Private and Secure Over-the-Air Multi-Party Communication

Jan Jonas Brune*, Matthias Frey§, Felix Klement‡, Igor Bjelaković*†, Stefan Katzenbeisser‡ and Sławomir Stańczak*†
*Fraunhofer Heinrich-Hertz-Institute, §University of Melbourne, ‡Universität Passau,
and †Technische Universität Berlin
matthias.frey@unimelb.edu.au, {felix.klement, stefan.katzenbeisser}@uni-passau.de
{jan.jonas.brune, igor.bjelakovic, slawomir.stanczak}@hhi.fraunhofer.de

***Abstract*—In this paper, we introduce Over-the-Air Multi-Party Communication, a novel approach to achieve efficiently scalable, private, secure, and dependable data aggregation using Over-the-Air computation. The main idea of our approach lies in a combination of techniques from lattice coding, Over-the-Air computation and secure Multi-Party Computation to securely and confidentially aggregate data over a multiple-access channel with additive white Gaussian noise. Our theoretical analysis of the proposed analog scheme developed in this work satisfies the necessary reliability, security, and privacy criteria. Among the potential applications of our approach are smart metering, distributed machine learning, and data aggregation in wireless sensor networks.**

I. Introduction

The proliferation of connected devices through the Internet of Things (IoT) has led to an explosion in the amount of data generated by these devices. To improve energy efficiency and achieve favorable scaling of communication capacity in the number of the devices deployed in this area, Over-the-Air (OTA) computation can be utilized for data aggregation [1]–[8]. This method allows for computation of a function of data distributed in a wireless network at a central terminal without the complete transmission of the distributed data. One promising application of OTA computation is the training of machine learning models (cf. [9]–[12]) in wireless sensor networks, where the computation of only a few features is needed, making the decoding of data from the individual sensors unnecessary. Smart meters are another example of wireless devices that can collect and process energy consumption data remotely, which has numerous benefits, such as reducing the cost and complexity of manual meter reading, improving billing accuracy, and enabling demand response programs to decrease energy consumption during peak periods. Nevertheless, the nature of data collected in these and similar applications poses new challenges, including privacy and security risks. To address these issues, private, secure, and reliable OTA data aggregation methods have gained in importance. Concepts and methods from information-theoretic secrecy and classical cryptography present a promising approach to surmount security challenges. In particular, the fusion of secure Multi-Party Computation (MPC) techniques [13] with OTA methodologies holds great potential to combine the privacy and security guarantees of the former with the efficiency advantages of the latter.

Drawing on methods from lattice coding, in this work we propose an Over-the-Air Multi-Party Communication (OTA-MPC) scheme which leverages techniques from OTA computation and MPC to aggregate data in an efficiently scalable, reliable, secure, and private way over a multiple-access channel with additive white Gaussian noise (AWGN). The suggested methodology comprises two distinct stages, namely, an offline phase and an online phase. During the offline phase, a suitable lattice is agreed upon, random keys at the transmitters are drawn and distributed and a key for the receiver is generated. After this setup, in the online phase, the stakeholders perform the pre-processing, transmission, and post-processing. Total omission of the offline phase in such a protocol is not always achievable. Therefore, an area for potential future research involves investigating various techniques and strategies that can mitigate the necessity for an offline phase, or alternatively, mitigate the complexity and resource demands of the offline phase. Although the proposed method is completely analog, it displays a clear parallel to the traditional schemes used in classical cryptography to provide privacy and security, for example, in secure MPC and smart metering.

The authors acknowledge the financial support by the Federal Ministry of Education and Research of Germany in the programme of "Souverän. Digital. Vernetzt." Joint project 6G-RIC, project identification number: 16KISK020K, 16KISK030, and 16KISK034.
SS and IB acknowledge the financial support by the German Research Foundation (DFG) under grant STA 864/15-1.

II. System Model

For OTA-MPC, we consider a system model as shown in Fig. 1. For individual measurements $s_k \in \mathcal{S}_k$, with $k \in \{1, \dots, K\}$, representing distributed data, the legitimate receiver's goal is to obtain a reliable estimate of a function $f\colon \mathcal{S}_1 \times \cdots \times \mathcal{S}_K \to \mathbb{R}$ of these measurements.

