

Algorithms for Overpredictive Signal Analytics in Federated Learning

Vijay Anavangot and Animesh Kumar
Department of Electrical Engineering
Indian Institute of Technology Bombay
Mumbai, India
email : {avijay, animesh}@ee.iitb.ac.in

Abstract—Distributed signal processing will play a major role in performing learning and inference tasks in a client-server model proposed in federated learning. Clients (IoT devices) having many signal samples can aid a data-center (server) in learning the global signal model, by pooling these distributed samples. The clients may have privacy concerns, and the pooling of distributed samples will require accounting of communication cost involved. As a result, a processed approximation of these samples may be desirable. This decentralized learning approach is termed as distributed signal analytics in this work. Overpredictive signal approximations may be desired to perform such distributed signal analytics, which are primarily motivated by applications in network demand (capacity) estimation and planning. In this work, we propose algorithms that calculate an overpredictive signal approximation at the client devices using an efficient convex optimization framework. A tradeoff between the number of bits communicated by clients to the server and the signal approximation error is quantified. An analysis of our approximations is presented on an available residential energy consumption dataset.

Index Terms—signal reconstruction, approximation methods

I. INTRODUCTION

In large scale distributed systems, consisting of several clients and a central server, a key problem is to learn useful user signal statistics, by considering the user privacy and client-server communication bottlenecks. The recently proposed federated learning architecture [1, 2] addresses these issues through client side signal processing and server level aggregation methods. Modern signal processing tools such as sparse representation, and signal approximations are expected to be useful in such distributed learning tasks.

In this paper, we consider the signal approximation problem in a federated learning based architecture, where the signals observed at the clients are primarily time series data samples. These decentralized signal samples spread across several clients need to be communicated to a central server, through a bandwidth constrained channel to learn a global statistical model at the server. To address the communication bottleneck, the client devices will send only an approximate signal which summarizes the relevant statistics useful in learning the server level global model. The global model to be learned is a function of the approximate signals reported by the clients, and will be termed as *signal analytics*.

An application example of such a client-server based federated learning model, is an electrical grid that consists of a cen-

tral power server and several households (clients) connected to it. Each household installed with an energy meter will record the instantaneous power consumption, which will be communicated to the server. At the server, the signal analytics such as statistical energy usage patterns (like hourly energy demand statistics), useful in grid planning are learned. The server may be interested in meeting the energy demand of the users, which will require each energy meter to communicate an approximate consumption pattern, that always overpredicts the true demand. Such an overprediction based approximation of signals in the federated learning setup, has not been researched previously as far as we know.

Prior research efforts have focussed on distributed mean-squared error approximation of scalar and vector parameters (not continuous-time signals) [3, 4], mostly in the context of sensor networks. An application of overprediction based scalar quantization has been proposed for TV whitespace spectrum allocation [5]. Recent research works in federated learning have addressed problems such as, communication efficient and privacy preserving distributed machine learning schemes [2, 6], distributed energy demand learning for electric vehicle charging [7], and quantized communication for distributed mean-estimation [8]. An information theoretic framework is used to address privacy issues in distributed non-intrusive load monitoring applications [9, 10]. Our work is different from the prior art, as it attempts to learn distributed signal analytics using sparse signal approximations and statistical learning, in a federated learning framework.

The main contributions of this work are summarized below.

- We propose an overprediction based approximation in the federated learning setup using the Fourier basis signal representation.
- We propose a federated aggregation procedure to learn global signal analytics at the server, using the statistical tools like empirical cumulative distribution function.
- We analyze the tradeoff between the communication cost and the approximation error, using experiments on a canned residential energy consumption dataset.

II. PROBLEM SETUP AND USEFUL RESULTS

A. Signal Model in Federated Learning

Consider a federated learning based architecture consisting of D clients and a central server, shown in Figure 1. Each

client device records a continuous-time signal represented as $f_1(t), f_2(t), \dots, f_D(t)$. Without loss of generality we will assume that the signals have a bounded support in the interval $[0, 1]$. For the purpose of analysis, we also assume that these signals are p -times differentiable, where $p \geq 1$. The signals at the client devices are to be communicated to the central server, so that a global model can be learned. However, due to bandwidth constraint, the devices only communicate an approximate signal, which will be denoted as $\hat{f}_1(t), \hat{f}_2(t), \dots, \hat{f}_D(t)$.

