

# 北京交通大学

## 博士学位论文

面向高效鲁棒联邦学习的传算联合优化方法研究

Research on Joint Optimization of Transmission and Computation for  
Efficient and Robust Federated Learning

作者：范新

导师：霍炎

北京交通大学

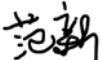
2022 年 12 月



## 学位论文版权使用授权书

本学位论文作者完全了解北京交通大学有关保留、使用学位论文的规定。特授权北京交通大学可以将学位论文的全部或部分内容编入有关数据库进行检索，提供阅览服务，并采用影印、缩印或扫描等复制手段保存、汇编以供查阅和借阅。同意学校向国家有关部门或机构送交论文的复印件和磁盘。学校可以为存在馆际合作关系的兄弟高校用户提供参考文献传递服务和交换服务。

(保密的学位论文在解密后适用本授权说明)

学位论文作者签名: 

导师签名: 

签字日期: 2022 年 12 月 8 日

签字日期: 2022 年 12 月 8 日



# 北京交通大学

## 博士学位论文

面向高效鲁棒联邦学习的传算联合优化方法研究

Research on Joint Optimization of Transmission and Computation for  
Efficient and Robust Federated Learning

作者姓名：范新

学 号：18111012

导师姓名：霍炎

职 称：教授

学位类别：工学

学位级别：博士

学科专业：信息与通信工程

研究方向：无线通信与机器学习

北京交通大学

2022 年 12 月



## 答辩委员会名单

答辩委员会	姓名	工作单位	职称
主席	艾 涠	北京交通大学	教授、博导
委员	王 莉	北京邮电大学	教授、博导
委员	赵雄文	华北电力大学	教授、博导
委员	赵友平	北京交通大学	教授、博导
委员	徐少毅	北京交通大学	教授、博导
秘书	高青鹤	北京交通大学	讲师



## 致谢

有知有觉，我已在交大十年了。十年踪迹十年心，回首学海入梦频。

十年交大路，那么漫长，已记不得了很多事儿，甚至回想起来，连我现在为什么会变成这样都已理不出头绪。但是这肯定与我当初的选择和做法一一相关。

高考失利，便放弃了努力。混日子的本科生活我猜应该和百分之八十的本科生一样吧。翘课、挂科、熬夜打游戏、睡觉、吃饭等等等，也没活出啥新意。本科即将毕业，却觉得自己啥都不会，害怕。于是十月一痛痛快快地玩了最后一周游戏，卸载游戏选择考研。过程也和大多数考研的一样，初试复试，分数尴尬，好歹是考上了。考上之后两三个月的日子才是最舒服的，毫无压力。那个时候才知道，原来没有目的的生活也是一种快乐，可以什么都不用想。快乐总是短暂的，暑假没放假被本科班主任叫去干活，美其名学点东西。开了学，我的硕导找我，也得干活。我当时还有十几门课，压力很大，太累了。虽然累，但是却感觉很充实，这种脚踏实地的感觉很迷人。干活的时候，我意识到了知识的用处，比如滑动滤波器之于美图秀秀。于是对一些东西产生了兴趣，开始想弄明白一些东西，也可以说是慢慢喜欢做科研吧。再加上硕士团队老师的劝说，我选择了读博。这个选择也考虑了很久，总之有自己的想法，有各种人的建议。现在想想其实当时并没有考虑得那么清楚，因为我根本不知道啥是读博，只是想跳出我当时的那种很累的状态，还不想早早毕业，依然害怕进入社会。

确定读博后，感觉确实不错，只剩上课和写论文。干劲也十足，成果也颇丰。硕士一年整了三篇论文，还和实验室其他人合作完成了几篇。这些成果仿佛印证了我适合读博的样子。以至于后来到我博二了，认识我的问我成果，我说没有，还被嘲笑，这些年还不如硕士一年呢。科研没成果的原因在于我的“眼高手低”，我想做出香农一样的成就，要打破香农极限。可惜能力不足，一些突破性的成果不是靠努力就能实现的。每天过得都一样，整天和论文打交道的生活也让心情很糟糕。眼高手低，心情不佳，科研自然停滞。

科研停滞，申请出国。来到了国外本以为可以学到新的东西后，心情会不一样。但发现还是没啥区别，依然还是枯燥的生活状态。只是换了个环境而已。又赶上了疫情，种族暴乱，变成了足不出户的生活。没来之前只听说美国，山好水好人寂寞。闭门生活，没让我体验到山水，只体验到了寂寞。一个人的时候总能引起人的思考，思考人生，活着的意义，思考将来。马上要毕业了，毕业了咋办呢。还有很多决定要做，还有很长却不知道多长的路要走。前路所有的选择导致了我的现状。

我一直认为，人生是由很多选择叠加组成的，每个选择都是一个正态分布。如果每次选得都是最好的，就会越变越好。但人不可能一直做出最好的选择。对于单

一选择，大多数人的选择会让自己趋于均值附近，就像考试的分数，均值附近的人最多，分数越高或越低都会导致人数变少。而人的一生要做很多的选择，就有很多的正态分布。这些正态分布相加结果还是正态分布。于是便简单地以为任何一个选择都不会决定一个人的一生。但是我却忽略了，有些选择前面的权重如果过大，足以确定整体正态分布走势，即人生走势。

看过很多电影，每一部电影都是一段人生。在美国的时候和室友们看了一部于谦老师的电影，叫《老师好》。里面于谦说过这样一句话，人这辈子会面对很多十字路口，而真正重要的其实就那么一两条。他这个答案其实是和我相悖的。后来想想确实如此，于谦在里面放弃北大中文系去读了师范，做出了选择。于是，一贫如洗，抱怨的妻子，不听话的学生，不公平的学校，小县城的一地鸡毛。看完电影，两室友笑道，“这可能就是将来的你”。

是的，我也想成为一名光荣的人民教师。之所以称之为光荣，是我认为有些职业是赋予特殊价值光环的。这种价值来源于使命。对于教师，这种使命是育人。我认为人活着不能只是为了自己，总要留下些什么。这种留下或许可以叫做传承。

活着不能只是为了自己，但前提是得好好活着，之后再去想其他的。人的选择必须基于方向或目的，否则就是盲目的。人很容易被自己当前所看所想所迷惑，以为自己以为的就是自己想要的。在一番经历过后，发现根本不是自己想要的样子。选择读博对于我是一种重大的决定，这或许不是我要的生活，但是我选择接受。

现在博士快结束了，回想起来，从本科到现在的十年路，真是一步步走到了现在。笑过哭过，发生了好多事儿。感觉除了没谈恋爱，我的大学生活基本圆满了。这一路走来遇到了很多人，很多事儿，值得我一生铭记感谢。

最该感谢的当然是我的博导霍炎教授。当初可真是手把手地教我如何搞科研。作为他的开山大弟子（第一个博士生），霍老师投入了太多的心血在我身上。可以说，没有霍老师就没有现在的我，可谓人生再造父母。按美国 TCU 的马利然老师的话说，一般开山弟子都是接班人来培养的。从零开始的艰辛，我和霍老师都不容易。霍老师对于我的帮助除了在科研上，还有为人处世上。霍老师肯于融入学生们的生活，每天和我们一起去食堂吃午饭，和我们一起运动健身。霍老师有一个豁达的胸怀，很多事在他眼中都不值一提。霍老师的言传身教之恩我将永远铭记于心。

感谢我在美国 GMU 访学时的导师 Zhi (Gerry) Tian 教授。我在她的身上看到了一个学术大佬该有的样子，对工作一丝不苟，论文已经是同意接收提交终稿阶段了还能改得一片红。工作科研用英语，只有生活琐事才会用中文。我学会了她对待科研和生活的态度。感谢她不计较论文署名排序问题，我才能满足北交大对博士发表论文署名的要求。感谢她能像对待她的亲学生一样对待我。由于疫情我回不了国，要申请延期，GMU 不批准，感谢她发长文和学校协商，帮我申请了延期并提供了

资助。

感谢 GMU 的王悦师兄。虽然他已经是老师了，但还是叫师兄更亲近些。王悦师兄是和我讨论科研工作最多的。毕竟导师在研究方向和内容上进行指导，具体细节还是要靠自己。实验室其他人又不熟悉我的研究内容，只有王悦师兄能帮我。也感谢王悦师兄在生活上给我的帮助，比如我在美国感染了新冠被室友赶出家门到了宾馆，他给我送了很多吃的和生活用品。在留学基金委的资助到期后，银行卡又被盗刷注销，感谢他借我钱帮我解决了一些经济问题。

感谢我的本科班主任侯亚丽老师。是她给我讲滑动滤波器如何用在美图秀秀的磨皮上，让我第一次感受到了学有所用，感受到了科研的魅力。

感谢我硕士实验室的侯建军老师，是他和我的硕导黄亮老师用激将法劝我读博，我才下定决心。他跟我说，你要是硕士毕业就结束了，你这辈子也就这样了，普普通通。要是读了博那就不一样了。你想想博士都读下来了，以后还有啥困难解决不了。我信了，于是才有了我的科研之路。虽然我读博之后，侯老师退休了，黄老师离职了，在科研上他们不能对我进行指导，但没有他们我可能不会读博，就不能体验这不一样的生活。

感谢国内博士实验室团队的荆涛老师和卢燕飞老师为实验室的付出，让我能有一个良好的科研环境。

感谢美国 TCU 的马利然老师在英文写作上对我的指导，感谢美国 GMU 的曾凯老师，王琨老师，David 叔叔和 Grace 阿姨对我的鼓励与关怀。感谢我在美国的第二任房东美云姐，在我感染新冠后找不到房子租，愿意收留我。

感谢国内实验室和我一起奋斗的李慧博士，高青鹤博士，温营坤博士，李学汉博士，吴玥博士，吴应臻博士，燕宇博士，杨兵博士，周宇迪博士，以及张伊慧、安茜雯、樊晶晶、吴元东等师妹师弟们。博士在读阶段没有过抱怨几乎不可能，聆听和倾诉很重要。感谢他们能和我一起分享快乐，诉说怨怒。

感谢美国 GMU 实验室和我一起奋斗的徐平博士，张宇博士，周冠强博士，张巍山博士。疫情期间，异国他乡求学十分不易。在组会上的简单交流也能让人感觉到不孤单。

感谢美国小哥 Josiah 帮我练习了一年多的英语口语和听力。

感谢我在美国时的室友让我见识了什么叫患难见真情。

感谢我在美国时遇到的韩煦，张晓荣，徐戈，单维嘉，Harry，Eden，Lilian，Aletheia，王明，曹仁杰，张筠，王可茵，王建平，尹莹，Carl，邬嘉诚，宁家琪，于凌帆等同学、好友们。感谢他们平时和我一起共度疫情的日子，感谢他们平时对我的帮助，带我去购物、参加活动等。

感谢我的高中同学彭跃，付文凯，黄宸，耿浩，贺庆，张帅，高凯，张亚光，

李亚超等人一路的陪伴。十多年了还能保持联系，这份友情得之不易，十分感激。

感谢那些关心我、支持我、不能一一提及的亲朋好友们。人的孤单大多来源于思想上的自我局限，有这么多关心我的人，我应心怀感恩，倍加珍惜。

感谢我的父母和家人，感谢他们对我的包容，理解，以及无条件地支持我所有的选择。在我感染新冠后，和我妈通电话，我很想哭，可我忍住了，我怕我哭她也哭。后来听说不光她哭了，我爸也哭了，她和我爸结婚这么多年，就见过我爸哭两次。第一次是我弟弟小时候得肺炎差点儿死掉，我爸跪下来哭着求医生给他治病。虽然我也曾抱怨过他们什么都帮不了我，可是我又为他们做了什么呢？这世上能为我哭的又有几人呢？父母养育之恩，感激涕零。

最后，感谢国家自然科学基金、北京市自然科学基金、国家教育部高等学校博士学科点专项科研基金，美国自然科学基金、教育部留学基金、及教育部研究生奖助学金对本文的资助。

## 摘要

近年来，各式智能终端随处可见，网络边缘产生了大量数据。传统的机器学习将多个设备的数据收集到中心服务器进行模型的训练，可以有效地利用不同设备的数据，从而训练出高准确率的预测模型。但是在海量设备参与的情况下，此种收集数据的集中式学习模式的通信代价很大，而且也存在数据隐私泄露的隐患。不同于集中式学习，联邦学习采用多个设备共同训练同一个模型。各个分布式设备只将训练好的模型参数传递给中心服务器，而原始数据保留在本地。中心服务器则将所有分布式设备的模型进行平均再反馈给各个分布式设备。通过模型参数的协作学习与共享（协作计算），联邦学习可以实现低时延的实时预测（本地执行预测）和模型训练的边缘化（个人数据隐私保护）。但是，当参与设备的数目过多或模型参数维度过高时，分布式设备之间的通信问题依然是联邦学习的瓶颈。

考虑到联邦学习关注所有设备更新的平均值而非单个设备的更新值，采用空中计算技术来传输设备的更新值更有利于节省通信带宽。空中计算技术是一种基于模拟聚合通信的传输技术，其允许多个设备使用相同的时频资源同时上传它们的更新，在接收端直接得到聚合值而非单体值。将空中计算技术与联邦学习相结合可以提升通信效率，但在理论分析、参数优化以及实际场景应用方面还存在着不限于以下问题尚未得到探索与解决：（1）空中计算技术涉及的模拟聚合传输对联邦学习性能影响的定量化描述问题，即模拟聚合通信与联邦学习之间的关系缺乏理论分析；（2）通信与学习联合优化问题，即相关的研究工作仅将通信和学习独立优化，并未考虑通信和学习两个过程的联合优化；（3）基于压缩感知技术的通信效率提高问题，即相关的研究工作在利用模拟聚合传输信号时一般采用的是发送全部内容，利用压缩感知技术来进一步提升通信效率的研究尚不充分；（4）恶意攻击场景中的鲁棒性增强问题，即相关的研究工作未考虑恶意用户的存在。

针对上述不足，本文针对基于空中计算技术的联邦学习（Federated Learning Over the Air, FLOA）中数据、学习以及通信所涉及的问题，以提升学习算法的准确性、高效性和鲁棒性为导向，首先推导模拟聚合通信对 FLOA 影响的定量化描述的理论结果，而后从带宽资源、通信效率以及对抗恶意攻击等三个方面提出解决方案。采用统计理论分析、建模和优化理论等数学方法，研究联邦学习和无线通信的联合优化方案，为通信高效的、鲁棒性增强的、隐私保护的联邦学习的推广与应用提供坚实保障。具体而言，本文主要研究内容和创新性成果总结如下：

1. 推导 FLOA 系统中的传输对学习性能影响的定量化描述的理论结果。当前已有的研究工作将模拟聚合传输和计算学习进行独立优化设计，缺少两者的联合优化。这是因为模拟聚合通信对 FLOA 的影响尚不明确，导致无法

将通信和学习过程联系起来。因此，本文通过推导 FLOA 在不同假设（如凸、非凸目标函数）下的闭合式收敛速率期望表达式，给出模拟聚合通信与联邦学习关系的定量化描述，其不仅揭示而且量化无线通信对联邦学习的收敛性和准确性的具体影响，为传算联合优化算法提供理论基础。