1) Each measurement s_k is encrypted in pre-processing using a key $U_k \in V_k$ by a pre-processing function

$$h_L^k : V_k \times \mathcal{S}_k \to \mathbb{R}^L, (U_k, s_k) \mapsto h_L^k(U_k, s_k)$$

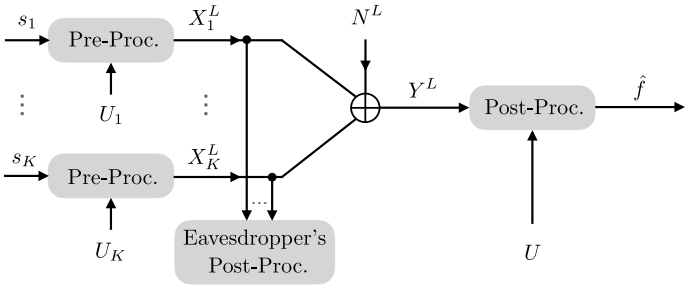

Fig. 1. System model of the considered scheme: The goal of the legitimate receiver is to obtain a reliable estimate $\hat{f}$ of a function $f : \mathcal{S}_1 \times \cdots \times \mathcal{S}_K \to \mathbb{R}$ that has a suitable nomographic representation. The legitimate receiver is subject to a privacy constraint regarding individual distributed data, while the eavesdropper should not be able to access either distributed data or the function $f(s_1, \ldots, s_K)$ thereof. The random variables $U_1, \ldots, U_K$ serve as keys for transmitters $k = 1, \ldots, K$, while the legitimate receiver has access to the random variable U which depends on $(U_1, \ldots, U_K)$.

which obeys the average power constraint $\left\|h_L^k(u,s)\right\|^2/L \leq \mathcal{P}$, for all $(u,s) \in V \times \mathcal{S}_k$, where $L \in \mathbb{N} \setminus \{0\}$, and $\mathcal{P} \in \mathbb{R}$ with $\mathcal{P} \geq 0$.

2) The pre-processed information $X_k^L := h_L^k(U_k, s_k) \in \mathbb{R}^L$, with $k \in \{1, \ldots, K\}$, is sent simultaneously over L uses of a wiretap channel with AWGN. The superposition property of the channel can be leveraged to perform OTA computation. Accordingly, the resulting received signal is

$$Y^L = \sum_{k=1}^{K} X_k^L + N^L \in \mathbb{R}^L, \tag{1}$$

with $N^L \sim \mathcal{N}\left(0, \sigma_N^2 \cdot \mathrm{id}_{L\times L}\right)$, where $\mathrm{id}_{L\times L}$ denotes the identity $L \times L$ matrix.

3) The aggregator has access to a key $U \in V$, which we assume can be generated by a central authority during an offline phase in dependence of $U_1, \ldots, U_K$. During post-processing, given by the function

$$G_L : V \times \mathbb{R}^L \to \mathbb{R}, \left(U, Y^L\right) \mapsto G_L\left(U, Y^L\right),$$

the aggregator decrypts the received signal to obtain an estimate $\hat{f} := G_L\left(U, Y^L\right)$ which is close (in the sense made precise below) to the desired function f, with high probability.

Based on this system model, the following objectives emerge:

1) (Reliability) We say that the OTA-MPC scheme is (ε, δ)-reliable if

$$\mathbb{P}\left(\left|f(s_1, \ldots, s_K) - \hat{f}\right| \geq \varepsilon\right) < \delta$$

uniformly for all $(s_1, \ldots, s_K) \in \mathcal{S}_1 \times \cdots \times \mathcal{S}_K$.

2) (Secrecy) We say that the secrecy requirement is satisfied if without knowledge of $U, U_1, \ldots, U_K$, the distribution of $(X_1^L, \ldots, X_K^L)$ is the same for every $(s_1, \ldots, s_K) \in \mathcal{S}_1 \times \cdots \times \mathcal{S}_K$.

3) (Privacy) We say that the privacy requirement is satisfied if without knowledge of $U_1, \ldots, U_K$, and for every $k \in \{1, \ldots, K\}$, the distribution of (X_k^L, U) is the same for every $(s_1, \ldots, s_K) \in \mathcal{S}_1 \times \cdots \times \mathcal{S}_K$.

Remark 1. *Note that we base our definition of privacy on the channel inputs instead of the channel output. We aim to incorporate alternative notions of privacy and explore their consequences in future works.*

The class of functions suitable for OTA approximation and which we will consider in this work is introduced in the following definition.