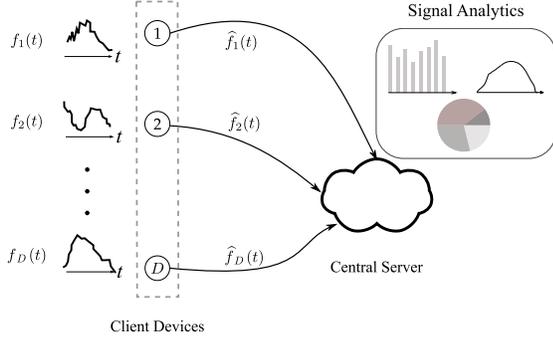


Fig. 1. The federated learning system model for distributed signal analytics. In this model, D client devices communicate an approximation of the recorded signals, and the central server pools these signals to obtain useful signal analytics such as the histogram, aggregate signal or a clustering of the devices.

B. Fourier Basis Representation

This paper, being the first exposition into the study of distributed signal approximation in a federated learning model, will use the Fourier basis representation for the signals at the client devices. That is, each recorded signal at the clients can be represented by a Fourier series,

$$f_i(t) = \sum_{k=-\infty}^{\infty} a_i[k] \exp(j2\pi kt), \quad t \in [0, 1],$$

where $i \in \{1, 2, \dots, D\}$. It will be assumed that $f_i(0) = f_i(1)$, which is due to a requirement of the Fourier series representation. Since the signals at the clients are assumed to p -times differentiable, we observe a polynomial decay in the Fourier series coefficients of the signals, which is stated in the fact below.

Fact 1 (Sec 2.3,[11]). *A signal $f(t), t \in [0, 1]$, with $f(0) = f(1)$, is p -times differentiable if its Fourier coefficient $a[k]$ satisfies the condition,*

$$|a[k]| \leq \frac{C}{|k|^{p+1+\varepsilon}} \quad \text{for some } C, p, \varepsilon > 0.$$

C. Mathematical Formulation and Related Definitions

Consider the L length bandlimited approximation of the signals, $f_i(t)$ for $i \in \{1, 2, \dots, D\}$, given by the Fourier series,

$$\hat{f}_i(t) = \sum_{k=-L}^L b_i[k] \exp(j2\pi kt) \quad \text{for } t \in [0, 1], \quad (1)$$

where the coefficients $b_i[k]$ for $k \in \{-L, \dots, L\}$, will be a function of $a_i[k], k \in \mathbb{Z}$. In particular, the approximation coefficients $b_i[k]$ are chosen such that it minimizes the approximation error measured with respect to distance metric $d(\cdot)$, such as \mathcal{L}_1 or \mathcal{L}_2 norm, subject to an overprediction constraint. Mathematically, this is stated as,

$$\arg \min_{\hat{f}_i(t)} d(f_i(t), \hat{f}_i(t)) \quad \text{subject to } \hat{f}_i(t) \geq f_i(t). \quad (2)$$

Since the Fourier basis is orthogonal, the problem now translates into an equivalent optimization in terms of the coefficients $a_i[k]$'s and $b_i[k]$'s, when the approximation error is measured with the \mathcal{L}_2 norm. That is,

$$\arg \min_{b_i[k]} \sum_{k=-L}^L |a_i[k] - b_i[k]|^2 \quad \text{subject to } \hat{f}_i(t) \geq f_i(t). \quad (3)$$

The constraint in the considered optimization problem always ensures an over-prediction of the true signal, hence the signal approximation obtained by this method will be called the *envelope approximation*. For notational simplicity, we will denote the envelope approximation of a signal $f(t)$ by $\hat{f}_{\text{env}}(t)$.