2. 提出一个基于计算收敛性的 FLOA 系统的传算联合优化框架。本文从问题制定、解决方案开发和算法实现的角度对 FLOA 进行全面研究，形成一个传算联合优化理论与方法框架。具体而言，借助模拟聚合通信与联邦学习关系的定量化描述的理论结果，本文首先建立一个学习、设备选择和功率控制的联合优化问题；接着，本文提出一种替代性的问题重构方法，将原始无法实现的优化问题近似为模拟聚合操作约束下的可行优化问题；最后，通过探索设备数量与功率缩放之间的关系，本文确定一个紧密的解空间，从而提出一种简单的离散枚举方法来有效地找到优化问题的全局最优解。
3. 提出一个基于 1 比特压缩感知的高效的 FLOA 方案。虽然模拟聚合通信可以减少带宽消耗，但其通信量依然与神经网络模型参数数目成正比。因此，本文引入 1 比特压缩感知来进一步提高 FLOA 的通信效率。具体而言，本文首先提出一个集稀疏化、降维、量化、模拟聚合传输和信号重建为一体的高效 FLOA 方案；接着，本文推导出所提方案的期望收敛行为的闭式表达式，其可衡量由稀疏化、降维、量化、信号重建和加性高斯白噪声引起的聚合误差导致的性能折衷；进而在上述理论结果的指导下，本文建立传算联合优化问题，以优化设备选择和功率控制来减轻聚合错误；最后，为解决此非凸优化问题，本文分别针对小型网络和大型网络场景提出两种解决方案：基于枚举的方法和交替乘子方向法。
4. 提出一个基于功率控制的鲁棒的 FLOA 方案。由于模拟聚合通信仅能得到设备更新的平均值而无法得到每个设备的单个更新，一些恶意节点会利用这一特点进行拜占庭攻击使得学习算法瘫痪，而且服务器无法检测出恶意节点。因此，本文提出一种最大努力投票机制，从功率控制角度来对抗拜占庭攻击。具体而言，本文首先分析恶意节点的行为对 FLOA 的影响；接着本文提出一种最大努力投票功率控制机制来对抗拜占庭攻击；最后，本文从收敛性的角度对抵御效果进行分析，从理论上证明所提鲁棒性方案的有效性。

全文共 5 章，图 27 幅，表 3 个，参考文献 128 篇。

**关键词：**联合优化；联邦学习；空中计算技术；高效性；鲁棒性；计算收敛性。

## ABSTRACT

In recent years, various intelligent terminals can be seen everywhere, generating a large amount of data at the edge of networks. In traditional machine learning, a server collects data from multiple devices for model training, which benefits to train a prediction model with high accuracy. However, with the participation of massive devices, the communication cost of such a machine learning mode of collecting data is very high, and there is also a risk of data privacy leakage. As an alternative, federated learning is a distributed learning paradigm in which multiple devices jointly train the same model instead of centralized training. In the process of federated learning, each distributed device only uploads the trained local model parameters to the central server, while the raw data is kept locally. The central server averages the models of all distributed devices and feeds them back to each distributed device. Through collaborative learning and sharing of model parameters (collaborative computing), federated learning can achieve real-time prediction with low delay (local execution prediction) and localization of model training (personal data privacy protection). However, when the number of participating devices is too large and the dimension of model parameters is too high, the communication between distributed devices and the central server is still the bottleneck of federated learning.

Considering that federated learning focuses on the average value of all local updates rather than the update value of a single device, using over the air computation technology for transmitting the update values of devices is more conducive to saving communication bandwidth. Over the air computation technology is a transmission technology based on analog aggregation communication, which allows multiple devices to upload their updates simultaneously using the same time-frequency resources, and directly obtain the aggregated value at the receiver. The combination of over the air computation technology and federated learning can improve communication efficiency, but there are still not limited to the following problems in theoretical analysis, parameter optimization and practical application that have not been explored and solved: (1) the quantitative description of the impact of analog aggregation transmission on federated learning performance involved in over the air computation technology, that is, the relationship between analog aggregation transmission and federated learning lacks theoretical analysis; (2) the joint optimization of communication and learning, that is, the relevant research work only optimizes communication or learning separately; (3) improving

communication efficiency by compressive sensing technology, that is, the relevant research work generally adopts the transmission of all content when transmitting signals by analog aggregation transmission, but the research on how to further improve communication efficiency by using compressive sensing technology is not sufficient; (4) robustness enhancement in malicious scenarios, that is, the related research work does not consider the existence of malicious users.

In view of the above deficiencies, facing the problems involved in data, learning and communication in federated learning over the air (FLOA), guided by the accuracy, efficiency and robustness of the learning algorithm, this dissertation first derives a quantitative description of the impact of analog aggregation communication on FLOA, and then proposes solutions from three aspects: joint optimization framework, communication efficiency and countering malicious attacks. Using mathematical methods such as statistical theory analysis, modeling and optimization theory, the joint optimization scheme of federated learning and wireless communication is studied to provide a solid guarantee for the promotion and application of federated learning with efficient communication, enhanced robustness and privacy protection. Specifically, the main research contents and innovative achievements of this dissertation are summarized as follows.

1. This dissertation proposes the methods for the quantitative description of the impact of transmission on learning performance in FLOA. At present, the existing research work focuses on the independent optimization design of analog aggregation transmission and computational learning, and lacks the joint optimization of the two. This is because the impact of analog aggregation communication on FLOA is not clear, resulting in the inability to link communication and learning processes. Therefore, by deriving the expected closed-form expressions of the convergence rate of FLOA under different assumptions (such as convex and non-convex objective functions), this dissertation provides a quantitative description of the relationship between analog aggregation transmission and federated learning, which not only reveals but also quantifies the specific impact of wireless communication on the convergence and accuracy of federated learning, and provides a theoretical basis for the joint optimization algorithm of transmission and computation.
2. This dissertation proposes a framework for the joint optimization of transmission and computation in FLOA. This dissertation aims to conduct a comprehensive

study of FLOA from the perspective of problem formulation, solution development and algorithm implementation, and form a theoretical and methodological framework of joint optimization of transmission and computation. Specifically, with the help of the theoretical results of quantitative description of the relationship between analog aggregation communication and federated learning, this dissertation first establishes a joint optimization problem of learning, device selection and power control; Then, this dissertation proposes an alternative problem reconstruction method, which approximates the original impossible optimization problem to a feasible optimization problem under the constraint of analog aggregation operation; Finally, by exploring the relationship between the number of devices and power scaling, this dissertation identifies a tight solution space, thereby proposing a simple discrete enumeration method to efficiently find the global optimal solution of the optimization problem.

3. This dissertation proposes a communication-efficient FLOA through 1-bit compressive sensing. Although analog aggregation communication reduces bandwidth consumption in FLOA, its traffic is still proportional to the amount of the parameters of the neural network model. Therefore, this dissertation introduces 1-bit compressive sensing to further improve the communication efficiency of FLOA. Specifically, this dissertation first proposes an efficient federated learning scheme that integrates sparsification, dimensionality reduction, quantization, analog aggregation transmission and signal reconstruction; Then, a closed-form expression of the expected convergence behavior of the proposed scheme is derived, which can measure the performance degradation caused by aggregation errors; Further, guided by the above theoretical results, this dissertation establishes a joint optimization problem to optimize device selection and power control to reduce aggregation errors; Finally, in order to solve this non-convex optimization problem, this dissertation proposes two solutions for small-scale and large-scale network scenarios respectively: an enumeration based method and an alternating multiplier direction method.
4. This dissertation proposes a robust FLOA through power control. Since the analog aggregation transmission can only get the average value of all local updates but cannot get a single update of each device, some malicious nodes will make full use of this character to carry out Byzantine attacks to paralyze the

learning algorithm, and the server cannot detect malicious nodes. Therefore, this dissertation proposes a best effort voting mechanism to combat Byzantine attacks from the perspective of power control. Specifically, this dissertation first analyzes the impact of the behavior of malicious nodes on FLOA; Then this dissertation proposes a best effort voting power control mechanism to combat Byzantine attacks; Finally, this dissertation analyzes the resistance effect from the perspective of convergence, which theoretically proves the effectiveness of the proposed robust scheme.

This dissertation consists of 5 chapters, including 27 figures, 3 tables and 128 references.

**KEYWORDS :** Joint optimization; Federated learning; Over the air computation; Efficiency; Robustness; Convergence of computation.

## 目录

<b>摘要 .....</b>	<b>IX</b>
<b>ABSTRACT.....</b>	<b>XI</b>
<b>图目录 .....</b>	<b>XIX</b>
<b>表目录 .....</b>	<b>XXI</b>
<b>缩略语 .....</b>	<b>XXIII</b>
<b>1 绪论.....</b>	<b>1</b>
1.1 研究背景及意义.....	1
1.2 国内外研究现状.....	4
1.2.1 基于信息压缩的通信效率增强方法研究现状.....	4
1.2.2 基于物理层技术的通信效率提升方法研究现状.....	4
1.2.3 联邦学习鲁棒性研究现状.....	6
1.2.4 当前研究尚存的问题.....	6
1.3 论文的主要工作与内容安排 .....	8
1.3.1 主要研究工作.....	8
1.3.2 论文的章节安排.....	10
<b>2 FLOA 传算联合优化框架研究.....</b>	<b>13</b>
2.1 系统模型 .....	14
2.1.1 联邦学习模型.....	14
2.1.2 基于模拟聚合的模型参数传输.....	16
2.2 模拟聚合通信对联邦学习影响的定量化描述 .....	18
2.2.1 基本假设.....	18
2.2.2 强凸损失函数假设下收敛性分析.....	18
2.2.3 非凸损失函数假设下收敛性分析.....	21
2.2.4 随机梯度下降法场景收敛性分析.....	22
2.3 通信与学习联合优化方法.....	24
2.3.1 优化问题建模.....	24
2.3.2 基于离散规划的求解方案.....	26
2.4 性能仿真与分析.....	28
2.4.1 系统设定.....	28
2.4.2 线性回归实验性能分析.....	29
2.4.3 图像识别实验性能分析.....	33
2.5 本章小结.....	34
<b>3 基于FLOA 传算联合优化框架的高效性方案研究 .....</b>	<b>35</b>
3.1 系统模型 .....	36
3.1.1 联邦学习模型.....	37
3.1.2 1比特压缩感知和模拟聚合传输过程.....	38
3.2 压缩和通信对 FLOA 的定量化描述 .....	41

3.2.1 基本假设 .....	41
3.2.2 收敛性分析 .....	42
3.3 通信与学习联合优化方法 .....	44
3.3.1 联合优化问题建模 .....	44
3.3.2 基于遍历算法的最优解 .....	45
3.3.3 基于 ADMM 算法的次优解 .....	46
3.4 SGD 算法下的拓展方案研究 .....	49
3.4.1 收敛性分析 .....	50
3.4.2 性能优化 .....	50
3.5 性能仿真与分析 .....	51
3.5.1 系统设定 .....	51
3.5.2 仿真结果与分析 .....	52
3.6 本章小结 .....	56
<b>4 基于 FLOA 传算联合优化框架的鲁棒性方案研究 .....</b>	<b>59</b>
4.1 系统模型 .....	60
4.1.1 联邦学习模型 .....	60
4.1.2 模拟聚合传输过程 .....	62
4.2 收敛性对比与分析 .....	65
4.2.1 基本假设 .....	65
4.2.2 最强拜占庭攻击 .....	66
4.2.3 信道反转方案下的收敛性分析 .....	66
4.2.4 最大努力投票方案下的收敛性分析 .....	68
4.3 性能仿真与分析 .....	69
4.3.1 系统参数设定 .....	69
4.3.1 无攻击场景下性能仿真与分析 .....	70
4.3.2 存在单个信道增益较弱的攻击者的性能仿真与分析 .....	70
4.3.3 存在单个信道增益较强的攻击者的性能仿真与分析 .....	72
4.3.4 存在多个随机选择的攻击者的性能仿真与分析 .....	73
4.4 本章小结 .....	74
<b>5 总结与展望 .....</b>	<b>75</b>
5.1 本文研究总结 .....	75
5.2 下一步研究展望 .....	77
<b>参考文献 .....</b>	<b>79</b>
<b>附录 A .....</b>	<b>89</b>
A1. 定理 2.1 的证明 .....	89
A2. 定理 2.2 的证明 .....	93
A3. 定理 2.3 的证明 .....	94
A4. 定理 2.4 的证明 .....	97
A5. 引理 3.1 的证明 .....	98
A6. 定理 3.1 的证明 .....	99
A7. 定理 3.2 的证明 .....	102

A8. 定理 4.1 的证明 .....	105
A9. 定理 4.2 的证明 .....	107
A10. 定理 4.3 的证明 .....	111
<b>作者简历及攻读博士学位期间取得的研究成果 .....</b>	<b>117</b>
<b>答辨决议书 .....</b>	<b>121</b>
<b>独创性声明 .....</b>	<b>123</b>
<b>学位论文数据集 .....</b>	<b>125</b>



## 图目录

图 1.1 联邦学习基础模型 .....	2
图 1.2 论文研究内容与结构示意图 .....	10
图 2.1 FLOA 网络模型示意图 .....	15
图 2.2 联邦学习算法的线性回归实验结果 .....	29
图 2.3 MSE 随迭代变化 .....	30
图 2.4 MSE 随分布式设备总数变化 .....	31
图 2.5 MSE 随每个设备的平均样本数变化 .....	31
图 2.6 MSE 随噪声方差变化 .....	32
图 2.7 交叉熵随迭代次数变化 .....	32
图 2.8 测试准确率随迭代次数变化 .....	33
图 3.1 基于 1 比特压缩感知的模拟聚合联邦通信模型框图 .....	37
图 3.2 (a) 不同稀疏化算子下的训练损失 .....	52
图 3.2 (b) 不同稀疏化算子下的测试准确率 .....	53
图 3.3 (a) 不同压缩后的维度大小下的训练损失 .....	53
图 3.3 (b) 不同稀疏化算子下的测试准确率 .....	54
图 3.4 (a) 不同联合优化算法在不同设备总数下的训练损失 .....	54
图 3.4 (b) 不同联合优化算法在不同设备总数下的测试准确率 .....	55
图 3.5 不同联邦学习方案在不同设备总数下的测试准确率 .....	55
图 3.6 学习准确率和通信开销之间的权衡 .....	56
图 4.1 不同功率控制策略下的 FLOA 模型 .....	61
图 4.2 (a) 没有攻击者的情况下训练损失 .....	70
图 4.2 (b) 没有攻击者的情况下测试准确率 .....	71
图 4.3 (a) 存在单个信道增益较弱的攻击者的训练损失 .....	71
图 4.3 (b) 存在单个信道增益较弱的攻击者的测试准确率 .....	72
图 4.4 (a) 存在单个信道增益较强的攻击者的训练损失 .....	72
图 4.4 (b) 存在单个信道增益较强的攻击者的测试准确率 .....	73
图 4.5 (a) 不同数量的攻击者下的训练损失 .....	73
图 4.5 (b) 不同数量的攻击者下的测试准确率 .....	74