Definition 1. *(Function class $\mathcal{F}_{\mathrm{mon}}$) [6, Definition 3] Let $\mathcal{S}_1, \ldots, \mathcal{S}_K \subseteq \mathbb{R}$ be sets. Then, a Borel measurable function $f\colon \mathcal{S}_1 \times \cdots \times \mathcal{S}_K \to \mathbb{R}$ is said to belong to $\mathcal{F}_{\mathrm{mon}}$ if there exist bounded and measurable functions $(f_k)_{k\in\{1,\ldots,K\}}$, a measurable set $D \subseteq \mathbb{R}$ with the property $f_1(\mathcal{S}_1) + \cdots + f_K(\mathcal{S}_K) \subseteq D$ and a measurable function $F\colon D \to \mathbb{R}$ such that for all $(s_1, \ldots, s_K) \in \mathcal{S}_1 \times \cdots \times \mathcal{S}_K$ we have*

$$f(s_1, \ldots, s_K) = F\left(\sum_{k=1}^{K} f_k(s_k)\right),$$

where for F there exists an increment majorant*, which is a strictly increasing function $\Phi\colon [0,\infty) \to [0,\infty)$ with $\Phi(0) = 0$ and*

$$|F(x) - F(y)| \leq \Phi(|x - y|)$$

for all $x, y \in D$.

This means that every $f \in \mathcal{F}_{\mathrm{mon}}$ has a nomographic representation, i.e. $f(s_1, \ldots, s_K) = F\left(\sum_{k=1}^{K} f_k(s_k)\right)$ for suitably chosen functions $f_k(s_k) : \mathcal{S}_k \to \mathbb{R}$, $k = 1, \ldots, K$ and $F\colon \mathbb{R} \to \mathbb{R}$. Examples of functions contained in $\mathcal{F}_{\mathrm{mon}}$ are weighted sums of bounded functions and p-norms for $p \geq 1$ among other practically important function types [6].

III. Proposed Approach

Our proposed approach is based on the construction of a suitable lattice. Λ is called an L-dimensional *lattice* if it is a discrete subgroup of the additive group $(\mathbb{R}^L, +)$. If its linear span is $\mathbb{R}^L$, the lattice is called *non-degenerate*. With growing values of the dimension L, both the resource consumption and the reliability of the scheme increase. Therefore, the choice of L represents a tradeoff between the goals of efficiency and reliability.

For every $\lambda \in \Lambda$, the *Voronoi cell* $\mathcal{V}_\lambda^\Lambda$ is defined as

$$\mathcal{V}_\lambda^\Lambda := \left\{x \in \mathbb{R}^L \ : \ \forall \lambda' \in \Lambda : \ \|\lambda - x\| \leq \|\lambda' - x\|\right\}.$$

In the following, $\mathcal{V}^\Lambda := \mathcal{V}_0^\Lambda$ denotes the *fundamental Voronoi cell* located at the origin. Ties are resolved in a systematic way such that the Voronoi cells of a non-degenerate lattice Λ are congruent and form a partition of $\mathbb{R}^L$. A lattice Λ induces a modulo operation

$$\mathrm{mod}\,\Lambda : \mathbb{R}^L \to \mathcal{V}^\Lambda,$$

which has some key properties that are essential to our scheme. Reliability will later be argued with the help of the *distributive law* [14, Proposition 2.3.1]

$$\forall x, y \in \mathbb{R}^L : \ (x \bmod \Lambda + y) \bmod \Lambda = (x + y) \bmod \Lambda.$$

For our proof, we require the definition of two radii concerning lattices. The *effective radius* r_{eff}^{Λ} of Λ is defined as the radius of a ball with the same volume as $\mathcal{V}^{\Lambda}$ [14, Definition 3.1.1]. The *covering radius* r_{cov}^{Λ} of the lattice Λ is defined as the minimum radius needed to cover $\mathbb{R}^L$ by balls of radius r centered at the lattice points:

$$r_{\text{cov}}^{\Lambda} \coloneqq \min \left\{ r : \Lambda + \mathcal{B}_r \left(\mathbf{0}_L \right) \right\},$$

where $\mathcal{B}_r \left(\mathbf{0}_L \right)$ is a ball of radius r centered at the origin $\mathbf{0}_L \in \mathbb{R}^L$. Thereby, the covering radius can be viewed as the minimum radius of a closed ball centered at $\mathbf{0}_L$ which contains $\mathcal{V}^{\Lambda}$. We proceed to formulate our main result, which states that we are able to find pre- and post-processing functions for our scheme that fulfill a reconstruction quality criterion as well as a secrecy and a privacy criterion.