D. Distance measures of interest

For the overpredictive signal approximation, we will construct an envelope for the signal recorded at each client device. The envelope approximation $\hat{f}_{\text{env}}(t)$ of the signal $f(t)$ is obtained by solving the optimization problem described in (2), with respect to a distance metric $d(\cdot)$ defined over the signal space. In this paper, the \mathcal{L}_1 and \mathcal{L}_2 distance measures will be analyzed. Using the Fourier representations of $\hat{f}_{\text{env}}(t)$ (akin to (1), with the subscript indices dropped) and $f(t)$, and the envelope property $\hat{f}_{\text{env}}(t) \geq f(t)$ it can be seen that,

$$\|\hat{f}_{\text{env}} - f\|_1 = b[0] - a[0], \quad (4)$$

$$\|\hat{f}_{\text{env}} - f\|_2^2 = \sum_{|k| \leq L} |b[k] - a[k]|^2 + \sum_{|k| > L} |a[k]|^2 \quad (5)$$

For ease of notation, hereafter the signal approximation error corresponding to the \mathcal{L}_1 and \mathcal{L}_2 distance measures will be denoted by SA_1 and SA_2 respectively.

III. OVERPREDICTIVE SIGNAL ANALYTICS IN FEDERATED LEARNING

Recall the distributed signal approximation application using federated learning illustrated in Fig. 1, where there are D devices and a central server. In this section we describe the algorithmic procedure for doing signal analytics in a federated learning setting, where each device reports an envelope approximation of the signal observed. The optimization program implemented at the client devices and the signal analytics performed at central server are discussed below.

At the clients: The analytics problem of interest at the client devices, is to determine the best possible approximation, which forms an envelope of the true signal. This can be stated as,

$$\min \text{SA}_q := \int_0^1 |\hat{f}(t) - f(t)|^q dt \quad \text{subject to } \hat{f}(t) \geq f(t)$$

for $q = 1, 2$, where the above minimizations are over $\hat{f}_1(t), \dots, \hat{f}_D(t)$. These device level analytics problems can be efficiently solved, by using the equivalent forms of the function norms, discussed in (4) and (5) respectively.

For $q = 1$, the above envelope approximation formulation will become a *linear program*, and when $q = 2$ it is equivalent to a *quadratic program with linear constraints*. Since linear programs are relatively easier to implement, the client devices with limited hardware can choose to solve \mathcal{L}_1 optimization problem over the \mathcal{L}_2 problem.

At the server: The signal approximations obtained from the clients will be combined at the server to learn the signal analytics of interest. Examples of such analytics will include the aggregate function, $\hat{s}(t) := \sum_i \hat{f}_i(t)$, or a statistical quantity such as the empirical CDF,

$$\hat{F}(x) := \frac{1}{N} \sum_{n=1}^N \mathbb{1}_{(-\infty, x]}(\hat{f}(t_n)), \quad (6)$$

where $\mathbb{1}_{(-\infty, x]}(Y)$ represents the 0–1 indicator function for the probability event $\{Y \leq x\}$, and t_n for $n \in \{1, 2, \dots, N\}$ represents the time samples in the interval $[0, 1]$. It is known that several statistical properties can be inferred using this classical Glivenko-Cantelli estimate of the CDF [12]. Since the Glivenko-Cantelli estimate satisfies the uniform convergence property, at the server there exists an implicit minimization of an empirical loss function involving the true and estimated CDFs. This minimization task conforms with the definition of federated optimization [2]. Application of CDF estimation based signal analytics will be discussed in the section V.

A. Algorithm sketch for over-predictive signal analytics

It is assumed that each client device in the federated learning model works in a distributed manner. To ensure $\hat{f}_i(t) \geq f_i(t)$ for $i \in \{1, 2, \dots, D\}$, we propose that each device can perform envelope approximation of its observed signal. The following steps are proposed for obtaining a bandwidth- L approximation $\hat{f}_i(t)$ of $f_i(t)$:

- 1) Each device i records its individual signal $f_i(t)$, calculates its envelope $\hat{f}_{i,\text{env}}(t)$, and communicates its $(2L + 1)$ Fourier coefficients to the server.
- 2) Using Fourier coefficients from each client device, the server calculates a global model of interest, denoted as $G_{\text{server}}(\hat{f}_{1,\text{env}}, \hat{f}_{2,\text{env}}, \dots, \hat{f}_{D,\text{env}})$.