## 表目录

表 2.1 INFLOTA 算法 .....	28
表 3.1 基于枚举的最优解算法 .....	45
表 3.2 基于 ADMM 的次优解算法.....	49



## 缩略语

缩略语	英文全称	中文全称
ADMM	Alternating Direction Method of Multipliers	交替方向乘子法
AirComp	Over-the-Air Computation	空中计算技术
AirShare	Over-the-Air Share	空中共享
AWGN	Additive White Gaussian Noise	加性高斯白噪声
BEV	Best Effort Voting	最大努力投票
BGD	Batched Gradient Descent	批量梯度下降法
BIHT	Binary Iteration Hard Thresholding	二进制迭代硬阈值
CI	Channel Inversion	信道反转
CNN	Convolutional Neural Network	卷积神经网络
CS	Compressive Sensing	压缩感知
CSI	Channel State Information	信道状态信息
DNN	Deep Neural Network	深度神经网络
DP	Discrete Programming	离散规划
EF	Error-free	无差错的
FedAvg	Federated Averaging	联邦平均算法
FL	Federated Learning	联邦学习
FLOA	Federated Learning Over the Air	基于空中计算技术的联邦学习
GD	Gradient Descent	梯度下降法
i.i.d.	Independent and Identically Distributed	独立同分布的
INFLOTA	Joint Optimization for FL Over the Air	基于空中计算技术的联邦学习的联合优化算法
IoT	Internet of Things	物联网
MAC	Multiple-access Channel	多址信道

---

MBGD	Mini-batched Gradient Descent	小批量梯度下降法
MIMO	Multiple Input Multiple Output	多输入多输出
MIP	Mixed Integer Programming	混合整数规划
MLP	Multi-layer Perceptron	多层感知机
MNIST	Mixed National Institute of Standards and Technology database	美国国家标准与技术研究院混合数据集
MSE	Mean Square Error	均方误差
OBCSAA	One Bit CS Analog Aggregation	1比特压缩感知和模拟聚合技术
OBDA	One-bit Broadband Digital Aggregation	1比特宽带数字聚合
PS	Parameter Server	参数服务器
ReLU	Rectified Linear Unit	线性整流函数
RIP	Restricted Isometry Property	有限等距性质
SGD	Stochastic Gradient Descent	随机梯度下降法
SNR	Signal-to-noise Ratio	信噪比
VGGNet	Visual Geometry Group Network	视觉几何组网络

# 1 绪论

本论文从学习算法的收敛性、无线通信的高效性、对抗恶意攻击的鲁棒性等角度凝练分布式学习中的通信瓶颈问题，确立研究内容为——高效鲁棒的联邦学习和模拟聚合通信联合优化理论与方法。本章具体内容安排如下：首先，1.1 节介绍研究背景和研究意义。然后，1.2 节对联邦学习中的通信问题的研究现状进行整理和总结，指出当前研究的不足之处。最后，1.3 节阐述本论文的主要研究工作和创新点，并给出本论文的具体章节安排。

## 1.1 研究背景及意义

近年来，物联网的发展动能尤为强劲，市场潜力得到产业界的普遍认可，技术和应用创新层出不穷。伴随着蓬勃发展的市场，物联网高速拓展已成为必然之势。根据 IoT Analytics 的调研表明，2020 年物联网连接（如联网汽车、智能家居设备、联网工业设备）的数量达到了 117 亿，在全球活跃连接设备总数中占比达到 54%，首次超过了非物联网连接（如智能手机、笔记本电脑和计算机）。到 2025 年，预计将有超过 300 亿的物联网连接，平均可达到每人近 4 台物联网设备<sup>[1]</sup>。中国信息通信研究院在 2020 年 12 月也发布了物联网白皮书，指出 2019 年我国的物联网连接数为 36.3 亿，其中移动物联网连接数占比 28.5%，并且预计到 2025 年我国物联网连接数将达到 80.1 亿<sup>[2]</sup>。国际数据公司的报告同样指出，在未来大规模海量连接的物联网设备会产生超过 79ZB 的数据<sup>[3]</sup>，并且随着联网设备的持续增加，由此产生的数据量也会不断增加。

为了有效地利用这些可用的大数据，传统的方式是收集各分布式设备的数据于某一中心节点，然后利用机器学习对大数据进行计算，进而形成有效的基于数据驱动的分析或推理模型，并给出更优的决策和战略性业务建议，实现诸如智慧医疗、环境监测等典型应用范式。然而，在万物互联和智能计算广泛实施部署的背景下，这种基于集中式机器学习的数据利用模式成为制约物联网持续发展的瓶颈<sup>[4]</sup>。主要原因包括：首先，该模式存在较大的传输延迟，难以应用于实时决策的应用场景（例如，车联网和自动驾驶汽车系统）；第二，将数据传送到中心节点进行学习或计算不仅会增大骨干网的负荷，更极大地依赖于无线通信网络质量，这对于非结构化数据任务（如视频分析与处理）尤是如此；最后，由于机器学习需要收集样本数据，而数据所有者需要高度的隐私安全，将用户隐私数据上传至某一中心节点进行处理会增大用户隐私泄露的风险。因此各个地区和国家都制定了相应的隐私安全法，例如欧盟通过了通用数据保护条例(General Data Protection Regulation, GDPR)，

美国颁布了消费者隐私权利法案（Consumer Privacy Bill of Rights, CPBR），我国也于 2021 年 8 月通过了“个人信息保护法”。

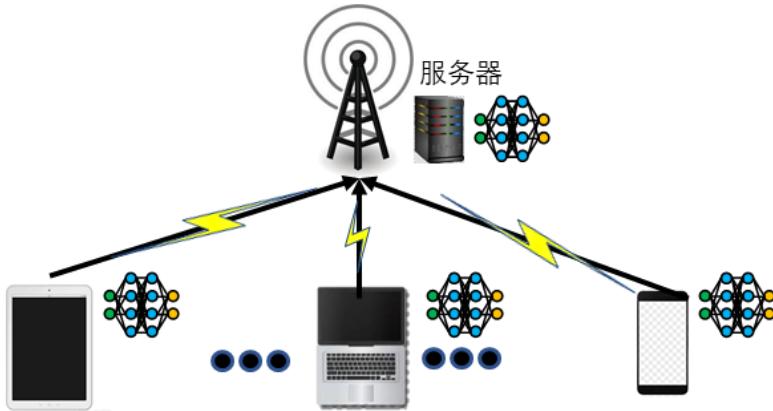


图 1.1 联邦学习基础模型

Figure 1.1 Basic model of federated learning

集中式机器学习应用分布式数据的模式存在局限性，因而催生出基于联邦学习的边缘计算新范式<sup>[5-10]</sup>。联邦学习的主要思想是在不共享原始数据的情况下使得多个设备共同学习。如图 1.1 所示，联邦学习模型包含一个服务器和多个分布式设备。在学习过程中，服务器首先初始化模型参数，然后将其下发给分布式设备。分布式设备在接收到后，使用本地数据集来更新模型参数，并将其上传给服务器。服务器收到所有分布式设备的参数反馈后，对它们进行平均聚合以得到全局模型，然后再将此结果作为当前的模型参数，广播给各设备用于下一轮迭代。上述过程可以概括为计算和通信两个环节。其中，计算包括分布式设备的模型参数更新，以及中心服务器处的模型参数平均；通信包括中心服务器传输当前的模型参数给分布式设备，而分布式设备则上传本轮计算所得的模型更新给中心服务器。

相较于传统集中式的机器学习，联邦学习存在诸多优势。首先，边缘智能设备仅上传更新的模型参数，这样不仅高效地利用了网络带宽资源，而且缓解了通信负荷；其次，设备不发送原始数据，这在一定程度上保护了用户隐私；最后，相比于在中心节点决策再传送到终端设备来说，边缘智能节点在本地进行实时决策具备更小的延迟。鉴于以上优势，面向边缘智能服务的联邦学习受到了广泛的关注。谷歌边缘网络（Google Edge Network, GEN）、英特尔边缘虚拟化（Intel Network Edge Virtualization, INEV）等印证了它的巨大优势，并逐步在智能边云协同内容分发网络、智能家居、智慧社区、智慧交通、智慧医疗等典型物联网场景中得到应用。

然而，基于无线网络通信实施联邦学习还存在诸多挑战。首先，**无线链路特征影响计算学习性能的问题**。虽然联邦学习方法避免了由于直接传输原始数据所导致的巨大通信开销，但对于海量本地用户的情况，将大规模模型参数传输到中心节

点成为了联邦学习的瓶颈<sup>[8,11,12]</sup>。因此，针对联邦学习中的通信问题的研究具有十分重大的现实意义。国内外相关文献考虑采用模型更新参数预处理的方法降低通信负载，比如稀疏化<sup>[13,14]</sup>、量化<sup>[15-17]</sup>和通信审查<sup>[18-22]</sup>等。这些方法是从计算机科学角度解决联邦学习中的通信问题，它将通信视为可靠传输通道但忽略了真实通信信道的复杂属性（即忽视了传输过程问题）。但在实际中，无线通信的有限带宽资源和链路的不可靠性会造成传输错误和时延问题，从而降低联邦学习的准确率和训练速度。因此，全面考虑通信链路物理层特征，优化无线通信传输过程来提升联邦学习性能是必要的。

**其次，庞大复杂的移动边缘智能网络导致的超高传输开销问题。**基于数字传输的联邦学习中的通信开销和时延正比于参与联邦学习的分布式设备数量，这对于大规模边缘智能网络来说，具有极高的挑战度。考虑到联邦学习重点关注的是所有本地用户的更新平均值而非个体节点的更新值，采用空中计算技术<sup>[23-27]</sup>，利用无线信道波形叠加特性，通过模拟聚合实现联邦学习中的数据传输，有利于节省通信带宽，降低通信时延<sup>[28-32]</sup>。然而，空中计算技术是一种基于模拟聚合传输的通信技术，其对联邦学习的具体影响尚不明确。如何联合考虑学习和通信过程进行资源管理和优化亟待研究。

**最后，联邦学习的鲁棒性和隐私增强问题。**由于联邦学习是一种合作学习模式，是一个多方参与学习的过程。倘若这一过程中存在节点合作不积极、不合作乃至恶意合作的情况，会影响上传的模型参数进而破坏学习过程，因此增强联邦学习的鲁棒性是必须的。此外，恶意用户可能窃听边缘节点上传的模型参数，并通过这些模型参数推演出原始数据，从而导致隐私信息的泄露。因此，针对恶意用户参与的场景，如何从通信角度保证学习质量并增强隐私保护性能对于联邦学习的广泛应用具有重要意义。

总体来说，传统集中式机器学习不适用于大规模网络场景，因为其通信开销过大且不能保证用户数据的隐私。而联邦学习可在一定程度上降低通信开销，但在网络规模过大时，其应用依然存在一定的挑战。基于压缩、量化等方法虽然能够在一定程度上降低联邦学习的通信开销，但真实无线通信的复杂属性，有限的带宽资源和无线链路的不可靠性等仍会造成传输差错和时延问题，影响学习性能。此外，网络中参与学习的节点存在个体属性特征，其合作程度和节点身份会影响联邦学习的鲁棒性与隐私安全。因此，本文将围绕联邦学习中存在的问题，在联邦学习系统中引入空中计算技术，研究传算联合优化理论分析与模型构建、高效传算联合优化方法设计、传算联合优化的鲁棒性分析与强化方案设计，实现降低系统时延、提高通信带宽利用率、提升联邦学习的准确率、增强数据隐私安全的最终目的。

## 1.2 国内外研究现状

联邦学习是一种合作机制，其最为基础的问题之一是如何节省通信开销<sup>[7, 33]</sup>。当前，针对联邦学习中通信开销问题的研究主要包含“信息压缩”和“物理层传输”两方面。此外，除了通信效率，联邦学习中的鲁棒性问题也是研究的一个热点。下面本文对国内外相关的研究现状进行概述。

### 1.2.1 基于信息压缩的通信效率增强方法研究现状

从压缩传输的信息量的角度来提升通信效率的研究工作主要分下述三类：

一是如何处理由于个别分布式设备计算能力差所导致整体时延的增加<sup>[34-37]</sup>。由于设备计算能力或网络通信环境不同，会导致每个设备上传更新参数的不同步。如果中心节点等到所有分布式设备完成上传参数后再进行聚合求平均，则会存在等待时延。因此有些研究者提出差异化处理不同分布式设备，如丢弃计算能力差的设备的更新参数来实现快速更新同步。

二是减少分布式设备数目<sup>[20, 38]</sup>。通信效率与参与学习计算的分布式设备的个数有直接关系。参与学习计算的分布式设备越多，所需要的通信带宽就越多，时延也会相应增长。若根据分布式设备提供的更新参数来选择比较重要的分布式设备，则可以提高通信效率。文献[20]和[38]表明，如果多个分布式设备的模型更新参数较为相近，则可达到训练模型快速收敛的效果。因此，根据分布式设备所需模型更新参数的差异性或更新梯度的差异性来选择分布式设备，可在一定程度上提高通信效率。

三是减少更新参数的通信量，例如采用稀疏化<sup>[13,14]</sup>、量化<sup>[15-17]</sup>和通信审查<sup>[18-22]</sup>等方法压缩分布式设备所需传输的信息。其中，稀疏方法仅传输模型更新参数向量中数值较大的元素；量化方法则将连续更新值量化成有限比特值，实现数字信道中的高效传输；通信审查方法则要求在通信前对更新信息进行评价，避免传输信息量较小的更新值。

上述方法是基于“传什么”的思想来降低通信开销，均假设分布式设备和服务器之间存在可靠的链接，其中无线信道仅抽象为“比特管道”而忽略其复杂信道特征（如衰落、多址、广播、空间复用等）。显然，无线链路的特性会对学习系统造成不可忽略的影响。因此，解决通信问题更为直接可行的方法是设计新型无线通信技术以支持低延迟和节省通信开销，即从物理层技术的角度来研究“如何传”。

### 1.2.2 基于物理层技术的通信效率提升方法研究现状

文献[7]和[39]较为全面地概述了联邦学习中无线通信策略所存在的挑战。为了

减少联邦学习的通信开销，研究者们陆续设计了一些通信、计算资源的分配和调度方法<sup>[40-44]</sup>。例如，文献[40]针对比例公平、循环和随机调度等策略，分析了收敛速率；文献[41]提出了一种分布式设备调度方案，其目标是最大化参与计算的分布式设备数量，同时通过带宽分配确保能量效率；文献[42]-[44]研究了联邦学习中的通信资源分配和分布式设备调度的优化问题。这些关于联邦学习在无线网络中的开创性工作都是围绕数字通信系统开发的，也取得了良好的结果。但实际上，随着参与计算的分布式设备数量持续性增加，其通信开销和延迟也会累积增大，因此大规模无线网络场景中联邦学习的通信问题仍属于严峻挑战之一。

针对上述大规模无线网络中联邦学习的通信开销过大的问题，可采用空中计算技术（Over the Air Computation, AirComp）进行联邦学习中的数据传输。这是因为联邦学习中用于全局聚合的本地更新参数仅依赖于所有本地更新的平均值，而不是每个本地设备的单独更新值，因此传输单独节点的更新值是冗余的。空中计算技术利用了无线多址信道的波形叠加特性，基于非编码线性调制的模拟聚合通信可以自动在接收端直接获得多用户发送信号的平均值，能够在联邦学习中实现与本地设备数量无关的高效输出传输<sup>[23]</sup>。

空中计算技术可追溯到无线传感器网络中数据源函数的计算<sup>[23]</sup>，它的设计依赖于结构化编码来应对多址信道引入的信道失真。这项工作的意义在于它改变了常规认知，利用“干扰”来辅助计算。随后，文献[45]发现如果数据源是独立同分布的高斯分布，则无需编码，具有信道预均衡的简单模拟传输便可实现最小失真。同时，如果信源服从更复杂的分布（如双变量高斯<sup>[46]</sup>和相关高斯<sup>[47]</sup>），编码对于通信系统的设置仍然有效。正是由于空中计算技术具备如此的优越性和简洁性，使得研究人员不遗余力地研究该技术的鲁棒性和准确性。文献[48]和文献[49]分别提出了用于分布式功率控制和鲁棒的空中计算技术对抗信道估计误差；文献[50]和[51]则分别提供了失真约束下空中计算技术的中断性能和计算速率的理论分析；文献[52]通过原型设计将空中计算技术从理论转化为实践，并在文献[53]中提出将数据调制成发射功率以放宽同步要求（即只需要粗略的块同步），而文献[54]中，作者设计了空中分享方案，该方案向所有设备广播共享时钟以强制同步。

