Privacy-Preserving Distributed Optimisation using Stochastic PDMM

Jordan, Sebastian O.; Li, Qiongxiu; Heusdens, Richard

PRIVACY-PRESERVING DISTRIBUTED OPTIMISATION USING STOCHASTIC PDMM

Sebastian O. Jordan[1], Qiongxiu Li[2] and Richard Heusdens[1,3]

[1] Delft University of Technology, the Netherlands
[2] Tsinghua University, China
[3] Netherlands Defence Academy, the Netherlands

ABSTRACT

Privacy-preserving distributed processing has received considerable attention recently. The main purpose of these algorithms is to solve certain signal processing tasks over a network in a decentralised fashion without revealing private/secret data to the outside world. Because of the iterative nature of these distributed algorithms, computationally complex approaches such as (homomorphic) encryption are undesired. Recently, an information theoretic method called subspace perturbation has been introduced for synchronous update schemes. The main idea is to exploit a certain structure in the update equations for noise insertion such that the private data is protected without compromising the algorithm's accuracy. This structure, however, is absent in asynchronous update schemes. In this paper we will investigate such asynchronous schemes and derive a lower bound on the noise variance after random initialisation of the algorithm. This bound shows that the privacy level of asynchronous schemes is always better than or at least equal to that of synchronous schemes. Computer simulations are conducted to consolidate our theoretical results.

Index Terms— Privacy, distributed optimisation, convex optimisation, PDMM, asynchronous algorithms

1. INTRODUCTION

In recent years, the large increase in connected devices and the data that is collected by these devices has caused a heightened interest in distributed signal processing. As many practical distributed networks are of heterogeneous nature, algorithms operating within these networks need to be simple, robust against network dynamics, asynchronous and resource efficient. Popular applications of distributed signal processing include telecommunication, wireless sensor networks, cloud computing and machine learning [1, 2, 3]. Due to the absence of a central processing point (fusion center), participants/agents/nodes use their own processing ability to locally carry out simple computations and transmit only the required and partially processed data to neighbouring nodes. The transmission of the data could potentially lead to information leakage. In certain applications like medical systems, financial analysis, and smart grids [4], the data held by the agents/participants is privacy sensitive and needs to be protected from being revealed to the outside world.

One way to protect the data from being revealed is to apply some form of encryption, like homomorphic encryption [5], to encrypt the private data such that the adversary cannot decrypt the encrypted data within a certain period of time. Such approaches e.g. [6, 7] are often computationally complex, because the encryption functions require high computation overhead. Alternatively, we could rely on techniques based on information-theoretical principles where the adversary is assumed to have unbounded compute power but does not have sufficient information for inferring the secret/private data. In this paper we will focus on information-theoretic approaches because these methods are usually computationally lightweight as they do not require heavy encryption functions.

Existing information-theoretic approaches for privacy-preserving distributed optimization can be classified into three classes. First, one can use differential privacy [8] to ensure user privacy, see e.g. [9, 10, 11, 12]. It guarantees privacy in the worst case that all nodes in the network except one are not trustworthy. That is, a node's privacy is ensured to be protected even though all other nodes in the network can collaborate. While this gives strong privacy guarantees, the downside is that it compromises accuracy. Second, for certain distributed optimization problems one can use secret-sharing based techniques [13], like the work proposed in [14, 15, 16]. Secret sharing works by splitting the sensitive information into a number of shares, and the privacy is guaranteed under the assumption that the sensitive information cannot be reconstructed unless a sufficient number of nodes cooperate together, which is assumed to be infeasible. The main advantage is that it achieves privacy at no accuracy cost. However, it is shown to be only applicable to certain types of problems, and for each problem it requires a tailored adoption. In addition, it often requires the underlying graph to be fully-connected, which limits the scalability of the network in practice. The third privacy-preserving method is the subspace perturbation method [17, 18, 19], which provides a lightweight yet general solution to ensure privacy by inserting noise in a subspace that is not affected during the iterations. The noise insertion will obfuscate the private data, but does not affect the optimization result. Moreover, the method is generally applicable to a variety of distributed optimization problems.

Although subspace perturbation seems to be a promising method for preserving the privacy, the results in [17, 18, 19] were derived for a synchronous update setting. For asynchronous update schemes, however, there does not exist a subspace that is not affected during the iterations. In this paper we will investigate such asynchronous schemes and derive a lower bound on the noise variance that can be used to obfuscate the private data when the probability of a node being updated is uniform. The bound guarantees that the obtained privacy level of asynchronous schemes will be higher than (at least equal to) the original subspace perturbation approach proposed for synchronous schemes, while the algorithm's accuracy still remains uncompromised.