Signal envelope calculation in Step 1 above is outlined next. For \mathcal{L}_1 distance (see (4)), it will be calculated as,

$$\begin{aligned} & \text{minimize } b[0] - a[0] \\ & \text{subject to } \vec{b}^T \Phi(t) \geq f(t), \end{aligned} \quad (7)$$

where $\Phi(t) = [\exp(-2\pi Lt), \dots, \exp(2\pi Lt)]^T$ and $\vec{b} = (b[-L], \dots, b[L])^T$ are the Fourier series coefficients of the envelope approximation. The above linear program with linear constraints is solvable efficiently [13]. For \mathcal{L}_2 , the cost function $b[0] - a[0]$ is replaced by the quadratic cost in (5).

As L is increased, the envelopes $\hat{f}_{i,\text{env}}(t)$ become more proximal to their target $f_i(t)$. It is expected that the approximation errors, SA_1 and SA_2 will decrease as L increases. However, analyzing the dependence of SA_q , $q = 1, 2$ versus L is difficult. Accordingly, a naïve envelope approximation will be used to analyze the fundamental bounds on their tradeoff.

B. A Naïve Envelope Approximation Scheme

First consider a single client device, $i = 1$ in isolation. Let $f_{1,\text{proj}}(t)$ be the orthogonal projection of $f_1(t)$ on the span of $\exp(j2\pi kt)$ for $|k| \leq L$. Then $f_{1,\text{proj}}(t) = \sum_{|k| \leq L} a_1[k] \exp(j2\pi kt)$. The naïve envelope approximation scheme is as follows [5]:

$$f_{1,\text{env}}(t) = f_{1,\text{proj}}(t) + C_0, \quad (8)$$

where $C_0 = \|f_1 - f_{1,\text{proj}}\|_\infty$. Using the triangle inequality,

$$C_0 \leq \sum_{|k| > L} |a_1[k]| \leq \sum_{|k| > L} \frac{C}{|k|^p}, \quad p > 1. \quad (9)$$

For $p > 1$ [14, Sec. 2.2], we can show that, $C_0 = O\left(\frac{1}{L^{p-1}}\right)$.

IV. ENVELOPE APPROXIMATION ANALYSIS

In this section we answer the question – for what class of signals is the naïve approximation scheme good? The result stated below shows that there exist certain signals for which the naïve approximation is order optimal to the optimal envelope scheme, which used \mathcal{L}_1 or \mathcal{L}_2 norm minimization.

Theorem 1. *Let SA_q be the optimal envelope approximation error computed with \mathcal{L}_q -norm based distance metric, and SA'_q be the approximation error corresponding to the naïve approximation in (8). Then we show that, there exist a class of signals such that for $q = 1, 2$,*

$$C_1(q) \leq \frac{\text{SA}'_q}{\text{SA}_q} \leq C_2(q) \text{ for } 1 \leq C_1(q) \leq C_2(q) < \infty,$$

where $C_1(q)$ and $C_2(q)$ are real valued constants.

Proof. : Using (9) and the triangle inequality on the signals and their envelopes, we determine bounds on SA_q and SA'_q . These are shown in Table I, and the desired result for $q = 2$ follows by taking ratios of SA_q and SA'_q . The detailed steps are omitted here due to space constraints. This result shows that the naïve approximation is *order optimal* for the \mathcal{L}^2 error, if the signals at the client devices are p -times differentiable.

The optimality of the naïve envelopes for the SA_1 distance holds for a signal class with the following properties: (i) the Fourier coefficients $a_1[k] \geq 0$, and (ii) $f_1(t)$ is real and even, that is $a_1[k] = a_1[-k]$. From these symmetry assumptions, it follows that $b_1[k] = b_1[-k]$. Restricted to this signal class, the SA_1 envelope approximation is re-stated as:

$$\begin{aligned} & \arg \min_{b_1[k], |k| \leq L} b_1[0], \quad \text{subject to} \\ & b_1[0] + 2 \sum_{1 \leq |k| \leq L} b_1[k] \cos(2\pi kt) \geq \sum_{|k| > L} a_1[k] e^{j2\pi kt} \end{aligned} \quad (10)$$