除了标量值函数计算之外，空中计算技术还引入了多输入多输出（Multiple Input Multiple Output, MIMO）技术，以实现向量值函数计算<sup>[26,55,56]</sup>，即 MIMO AirComp。文献[26]使用波束赋形优化和匹配的有限反馈设计，提出了 MIMO AirComp 综合框架。该框架在后续工作中扩展到无线供电的 AirComp 系统<sup>[55]</sup>，通过波束赋形器与无线功率控制的联合优化进一步减少空中计算技术的失真。该框架也实现向大规模 MIMO AirComp 系统的扩展<sup>[56]</sup>，通过利用集群信道结构减少了信道反馈开销，降低了信号处理的复杂性，设计了缩减版的两层波束赋形器。虽然

空中计算技术主要部署在以计算为中心的无线传感器网络，但它也常用于速率最大化的蜂窝系统，如双向中继系统<sup>[57]</sup>和 MIMO 点阵解码系统<sup>[58]</sup>。

针对空中计算技术的研究掀开了解决联邦学习中通信开销问题的新篇章。关于基于空中计算技术的联邦学习（Federated Learning Over the Air, FLOA）的相关研究得到了广泛关注，它涉及通信信号处理、数学优化方法、机器学习等众多研究领域，如功率控制<sup>[59, 60]</sup>、本地设备调度<sup>[31, 61]</sup>、梯度压缩<sup>[28, 29, 62]</sup>、波束赋形设计<sup>[30, 32, 63]</sup>、学习速率优化<sup>[64]</sup>等。例如，文献[61]提出了一种功率控制和设备调度的宽带模拟聚合方案，讨论了通信和学习之间的平衡参数；考虑到能量受限的本地设备，文献[31]提出了一种基于能量感知的设备调度策略，以最大化参与参数更新的平均本地设备数量；文献[59]提出了最小化聚合信号的均方误差的方法；而文献[60]则针对梯度更新值的随机特征进行功率设计。在多天线场景中，文献[32]提出了设备调度和波束赋形的联合设计方法，在给定均方误差前提下，最大化所选的本地设备数量；基于 1 比特梯度量化，文献[62]提出了一种数字版本的宽带空中聚合方法，并分析了无线信道对其收敛速度的影响。文献[28]和文献[29]分别利用梯度稀疏化和随机线性投影以降低窄带信道中梯度向量的维度，来降低通信需求。考虑到超参数对学习性能的影响，文献[64]针对多天线系统设计了一种学习率优化方案，进一步提高聚合信号均方误差性能和测试精度。

### 1.2.3 联邦学习鲁棒性研究现状

FLOA 不仅提高了通信效率，并且由于其具有对局部梯度的不可访问性，增强了数据隐私，防止了潜在的模型反转攻击（如梯度的深度泄漏<sup>[65]</sup>）。虽然 FLOA 能够抵抗梯度深度泄漏，但它也为攻击者提供了进行拜占庭攻击的机会。所谓拜占庭攻击问题，就是在不可信的分布式节点存在的情况下，如何保证学习算法的正常实施。事实上，即使是单一的拜占庭式故障也可以摧毁 FLOA。对于传统联邦学习，拜占庭鲁棒聚合方案已经得到了广泛的研究<sup>[66-78]</sup>，大部分方案使用了筛选方案，如几何中值<sup>[70-73]</sup>、坐标中值和坐标修剪均值<sup>[67]</sup>、Krum/Multi-Krum<sup>[74]</sup>、Bulyan<sup>[75, 76]</sup>、Zeno/Zeno++<sup>[77, 78]</sup>等。这些方法的基本思想是在聚合本地梯度的同时剔除异常值。这些方法取决于本地梯度的个体值信息，但由于所有本地梯度值进行了模拟叠加，这致使本地梯度值在 FLOA 系统中不可访问。因此，现有的拜占庭鲁棒方法难以应用于 FLOA 之中，这促使研究者们必须设计新型的拜占庭鲁棒方法。

### 1.2.4 当前研究尚存的问题

通过上述分析可知，虽然针对 FLOA 已存在一些研究，并取得一定的研究成果

果,但由于该领域的研究仍然处于初期,通过空中计算技术实现高效的通信传输和高性能的联邦学习还存在以下尚未解决的问题:

- (1) **模拟聚合通信对联邦学习的影响尚不明确:** 空中计算技术所涉及的模拟聚合通信与联邦学习之间关系尚缺少定量理论分析与描述; 参与计算的本地设备数量简单最大化与学习性能没有直接关系, 相应的结果不一定是最佳的; 将计算和通信的优化问题解耦, 则无法衡量参与计算的设备对引入误差或有效信息的贡献大小。
- (2) **通信传输和计算学习的独立优化设计, 缺少两者的联合优化:** 通信传输参数对学习过程会产生影响, 不能将这两个阶段分别进行优化设计。在通信可靠的情况下, 若仅优化学习参数, 而忽略通信误差和时延带来的负面影响, 会造成联邦学习系统在实际应用中不可靠, 因此如何设计传算协同优化方法十分重要。
- (3) **缺少压缩/量化技术对联邦学习影响的定量化描述:** 已有工作在利用模拟聚合通信技术传输信号时通常发送全部更新参数, 鲜有利用压缩感知技术进一步降低通信开销和时延的研究。而在模拟聚合通信中使用压缩感知、量化等技术可进一步提升通信效率, 但由于引入了有损压缩, 相应的压缩量化误差对联邦学习性能的影响也需要进行定量化的理论分析。进而进行相关系统参数的优化设计。
- (4) **FLOA 的鲁棒性有待加强:** 已有研究很少考虑边缘节点对计算学习的干扰, 如果这种多方学习中存在合作不积极、不合作乃至恶意合作的节点, 有可能发动中毒攻击来破坏学习过程, 并降低联邦学习效果。因此需要量化这种非正常行为对联邦学习的影响, 设计抵抗该行为的措施以提升联邦学习的鲁棒性。

综上所述, 本文着眼于大规模网络下的 FLOA 系统传算联合优化算法研究, 重点解决上述亟待突破的问题。这对于面向边缘智能服务的联邦学习技术的广泛应用具有重要的理论意义和实际价值。本文在提出传算联合优化模型架构与基本优化方法的基础上, 结合凸优化理论、概率论、信号检测与估计、机器学习等技术, 从通信信号传输的角度进一步提出基于传算联合优化框架的高效性和鲁棒性方案。论文研究成果不仅可为边缘智能计算提供理论基础, 还可延伸至智能交通、智能电网、智慧城市、智能制造等边缘智能服务与应用之中。

## 1.3 论文的主要工作与内容安排

### 1.3.1 主要研究工作

本文以通信传输和分布式计算(即“传算一体化”)的视角研究无线网络资源(包括通信资源和学习资源)联合最优分配的高效性、鲁棒性与隐私安全增强的理论与方法。具体来说,本文首先围绕 FLOA, 构建联邦学习的收敛性量化描述, 建立面向无线通信与联邦学习的联合优化模型架构; 其次针对高维神经网络模型场景中传输效率低下的问题, 进一步引入压缩感知等高效通信技术, 分析无线通信传输效率和计算学习性能之间的博弈关系, 以提升通信高效性的同时满足学习性能需求为目标, 实现高效的传算联合优化方法; 最后针对 FLOA 存在的鲁棒性差的问题, 利用无线链路特征与边缘节点特征分析结果, 通过物理层传输技术来改善联邦学习方案设计, 实现鲁棒性增强的传算联合优化方法。本文研究的具体内容和创新点总结如下:

- (1) (创新点一: 推导了 FLOA 收敛性量化描述的闭式表达式) 当前关于 FLOA 的研究大都只针对通信传输或学习的某一方面。然而仅考虑通信问题则会忽视传输参数和交付结果对计算学习性能的影响; 仅考虑计算学习性能时则时常假设通信传输能够提供可靠的交付。由此可见, 将通信传输和计算学习分成两个独立的阶段分别进行优化不一定是最优的。本文从计算学习算法的收敛性入手, 分析无线通信和计算学习之间的可量化关系。针对不同学习网络的损失函数(如凸损失函数、非凸损失函数等)和不同计算学习算法(如标准梯度下降法、随机梯度下降法等), 全面推导联邦学习的收敛性的闭式表达式。该收敛性的分析可以量化描述系统参数对联邦学习算法的具体影响, 为传算联合优化算法提供理论基础。
- (2) (创新点二: 提出了基于收敛性分析的 FLOA 系统传算联合优化框架) 本文在利用收敛性分析通信传输和计算学习间隐含关系的基础上, 探讨通信和学习资源的联合优化问题。根据收敛性分析, 针对不同损失函数和多种学习算法, 本文探究可以被优化的无线网络参数, 设计无线网络资源与计算学习资源的联合优化问题。考虑到优化问题中目标函数的复杂性(融合了通信参数和学习参数)以及网络参数的离散性(网络资源的离散性和用户选择的离散性), 这些问题通常是非凸、非线性、非连续的高复杂度优化问题, 本文提出低复杂度的求解方法来满足边缘智能应用中低时延传输的需求。最终建立面向 FLOA 的无线传输和联邦学习联合优化框架。
- (3) (创新点三: 提出了基于 FLOA 传算联合优化框架的高效性提升方案) FLOA 将多用户网络参数在相同时频资源条件下传输, 可大大节省通信资

源。但随着深度学习的发展和广泛应用，神经网络的规模越来越大，如视觉几何组网络（Visual Geometry Group Network，VGGNet）架构高达 1.38 亿网络参数<sup>[79]</sup>，而这些大规模网络参数的持续性传输势必造成极大的网络时延，这不满足边缘智能的低时延需求。针对高维神经网络场景，本文将 1 比特压缩感知技术引入 FLOA 中，提出集稀疏、降维、量化、模拟聚合传输以及信号重建于一体的高效联邦学习方案。由于量化和压缩感知是一种有损压缩方式，这种高效通信模式势必会降低学习性能。因此，为了定量化描述由于稀疏、降维、量化、模拟聚合传输、信号重建以及信道噪声引入的误差对联邦学习性能的负面影响，在不同网络损失函数和多种学习算法下，本文推导基于 1 比特压缩感知技术的 FLOA 方案的学习性能的收敛性。最后，根据所推导的收敛性闭式表达式，本文建立以降低收敛误差为目标，以通信和学习资源为约束的联合优化方程，并通过凸优化理论对该优化问题进行求解，提出低复杂度的求解算法。

(4) **(创新点四：提出了基于 FLOA 传算联合优化框架的鲁棒性增强方案)** 无线信道的开放性和联邦学习对于模型数据的需求导致传输的信号存在被恶意用户窃听的可能，致使隐私数据的泄露。与此同时，联邦学习依赖于多用户合作，其本身就容易遭受拜占庭攻击。在 FLOA 系统中，中心节点仅能获取所有参与联邦学习用户数据值的平均值，这关闭了模型反演等隐私泄露的大门，解决了联邦学习的隐私泄露问题，但是却为恶意用户发动拜占庭攻击提供了便利。由于无法获得每个分布式设备的具体更新数据而无法筛选出恶意用户，则传统基于恶意用户检测的拜占庭攻击防护策略在该场景下失去作用。本文针对 FLOA 系统，从理论上推导拜占庭攻击对联邦学习方案的影响的闭式表达式，从通信角度提出相应的对抗方案，并对所提出的对抗性方案进行理论分析和实际测试。具体而言，首先本文讨论拜占庭攻击者的攻击方式。针对梯度下降法，拜占庭攻击者若想最大程度地破坏联邦学习算法，应发送与正确梯度相反的向量。若攻击效果较强，则梯度值不降反升，梯度下降法则会失效，即联邦学习遭到破坏。本文从理论上证明此攻击的效果，并证明拜占庭攻击者的攻击存在一种最强的攻击方式。其次，由于不同功率控制策略对拜占庭攻击的抵御能力不同，本文进一步探讨各种功率控制策略下，拜占庭攻击者的最强攻击效果。从联邦学习算法收敛性的角度分析，本文从理论上推导出学习算法在攻击下的性能边界。最后，根据上述对拜占庭攻击者的攻击特点分析，本文提出一种对抗拜占庭攻击的新型功率控制策略，并对所提算法进行收敛性推导，进而比较所提出的功率控制策略和现存策略的性能差异。

### 1.3.2 论文的章节安排

综上所述，本文面向 FLOA 系统，围绕传算联合优化理论与方法进行研究和探讨，包括传算联合优化基础框架研究，通信高效性提升方案研究，鲁棒性增强方案研究等三个方面。全文共分五个章节。各章节内容安排如图 1.2 所示：

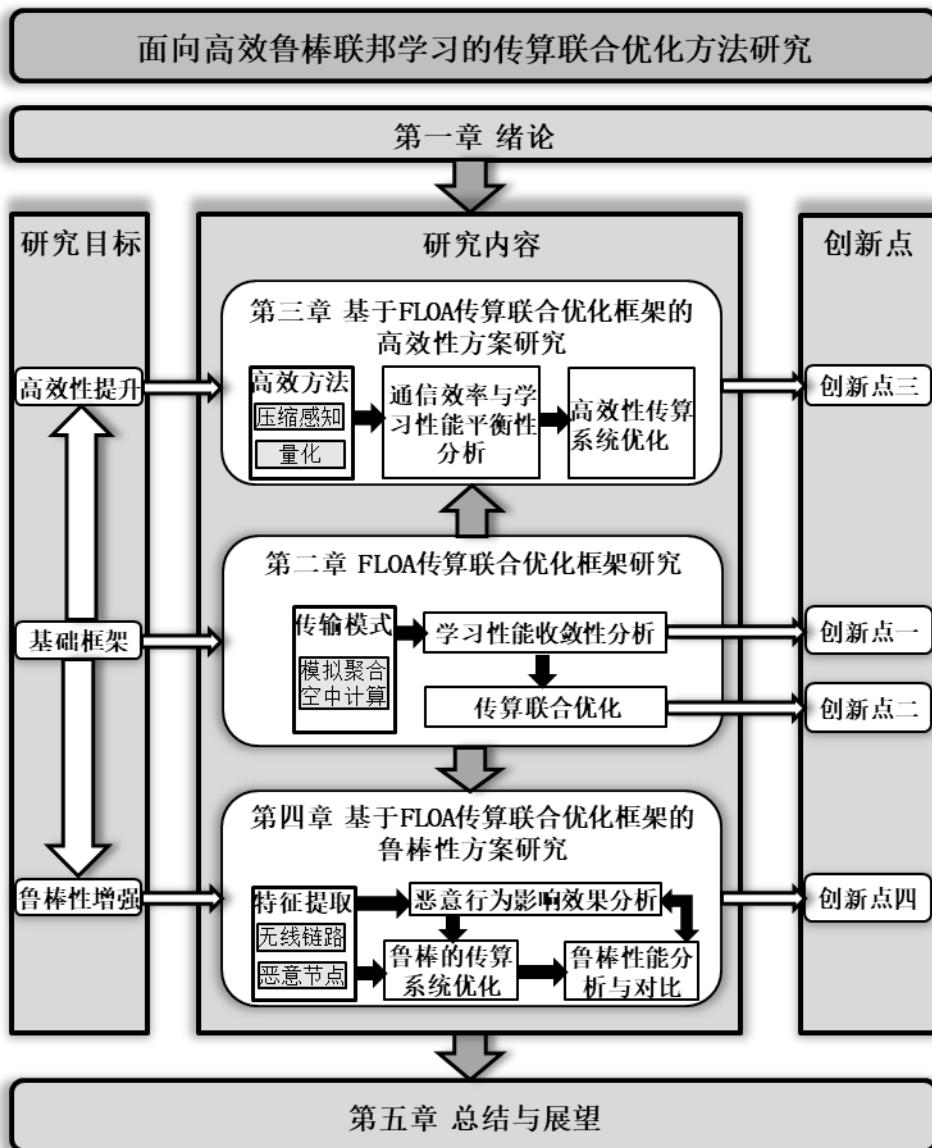


图 1.2 论文研究内容与结构示意图

Figure 1.2 The structure and content of this dissertation

第一章介绍本文的研究背景及意义，对国内外针对本文研究主题的研究现状进行概述，并指出现存工作的不足之处，然后归纳本文的主要研究内容和创新点。本章仅提供本文所涉及的联邦学习和空中计算技术的基本概念，详细基础理论知识可参考文献[11,23,27,80-82]。