2. NOTATIONS AND PROBLEM STATEMENT

Consider a graphical model $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of vertices representing the nodes/agents in a network and $\mathcal{E} = \{(i, j) : i, j \in \mathcal{V}\}$ is the set of (undirected) edges in the graph representing

the communication links in the network. We use $\mathcal{E}_d$ to denote the set of all directed edges (ordered pairs). Therefore, $|\mathcal{E}_d| = 2|\mathcal{E}|$. Hence, given a graph $G = (\mathcal{V}, \mathcal{E})$, only neighbouring nodes are allowed to communicate with each other directly. We use $\mathcal{N}_i$ to denote the set of all neighbouring nodes of node i, i.e., $\mathcal{N}_i = \{j : (i,j) \in \mathcal{E}\}$ and $d_i = |\mathcal{N}_i|$ denotes the number of neighbouring nodes of node i.

Given a vector x, we use $\|x\|$ to denote its ℓ_2-norm. When x is updated iteratively, we write $x^{(k)}$ to indicate the update of x at the kth iteration. When we consider $x^{(k)}$ as a realisation of a random variable, the corresponding random variable will be denoted by $X^{(k)}$ (corresponding capital). Furthermore, we use the following notational conventions: a variable z_i is related to node i; a variable $z_{i|j}$ is related to edge (i,j) but held by node i.

We would like to solve the following optimisation problem:

$$
\begin{aligned}
\text{minimise} \quad & \sum_{i \in \mathcal{V}} f_i(x_i) \\
\text{subject to} \quad & \mathrm{A}_{ij} x_i + \mathrm{A}_{ji} x_j = \mathrm{b}_{ij}, \quad (i,j) \in \mathcal{E},
\end{aligned} \tag{1}
$$

where $f_i : \mathbb{R}^{n_i} \mapsto \mathbb{R} \cup \{\infty\}$ are CCP functions, $\mathrm{A}_{ij} \in \mathbb{R}^{m_{ij} \times n_i}$ and $\mathrm{b}_{ij} \in \mathbb{R}^{m_{ij}}$. Let $n = \sum_i n_i$ and $m = \sum_{(i,j)} m_{i,j}$. For simplicity, we will assume that $n_i = n_j = m_{ij}$, $\mathrm{A}_{ij} = -\mathrm{A}_{ji} = \mathrm{I}$ if $i > j$, and $\mathrm{b}_{ij} = 0$ for all $(i,j) \in \mathcal{E}$. This corresponds to simple (consensus) edge constraints of the form $(\forall (i,j) \in \mathcal{E}) \; x_i = x_j$. The results, however, can straightforwardly be generalised to arbitrary dimensions and arbitrary (linear) edge constraints.

The local objective functions f_i are a function of the primal optimisation variable x_i, but also of data related to node/agent i. Examples of such cost functions are $f_i(x_i) = \|x_i - s_i\|^2$ in the case of average consensus, where s_i is the local measurement data that needs to be averaged among the nodes/agents, or $f_i(x_i) = \|\mathrm{A}_i x_i - \mathrm{b}_i\|^2$ in the case of linear regression, where the pair $s_i = (\mathrm{A}_i, \mathrm{b}_i)$ are the independent (explanatory) and dependent (response) variables at node i. In certain applications like medical systems, financial analysis, and smart grids, the data is privacy sensitive and needs to be protected for being revealed to the outside world. In this paper we will focus on solving (1) while protecting the secret/private data s_i from being revealed to the outside world.

3. BACKGROUND

In this section, we formulate the mathematical background of privacy-preserving distributed optimisation and introduce the primal-dual method of multipliers (PDMM) updating equations that will be used in the proof in our work. For more details regarding PDMM we refer to [20, 21].