Theorem 1. *For all $\sigma_N < 1$, $\varphi \in (\sigma_N, 1)$, $K \in \mathbb{N} \setminus \{0\}$, $f \in \mathcal{F}_{\text{mon}}$ with $f \colon \mathbb{R}^K \mapsto \mathbb{R}$ and a fixed nomographic representation $f_1, \ldots, f_K, F, \Phi$, for which it holds that for all transmitters $k \in \{1, \ldots, K\}$ holding $s_k \in \mathcal{S}_k$ we have $\vartheta \coloneqq f_1(s_1) + \cdots + f_K(s_K) \in [-1, 1]$, and (given σ_N and φ) for sufficiently large $L \in \mathbb{N} \setminus \{0\}$, there is a lattice $\Lambda_L \subset \mathbb{R}^L$ and a real number $\gamma > 0$ such that for all $k \in \{1, \ldots, K\}$, there are pre-processing schemes*

$$h_{\Lambda_L, \varphi}^k : \mathcal{V}^{\Lambda_L} \times \mathbb{R} \to \mathbb{R}^L$$

satisfying the average power constraint of 1; and there is a post-processing scheme

$$G_{\Lambda_L, \varphi} : \mathcal{V}^{\Lambda_L} \times \mathbb{R}^L \to \mathbb{R}$$

such that the following holds:

Let $\hat{f}$ be the estimate of $f(s_1, \ldots, s_K)$ at the receiver after post-processing.

Then, we have

1) *(Reliability criterion). The schemes of pre- and post-processing are (ε, δ)-reliable, i.e.*

$$\mathbb{P}_{G_{\Lambda_L, \varphi}(U, Y^L)} \left(\left| f(s_1, \ldots, s_K) - \hat{f} \right| \geq \varepsilon \right) \leq \delta, \quad (2)$$

with $\varepsilon > 0$, $\delta = 2 \cdot \frac{\sigma_N}{\sqrt{L}(\varphi - \sigma_N)\Phi^{-1}(\epsilon)} \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{-L(\varphi-\sigma_N)^2 \left(\Phi^{-1}(\epsilon)\right)^2}{2\sigma_N^2}} + \exp(-L\beta_1) + \exp(-L\gamma)$, where Y^L is the channel output.

2) *(Secrecy criterion). Without knowledge of $U, U_1, \ldots, U_K$, the distribution of $(X_1^L, \ldots, X_K^L)$ is the same for every $(s_1, \ldots, s_K) \in \mathcal{S}_1 \times \cdots \times \mathcal{S}_K$.*

3) *(Privacy criterion). If $K > 1$, then without knowledge of $U_1, \ldots, U_K$, and for every $k \in \{1, \ldots, K\}$, the distribution of (X_k^L, U) is the same for every $(s_1, \ldots, s_K) \in \mathcal{S}_1 \times \cdots \times \mathcal{S}_K$.*

Remark 2. *Theorem 1 also holds for $\vartheta \coloneqq f_1(s_1) + \cdots + f_K(s_K) \in [a, b]$ with $a, b \in \mathbb{R}$, when normalizing with the spread of f as done in [5]. For better readability, we restrain ϑ to $[-1, 1]$ in this paper. In the same spirit, we assume the average power constraint $\mathcal{P} = 1$, as one can always apply suitable scaling in the pre- and post-processors.*

We proceed to prove *Theorem 1* through Proposition 1 later introduced in this section. First, we define the pre- and post-processing schemes.

Definition 2. *Given any specific lattice Λ of dimension L and $\varphi \in (\sigma_N, 1)$, we define pre- and post-processing schemes induced by (Λ, φ) as follows. The pre-processor at user k is defined as*

$$h_{\Lambda, \varphi}^k : \\ (U_k, s_k) \mapsto \begin{cases} \left(\frac{f_k(s_k) r_{sig}}{\sqrt{L}} \mathbb{1}_L + U_k \right) \bmod \Lambda, & \text{if } r_{cov} \leq \sqrt{L}, \\ \mathbf{0}_L, & \text{else}, \end{cases} \quad (3)$$

where $\mathbf{0}_L, \mathbb{1}_L \in \mathbb{R}^L$ are the all-0 and all-1 vectors respectively, $r_{sig} \coloneqq \sqrt{L}(\varphi - \sigma_N)$ and the key U_k is uniformly drawn from the fundamental Voronoi cell $\mathcal{V}$ of Λ. The post-processor is defined in three steps:

1) *Subtract the key using the modulo operation and rescale the signal by applying*

$$g_{\Lambda, \varphi} : (U, Y^L) \mapsto \frac{\left(Y^L - U \right) \bmod \Lambda}{r_{sig}} \cdot \sqrt{L}, \quad (4)$$

where $U = \left(\sum_{k=1}^K U_k \right) \bmod \Lambda$.