第二章首先推导 FLOA 的收敛性的封闭表达式，为参数优化提供指导。然后根据关于收敛性的理论结果，建立联合优化问题，并将其转化为实践中可行的优化问题。进而借助凸优化技术对该优化问题进行求解。最后，通过仿真分析证明本章所提出的优化方案的有效性。

第三章将 1 比特压缩感知技术引入 FLOA 中，进一步提高通信效率。首先，本章设计一套集稀疏化、降维、量化、模拟聚合传输以及信号重构于一体的联邦学习方案。然后，对所提出的方案进行收敛性分析，分析系统参数对学习收敛性的影响。接着，根据收敛性的理论指导，建立新的联合优化问题，并给出具体的解决方案。最后，性能仿真与分析说明本章所提方案的有效性。

第四章研究 FLOA 中的鲁棒性问题。首先分析当前存在的方案对恶意拜占庭用户的抵抗能力。然后提出本文的对抗性方案。接着对本文所提鲁棒性方案进行性能分析。最后通过仿真实验证实所提方案的鲁棒性。

第五章总结全文，并对未来的的研究工作进行展望。



## 2 FLOA 传算联合优化框架研究

随着物联网和社交网络的发展，网络边缘产生了大量的数据<sup>[83,84]</sup>。为了从大数据中获取有用的信息，联邦学习被广泛应用于处理新兴数据驱动应用中的复杂模型和任务<sup>[85]</sup>。现有的联邦学习研究主要集中在理想化通信链接假设下的学习算法。然而，在实际无线系统中部署的联邦学习应考虑无线环境对学习性能的影响。否则，这种影响可能会引入不必要的训练错误，从而在准确性和收敛速度方面显著降低学习性能<sup>[86]</sup>。为了解决这个问题，研究人员致力于优化联邦学习中的网络资源<sup>[44,87]</sup>。但随着活跃的分布式设备数量的增加，基于数字通信的联邦学习的通信开销和传输延迟会变大。另一方面，空中计算技术为联邦学习中本地更新的传输与均值计算提供了一种直接有效的方式<sup>[28-32]</sup>。空中计算技术是一种基于模拟聚合通信的技术，采用的是非编码的模拟调制，利用无线多址信道（Multiple-Access Channel, MAC）中波形叠加的原理<sup>[23-27]</sup>，可以在通信的过程中实现“空中计算”。通过多个发送端的联合设计，可以使得多个发射端利用相同的时频资源同时发送它们的数据信号。而在接收端统一解调，接收到的是所有发送端发来的数据信号的和，即所有发送的数据在空中实现了一个和函数的功能。作为一种联合传输和聚合的策略，空中计算技术可以使所有分布式设备使用相同的时频资源同时上传它们的本地模型更新，从而大大减少了联邦学习的无线通信开销。

FLOA 的研究仍处于早期阶段，一些基本问题尚未探索，例如其收敛行为和高效算法的设计。鉴于分布式设备的传输功率和通信带宽有限，分布式设备在将本地更新传输到参数服务器（Parameter Server, PS）时可能不得不争夺通信资源。这些未被探索的问题产生了对高效传输范式的需求，以及在分布式设备选择和传输功率控制方面的网络资源分配的需求。所有这些实际问题促使从同时考虑无线通信和机器学习的角度研究联邦学习。借助一些定义和假设，本章量化模拟聚合传输对联邦学习的收敛行为和性能的影响。这样的量化结果对于指导通信和计算资源的联合优化至关重要。本章旨 在全面研究无线通信和联邦学习的联合设计和优化的理论推导、问题制定、解决方案开发和算法实现，形成一个 FLOA 传算联合优化框架。具体而言，本章的主要研究内容和贡献总结如下：

- (1) 在凸损失函数和非凸损失函数的情况下，本章分别推导 FLOA 的预期收敛速率的闭式表达式，它不仅可以解释而且可以量化无线通信对联邦学习的收敛和准确性的影响。此外，这项工作还考虑梯度下降（Gradient Descent, GD）和小批量统计梯度下降（Stochastic Gradient Descent, SGD）方法。这些封闭形式的表达式可以揭示模拟无线通信和具有模拟聚合的联邦学习之间的基本联系，这为衡量模拟无线系统的参数设计如何影响无线联邦学习

的性能提供一个全新的视角。

- (2) 基于封闭形式的理论结果, 本章建立一个学习、设备选择和功率控制的联合优化问题, 目标是在有限的发射功率和带宽的情况下最小化全局联邦学习损失函数。事实证明, 优化公式对于凸和非凸损失函数情况下采用 GD 和 SGD 的联邦学习是通用的。此外, 为了在存在一些不可观察参数的情况下实际实现联合优化问题, 本章提出一种替代重构策略, 将原始不可实现的问题近似为模拟聚合操作约束下的可行优化问题。
- (3) 为了有效地解决近似问题, 本章通过探索分布式设备数量与功率缩放之间的关系来确定一个紧密的解决方案空间。由于减少了搜索空间, 本章提出一种简单的离散枚举方法来有效地找到全局最优解。

此外, 本章分别在解决线性回归和图像分类问题上评估提出的 FLOA 传算联合优化方案。仿真结果表明, 本章提出的联邦学习优于使用随机分布式设备选择和功率控制的基准方案, 并且实现了与在无噪声无线信道上实施联邦学习的理想情况相当的性能。

本章具体内容安排如下: 2.1 节对 FLOA 进行介绍, 并详细地描述节点传输与计算的过程。2.2 节推导出联邦学习在模拟聚合通信下的预期收敛速率的封闭式表达式, 作为算法设计和性能分析的理论基础。2.3 节提供通信和联邦学习的传算联合优化的框架, 并提出相应的求解算法。2.4 节通过仿真对提出的方案进行性能分析, 2.5 节对本章进行总结。

## 2.1 系统模型

如图 2.1 所示, 本章考虑由基站处的单个参数服务器和作为分布式设备的  $U$  个用户设备组成的单跳无线网络。通过联邦学习, 参数服务器和所有分布式设备协作训练一个用于监督学习和数据推理的通用模型, 而无需共享本地数据。

### 2.1.1 联邦学习模型

定义  $\mathcal{D}_i = \{\mathbf{x}_{i,k}, \mathbf{y}_{i,k}\}_{k=1}^{K_i}$  为第  $i$  个分布式设备的本地数据集, 其中  $\mathbf{x}_{i,k}$  为输入数据向量,  $\mathbf{y}_{i,k}$  为标签输出向量,  $k = 1, 2, \dots, K_i$ ,  $i = 1, \dots, U$ 。第  $i$  个分布式设备的可用样本数量为  $K_i = |\mathcal{D}_i|$ 。通过使用样本总数  $K$  为所有本地样本数之和的样本, 这些分布式设备寻求合作最小化一个全局损失函数来学习出一个全局神经网络模型, 该神经网络模型可以被参数  $\mathbf{w} = [w^1, \dots, w^D] \in \mathcal{R}^D$  表征, 其中  $D$  为参数维度。全局损失函数定义如下:

$$(全局损失函数) \quad F(\mathbf{w}; \mathcal{D}) = \frac{1}{K} \sum_{i=1}^U \sum_{k=1}^{K_i} f(\mathbf{w}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k}), \quad (2-1)$$

其中  $\mathcal{D} = \bigcup_i \mathcal{D}_i$  为全局数据集，全局损失函数  $F(\mathbf{w}; \mathcal{D})$  是  $K$  个数据相关分量的总和，每个分量  $f(\mathbf{w}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})$  是一个局部采样损失函数，用于量化由全局模型参数  $\mathbf{w}$  参数化的数据模型的模型预测误差。

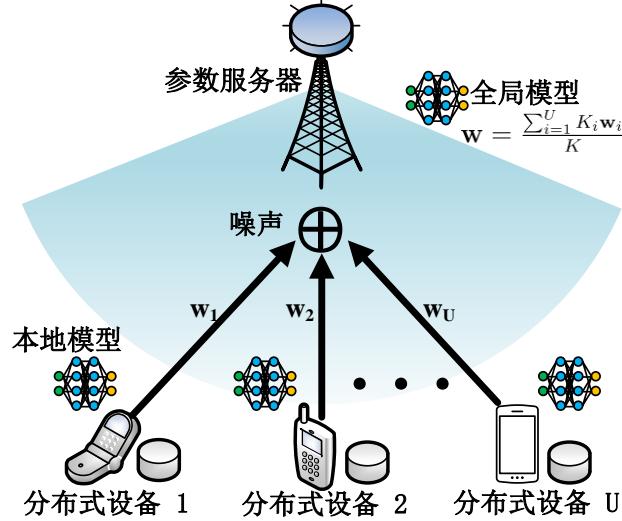


图 2.1 FLOA 网络模型示意图

Figure 2.1 Schematic diagram of federated learning over the air

在分布式学习中，每个分布式设备使用它们的本地数据集  $\mathcal{D}_i$  来训练一个本地神经网络模型  $\mathbf{w}_i$ ，其中  $\mathbf{w}_i$  可视为全局模型  $\mathbf{w}$  的本地副本。因此，本地损失函数被定义为：

$$(本地损失函数) \quad F_i(\mathbf{w}_i; \mathcal{D}_i) = \frac{1}{K_i} \sum_{k=1}^{K_i} f(\mathbf{w}_i; \mathbf{x}_{i,k}, \mathbf{y}_{i,k}), \quad (2-2)$$

其中  $\mathbf{w}_i = [w_i^1, \dots, w_i^D] \in \mathcal{R}^D$  是局部模型参数。通过协作，分布式学习希望达到  $\mathbf{w}_i = \mathbf{w} = \mathbf{w}^*, \forall i$ ，使所有分布式设备的本地模型达到全局最优模型  $\mathbf{w}^*$ 。这种分布式学习可以通过共识优化表示为<sup>[85,88]</sup>：

$$\text{P2.1: } \min_{\mathbf{w}} \quad \frac{1}{K} \sum_{i=1}^U \sum_{k=1}^{K_i} f(\mathbf{w}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k}). \quad (2-3)$$

为了求解 P2.1，本章采用联邦学习中的模型平均算法<sup>[85,88]</sup>。它本质上是一个迭代过程，包括每次迭代的计算和通信步骤。具体而言，在每一轮通信中，参数服务器将当前的全局模型  $\mathbf{w}$  广播给所有分布式设备。然后，第  $i$  个分布式设备使用优化算法根据当前  $\mathbf{w}$  更新它的本地模型  $\mathbf{w}_i$ 。本章采用基本梯度下降法 GD 来求解 P2.1（本章以基本梯度下降为例，而所提出的方法也可以扩展到小批量梯度下降法，具体拓展方法后文也将描述），其中局部模型在第  $i$  个分布式设备的更新为

$$\begin{aligned} \mathbf{w}_i &= \mathbf{w} - \alpha \nabla F_i(\mathbf{w}_i; \mathcal{D}_i) \\ (\text{本地模型更新}) \quad &= \mathbf{w} - \frac{\alpha}{K_i} \sum_{k=1}^{K_i} \nabla f(\mathbf{w}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k}), \end{aligned} \quad (2-4)$$

其中， $\alpha$ 是学习速率，而 $\nabla f(\mathbf{w}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})$ 是样本损失函数 $f(\mathbf{w}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})$ 相对于模型参数 $\mathbf{w}$ 的梯度。

当本地更新完成后，每个分布式设备通过无线上行链路将其更新后的本地模型参数 $\mathbf{w}_i$ 发送给参数服务器，以更新全局模型参数 $\mathbf{w}$ 为

$$(\text{全局模型更新}) \quad \mathbf{w} = \frac{\sum_{i=1}^U K_i \mathbf{w}_i}{K}. \quad (2-5)$$

进而，参数服务器将公式(2-5)中的全局模型参数 $\mathbf{w}$ 广播给所有参与联邦学习的分布式设备，作为下一轮的初始值。联邦学习迭代地在公式(2-4)中实现局部模型更新和在公式(2-5)中实现全局模型平均，直到收敛。在文献[85,88]中已经证明，在全局损失函数 $F(\mathbf{w}; \mathcal{D})$ 是凸函数且中心参数服务器和分布式设备之间的数据传输无差错的条件下，联邦学习算法收敛到P2.1中原问题的全局最优解 $\mathbf{w}^*$ 。

值得注意的是，公式(2-4)和公式(2-5)中的实现步骤只涉及联邦学习的计算方面，并假设了全局模型参数 $\mathbf{w}$ 和本地模型参数 $\mathbf{w}_i$ 在中心参数服务器和所有分布式设备之间的通信为完美无误差的。然而，无线通信对联邦学习的性能的影响不容忽视。尤其是在实际的无线网络环境中，由于无线信道特性的不完善，在更新传输过程中不可避免地会引入一定的误差。

### 2.1.2 基于模拟聚合的模型参数传输

为了避免繁重的通信开销并节省联邦学习参数在无线信道上的传输带宽，本章采用无需编码的模拟聚合传输技术，这允许多个分布式设备在相同的时频资源上同时将它们的本地模型更新上传到参数服务器。所有分布式设备以模拟形式传输它们本地的模型参数 $\mathbf{w}_i$ ，并假设它们之间具有完美的时间同步<sup>1</sup>。通过这种方式，所有本地更新的模型参数 $\mathbf{w}_i$ 在空中聚合以实现公式(2-5)中的全局模型更新步骤。这种模拟聚合是以逐项方式进行的。换言之，对于任何 $d \in [1, D]$ ，来自所有分布式设备的本地模型参数 $\mathbf{w}_i$ 的第 $d$ 个条目 $w_i^d$ ， $i = 1, \dots, U$ ，在传输中将被聚合以计算公式(2-5)中的 $w^d$ 。

令 $\mathbf{p}_{i,t} = [p_{i,t}^1, \dots, p_{i,t}^d, \dots, p_{i,t}^D]$ 表示在第 $t$ 次迭代中第 $i$ 个分布式设备的功率控制向