3.1. PDMM

PDMM is an iterative algorithm that can solve the optimisation problem stated in (1). An insightful derivation of the algorithm using monotone operator theory is given in [21]. As derived in [21], PDMM can be formulated as a monotone inclusion problem that can be iteratively solved using Peaceman-Rachford splitting (see [22, Sec. 26.4]). This leads to the following PDMM update equations:

$$
x^{(k+1)} = \arg \min_x \left(f(x) + z^{(k)\mathrm{T}} \mathrm{C} x + \frac{c}{2} \|\mathrm{C}x\|^2 \right), \tag{2}
$$

$$
z^{(k+1)} = \mathrm{P} z^{(k)} + 2c\mathrm{PC} x^{(k+1)}, \tag{3}
$$

where $x \in \mathbb{R}^n$ is the primal variable, $z \in \mathbb{R}^{2m}$ an auxiliary variables, and $c > 0$ is a constant that controls the convergence rate. The matrix $\mathrm{C} \in \mathbb{R}^{2m \times n}$ is a constraint matrix constructed as $\mathrm{C} =$

$(c_1, c_2, \ldots, c_{|\mathcal{V}|})$, $c_i \in \mathbb{R}^{2m \times n_i}$, where $c_i(l) = \mathrm{A}_{ij}$ if and only if $e_l = (i,j) \in \mathcal{E}$ and $i < j$, and $c_i(l+m) = \mathrm{A}_{ij}$ if and only if $e_l = (i,j) \in \mathcal{E}$ and $i > j$. $\mathrm{P} \in \mathbb{R}^{2m \times 2m}$ is a symmetric permutation matrix exchanging the first m with the last m rows.

Due to the construction of the PDMM algorithm, the update equations are separable across the nodes of the network and thus can be performed in a distributed/parallel manner. At every iteration, each node i updates its local optimisation variable x_i, and computes, for every neighbouring node $j \in \mathcal{N}_i$, an auxiliary variable $z_{j|i}$. In particular, (2)-(3) can be computed as

$$
x_i^{(k+1)} = \arg \min_{x_i} \left(f_i(x_i) + \sum_{j \in \mathcal{N}_i} z_{i|j}^{(k)\mathrm{T}} \mathrm{A}_{ij} x_i + \frac{c}{2} d_i x_i^{(k+1)} \right), \tag{4}
$$

$$
(\forall j \in \mathcal{N}_i) \quad z_{j|i}^{(k+1)} = z_{i|j}^{(k)} + 2c\mathrm{A}_{ij} x_i^{(k+1)}. \tag{5}
$$

PDMM converges for strongly convex and differentiable cost functions. This holds for both synchronous [21] and asynchronous [23] update schemes. To guarantee convergence for arbitrary closed, convex and proper (CCP) cost functions, we can apply operator averaging, resulting in the θ-averaged auxiliary vector update

$$
z_{j|i}^{(k+1)} = (1-\theta)z_{j|i}^{(k)} + \theta\big(z_{i|j}^{(k)} + 2c\mathrm{A}_{ij} x_i^{(k+1)}\big), \tag{6}
$$

with $\theta \in (0,1)$.

3.2. Threat model and privacy metric

In this paper we will consider two widely-used adversary models: the eavesdropping and the passive (or honest-but-curious) adversary model. The passive adversary consists of a number of colluding nodes, referred to as *corrupt nodes*, which will follow the algorithm instructions but will use received information to infer the input data of the other, non-corrupt nodes. We will refer to the latter as *honest nodes*. The eavesdropping adversary is usually neglected in existing schemes since eavesdropping can be prevented by using channel encryption [24]. However, channel encryption is computationally expensive. For iterative algorithms where the communication channels are used many times, channel encryption is less attractive. We thus assume that the communication in the network is performed through non-secure channels, except for the communication during the initialisation of the network, as will be explained later.

Two commonly used metrics for measuring the privacy loss are ϵ-differential privacy and mutual information [25]. In this paper we choose mutual information as the individual privacy metric. The reason for it is twofold. Firstly, it has been proven effective in the context of privacy-preserving distributed processing [26], and has been applied in various applications [27]. Secondly, mutual information is closely related to ϵ-differential privacy (see [28] for more details), and is easier to realise in practice [29]. Let S_i be the secret/private continuous random variable that relates to node i having variance $\sigma_{S_i}^2$. Let $Y_i = S_i + N_i$ be what the adversaries can observe related to S_i, where N_i is a random variable that is statistically independent of S_i, acting as noise obfuscating S_i. Using $I(S_i; Y_i)$ to denote the mutual information between S_i and N_i, we can define δ-level privacy as $I(S_i; Y_i) \leq \delta$.

4. PRIVACY PRESERVATION

To analyse what information the adversaries can deduce, we can express the primal update equation (4) as

$$
0 \in \partial f_i(x_i^{(k+1)}) + \sum_{j \in \mathcal{N}_i} \mathrm{A}_{ij} z_{i|j}^{(k)} + cd_i x_i^{(k+1)}. \tag{7}
$$

8572

The private data of the nodes is present in the objective function f_i, so the only term in (7) that contains private data is $\partial f_i(x_i^{(k+1)})$. Through knowledge of the PDMM algorithm, adversaries can use the inclusion stated above to deduce information about the term $\partial f_i(x_i^{(k+1)})$. The data that is known to the adversaries depends on the implementation of PDMM.