2) *Average the elements of the L-dimensional vector $g_{\Lambda, \varphi}\left(U, Y^L\right) = \left(g_1\left(U, Y^L\right), \ldots, g_L\left(U, Y^L\right) \right)$ by*

$$\tilde{g}_{\Lambda, \varphi}\left(U, Y^L\right) = \frac{1}{L} \sum_{l=1}^{L} g_l\left(U, Y^L\right). \quad (5)$$

3) *Apply the outer function F by*

$$G_{\Lambda, \varphi}\left(U, Y^L\right) = F\left(\tilde{g}_{\Lambda, \varphi}\left(U, Y^L\right) \right). \quad (6)$$

The estimate $\hat{f}$ of the function f is given by

$$\hat{f} = G_{\Lambda, \varphi}\left(U, Y^L\right). \quad (7)$$

Remark 3. *The reason we choose the all-zero vector in (3) in case $r_{cov} > \sqrt{L}$ is so we can assume that the average power constraint is satisfied (see (15)).*

In the following, let $\vartheta \coloneqq \sum_{k=1}^K f_k$ and let $\mathbb{P}_{N^L}$ denote the probability distribution of the random vector

$$Z^L \coloneqq \vartheta \mathbb{1}_L + \frac{N^L}{(\varphi - \sigma_N)}, \quad (8)$$

where $N^L \sim \mathcal{N}\left(0, \sigma_N^2 \cdot \mathrm{id}_{L \times L}\right)$. Moreover, let $\mathbb{P}^{(t)}$ denote the probability distribution of the random vector $g\left(U, Y^L\right)$ given in (4). We proceed to choose in a systematic manner several parameters that we will use in the following definitions and proofs. For every $\varphi \in (\sigma_N, 1)$, we pick φ_1, φ_2 such that

$\varphi < \varphi_1 < \varphi_2 < 1$ and φ_3 such that $\varphi_1/\varphi_2 < \varphi_3 < 1$. The following events will be crucial for the proof of our main result in Theorem 1:

$$\mathcal{E}_1 := \left\{ \frac{r_{\text{cov}}^{\Lambda_L}}{r_{\text{eff}}^{\Lambda_L}} > \varphi_2^{-1} \right\} \tag{9}$$

and

$$\mathcal{E}_2 := \left\{ \frac{\vartheta r_{\text{sig}}^{\Lambda_L}}{\sqrt{L}} \mathbb{1}_L + N^L \notin \mathcal{V}^{\Lambda_L} \right\}. \tag{10}$$

Proposition 1. *There is a sequence of lattices such that the following hold:*

1) *Conditioned under $\mathcal{E}_1^c \cap \mathcal{E}_2^c$, the post-processor output is of the form*

$$g_{\Lambda_L,\varphi}\left(U, Y^L\right) = \left(\vartheta + \frac{N_1}{\varphi - \sigma_N}, \ldots, \vartheta + \frac{N_L}{\varphi - \sigma_N}\right), \tag{11}$$

where $N_1, \ldots, N_L$ are the components of N^L.

2) *For a fixed absolute constant $\beta_1 > 0$ and sufficiently large L, it holds that*

$$\mathbb{P}_{N^L}(\mathcal{E}_1 \cup \mathcal{E}_2) \leq \exp(-L\beta_1). \tag{12}$$

3) *There is $\gamma > 0$, such that for all sufficiently large $L \in \mathbb{N} \setminus \{0\}$ we have*

$$\left\| \mathbb{P}_{N^L} - \mathbb{P}^{(t)} \right\|_{\text{TV}} \leq \exp\left(-L \cdot \gamma\right). \tag{13}$$

For the proof of Proposition 1, we refer the interested reader to [15, Section 4].