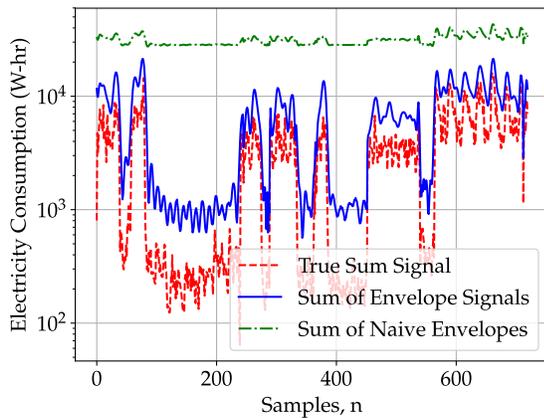


Fig. 2. The sum of signals obtained by distributed envelope approximation schemes are compared with the ground truth signal, when the number of coefficients communicated are $L = 324$. In the plot, the optimal envelope signal is seen to estimate the peak energy demand regions.

The above optimization can be shown to result in $b_{1,\text{opt}}[0] = \sum_{|k|>L} a_1[k]$ if $a_1[k] = 0$ for $|k| < L$. For this signal class, the naive approximation error, $SA'_1 = C_0 = \|f_1 - f_{1,\text{proj}}\|_\infty \leq \sum_{|k|>L} a_1[k]$ as all the Fourier coefficients, $a_1[k]$ are positive. Since $SA'_1 \geq SA_1$, we get $C_1(1) = C_2(1) = 1$. \square

TABLE I
BOUNDS ON THE APPROXIMATION ERRORS

\mathcal{L}_q norm	SA_q	SA'_q
$q = 1$	$b_1[0] - a_1[0]$	$\sum_{ k >L} a_1[k] $
$q = 2$	$\frac{2}{2p-1} \frac{1}{(L+1)^{2p-1}}$	$\frac{2}{2p-1} \frac{1}{L^{2p-1}}$

Remark: The presented result is for one (that is $D = 1$) client device. For D clients with a sum signal analytic at the server (that is $\hat{s}(t) := \sum_{i=1}^D \hat{f}_i(t)$), SA_1 as well as SA'_1 scale linearly with D . In contrast, SA_2 and SA'_2 will scale quadratically with D . Thus in both \mathcal{L}_1 and \mathcal{L}_2 norm based error, the ratio between the approximation errors will remain the same, as in the $D = 1$ case discussed in the proof.

V. EXPERIMENTS ON AN OFF-THE-SHELF RESIDENTIAL ENERGY CONSUMPTION DATASET

To analyze the effect of envelope approximation in a federated learning model, we have conducted experiments on an available electrical energy consumption dataset [15], consisting of hourly electricity consumption of 39 users from a residential building. The experimental results discussed here, are conducted on those users having atleast 30 days of data with synchronized timestamps – which turned out to be 37 users. In the original dataset, energy (in W-hr units) measurements corresponding to three phases were available. However, for the doing the experimental study we have considered only a

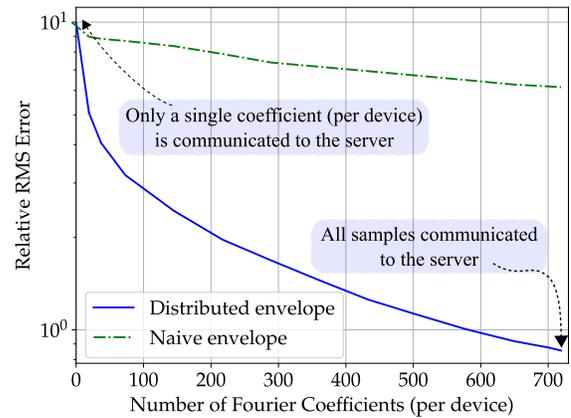


Fig. 3. This plot depicts the tradeoff between normalized root mean-squared error and the number of Fourier coefficients transmitted per device. The distributed envelope approximation scheme based on \mathcal{L}_2 -norm converges to the zero error as the number of communicated coefficients increase.

single phase (represented as W3 in the dataset). Due to space constraints, we have restricted the experimental results and discussions to the \mathcal{L}_2 norm based envelope approximation.