Since the adversary can eavesdrop all communication channels, by inspection of (7), transmitting the auxiliary variables $z_{i|j}^{(k)}$ would reveal $\partial f_i(x_i^{(k+1)})$, as $x_i^{(k+1)}$ can be determined from (6). Encrypting $z_{i|j}^{(k)}$ at every iteration would solve the problem but is computationally too demanding. To overcome this problem, only initial values $z_{i|j}^{(0)}$ are securely transmitted and $\Delta z_{i|j}^{(k+1)} = z_{i|j}^{(k+1)} - z_{i|j}^{(k)}$ are transmitted (without any encryption) during subsequent iterations [19, 30]. As a consequence, $z_{i|j}^{(k+1)}$ can only be determined whenever $z_{i|j}^{(0)}$ is known. Let $\mathcal{N}_{i,h}$ denote the set of honest neighbours of node i and assume that $\mathcal{N}_{i,h} \neq \emptyset$. That is, we assume that each honest node has at least one honest neighbour. After deducing the known terms from (7), the adversaries can observe

$$\partial f_i(x_i^{(k+1)}) + \sum_{j \in \mathcal{N}_{i,h}} A_{ij} z_{i|j}^{(k)}. \tag{8}$$

By inspection of (8), we conclude that we can obfuscate the secret/private data s_i by making the variance of $Z_{i|j}$ sufficiently large. This can be achieved by initialising the auxiliary variables $Z_{i|j}^{(0)}$ with sufficiently large variance, assuming this variance remains sufficiently large during the iterations. In [17, 18, 19] it has been shown that for a synchronous update scheme this is indeed the case. The reason for this is that at every two synchronous PDMM updates the auxiliary variables are only affected in the subspace $\Psi = \mathrm{ran}(C) + \mathrm{ran}(PC)$ and leave $\Psi^\perp = \ker(C^T) \cap \ker((PC)^T)$, the orthogonal complement of Ψ, unchanged. As shown in [21], $\Pi_\Psi z^{(k)} \to z^*$ for any $z^{(0)}$, where $z_\Psi^{(k)} = \Pi_\Psi z^{(k)}$ denotes the orthogonal projection of $z^{(k)}$ onto Ψ. Because of this the term $\Pi_\Psi z^{(k)}$ will eventually always have the same value so that $Z_\Psi^{(k)}$ will have zero variance, and cannot be used to obfuscate the private data. For the component $z_{\Psi^\perp}^{(k)}$ in the orthogonal subspace $\Psi^\perp$, however, it has been shown in [30] that

$$\mathbb{E}\left(Z_{\Psi^\perp}^{(k)} Z_{\Psi^\perp}^{(k)\mathrm{T}} \right) = \Pi_{\Psi^\perp} \left(\frac{\sigma^2}{2} \left((I+P) + |1-2\theta|^{2k}(I-P) \right) \right).$$

This expression indicates that the variance of $Z_{\Psi^\perp}^{(k)}$ has a nonzero asymptotic component that depends on the initialisation variance, and can be made arbitrarily large to obtain δ-level privacy.

For a stochastic update scheme, however, the z-updates are not restricted to the subspace Ψ anymore [23], and deriving a bound for the variance of the auxiliary variable z in this case is more involved than in the synchronous case. In the next section we will derive a lower bound for the variance of $Z_{\Psi^\perp}^{(k)}$ in a stochastic setting[1].

5. PRIVACY PRESERVATION WITH STOCHASTIC PDMM

Following [23], stochastic updates can be defined by assuming that each auxiliary variable $z_{i|j}$ can be updated based on a Bernoulli random variable $U_{i|j} \in \{0, 1\}$. Collecting all random variables $U_{i|j}$ in

---

[1]Since in this framework we transmit differential auxiliary variables $\Delta z_{i|j}$, we assume that there are no transmission losses. With such losses the algorithm as presented here will fail to converge.

the random vector $U \in \mathbb{R}^{2m}$, following the same ordering as the entries of z, let $(U^{(k)})_{k \in \mathbb{N}}$ denote an i.i.d. random process defined on a common probability space $(\Omega, \mathcal{A})$. Hence, $U^{(k)}(\omega) \subseteq \{0, 1\}^{2m}$ indicates which entries of $z^{(k)}$ will be updated at iteration k, where we assume that at every iteration, entry $z_{i|j}^{(k)}$ has nonzero probability to be updated. With this, a stochastic PDMM iteration can be expressed as

$$Z^{(k+1)} = \left(I - \theta U^{(k+1)} \right) Z^{(k)} + \theta U^{(k+1)} (PZ^{(k)} + 2cPCX^{(k+1)}). \tag{9}$$

We will utilise the following lemma and some properties of the conditional expectation to derive a lower bound on the variance of $Z_{\Psi^\perp}^{(k)}$ for stochastic PDMM.