Theorem 1 follows from Proposition 1. In the proof of Theorem 1, we require the *total variation distance* between two probability distributions $\mathbb{Q}$ and $\mathbb{P}$ on a probability space Ω and and an event space Σ, which is defined as [16, Example 3.17]:

$$\|\mathbb{Q} - \mathbb{P}\|_{\text{TV}} = \sup_{A \in \Sigma} |\mathbb{Q}(A) - \mathbb{P}(A)|. \tag{14}$$

To prove secrecy and privacy for our scheme, we apply the Crypto Lemma:

Lemma 1. *(Crypto Lemma) [14, Lemma 4.1.1]*
Let Λ be a lattice and U be a random variable distributed uniformly over the fundamental cell $\mathcal{V}^\Lambda$ of Λ. Then $(s + U) \bmod \Lambda$ is distributed uniformly over $\mathcal{V}^\Lambda$, independent of the value of s.

Proof of Theorem 1. First, we argue that the pre-processing operation obeys the power constraint. This is immediate in case $r_{\text{cov}}^{\Lambda_L} > \sqrt{L}$, since in this case an all-0 signal is transmitted. If, on the other hand, $r_{\text{cov}}^{\Lambda_L} \leq \sqrt{L}$, then a signal in $\mathcal{V}^L$ is transmitted. In this case, we have

$$(X_1, \ldots, X_L) \in \mathcal{V}^L \subseteq \mathcal{B}_{\sqrt{L}}(\mathbf{0}_L), \tag{15}$$

and therefore, the average power constraint is satisfied.

Next, we show that the reliability criterion of Theorem 1 follows from Proposition 1. We note that

$$\begin{aligned}
\left|\hat{f} - f\right| &\geq \epsilon \\
\Leftrightarrow \left|G_{\Lambda_L,\varphi}\left(U, Y^L\right) - f\right| &\geq \epsilon \\
\Leftrightarrow \left|F\left(\tilde{g}_{\Lambda_L,\varphi}\left(U, Y^L\right)\right) - F\left(\sum_{k=1}^{K} f_k\left(s_k\right)\right)\right| &\geq \epsilon \\
\Rightarrow \Phi\left(\left|\tilde{g}_{\Lambda_L,\varphi}\left(U, Y^L\right) - \sum_{k=1}^{K} f_k\left(s_k\right)\right|\right) &\geq \epsilon \\
\Leftrightarrow \left|\tilde{g}_{\Lambda_L,\varphi}\left(U, Y^L\right) - \sum_{k=1}^{K} f_k\left(s_k\right)\right| &\geq \Phi^{-1}(\epsilon),
\end{aligned}$$

where the fourth line follows from the definition of the increment majorant Φ, and the fifth line follows from the fact that Φ is strictly increasing. Thereby,

$$\begin{aligned}
&\mathbb{P}^{(t)}\left(\left|\hat{f} - f\right| \geq \epsilon\right) \\
&\leq \mathbb{P}^{(t)}\left(\left|\tilde{g}_{\Lambda_L,\varphi}\left(U, Y^L\right) - \sum_{k=1}^{K} f_k\left(s_k\right)\right| \geq \Phi^{-1}(\epsilon)\right).
\end{aligned} \tag{16}$$

By definition of $\|\cdot\|_{\text{TV}}$ in (14), we have for every event A

$$\mathbb{P}^{(t)}(A) \leq \mathbb{P}_{N^L}(A) + \left\| \mathbb{P}_{N^L} - \mathbb{P}^{(t)} \right\|_{\text{TV}}.$$

Using this and (16), we obtain

$$\begin{aligned}
&\mathbb{P}^{(t)}\left(\left|\hat{f} - f\right| \geq \epsilon\right) \\
&\quad \leq \mathbb{P}_{N^L}\left(\left|\frac{1}{L}\sum_{l=1}^{L} g_l\left(U, Y^L\right) - \sum_{k=1}^{K} f_k\left(s_k\right)\right| \geq \Phi^{-1}(\epsilon)\right) \\
&\qquad + \left\| \mathbb{P}_{N^L} - \mathbb{P}^{(t)} \right\|_{\text{TV}}
\end{aligned} \tag{17}$$

Let $\beta_1 > 0$ and $\gamma \in (0, \beta_1)$. By (13) in Proposition 1, we have $\left\| \mathbb{P}_{N^L} - \mathbb{P}^{(t)} \right\|_{\text{TV}} \leq \exp\left(-L\gamma\right)$ and therefore by (17), we obtain

$$\begin{aligned}
&\mathbb{P}^{(t)}\left(\left|\hat{f} - f\right| \geq \epsilon\right) \\
&\quad \leq \mathbb{P}_{N^L}\left(\left|\frac{1}{L}\sum_{l=1}^{L} g_l\left(U, Y^L\right) - \sum_{k=1}^{K} f_k\left(s_k\right)\right| \geq \Phi^{-1}(\epsilon)\right) \\
&\qquad + \exp(-L\gamma).
\end{aligned} \tag{18}$$