Hardware/Software Specifications: The simulations on the dataset were performed in a PC with the Processor model – Intel(R) Core(TM) i3-2310M CPU @ 2.10GHz, 2100 Mhz, 2 Core(s), RAM 6 GB; and implemented in MATLAB 2015b (Windows platform) using the standard curve fitting toolbox and CVX package (ver 2.1).

A. Communication Cost and Approximation Error Tradeoff

The tradeoff between the number of communicated Fourier coefficients and the \mathcal{L}_2 norm based approximation error, is studied by considering a sum signal analytic (or the sum of user energy consumptions) to be learned at the central server. Each Fourier coefficient communicated to the server is a real valued floating point number which is typically 4 Bytes (or 32 bits). In this setting, each client device (i.e. the 37 users) sends the Fourier coefficients of the envelope approximation, representing 30 days of hourly energy consumption, or 720 signal samples (refer Fig. 2 for the time-series plot of the sum signal analytic). In the tradeoff plot shown in Fig. 3, we notice that there is a graceful degradation of the relative root mean-squared (RMS) error of the optimal envelope approximation scheme, as the number of Fourier coefficients transmitted are reduced. It is observed that the distributed signal approximation scheme, represented by the solid blue line, approaches an error of zero when 720 Fourier coefficients are transmitted. However, the naive approximation, which adds only a constant to the projection based Fourier representation, fails to capture the envelope trend for the considered dataset. For reference, we have marked two extremum points – first corresponding to the least communication scheme (that is $L = 1$), and the second where all coefficients are communicated (that is $L = 360$). Further, in our experiments we observe that the relative approximation error of classical MSE minimizer (that

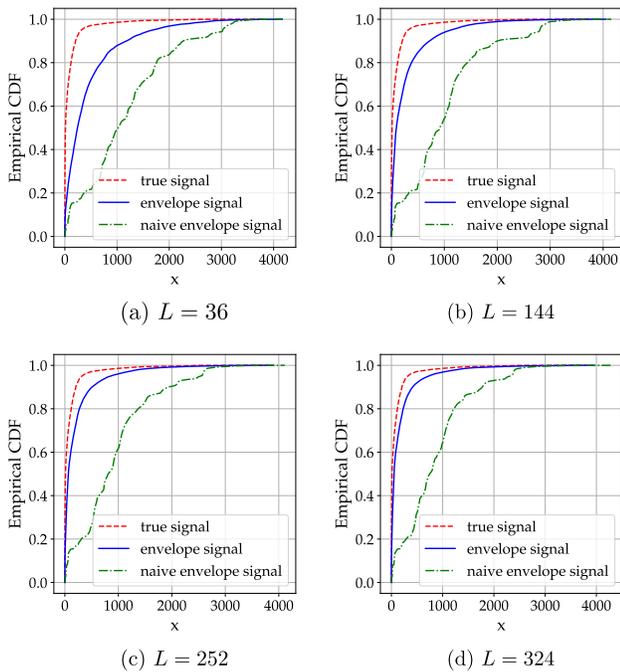


Fig. 4. Plot illustrating the convergence of the empirical CDFs of the envelope signal to the true CDF (obtained by pooling the raw signal samples at server). The optimum as well as the naïve envelope signal CDFs always appears towards the right of the true CDF, due to the envelope constraint.

is, without the envelope constraint) is lower than distributed envelope by one order of magnitude.

B. Cumulative Distribution Function based Signal Analytics

In this experimental study, we allow the central server to learn the CDF of the signal samples from the approximate signals reported by the individual clients. From the learned CDF, the central server can infer statistical properties, which are symmetric functions of the data, to measure the electricity usage patterns at the client devices. In Fig. 4, we illustrate the convergence of the CDF estimate obtained from the envelope approximation samples to the baseline CDF, obtained by a model which access to all the raw signal samples from all clients. At low approximation levels, for instance $L = 36$, the estimated CDF of the envelope scheme is much away from the true CDF. As we increase the number of communicated Fourier coefficients, the gap between the CDFs is observed to reduce. Another important observation is that the overprediction based CDFs (envelope as well as the naïve approximation schemes) always appears to the right of the true CDF. This is because of the envelope constraint at the devices.