Lemma 5.1. *Let $\Pi_{\Psi^\perp}$ denote the projection onto the subspace $\Psi^\perp$ and P denote the PDMM permutation matrix. Then*

  a) $\Pi_{\Psi^\perp} P = P \Pi_{\Psi^\perp}$,

  b) $\Pi_{\Psi^\perp} \mathbb{E}(X) = \mathbb{E}\left(\Pi_{\Psi^\perp} X \right)$,

  c) $\Pi_{\Psi^\perp} \left(PC x^{(k+1)} \right) = 0$.

Proof. For a), see [30, Lemma 5.2]. Item b) follows from the linearity of the expectation operator. Item c) follows from the fact that $PC x^{(k+1)} \in \Psi$. $\square$

To simplify the analysis, we make use of the observation that in expected value stochastic PDMM with uniform updating probabilities, that is, $(\forall i \in \mathcal{V})(\forall j \in \mathcal{N}_i) \Pr(\{U_{i|j}^{(k)} = 1\}) = \mu$, is equivalent to μ-averaged PDMM. This assumption holds, for example, in the case of asynchronous PDMM where one node is selected uniformly at random and updated each iteration, so that $\mu = 1/|\mathcal{V}|$.

Let $(\mathcal{A}_k)_{k \geq 1}$ be a filtration on $(\Omega, \mathcal{A})$ such that $\mathcal{A}_k := \sigma\{U^{(m)} : m \leq k\}$, the σ-algebra generated by the random variables $U^{(1)}, \dots, U^{(m)}$ and thus $\mathcal{A}_k \subseteq \mathcal{A}_l$ for $k \leq l$. By taking the conditional expectation of (9) with respect to $\mathcal{A}_0$, and using the fact that $U^{(k+1)}$ is independent of $Z^{(k)}$ and $X^{(k+1)}$, we have that

$$\mathbb{E}(Z^{(k+1)}|\mathcal{A}_0) = \left(I - \theta \mathbb{E}(U^{(k+1)}) \right) \mathbb{E}(Z^{(k)}|\mathcal{A}_0) +$$
$$\theta \mathbb{E}(U^{(k+1)}) \left(P \mathbb{E}(Z^{(k)}|\mathcal{A}_0) + 2cPC \mathbb{E}(X^{(k+1)}|\mathcal{A}_0) \right)$$
$$= (1 - \theta\mu) \mathbb{E}(Z^{(k)}|\mathcal{A}_0) +$$
$$\theta\mu \left(P \mathbb{E}(Z^{(k)}|\mathcal{A}_0) + 2cPC \mathbb{E}(X^{(k+1)}|\mathcal{A}_0) \right),$$

where for the last step we assume equal updating probabilities. As discussed in Section 4, the part of Z in the subspace $\Psi^\perp$ is relevant for privacy preservation. Hence, we have by Lemma 5.1 that

$$\mathbb{E}(Z_{\Psi^\perp}^{(k)}|\mathcal{A}_0) = ((1-\theta\mu)I + \theta\mu P) \mathbb{E}(Z_{\Psi^\perp}^{(k-1)}|\mathcal{A}_0)$$
$$= ((1-\theta\mu)I + \theta\mu P)^k Z_{\Psi^\perp}^{(0)}$$
$$= \frac{1}{2} \left((I+P) + (1-2\theta\mu)^k(I-P) \right) Z_{\Psi^\perp}^{(0)}, \tag{10}$$

where the last equality is shown in [30] using the eigenvalue decomposition of $(1-\theta\mu)I + \theta\mu P$. Hence, we have