We define the shorthand notation

$$A := \left\{ \left|\frac{1}{L}\sum_{l=1}^{L} g_l\left(U, Y^L\right) - \sum_{k=1}^{K} f_k\left(s_k\right)\right| \geq \Phi^{-1}(\epsilon) \right\},$$

and obtain

$$\begin{aligned}
\mathbb{P}_{N^L}(A) &= \mathbb{P}_{N^L}\left(A \cap (\mathcal{E}_1^c \cap \mathcal{E}_2^c)\right) + \mathbb{P}_{N^L}\left(A \cap (\mathcal{E}_1 \cup \mathcal{E}_2)\right) \\
&\leq \mathbb{P}_{N^L}\left(A \cap (\mathcal{E}_1^c \cap \mathcal{E}_2^c)\right) + \mathbb{P}_{N^L}\left((\mathcal{E}_1 \cup \mathcal{E}_2)\right) \\
&\leq \mathbb{P}_{N^L}\left(A \cap (\mathcal{E}_1^c \cap \mathcal{E}_2^c)\right) + \exp\left(-L\beta_1\right),
\end{aligned} \tag{19}$$

by (12) in Proposition 1. Note that

$$\begin{aligned}
&\mathbb{P}_{N^L}\left(A \cap \left(\mathcal{E}_1^c \cap \mathcal{E}_2^c\right)\right) \\
&\leq \mathbb{P}_{N^L}\left(\left|\frac{1}{L}\sum_{l=1}^{L}\left(\vartheta + \frac{N_l}{\varphi - \sigma_N}\right) - \sum_{k=1}^{K} f_k\left(s_k\right)\right| \geq \Phi^{-1}\left(\epsilon\right)\right) \\
&= \mathbb{P}_{N^L}\left(\left|\underbrace{\frac{1}{\sigma_N\sqrt{L}}\sum_{l=1}^{L} N_l}_{\sim\mathcal{N}(0,1)}\right| \geq \frac{\sqrt{L}\left(\varphi - \sigma_N\right)\Phi^{-1}\left(\epsilon\right)}{\sigma_N}\right) \qquad (20)
\end{aligned}$$

holds, because $\vartheta \coloneqq \sum_{k=1}^{K} f_k$. Therefore, it is sufficient to find an upper bound on the right hand side of (20). By [17, Proposition 2.1.2], we know that for a random variable $\hat{A} \sim \mathcal{N}\left(0,1\right)$ and all $z > 0$ it holds that

$$\mathbb{P}\left(\hat{A} \geq z\right) \leq \frac{1}{z} \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}.$$

Note that since $N^L \sim \mathcal{N}\left(0, \sigma_N^2 \cdot \mathrm{id}_{L\times L}\right)$, we have that $\frac{1}{\sigma_N\sqrt{L}}\sum_{l=1}^{L} N_l \sim \mathcal{N}\left(0,1\right)$. Thereby, we can apply [17, Proposition 2.1.2] and obtain

$$\begin{aligned}
&\mathbb{P}_{N^L}\left(\left|\frac{1}{\sigma_N\sqrt{L}}\sum_{l=1}^{L} N_l\right| \geq \frac{\sqrt{L}\left(\varphi - \sigma_N\right)\Phi^{-1}\left(\epsilon\right)}{\sigma_N}\right) \\
&\leq 2 \cdot \frac{\sigma_N}{\sqrt{L}\left(\varphi - \sigma_N\right)\Phi^{-1}\left(\epsilon\right)} \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{-\left(\sqrt{L}\left(\varphi-\sigma_N\right)\Phi^{-1}\left(\epsilon\right)\right)^2}{2\sigma_N^2}} \\
&= 2 \cdot \frac{\sigma_N}{\sqrt{L}\left(\varphi - \sigma_N\right)\Phi^{-1}\left(\epsilon\right)} \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{-L\left(\varphi-\sigma_N\right)^2\left(\Phi^{-1}\left(\epsilon\right)\right)^2}{2\sigma_N^2}}
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{P}^{(t)}\left(\left|\hat{f} - f\right| \geq \epsilon\right) \\
&\leq 2 \cdot \frac{\sigma_N}{\sqrt{L}\left(\varphi - \sigma_N\right)\Phi^{-1}\left(\epsilon\right)} \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{-L\left(\varphi-\sigma_N\right)^2\left(\Phi^{-1}\left(\epsilon\right)\right)^2}{2\sigma_N^2}} \\
&\quad + \exp(-L\beta_1) + \exp(-L\gamma) =: \delta,
\end{aligned}$$