Using these CDF plots, the quantile estimates can also be inferred. In Table II, we compare the quantiles of the true CDF with the quantiles of the CDF estimates. It is noted that the higher order quantiles deviate much as compared to the lower ones, which is attributed to the overprediction constraint at the devices. The accuracy of the quantile estimates in Table II can be improved by an appropriate choice of the basis representation. This shall be addressed in a future work.

TABLE II
A COMPARISON OF THE QUANTILES OF THE TRUE SIGNAL WITH THE ENVELOPE APPROXIMATION SIGNAL

Quantile	True Signal	Envelope Approximation		
		$L = 36$	$L = 180$	$L = 324$
10%	0	7.91	2.47	1.19
50%	6.09	168.26	53.192	33.13
90%	204.53	1181	631.81	413.89

VI. CONCLUSIONS

We considered the problem of overpredictive signal approximation in a client-server based federated learning model. Signal analytics considering this overpredictive approximations were discussed. Based on convex optimization methods, we proposed algorithms that can compute an overpredictive approximation of the signal, by minimizing the approximation error. The tradeoff between the number of signal samples communicated and the signal approximation error incurred was analyzed. These results were demonstrated on an off-the-shelf electrical energy consumption dataset. We envisage to extend the analysis to vector signals in a future work.

Acknowledgement: Part of this work was supported by Bharti Centre for Communication in IIT Bombay.

REFERENCES

- [1] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," in *NIPS Workshop*, 2016.
- [2] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial Intelligence and Statistics*, 2017.
- [3] D. P. Spanos, R. Olfati-Saber, and R. M. Murray, "Approximate distributed kalman filtering in sensor networks with quantifiable performance," in *IPSN 2005.*, April 2005, pp. 133–139.
- [4] S. Kar, J. M. F. Moura, and K. Ramanan, "Distributed parameter estimation in sensor networks: Nonlinear observation models and imperfect communication," *IEEE Transactions on Information Theory*, 2012.
- [5] G. Maheshwari and A. Kumar, "Optimal quantization of tv white space regions for a broadcast based geolocation database," in *2016 24th European Signal Processing Conference (EUSIPCO)*, Aug 2016.
- [6] M. M. Amiri and D. Gündüz, "Federated learning over wireless fading channels," *IEEE Transactions on Wireless Communications*, 2020.
- [7] Y. M. Saputra, D. T. Hoang, D. N. Nguyen, E. Dutkiewicz, M. D. Mueck, and S. Srikanteswara, "Energy demand prediction with federated learning for electric vehicle networks," *arXiv:1909.00907*, 2019.
- [8] A. T. Suresh, F. X. Yu, S. Kumar, and H. B. McMahan, "Distributed mean estimation with limited communication," in *Proceedings of the 34th International Conference on Machine Learning*, 2017.
- [9] L. Sankar, S. R. Rajagopalan, S. Mohajer, and H. V. Poor, "Smart meter privacy: A theoretical framework," *IEEE Trans. on Smart Grid*, 2013.
- [10] G. Giacconi, D. Gunduz, and H. V. Poor, "Privacy-aware smart metering: Progress and challenges," *IEEE Signal Processing Magazine*, 2018.
- [11] S. Mallat, *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*, Academic Press, 3rd edition, 2008.
- [12] M. J. Wainwright, *High-dimensional statistics: A non-asymptotic viewpoint*, vol. 48, Cambridge University Press, 2019.
- [13] A. Kumar, "Optimal envelope approximation in fourier basis with applications in TV white space," *arXiv:1706.00900*, 2017.
- [14] R. Bhatia, *Fourier Series*, Hindustan Book Agency, 2nd edition, 1993.
- [15] P. M. Mammen, H. Kumar, K. Ramamritham, and H. Rashid, "Want to reduce energy consumption, whom should we call?," in *Proceedings of the Ninth Intl. Conference on Future Energy Systems*. ACM, 2018.