$$\mathbb{E}\left(Z_{\Psi^\perp}^{(k)} Z_{\Psi^\perp}^{(k)\mathrm{T}} \right)_{(i|j),(i|j)} = \mathbb{E}\left(\left(Z_{\Psi^\perp}^{(k)} \right)_{(i|j)}^2 \right)$$
$$= \mathbb{E}\left(\mathbb{E}\left(\left(Z_{\Psi^\perp}^{(k)} \right)_{(i|j)}^2 \Big| \mathcal{A}_0 \right) \right)$$
$$\geq \mathbb{E}\left(\mathbb{E}\left(\left(Z_{\Psi^\perp}^{(k)} \right)_{(i|j)} \Big| \mathcal{A}_0 \right)^2 \right),$$
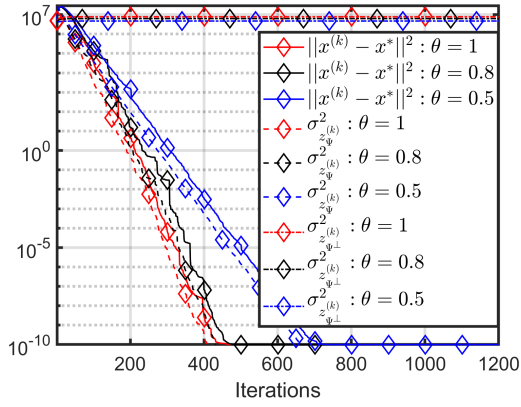
Fig. 1. Convergence of the variables $x^{(k)}$, $Z_\Psi^{(k)}$, and $Z_{\Psi\perp}^{(k)}$ using asynchronous PDMM for $\theta = 1, 0.8$, and 0.5.

Fig. 2. Variance of $Z_{\Psi\perp}^{(k)}$ using synchronous and asynchronous PDMM for $\theta = 1, 0.8$, and 0.5.

where the last inequality follows from Jensen's inequality. Hence, if the auxiliary variable is randomly initialised in such a way that $\mathbb{E}(Z^{(0)}Z^{(0)\mathrm{T}}) = \sigma^2 \mathrm{I}$, we have that

$$\mathrm{diag}\Big(\mathbb{E}\left(Z_{\Psi\perp}^{(k)} Z_{\Psi\perp}^{(k)\mathrm{T}}\right)\Big)$$

$$\geq \mathrm{diag}\Big(\mathbb{E}\left(\mathbb{E}\left(Z_{\Psi\perp}^{(k)}|\mathcal{A}_0\right)\mathbb{E}\left(Z_{\Psi\perp}^{(k)}|\mathcal{A}_0\right)^{\mathrm{T}}\right)\Big)$$

$$= \mathrm{diag}\Big(\Pi_{\Psi\perp}\Big(\frac{\sigma^2}{2}\left((\mathrm{I+P}) + |1 - 2\theta\mu|^{2k}(\mathrm{I-P})\right)\Big)\Big),$$

where the last equality follows from substitution of (10), Lemma 5.1, and the fact that $\Pi_{\Psi\perp}\Pi_{\Psi\perp}^{\mathrm{T}} = \Pi_{\Psi\perp}$. This inequality provides a lower bound for the variance of the auxiliary variable that is based on the initialisation variance σ^2. Note that this bound is equal to the variance of the auxiliary variables for synchronous $\theta\mu$-averaged PDMM. Since for most practical networks $\theta\mu \ll 1$, the variance will be lower bounded by the variance of synchronous PDMM.

6. NUMERICAL RESULTS

To validate the variance bound derived in Section 5, we perform distributed averaging simulations using asynchronous PDMM over a random geometric graph with $n = 10$ nodes having a communication radius $r = \sqrt{2 \log n / n}$ so that the graph is connected with probability at least $1 - 1/n^2$ [31]. The nodes are being activated each iteration with uniform probability $\mu = 1/|\mathcal{V}|$. The values in the private measurement data s were randomly generated from a zero-mean unit-variance Gaussian distribution and the auxiliary data $Z^{(0)}$ was initialized with zero mean and $\sigma^2_{Z^{(0)}} = 10^8$. To analyse the variance of $Z^{(k)}$, we performed 100 Monte Carlo runs with different initialisation $z^{(0)}$ where, for consistency, the sequence of selected nodes was kept constant for each Monte Carlo run.