where we have used (18), (19), (20). For the secrecy and criterion, we observe that depending on L, the pre-processor either outputs 0 deterministically or it has the form

$$\left(\frac{f_k\left(s_k\right) r_{\text{sig}}}{\sqrt{L}}\,\mathbb{1}_L + U_k\right) \bmod \Lambda$$

as defined in (3), in which case its output is uniformly distributed in $\mathcal{V}^{\Lambda_L}$ regardless of s_k according to the Crypto Lemma in Lemma 1. For the proof of the privacy criterion, note that the k channel inputs $X_k^L \coloneqq h_L^k\left(U_k, s_k\right)$ are dependent of U_k and s_k. However, due to the Crypto Lemma and since $K > 1$, the key U at the receiver is independent of U_k and therefore $\left(X_k^L, U\right)$ follows a product distribution. Again, by the Crypto Lemma in Lemma 1, for each k, both X_k^L and U are uniformly distributed over $\mathcal{V}^{\Lambda_L}$ regardless of the value of $(s_1, \ldots, s_K) \in \mathcal{S}_1 \times \cdots \times \mathcal{S}_K$. □

REFERENCES

[1] M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," in *Information Processing in Sensor Networks*. Springer, 2003, pp. 162–177.

[2] M. Goldenbaum, H. Boche, and S. Stańczak, "Harnessing interference for analog function computation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 61, no. 20, pp. 4893–4906, 2013.

[3] M. Goldenbaum and S. Stanczak, "Robust analog function computation via wireless multiple-access channels," *IEEE Transactions on Communications*, vol. 61, no. 9, pp. 3863–3877, 2013.

[4] M. Goldenbaum, H. Boche, and S. Stańczak, "Nomographic functions: Efficient computation in clustered gaussian sensor networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 4, pp. 2093–2105, 2014.

[5] I. Bjelaković, M. Frey, and S. Stańczak, "Distributed approximation of functions over fast fading channels with applications to distributed learning and the max-consensus problem," in *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2019, pp. 1146–1153.

[6] M. Frey, I. Bjelaković, and S. Stańczak, "Over-the-air computation in correlated channels," *IEEE Transactions on Signal Processing*, vol. 69, pp. 5739–5755, 2021.

[7] P. Park, P. D. Marco, and C. Fischione, "Optimized over-the-air computation for wireless control systems," *IEEE Communications Letters*, vol. 26, pp. 424–428, 2022.

[8] H. Jung and S.-W. Ko, "Performance analysis of uav-enabled over-the-air computation under imperfect channel estimation," *IEEE Wireless Communications Letters*, vol. 11, pp. 438–442, 2022.

[9] M. M. Amiri and D. Gündüz, "Machine learning at the wireless edge: Distributed stochastic gradient descent over-the-air," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2155–2169, 2020.

[10] J.-H. Ahn, O. Simeone, and J. Kang, "Wireless federated distillation for distributed edge learning with heterogeneous data," in *2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. IEEE, 2019, pp. 1–6.

[11] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proceedings of the 20 th International Conference on Artificial Intelligence and Statistics (AISTATS) 2017*, 2017, pp. 1273–1282.

[12] J. Konečnỳ, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," *arXiv preprint arXiv:1610.05492*, 2016.

[13] R. Cramer, I. B. Damgård, and J. B. Nielsen, *Secure Multiparty Computation and Secret Sharing*. Cambridge University Press, 2015.

[14] R. Zamir, *Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation, and Multiuser Information Theory*. Cambridge University Press, 2014.

[15] J. J. Brune, M. Frey, F. Klement, I. Bjelaković, S. Katzenbeisser, and S. Stańczak, "Private and secure over-the-air multi-party communication," *To appear. Will be made available on arXiv*.

[16] M. J. Wainwright, *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*, ser. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2019.

[17] R. Vershynin, *High Dimensional Probability: An Introduction with Applications in Data Science*, ser. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2018, vol. 47.