}\\ =&\mathcal{C}(\mathbf{f}, \{P(\mathbf{x}_k)\}_{k=1}^{K}, \{P(\bar{g}_k)\}_{k=1}^{K}, \{P(\nu_k)\}_{k=1}^{K})\end{aligned} \quad (6)$$

where $F(\mathbf{w}^*)$ is the ground truth; $P(\mathbf{x}_k)$, $P(\bar{g}_k)$ and $P(\nu_k)$ respectively stand for the transmit power of the centralized version, the mean, and the gradient deviation at the device k. The optimal transceiver design can be subsequently found by solving the following optimization problem:

$$\min_{P,\mathbf{f}} \quad \mathcal{C}(\mathbf{f}, \{P(\mathbf{x}_k)\}_{k=1}^{K}, \{P(\bar{g}_k)\}_{k=1}^{K}, \{P(\nu_k)\}_{k=1}^{K}) \quad (7a)$$

$$s.t. \quad |\mathbf{f}| = 1 \quad (7b)$$

$$P(\mathbf{x}_k) + P(\bar{g}_k) + P(\nu_k) \leq P_{max}, \forall k. \quad (7c)$$

3. PERFORMANCE EVALUATION

To test the performance of the integrated wireless sensing with OTA-FL system, we simulate a scenario where 9 edge devices take part in an image classification task with the Fashion-MNIST dataset. Each device locally trains a convolution neural network consisting of 21,921 parameters.

We set the noise floor at -10dB, and compute the correlation gain of the transmit sequence with the noisy echo. As depicted in Fig. 1, the gain is 12dB higher above the noise floor, even when the echo is 20dB weaker than the noise. Also, the strongest side lobe remains 15dB lower than the peak in

[1]More detailed derivations will be presented in the full-version journal paper due to space limitation.

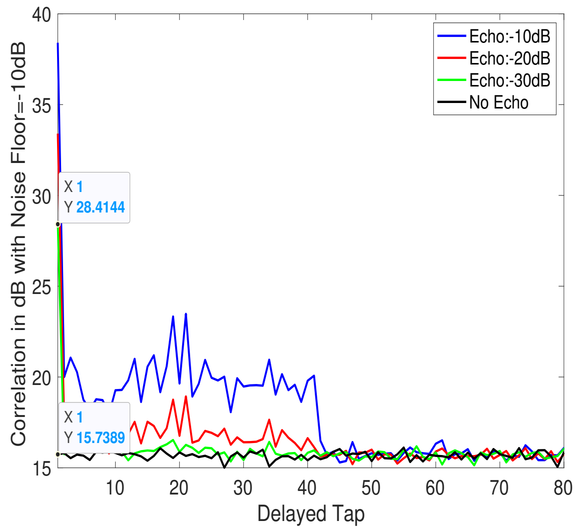

Fig. 1: Correlation gain of transmit sequence with noisy echo

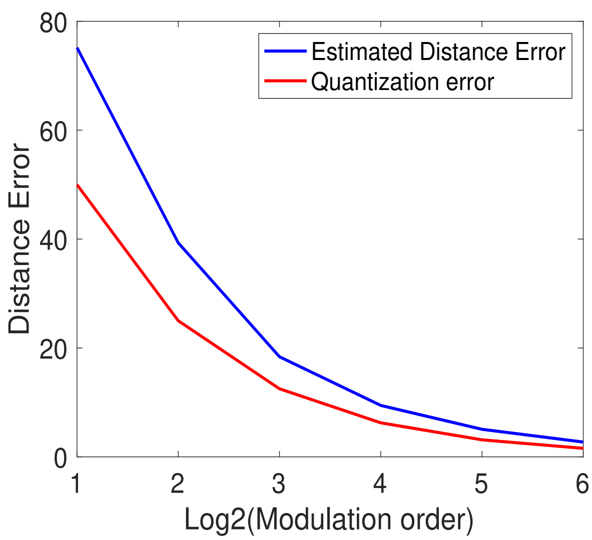

Fig. 2: The object positioning error (in meters) via sensing embedded OTA-FL under different modulation order.

all cases, signifying that detection and rage estimation can be conveniently accomplished with high reliability.

In Fig. 2, we assume that 9 devices are evenly distributed along the x-axis with a 10-meter gap between each device. The object is uniformly positioned within a 200×200 square. As can be seen, the positioning error obtained through least-square estimation from the quantized angles can be very close to the lower bound of ideal quantization. This observation suggests that involving multiple users can significantly reduce the positioning error.