Figure 1 shows convergence results of both the primal variable x and the auxiliary variable z as a function of the number of iterations for different values of θ. The primal variable $x^{(k)}$ converges to the correct solution and has, as expected, a decreasing convergence rate for increasing values of θ. As for the auxiliary variable z, we can see that the variance of $Z_\Psi^{(k)}$ goes to zero as $k \to \infty$, whereas the variance of $Z_{\Psi\perp}^{(k)}$ remains nonzero. Figure 2 shows the evaluation of $Z_{\Psi\perp}^{(k)}$ in more detail and shows a comparison with the results obtained with synchronous PDMM [30]. The results show
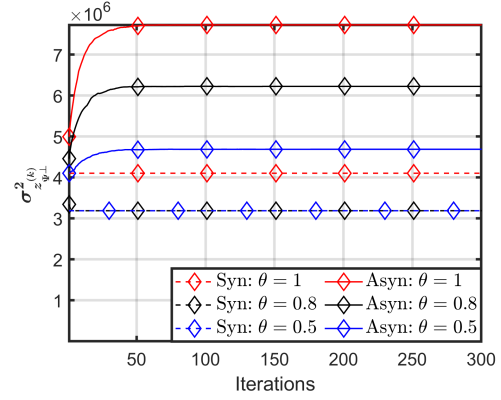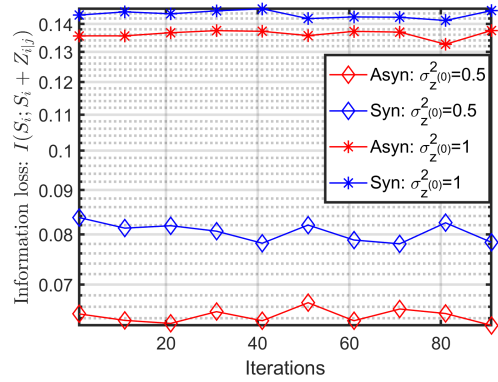

Fig. 3. Information loss in asynchronous and synchronous PDMM for $\sigma^2_{Z^{(0)}} = 0.5$ and $\sigma^2_{Z^{(0)}} = 1$.

that in both asynchronous and synchronous PDMM the variance of $Z_{\Psi\perp}^{(k)}$ is bounded below by a value depending on the variance of the initialisation $Z_{\Psi\perp}^{(0)}$, and that this variance is higher in the case of an asynchronous update scheme.

Figure 3 shows the normalized mutual information based on (8) by assuming there is one honest neighbor j in the neighborhood of node i for different values of the variance of $Z^{(0)}$. The results show that the larger $\sigma^2_{Z^{(0)}}$ is, the smaller the mutual information is between the private data and the observations of the adversary. It also confirms the fact that asynchronous update schemes reveal less information about the private data during the iterations than synchronous schemes.

7. CONCLUSION

In this paper we investigated the privacy preservation of asynchronous distributed optimisation schemes. We extended the results of subspace perturbation, originally introduced for synchronous schemes, to stochastic update schemes. We derived a lower bound on the noise variance that can be used to obfuscate the private data when the probability of a node being updated is uniform. The bound guarantees that the obtained privacy level of asynchronous schemes will be higher than the original subspace perturbation approach, while the algorithm's accuracy still remains uncompromised. Computer simulations confirmed our theoretical findings.

8574

References

[1] T. Richardson and R. Urbanke, *Modern Coding Theory*, Cambridge University Press, 2008.

[2] G. Zhang, R. Heusdens, and W. Bastiaan Kleijn, "Large Scale LP Decoding with Low Complexity," *IEEE Communications Letters*, vol. 17, no. 11, pp. 2152–2155, 2013.

[3] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Information Theory*, vol. 52, no. 6, pp. 2508–2530, 2006.

[4] G. Giaconi, D. Gündüz, H. V. Poor, "Privacy-aware smart metering: Progress and challenges," *IEEE Signal Process. Mag., vol. 35, no. 6, pp. 59-78*, 2018.

[5] C. Gentry, "Fully homomorphic encryption using ideal lattices," in *STOC, pp.169–178*, 2009.

[6] Kaihe Xu, Haichuan Ding, Linke Guo, and Yuguang Fang, "A secure collaborative machine learning framework based on data locality," in *2015 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2015, pp. 1–5.

[7] C. Zhang, M. Ahmad, and Y. Wang, "ADMM based privacy-preserving decentralized optimization," *IEEE Trans. Inf. Forensics Security, vol. 14, no. 3, pp. 565–580*, 2019.

[8] C. Dwork, F. McSherry, K. Nissim, A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Proc. Theory of Cryptography Conf. , pp. 265-284*, 2006.

[9] Z. Huang, S. Mitra, and N. Vaidya, "Differentially private distributed optimization., pp. 1–10," in *Proc. Int. Conf. Distrib. Comput. Netw*, 2015.