---

<sup>1</sup> 时间同步的实现和不完美同步的影响超出了本章的研究范围。感兴趣的读者可以参考文献[53,54]。

量。值得注意的是，在基于模拟聚合通信技术的联邦学习中  $\mathbf{p}_{i,t}$  的选择不仅要有效地实现公式（2-5）中的聚合规则，还要适当地满足用于网络资源分配的需要。因此，本章将功率控制策略设置为

$$p_{i,t}^d = \frac{\beta_{i,t}^d K_i b_t^d}{h_{i,t}^d}, \quad (2-6)$$

其中  $h_{i,t}$  表示在第  $t$  次迭代中第  $i$  个分布式设备和参数服务器之间的信道增益（在本章中，信道状态信息 (Channel State Information, CSI) 被假设为在每次迭代中保持不变，但可能会随着迭代而变化。并且本章还假设 CSI 在各节点上是完全已知的）， $b_t^d$  定义为功率缩放因子， $\beta_{i,t}^d$  代表传输调度指示符。换言之， $\beta_{i,t}^d = 1$  表示第  $i$  个分布式设备的本地模型参数  $\mathbf{w}_{i,t}$  的第  $d$  项参数计划在第  $t$  次迭代时为联邦学习算法做出贡献，否则  $\beta_{i,t}^d = 0$ 。第  $i$  个分布式设备上传第  $d$  项参数的发射功率不应超过其最大功率限制  $P_i^{d,\max} = P_i^{\max}$ ，如下所示：

$$|p_{i,t}^d w_{i,t}^d|^2 = \left| \frac{\beta_{i,t}^d K_i b_t^d}{h_{i,t}^d} w_{i,t}^d \right|^2 \leq P_i^{\max}. \quad (2-7)$$

在参数服务器侧，第  $t$  次迭代的接收信号为

$$\begin{aligned} \mathbf{y}_t &= \sum_{i=1}^U \mathbf{p}_{i,t} \odot \mathbf{w}_{i,t} \odot \mathbf{h}_{i,t} + \mathbf{z}_t \\ &= \sum_{i=1}^U K_i \mathbf{b}_t \odot \boldsymbol{\beta}_{i,t} \odot \mathbf{w}_{i,t} + \mathbf{z}_t, \end{aligned} \quad (2-8)$$

其中  $\odot$  代表 Hadamard 积， $\mathbf{h}_{i,t} = [h_{i,t}^1, h_{i,t}^2, \dots, h_{i,t}^D]$ ， $\boldsymbol{\beta}_{i,t} = [\beta_{i,t}^1, \beta_{i,t}^2, \dots, \beta_{i,t}^D]$ ， $\mathbf{b}_t = [b_t^1, b_t^2, \dots, b_t^D]$ ，而  $\mathbf{z}_t \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})$  是加性高斯白噪声 (Additive White Gaussian Noise, AWGN)。

给定接收到的  $\mathbf{y}_t$ ，参数服务器通过后处理操作来估计  $\mathbf{w}_t$  如下：

$$\begin{aligned} \mathbf{w}_t &= \left( \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot -1} \odot \mathbf{y}_t \\ &= \left( \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \right)^{\odot -1} \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{w}_{i,t} + \left( \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot -1} \odot \mathbf{z}_t, \end{aligned} \quad (2-9)$$

其中  $(\sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t)^{\odot -1}$  是一个缩放向量，它的选择是为了在公式 (2-9) 中使得参与的本地模型参数  $\mathbf{w}_i$  具有相等的权重，如公式 (2-5) 中所需要的一样。操作  $(\mathbf{X})^{\odot -1}$  表示  $\mathbf{X}$  的逆 Hadamard 运算，其计算向量中的元素的倒数。值得注意的是，为了在空中实现联邦学习中如公式 (2-5) 那样的平均，这样的后处理操作要求  $\mathbf{b}_t$  对于给定的  $t$  和  $d$  所有分布式设备都相同。这样的操作可以从公式 (2-9) 的第一项中消除  $\mathbf{b}_t$ 。

通过上述模拟聚合传输过程，联邦学习实现了空中计算的结果。比较公式 (2-

9) 和公式 (2-5)，由于无线通信的影响， $\mathbf{w}_t$  和  $\mathbf{w}$  之间存在差异。本章旨在通过优化分布式设备选择  $\beta_{i,t}$  和功率缩放因子  $b_i$  来缓解这种差距。为此，本章的下一步是揭示一个重要但尚未探索的基础，即无线通信如何影响 FLOA 的收敛行为。

## 2.2 模拟聚合通信对联邦学习影响的定量化描述

本节通过分析在凸和非凸损失函数情况下 FLOA 的收敛行为来研究模拟聚合传输对学习性能的定量化影响。通过对瞬时 SNR 进行平均，本章得出联邦学习在模拟聚合传输下的预期收敛速率，它量化无线模拟聚合传输对联邦学习的影响。

### 2.2.1 基本假设

为了方便进行收敛性分析，本节首先给出所采用的优化领域常见的一些基本假设<sup>[44,88-93]</sup>，如下：

**假设 2.1** (Lipschitz 连续，平滑) 损失函数  $F(\mathbf{w})$  的梯度  $\nabla F(\mathbf{w})$  关于  $\mathbf{w}$  是一致 Lipschitz 连续的，即

$$\|\nabla F(\mathbf{w}_{t+1}) - \nabla F(\mathbf{w}_t)\| \leq L\|\mathbf{w}_{t+1} - \mathbf{w}_t\|, \forall \mathbf{w}_t, \mathbf{w}_{t+1}, \quad (2-10)$$

其中， $L$  是一个正常数，称为函数  $F(\cdot)$  的 Lipschitz 常数。

**假设 2.2** (强凸)  $F(\mathbf{w})$  是强凸的，具有正参数  $\mu$ ，服从

$$F(\mathbf{w}_{t+1}) \geq F(\mathbf{w}_t) + (\mathbf{w}_{t+1} - \mathbf{w}_t)^T \nabla F(\mathbf{w}_t) + \frac{\mu}{2} \|\mathbf{w}_{t+1} - \mathbf{w}_t\|^2, \quad \forall \mathbf{w}_t, \mathbf{w}_{t+1}. \quad (2-11)$$

**假设 2.3** (样本梯度有界) 分布式设备处的本地样本梯度受限于它们的全局梯度值，即

$$\|\nabla f(\mathbf{w}_t)\|^2 \leq \rho_1 + \rho_2 \|\nabla F(\mathbf{w}_t)\|^2, \quad (2-12)$$

其中， $\rho_1, \rho_2 > 0$ 。

这些假设可以广泛使用于一些损失函数，例如均方误差、逻辑回归和交叉熵等。这些常见的损失函数可用于实现用于识别、预测和分类的实用联邦学习算法。

### 2.2.2 强凸损失函数假设下收敛性分析

根据文献[11]和文献[94]，在理想无线信道上应用的联邦学习算法能够解决原问题 P2.1 并收敛到全局最优解  $\mathbf{w}^*$ 。在模拟聚合传输的情况下，本小节推导出 FLOA 在强凸假设下的预期收敛速率，如下面定理 2.1 所示。

**定理 2.1** 采用假设 2.1，假设 2.2 和假设 2.3，将公式 (2-3) 中的全局最优学习模型记为  $\mathbf{w}^*$ 。基于模拟聚合传输的联邦学习方案的模型  $\mathbf{w}_t$ ,  $\forall t$  的更新规则由公式

(2-9) 给出。给定发射功率缩放因子  $\mathbf{b}_t$ , 分布式设备选择向量  $\boldsymbol{\beta}_{i,t}$ , 并将学习速率设置为  $\alpha = \frac{1}{L}$ , 则第  $t$  次迭代的全局模型参数  $\mathbf{w}_t$  与最优模型参数  $\mathbf{w}^*$  的预期性能差距  $\mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}^*)]$  由下式给出:

$$\mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}^*)] \leq B_t + A_t \mathbb{E}[F(\mathbf{w}_{t-1}) - F(\mathbf{w}^*)], \quad (2-13)$$

其中,

$$A_t = 1 - \frac{\mu}{L} + \rho_2 \sum_{d=1}^D \left( \frac{K}{\sum_{i=1}^U K_i \beta_{i,t}^d} - 1 \right), \quad (2-14)$$

$$B_t = \frac{\rho_1}{2L} \sum_{d=1}^D \left( \frac{K}{\sum_{i=1}^U K_i \beta_{i,t}^d} - 1 \right) + \left\| \left( \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot -1} \right\|^2 \frac{L\sigma^2}{2}, \quad (2-15)$$

所涉及期望关于零均值方差为  $\sigma^2$  的 AWGN。

**证明:** 参见附录 A.1 ■

根据定理 2.1, 为了更好地分析联邦学习算法的收敛行为, 本章节进一步推导出由于无线通信和分布式设备选择导致的整个联邦学习过程累积的性能差距, 总结为以下引理 2.1。

**引理 2.1** 给定一个初始全局模型  $\mathbf{w}_0$ , FLOA 的全局模型参数  $\mathbf{w}_t$  与最优模型参数  $\mathbf{w}^*$  累积性能差距  $\mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}^*)]$  经过第  $t$  次迭代后的上界为:

$$\mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}^*)] \leq \underbrace{\sum_{i=1}^{t-1} \left( \prod_{j=1}^i A_{t+1-j} \right) B_{t-i}}_{\Delta_t} + B_t + \prod_{j=1}^t A_j \mathbb{E}[F(\mathbf{w}_0) - F(\mathbf{w}^*)]. \quad (2-16)$$

**证明:** 鉴于定理 3.1 中第  $t$  次迭代的预期性能差距, 执行递归如下:

$$\begin{aligned} \mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}^*)] &\leq B_t + A_t \mathbb{E}[F(\mathbf{w}_{t-1}) - F(\mathbf{w}^*)] \\ &\leq B_t + A_t \left( B_{t-1} + A_{t-1} \mathbb{E}[F(\mathbf{w}_{t-2}) - F(\mathbf{w}^*)] \right) \\ &\leq \dots \\ &\leq \sum_{i=1}^{t-1} \left( \prod_{j=1}^i A_{t+1-j} \right) B_{t-i} + B_t \\ &\quad + \prod_{j=1}^t A_j \mathbb{E}[F(\mathbf{w}_0) - F(\mathbf{w}^*)]. \end{aligned} \quad (2-17)$$

由此得证。 ■

**引理 2.1** 揭示了联邦学习算法在一定温和条件下随着迭代次数 $t$ 渐近收敛，如下面**命题 2.1** 所述。

**命题 2.1** 给定学习率 $\alpha = \frac{1}{L}$ , FLOA 算法确保收敛于 $\lim_{t \rightarrow \infty} \mathbf{w}_t = \mathbf{w}^*$ , 只要公式 (2-12) 中的 $\rho_2$ 满足以下条件:

$$0 < \rho_2 < \frac{\mu}{(\frac{K}{K_{min}} - 1)DL}, \quad (2-18)$$

其中,  $K_{min} = \min\{K_i\}_{i=1}^U$ 。

**证明:** 当 $A_t < 1$ ,  $\forall t$ 时, 很明显可得  $\lim_{t \rightarrow \infty} \prod_{j=1}^{t+1} A_j = 0$ 。从**引理 2.1** 可以看出, 为了保证联邦学习算法的收敛性, 一个充分条件是确保  $A_{max} \triangleq \max\{A_t, t = 1, 2, \dots\} < 1$ 。给定公式 (2-14), 可得

$$\begin{aligned} A_t &= 1 - \frac{\mu}{L} + \rho_2 \sum_{d=1}^D \left( \frac{K}{\sum_{i=1}^U K_i \beta_{i,t}^d} - 1 \right) \\ &\leq 1 - \frac{\mu}{L} + \rho_2 \sum_{d=1}^D \left( \frac{K}{K_{min}} - 1 \right), \end{aligned} \quad (2-19)$$

其中,  $K_{min} = \min\{K_i\}_{i=1}^U$ 。如果所有的分布式设备拥有相同数量的样本数据, 即  $K_i = \frac{K}{U}$ ,  $\forall i$ , 则有  $A_t \leq 1 - \frac{\mu}{L} + \rho_2 D(U - 1)$ 。

为了确保 $A_{max} < 1$ , 须确保下列条件成立:

$$A_{max} \leq 1 - \frac{\mu}{L} + \rho_2 \sum_{d=1}^D \left( \frac{K}{K_{min}} - 1 \right) < 1. \quad (2-20)$$

由上式 (2-20) 可得,  $\rho_2 < \frac{\mu}{(\frac{K}{K_{min}} - 1)DL}$ 。另一方面来讲, 根据**假设 2.3** 中公式 (2-12), 有  $\rho_2 > 0$ 。综上, 可得出  $0 < \rho_2 < \frac{\mu}{(\frac{K}{K_{min}} - 1)DL}$  的收敛性条件。 ■

从**命题 2.1** 可以看出, FLOA 算法的收敛行为取决于与学习相关的参数, 即 $\mu$ ,  $L$ ,  $\rho_1$ ,  $\rho_2$ 以及与通信相关的参数, 包括 $\beta$ ,  $\mathbf{b}$ 和 $\sigma^2$ 。有趣的是, 信道噪声 $\sigma^2$ 和功率缩放因子 $\mathbf{b}$ 不影响 $A_t$ , 因此它们不影响联邦学习算法的收敛性, 但决定了联邦学习算法收敛到的稳态。

另一方面, **引理 2.1** 还提供了传输链路无差错时, 联邦学习算法的预期收敛速率。在这种无差错传输的理想情况下, 联邦学习算法可实现最快收敛速度, 如下述**引理 2.2** 所示。

**引理 2.2** 考虑一种资源不受约束且无错误的传输模式, 其中无线信道的影响以及噪声的影响已经得到消除或完全补偿。给定公式 (2-9) 中学习得到的全局模

型参数  $\mathbf{w}_t$  与全局最优模型参数  $\mathbf{w}^*$ ，该理想情况下的联邦学习的性能累积误差  $\mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}^*)]$  的上界由下式给出：

$$\mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}^*)] \leq (1 - \frac{\mu}{L})^t \mathbb{E}[F(\mathbf{w}_0) - F(\mathbf{w}^*)]. \quad (2-21)$$

**证明：** 在没有信道噪声或分布式设备选择（即所有分布式设备都参与联邦学习并完美交付它们的更新参数）的情况下，有  $\sigma^2 = 0$  和  $\sum_{d=1}^D \left( \frac{K}{\sum_{i=1}^U K_i \beta_{i,t}^d} - 1 \right) = 0$ 。然后，考虑公式 (2-14) 和公式 (2-15)，有  $B_t = B = 0$  和  $A_t = A = 1 - \frac{\mu}{L}$ ,  $\forall t$ 。因此，公式 (2-16) 被简化为公式 (2-21)。引理由此得证。 ■

值得注意的是引理 2.2 提供了理想情况下的收敛速率，该理想情况假设无线通信的影响，包括噪声、信道和受限资源，都得到了消除，从而在分布式设备和参数服务器之间实现更新参数的无差错传输。根据实际情况下的公式 (2-16)，联邦学习算法在每一步更新时，损失函数的期望值  $\mathbb{E}[F(\mathbf{w}_{t+1})]$  的轨迹表现出带有性能误差项  $\Delta_t$  的跳跃不连续性，如引理 2.1 中所定义：

$$\Delta_t = \sum_{i=1}^{t-1} \left( \prod_{j=1}^i A_{t+1-j} \right) B_{t-i} + B_t.$$