In Fig. 3, the comparison of the convergence gap among different power allocation schemes is conducted. Evenly distributing the power among all 21,923 elements (21,921 centralized gradient entries and 2 attached statistics) yields the worst performance. Dedicated optimization can reduce the gap by more than 50% compared to random allocation. Still, the performance gap is significant with respect to the ideal case, where perfect statistics are obtained. This observation, in conjunction with the lower performance of an even distribution, implies that neglecting the imperfections of statistics is improper for OTA-FL, and may render largely inferior learning performance. In summary, despite accounting for a tiny ratio of the entire transmission frame, gradient statistics play a critical role in OTA-FL, and their influence should be carefully considered to optimize the learning performance.

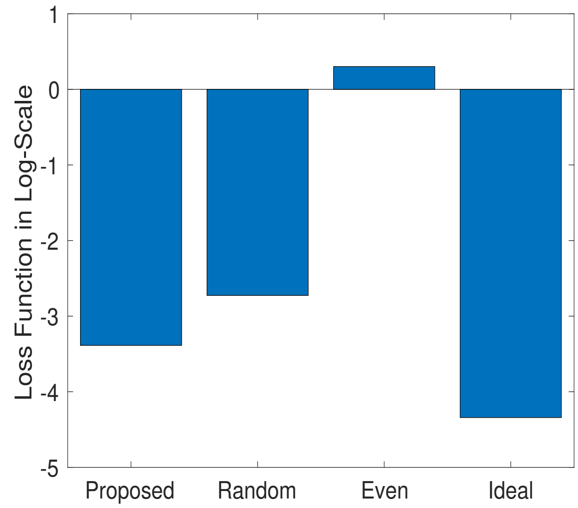

Fig. 3: Convergence gap under various power allocation schemes

4. CONCLUSIONS

This paper presented a novel approach to enhance OTA-FL by seamlessly integrating sensing functionality into the wireless communication framework. Leveraging the exceptional auto-correlation properties of the OTA-FL waveform, we embed sensing capabilities without requiring additional time or frequency resources. Through optimized transceiver design, we minimize the learning error bound while considering imperfections in learning statistics. Simulations demonstrate the remarkable sensing performance of our proposed method and its significant improvement to conventional OTA-FL in terms of system efficiency and robustness. Given the increasing demand for distributed learning and sensing in wireless networks, our work provides valuable insights and paves the way for future developments in this research area.

5. REFERENCES

[1] T. Chen, G. B. Giannakis, T. Sun, and W. Yin., “LAG: Lazily aggregated gradient for communication-efficient distributed learning,” in *Proc. of Advances in Neural Information Processing Systems*, December 2018.

[2] G. Zhu, Y. Wang, and K. Huang., “Broadband analog aggregation for low-latency federated edge learning,” *IEEE Trans. Wireless Communications*, vol. 19, no. 1, pp. 491–506, October 2019.

[3] X. Cheng, D. Duan, S. Gao, and L. Yang., “Integrated sensing and communications (ISAC) for vehicular communication networks (VCN),” *IEEE Internet of Things Journal*, vol. 9, no. 23, pp. 23 441–23 451, December 2022.

[4] H. Liu, X. Yuan, and Y.-J.-A. Zhang., “A reconfigurable intelligent surface enabled federated learning: A unified communication-learning design approach,” *IEEE Trans. Wireless Communications*, vol. 20, no. 11, pp. 7595–7609, June 2021.

[5] M. Chen, Z. Yang, W. Saad, C. Yin, H. Poor, and S. Cui., “A joint learning and communications framework for federated learning over wireless networks,” *IEEE Trans. Wireless Communications*, vol. 17, no. 1, pp. 269–283, January 2021.

[6] C. Berger, D. Demissie, J. Heckenbach, P. Willett, and S. Zhou., “Signal processing for passive radar using OFDM waveforms,” *IEEE Journal on Selected Topics in Signal Processing*, vol. 4, no. 1, pp. 226–238, February 2010.

[7] S. M. Kay., *Fundamentals of Statistical Signal Processing*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.

[8] M. A. Richards., *Fundamentals of Radar signal Processing*. New York, NY, USA: McGraw-Hill, 2005.

[9] F. Liu, Y. Cui, C. Masouros, J. Xu, T.-X. Han, Y. C. Eldar, and S. Buzzi., “Integrated sensing and communications: Towards dual-functional wireless networks for 6G and beyond,” *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 6, pp. 1728–1767, June 2022.