[10] E. Nozari, P. Tallapragada, and J. Cortés, "Differentially private distributed convex optimization via functional perturbation," *IEEE Trans. Control Netw. Syst., vol. 5, no. 1, pp 395-408*, 2018.

[11] X. Zhang, M. M. Khalili, and M. Liu, "Improving the privacy and accuracy of ADMM-based distributed algorithms," *Proc. Int. Conf. Mach. Lear. pp. 5796–5805*, 2018.

[12] Y. Xiong, J. Xu, K. You, J. Liu and L. Wu, "Privacy preserving distributed online optimization over unbalanced digraphs via subgradient rescaling," *IEEE Trans. Control Netw. Syst.*, 2020.

[13] R. Cramer, I. B. Damgård, and J. B. Nielsen, "Secure multiparty computation and secret sharing," 2015.

[14] N. Gupta, J. Katz, N. Chopra, "Privacy in distributed average consensus," *IFAC-PapersOnLine, vol. 50, no. 1, pp. 9515-9520*, 2017.

[15] Q. Li, I. Cascudo, and M. G. Christensen, "Privacy-preserving distributed average consensus based on additive secret sharing," in *Proc. Eur. Signal Process. Conf., pp. 1-5*, 2019.

[16] K. Tjell and R. Wisniewski, "Privacy preservation in distributed optimization via dual decomposition and ADMM," in *Proc. IEEE 58th Conf. Decis. Control., pp. 7203–7208*, 2020.

[17] Q. Li, R. Heusdens and M. G. Christensen, "Convex optimisation-based privacy-preserving distributed average consensus in wireless sensor networks," in *Proc. Int. Conf. Acoust., Speech, Signal Process., pp. 5895-5899*, 2020.

[18] Q. Li, R. Heusdens and M. G. Christensen, "Convex optimization-based privacy-preserving distributed least squares via subspace perturbation," in *Proc. Eur. Signal Process. Conf., pp. 2110-2114*, 2021.

[19] Q. Li, R. Heusdens and M. G. Christensen, "Privacy-preserving distributed optimization via subspace perturbation: A general framework," in *IEEE Trans. Signal Process., vol. 68, pp. 5983 - 5996*, 2020.

[20] G. Zhang and R. Heusdens, "Distributed optimization using the primal-dual method of multipliers," *IEEE Trans. on Signal and Information Processing over Networks*, vol. 4, no. 1, pp. 173 – 187, 2018.

[21] T. W. Sherson, R. Heusdens, and W. B. Kleijn, "Derivation and Analysis of the Primal-Dual Method of Multipliers Based on Monotone Operator Theory," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 5, no. 2, pp. 334–347, June 2019.

[22] Heinz H. Bauschke and Patrick L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics. Springer International Publishing, Cham, 2017.

[23] S. O. Jordan, T. W. Sherson, and R. Heusdens, "Convergence of stochastic pdmm," in *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1–5.

[24] D. Dolev, C. Dwork, O. Waarts, M. Yung, "Perfectly secure message transmission," *J. Assoc. Comput. Mach., vol. 40, no. 1, pp. 17-47,*, 1993.

[25] T. M. Cover and J. A. Thomas, *Elements of information theory*, John Wiley & Sons, 2012.

[26] Q. Li, J. S. Gundersen, R. Heusdens and M. G. Christensen, "Privacy-preserving distributed processing: Metrics, bounds, and algorithms," in *IEEE Trans. Inf. Forensics Security. vol. 16, pp. 2090–2103*, 2021.

[27] Y. Bu, S. Zou, and V. V. Veeravalli, "Tightening mutual information-based bounds on generalization error," *IEEE J. Sel. Areas Info. Theory. vol. 1, no. 1, pp. 121–130*, 2020.

[28] P. Cuff and L. Yu, "Differential privacy as a mutual information constraint," in *Proc. 23rd ACM SIGSAC Conf. Comput. Commun. Secur., pp 43–54*, 2016.

[29] M. Gtz, A. Machanavajjhala, G. Wang, X. Xiao, J. Gehrke, "Publishing search logs—a comparative study of privacy guarantees," *IEEE Trans. Knowl. Data. Eng. vol. 24, pp. 520 - 532*, 2011.

[30] Q. Li, R. Heusdens and M. G. Christensen, "Communication efficient privacy-preserving distributed optimization using adaptive differential quantization," *Signal Process.*, 2022.

[31] J. Dall and M. Christensen, "Random geometric graphs," *Physical review E, vol. 66, no. 1, pp. 016121*, 2002.