此性能误差项反映了无线通信因素通过分布式设备选择、发射功率缩放和 AWGN 对 FLOA 的影响。直观地说，由于选定分布式设备数量的增加使得  $A_t$  减小，上述误差项误差也随着设备数量的增加而缩小。同时，随着功率缩放因子  $b_t$  的增加， $B_t$  减小，这也导致了误差的缩小。因此，有必要优化发射功率缩放因子和分布式设备选择，以最小化现实无线网络上实现联邦学习算法在公式 (2-16) 中的性能误差。

### 2.2.3 非凸损失函数假设下收敛性分析

在实际应用中，损失函数  $F(w)$  有可能是非凸的，例如在卷积神经网络的情况下。为了使得本文研究内容应用具有一般普适性，本章节在没有假设 2.2 的非凸情况下，推导出 FLOA 的收敛行为，如下面定理 2.2 所示。

**定理 2.2** 在假设 2.1 和假设 2.3 成立的情况下，对于非凸损失函数情况，给定发射功率缩放因子  $b_t$ ，分布式设备选择向量  $\beta_{i,t}$ ，学习速率  $\alpha = \frac{1}{L}$ ，联邦学习算法在第  $t$  次迭代的收敛性表示如下：

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \|\nabla F(\mathbf{w}_{t-1})\|^2 &\leq \frac{2L \sum_{t=1}^T B_t}{T(1 - \rho_2 D(\frac{K}{K_{min}} - 1))} \\ &+ \frac{2L}{T(1 - \rho_2 D(\frac{K}{K_{min}} - 1))} \mathbb{E}[F(\mathbf{w}_0)] - F(\mathbf{w}^*). \end{aligned} \quad (2-22)$$

证明：参见附录 A.2 ■

从定理 2.2 可以看出，当迭代次数  $T$  足够大时，可得

$$\min_{0,1,\dots,T} \mathbb{E}[\|\nabla F(\mathbf{w}_{t-1})\|^2] \leq \frac{1}{T} \sum_{t=1}^T \|\nabla F(\mathbf{w}_{t-1})\|^2 \stackrel{T \rightarrow \infty}{\leq} \underbrace{\frac{2L \sum_{t=1}^T B_t}{T(1 - \rho_2 D(\frac{K}{K_{min}} - 1))}}_{\Delta_T^{NC}}, \quad (2-23)$$

上式 (2-23) 保证了联邦学习算法收敛到一个固定点<sup>[88,95]</sup>。类似地，在非凸损失函数情况下，由于无线通信和设备选择，FLOA 算法的性能误差由下式给出

$$\Delta_t^{NC} = \frac{2L \sum_{t=1}^T B_t}{T(1 - \rho_2 D(\frac{K}{K_{min}} - 1))}. \quad (2-24)$$

此处，非凸和凸损失函数情况下的联邦学习算法的收敛充分条件相同，如命题 2.1 中公式 (2-18) 所示。

#### 2.2.4 随机梯度下降法场景收敛性分析

本章节研究的方法与结果也可以扩展到 SGD 的版本。在这里，本章节提供具有恒定批量大小为  $K_b$  的批量 SGD 的收敛性分析，而结果通过设置  $K_b = 1$  可直接应用于标准 SGD。在强凸情况下，当应用批量 SGD 时，FLOA 的收敛行为总结如下面定理 2.3 所示。

**定理 2.3** 在假设 2.1, 2.2 和 2.3 成立的情况下，给定发射功率缩放因子  $\mathbf{b}_t$ ，分布式设备选择向量  $\beta_{i,t}$ ，最优全局联邦学习模型  $\mathbf{w}^*$ ，学习速率  $\alpha = \frac{1}{L}$ ，则 FLOA 算法在实施批量大小为  $K_b$  的 SGD 情况下的收敛行为由下式给出：

$$\begin{aligned} \mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}^*)] &\leq \underbrace{\sum_{i=1}^{t-1} \prod_{j=1}^i A_{t+1-j}^{SGD} B_{t-i}^{SGD} + B_t^{SGD}}_{\Delta_t^{SGD}} \\ &+ \prod_{j=1}^t A_j^{SGD} \mathbb{E}[F(\mathbf{w}_0) - F(\mathbf{w}^*)], \end{aligned} \quad (2-25)$$

其中，

$$\begin{aligned} A_t^{SGD} &= 1 - \frac{\mu}{L} + \rho_2 \left( \sum_{d=1}^D \left( \frac{(\sum_{i=1}^U K_b)^2 - 2K(\sum_{i=1}^U K_b)}{K^2} \right. \right. \\ &\quad \left. \left. + \frac{(\sum_{i=1}^U K_b)}{\sum_{i=1}^U K_b \beta_{i,t}^d} \right) + \frac{(\sum_{i=1}^U (K_i - K_b))^2}{K^2} \right), \end{aligned} \quad (2-26)$$

$$B_t^{SGD} = \frac{\rho_1}{2L} \left( \sum_{d=1}^D \left( \frac{(\sum_{i=1}^U K_b)^2 - 2K(\sum_{i=1}^U K_b)}{K^2} + \frac{(\sum_{i=1}^U K_b)}{\sum_{i=1}^U K_b \beta_{i,t}^d} \right) \right. \\ \left. + \frac{(\sum_{i=1}^U (K_i - K_b))^2}{K^2} \right) + \left\| \left( \sum_{i=1}^U K_i \beta_{i,t} \odot \mathbf{b}_t \right)^{\odot -1} \right\|^2 \frac{L\sigma^2}{2}. \quad (2-27)$$

证明：参见附录 A.3。 ■

从定理 2.3 可以看出，在实施批量 SGD 的第  $t$  次迭代之后，FLOA 算法的累积性能误差为

$$\Delta_t^{SGD} = \sum_{i=1}^{t-1} \left( \prod_{j=1}^i A_{t+1-j}^{SGD} \right) B_{t-i}^{SGD} + B_t^{SGD}. \quad (2-28)$$

通过定理 2.1 和定理 2.3 的对比可以发现，如果将批量大小  $K_b$  设置为  $K_i$ ，则 SGD 情形的定理 2.3 与 GD 情形的定理 2.1 相同。此外，由于批量大小一般不大于最小本地数据量大小，即  $K_b \leq K_{min} \leq \frac{K}{U}$ ，所以公式 (2-26) 中的  $A_t^{SGD}$  和公式 (2-27) 中的  $B_t^{SGD}$  均随着  $K_b$  的增加而减小，这导致公式 (2-28) 中的性能误差  $\Delta_t^{SGD}$  更小。换句话说，FLOA 在  $K_b$  较大的情况下具有更好的收敛性能。另一方面，这种性能的提高是以每轮通信中分布式设备的高计算负载为代价的，这反映了 SGD 情形中训练性能和计算复杂度之间的权衡。

同样地，本节也推导出批量 SGD 情形下收敛的条件，由以下命题 2.2 给出。

**命题 2.2** 给定学习速率  $\alpha = \frac{1}{L}$ ，若要保证 SGD 情况下 FLOA 算法的收敛性  $\lim_{t \rightarrow \infty} \mathbf{w}_t = \mathbf{w}^*$ ，只要 (2-12) 中的  $\rho_2$  满足以下条件：

$$0 < \rho_2 < \frac{\mu}{\left( \frac{2UK_b}{K} + \frac{U^2K_b^2}{K^2} + DU - \frac{2DUK_b}{K} + \frac{DU^2K_b^2}{K^2} \right) L}. \quad (2-29)$$

证明：与 GD 情况类似，为了保证收敛，充分条件仍然是保证  $A_{max}^{SGD} \triangleq \max\{A_t^{SGD}, t = 1, 2, \dots\} < 1$ 。因此可得

$$A_t^{SGD} = 1 - \frac{\mu}{L} + \rho_2 \left( \frac{(K - UK_b)^2}{K^2} + \sum_{d=1}^D \left( \frac{U^2K_b^2 - 2KUK_b}{K^2} + \frac{U}{\sum_{i=1}^U \beta_{i,t}^d} \right) \right) \\ \leq 1 - \frac{\mu}{L} + \rho_2 \left( \frac{(K - UK_b)^2}{K^2} + \frac{DU^2K_b^2 - 2DKUK_b}{K^2} + DU \right), \quad (2-30)$$

上式 (2-30) 的第二步成立是因为  $\sum_{i=1}^U \beta_{i,t}^d \geq 1$ 。

由上式 (2-30) 可得

$$A_{max}^{SGD} \leq 1 - \frac{\mu}{L} + \rho_2 \left( \frac{(K - UK_b)^2 + DU^2K_b^2 - 2DKUK_b + DUK^2}{K^2} \right) < 1. \quad (2-31)$$

进而由上式 (2-31) 可得

$$\rho_2 < \frac{\mu}{\left(1 - \frac{2UK_b}{K} + \frac{U^2K_b^2}{K^2} + DU - \frac{2DUK_b}{K} + \frac{DU^2K_b^2}{K^2}\right)L}. \quad (2-32)$$

考虑到假设 2.3 中  $\rho_2 > 0$ , 因而命题 2.2 得证。 ■

## 2.3 通信与学习联合优化方法

本章节首先建立一个联合优化问题, 以减少 FLOA 算法的性能误差。此联合优化问题不仅适用于凸和非凸损失函数的情况, 而且适用于使用 GD 或者 SGD 来实现联邦学习算法的情况。为了使该优化问题在参数服务器侧存在一些不可观察参数的情况下适用于实践, 本章节通过施加保守的功率约束将其重新表述为一个近似问题。为了有效地解决这样一个近似问题, 本章节首先确定一个紧密的解空间, 然后通过离散规划设计一个最优求解算法。

### 2.3.1 优化问题建模

由于关注收敛精度, 本章节所涉及优化问题归结为在相应的收敛条件下(即满足命题 2.1 和命题 2.2 中的条件)最小化不同情况下的每次迭代的累积性能误差(即  $\Delta_t$ 、 $\Delta_t^{NC}$  和  $\Delta_t^{SGD}$ )。

认识到求解原优化问题 P1 相当于在公式 (2-7) 中的发射功率约束下, 以迭代的方式最小化这些性能误差  $\Delta_t$ 、 $\Delta_t^{NC}$  和  $\Delta_t^{SGD}$ 。在第  $t$  次迭代中, 这三种情况下的目标函数由下式给出:

$$\Delta_t = B_t + A_t \Delta_{t-1}, \quad (2-33)$$

$$\Delta_t^{NC} = B_t, \quad (2-34)$$

$$\Delta_t^{SGD} = B_t^{SGD} + A_t^{SGD} \Delta_{t-1}^{SGD}, \quad (2-35)$$

其中,  $\Delta_0 = 0$ , 并且  $\Delta_0^{SGD} = 0$ 。注意, 当在第  $t$  次迭代中执行优化时,  $\Delta_{t-1}$  和  $\Delta_{t-1}^{SGD}$  可以被视为常量。

考虑到模拟聚合的逐项传输, 本章节删除了公式 (2-33)、(2-34) 和 (2-35) 中的不相关项, 并从中抽取第  $d$  项作为最小化目标, 具体由下式给出

$$R_t[d] = \frac{L\sigma^2}{2 \left( \sum_{i=1}^U \beta_{i,t}^d K_i b_t^d \right)^2} + \frac{K\rho_1 + 2KL\rho_2 \Delta_{t-1}}{2L \sum_{i=1}^U K_i \beta_{i,t}^d}, \forall d, \quad (2-36)$$

$$R_t^{NC}[d] = \frac{L\sigma^2}{2\left(\sum_{i=1}^U \beta_{i,t}^d K_i b_t^d\right)^2} + \frac{K\rho_1}{2L\sum_{i=1}^U K_i \beta_{i,t}^d}, \forall d, \quad (2-37)$$

$$R_t^{SGD}[d] = \frac{L\sigma^2}{2\left(\sum_{i=1}^U \beta_{i,t}^d K_i b_t^d\right)^2} + \frac{U(\rho_1 + 2L\rho_2\Delta_{t-1})}{2L\sum_{i=1}^U K_i \beta_{i,t}^d}, \forall d. \quad (2-38)$$

由于标识符  $d$  索引的所有条目相对于设计参数都是可分离的，因此本章节通过考虑  $\mathbf{w}_t$  和  $\mathbf{w}_{i,t}$  来逐个优化每个条目。这里省略了上标  $d$  和针对不同情形的索引。为了确定第  $t$  次迭代时的分布式设备选择向量  $\beta_{i,t}$  和功率缩放因子  $b_t$ ，参数服务器执行如下联合优化问题：

$$\begin{aligned} \mathbf{P2.2:} \quad & \min_{\{b_t, \beta_{i,t}\}_{i=1}^U} R_t \\ & \text{s.t. } \left| \frac{\beta_{i,t} K_i b_t}{h_{i,t}} w_{i,t} \right|^2 \leq P_i^{\max}, \\ & \quad \beta_{i,t} \in \{0, 1\}, i \in \{1, 2, \dots, U\}, \end{aligned} \quad (2-39)$$

其中，对于 SGD 情况，公式 (2-39) 中的  $K_i$  应改为  $K_b$ 。

然而，在公式 (2-39) 中，分布式设备的更新  $\{w_{i,t}\}_{i=1}^U$  的知识是优化问题所必需的，但由于模拟聚合通信，参数服务器无法提前获得。为了克服这个问题，本章节通过对本地更新参数  $\mathbf{w}_{i,t}$  的近似来重新制定一个实际的优化问题。根据定理 2.1 证明中的公式，每个局部参数  $\mathbf{w}_{i,t}$  都是从广播的上一轮迭代的全局更新  $\mathbf{w}_{t-1}$  沿其局部数据上的平均梯度方向更新  $\frac{\alpha}{K_i} \sum_{k=1}^{K_i} \nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})$  得来。因此，考虑到可以通过调整学习率或通过简单的裁剪来控制局部梯度，本章节对有界局部梯度做出以下常见合理的假设<sup>[15,88,96,97]</sup>。

**假设 2.4** (局部参数更新有界) 全局参数  $w_{t-1}$  和局部参数更新  $w_{i,t}, \forall i, t$  之间的差距上界表述如下

$$|w_{t-1} - w_{i,t}| \leq \eta, \quad (2-40)$$

其中  $\eta \geq 0$  且满足以下学习率  $\alpha$  相关条件的取值范围

$$\eta \geq \max \left\{ \left\{ \left| \frac{\alpha}{K_i} \sum_{k=1}^{K_i} \nabla f(w, \mathbf{x}_{i,k}, \mathbf{y}_{i,k}) \right| \right\}_{i=1}^U \right\}. \quad (2-41)$$

注意，公式 (2-41) 暗示了  $\eta$  的实际的取值范围。此外，对于 SGD 的情况， $\eta$  的取值范围为  $\eta \geq \max\{|\alpha \mathbb{E}_{\mathcal{D}_i}[\nabla f(w, \mathbf{x}_{i,k}, \mathbf{y}_{i,k})]| \}_{i=1}^U$ 。在实践中， $\eta$  可以取  $\eta = |w_{t-1} - w_{t-2}|$ 。

在假设 2.4 下，通过替换公式 (2-39) 限定条件中的  $w_{i,t}, \forall i, t$ ，本章节将优化问题 P2.2 重新公式化为其近似形式 P2.3，如下：

$$\begin{aligned}
 \text{P2.3: } & \min_{\{b_t, \beta_{i,t}\}_{i=1}^U} R_t \\
 \text{s.t. } & \left| \frac{\beta_{i,t} K_i b_t}{h_{i,t}} \right|^2 (|w_{t-1}| + \eta)^2 \leq P_i^{\max}, \\
 & \beta_{i,t} \in \{0, 1\}, \quad i \in \{1, 2, \dots, U\},
 \end{aligned} \tag{2-42}$$

其中，上式 (2-42) 中的功率约束条件是基于以下事实构建的

$$\left| \frac{\beta_{i,t} K_i b_t}{h_{i,t}} w_{i,t} \right|^2 = \left| \frac{\beta_{i,t} K_i b_t}{h_{i,t}} \right|^2 |w_{i,t}|^2 \leq \left| \frac{\beta_{i,t} K_i b_t}{h_{i,t}} \right|^2 (|w_{t-1}| + \eta)^2. \tag{2-43}$$

对于第  $t$  次迭代的优化，由于  $w_{t-1}$  在参数服务器处总是可用的，优化问题 P2.3 在实践中是可行的。接下来，本章通过离散规划来找到优化问题 P2.3 的最优解。

### 2.3.2 基于离散规划的求解方案

优化问题 P2.3 是一个混合整数规划 (Mixed Integer Programming, MIP) 问题，其直接解决方案具有较高的计算复杂度。为了以有效的方式解决该优化问题 P2.3，本章节通过确定一个紧密的搜索空间而不损失最优性来开发一个简单的解决方案。以下定理 2.4 给出的紧密搜索空间是问题 P2.3 中约束条件所满足的结果，其与目标函数无关。因此，它普遍适用于公式 (2-36)，公式 (2-37) 以及公式 (2-38) 中的任何  $R_t$ 。

**定理 2.4** 当优化问题 P2.3 中的所有必需参数，即  $\{P_i^{\max}, w_{t-1}, h_{i,t}, K_i, \eta\}_{i=1}^U$ ，在参数服务器处可获得，则优化问题 P2.3 中  $(b_t, \beta_{i,t})$  的解空间可以缩减到以下紧密搜索空间而不损失其最优性：

$$\begin{aligned}
 \mathcal{S} = \left\{ \left\{ \left( b_t^{(k)}, \beta_{i,t}^{(k)} \right) \right\}_{k=1}^U \middle| b_t^{(k)} = \left| \frac{\sqrt{P_k^{\max}} h_{k,t}}{K_k (|w_{t-1}| + \eta)} \right|, \right. \\
 \left. \beta_t^{(k)}(b_t^{(k)}) = [\beta_{1,t}^{(k)}, \dots, \beta_{U,t}^{(k)}], k = 1, \dots, U \right\}, \tag{2-44}$$

其中  $\beta_t^{(k)}$  是  $b_t^{(k)}$  的函数，其形式为：

$$\beta_{i,t}^{(k)} = H \left( P_i^{\max} - \left| \frac{K_i b_t^{(k)} (|w_{t-1}| + \eta)}{h_{i,t}} \right| \right), \tag{2-45}$$

$H(x)$  是 Heaviside 阶跃函数，即对于  $x > 0$ ， $H(x) = 1$ ，否则  $H(x) = 0$ 。

**证明：** 参见附录 A.4 ■

借助定理 3.4，优化问题 P3 可以等价地从 MIP 转换为离散规划 (Discrete Programming, DP) 问题 P2.4，如下所示

$$\text{P2.4: } \min_{(b_t, \beta_t) \in \mathcal{S}} R_t = R_t(b_t, \beta_t). \tag{2-46}$$

根据问题P2.4，目标函数 $R_t$ 只能取与 $b_t$ 的 $U$ 个可行值对应的可能值；同时，给定每个 $b_t$ ，参数 $\beta_t$ 的值是唯一确定的。因此，通过对公式(2-44)中 $U$ 个可行点 $(b_t, \beta_t)$ 的线性搜索来获得目标函数 $R_t$ 的最小值。请注意，公式(2-44)中的可行点由 $U$ 个分布式设备的信道增益、设备发送功率限制和本地样本数决定。因此，由优化问题P2.4决定的最优传输策略反映了分布式设备在信道质量、可用功率资源和样本数量方面的权衡。

值得注意的是，优化问题P2.4的解 $b_t^*$ 可能超过分布式设备允许的最大值，这是由于假设2.4中引入了近似项。为了严格遵守发送功率约束，每个分布式设备在发送其本地参数时需要采取以下修剪步骤：

- 1) 如果  $\left| \frac{K_i b_t^* w_{i,t}}{h_{i,t}} \right|^2 \leq P_i^{\max}$ ，那么第 $i$ 个分布式设备发送  $\frac{K_i b_t^* w_{i,t}}{h_{i,t}}$ ；
- 2) 否则，它发送  $\sqrt{P_i^{\max}} \operatorname{sgn}(w_{i,t})$ ，其中  $\operatorname{sgn}(\cdot)$  表示符号函数，其返回一个整型变量，指出参数的正负号。

综上所述，本章节为 FLOA 提出了一个联合优化算法(joint optimization for **FL over the air**, INFLOTA)，其具体步骤归纳如算法2.1所示。该算法是一种动态调度和功率放缩策略。通过使用不同的目标函数 $R_t$ ，本章节所提出的算法 INFLOTA 可以调整到所有考虑的情况，包括凸和非凸情况，以及使用 GD 或 SGD 实现的联邦学习。

优化问题P2.3等价地重新表述为优化问题P2.4，并通过线性搜索方法快速解决得到全局最优解。由于搜索空间的减少，搜索算法的计算复杂度大大降低。通过比较优化问题P2.3和P2.2，优化问题P2.3的约束条件比优化问题P2.2中的约束条件更严格。由于优化问题P2.3缩小了优化问题P2.2的可行域，优化问题P2.3的解不会优于优化问题P2.2的解。因此，优化问题P2.3的最优解是优化问题P2.2的上界，即通过求解优化问题P2.3计算得到的目标函数值 $R_t$ 大于等于实际值。

**算法 2.1** 为参数服务器和分布式设备双方的合作进行的联邦学习的实施提供了一个整体解决方案。它的计算复杂度主要由优化问题P2.4中的优化步骤所决定。该优化步骤的复杂度低至 $\mathcal{O}(U)$ ，因为搜索空间通过定理2.4减少到了 $U$ 个可行点。

为了在**算法 2.1**中实现 FLOA，参数服务器必须知道信道状态信息 CSI，每个分布式设备的数据样本的数量以及所有分布式设备的最大发射功率。当分布式设备最初连接到参数服务器时，参数服务器可以获取这些信息。在执行**算法 2.1**之前，参数服务器必须首先将全局模型信息广播给所有分布式设备。

值得注意的是，将优化问题P2.4带入联邦学习的具体实施中，一些分布式设备可能需要发送  $\operatorname{sgn}(w_{i,t}) \min\left(\frac{K_i b_t |w_{i,t}|}{h_{i,t}}, \sqrt{P_i^{\max}}\right)$  以满足其最大发射功率的要求。这样的边界修剪方法可以看作是一种量化措施，依然可以保证算法收敛<sup>[98]</sup>。

表 2.1 INFLOTA 算法

Table 2.1 The implementation of INFLOTA

**算法 2-1 INFLOTA 算法****输入:**系统参数  $\{P_i^{\max}, K_i, \eta\}_{i=1}^U$ 1. 参数服务器初始化  $\{\mathbf{w}_0, b_1^*, \beta_1^*\}$  并将它们广播给所有分布式设备;2. **For**  $t = 1 : T$  **do****分布式设备:**3. **计算:** 通过公式 (2-4) 更新本地模型, 其中  $\mathbf{w} = \mathbf{w}_{t-1}$  来自参数服务器;4. **通信:** 收到  $(b_t, \beta_t)$  后, 如果  $\beta_{i,t} = 1, \forall i, d$ , 发送  $\text{sgn}(w_{i,t}) \min\left(\frac{K_i b_t |w_{i,t}|}{h_{i,t}}, \sqrt{P_i^{\max}}\right)$  到参数服务器;**参数服务器:**5. 通过公式 (2-9) 中接收到的来自分布式设备的聚合参数计算全局模型  $\mathbf{w}_t$ ;6. **For**  $d = 1 : D$  **do**7. 根据公式 (2-44) 计算  $\mathcal{S}$ , 得到  $U$  个可行点  $\{(b_{t+1}^{(k)}, \beta_{t+1}^{(k)})\}_{k=1}^U$ ;8. 对于给定的  $d$  和  $t$ , 求解公式 (2-46) 中的问题 P2.4, 方法是在  $U$  个可行点上进行线性搜索以找到最优解  $\{b_{t+1}^*, \beta_{t+1}^*\}$ ;9. **End For**10. 发送  $w_t$  和  $(b_{t+1}^*, \beta_{t+1}^*)$  (包括所有  $D$  个最优的  $\{b_{t+1}^*, \beta_{t+1}^*\}$ ) 给所有的分布式设备;11. **End For**

## 2.4 性能仿真与分析

本节对所提出的 INFLOTA 算法分别在线性回归和图像分类任务中进行仿真性能分析, 这两种任务分别基于生成数据集和美国国家标准与技术研究院混合数据集 (Mixed National Institute of Standards and Technology database, MNIST)<sup>2</sup>。

### 2.4.1 系统设定

所考虑的无线网络中有  $U = 20$  个分布式设备, 对于任何  $i \in [1, U]$ , 其最大发射

<sup>2</sup> <http://yann.lecun.com/exdb/mnist/>

功率设置为  $P_i^{\max} = P^{\max} = 10 \text{ mW}$ 。参数服务器处的接收机噪声功率设置为  $\sigma^2 = 10^{-4} \text{ mW}$ , 即  $SNR = \frac{P^{\max}}{\sigma^2} = 5 \text{ dB}$ 。分布式设备和参数服务器之间的无线信道增益由瑞利衰落模型生成。对于不同的  $i$  和  $t$ , 信道增益  $h_{i,t}$  由服从单位均值的指数分布生成。

在本章节中, 两种基线方法和本章所提出的 IFLOTA 算法进行比较:

- (1) 一种假设理想化的无线传输具有无差错链路以实现完美聚合的联邦学习算法, 命名为“完美聚合”;
- (2) 一种随机确定功率缩放和分布式设备选择的联邦学习算法, 命名为“随机策略”。在随机策略中, 每个分布式设备被选中的概率是 50%, 并且功率缩放因子是利用具有单位均值的指数分布独立生成。

#### 2.4.2 线性回归实验性能分析

在线性回归实验中, 用于训练联邦学习算法的生成输入数据集  $x$  在  $[0, 1]$  之间随机生成。输入  $x$  和输出  $y$  遵循函数  $y = -2x + 1 + n \times 0.4$ , 其中  $n$  遵循高斯分布  $\mathcal{N}(0, 1)$ 。联邦学习算法用于建模输入  $x$  和输出  $y$  之间的关系。

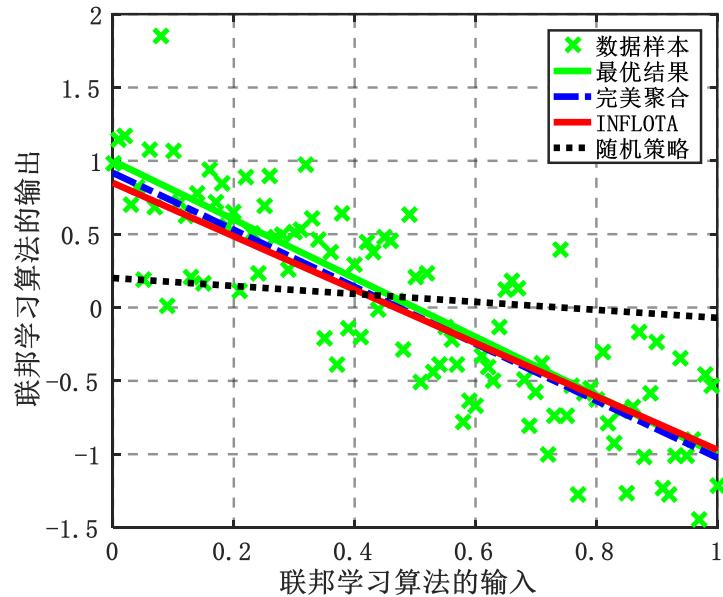


图 2.2 联邦学习算法的线性回归实验结果

Figure 2.2 An example of implementing federated learning for linear regression

由于线性回归只涉及两个参数, 仿真实验通过 20 个分布式设备共同训练了一个简单的两层神经网络, 每层神经网络有一个神经元, 层与层之间没有激活函数。损失函数是模型预测结果  $\hat{y}$  和标记的真实输出结果  $y$  的最小均方误差 MSE。算法所

使用的固定学习速率设置为 0.01。此仿真实验可以看作是在凸损失函数的场景下施行全批量梯度下降法的结果。

图 2.2 显示了使用各个联邦学习算法进行线性回归实验的不同示例。因为原始数据生成函数是  $y = -2x + 1 + n \times 0.4$ , 所以线性回归的最优结果是  $y = -2x + 1$ 。在图 2.2 中, 可以看出最准确的近似是通过完美聚合情况下的联邦学习算法实现的, 但其是不考虑无线通信以及设备选择的影响的理想情况 (即所有设备全部参与且它们上传的数据均是完美无误差的)。在随机策略下的联邦学习算法中考虑了无线通信的影响, 但却没有进行任何优化。因此, 它的性能是最差的, 由此可见无线通信和设备选择造成的负面影响可以显著降低联邦学习的性能。而本章所提出的 INFLOTA 算法共同考虑了无线通信的影响和学习算法本身, 其性能与理想情况非常接近。这是因为本章所提出的 INFLOTA 算法可以优化分布式设备选择和功率控制, 从而减少无线传输错误对联邦学习的负面影响。

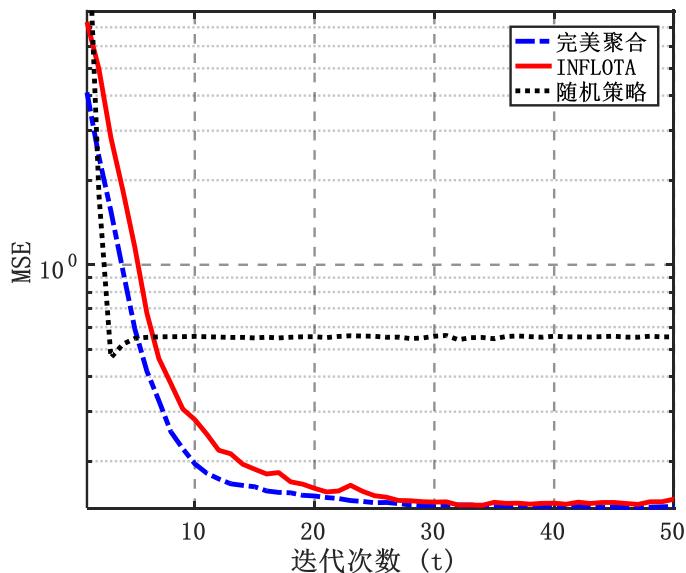


图 2.3 MSE 随迭代变化

Figure 2.3 MSE as the number of iterations varies

在图 2.3 中, 本章节展示了无线传输如何影响全局联邦学习模型训练的收敛行为。随着迭代次数的增加, 全局联邦学习模型的损失函数值最终保持不变, 这表明全局联邦学习模型收敛。正如所看到的, 随着迭代次数的增加, 所有考虑到的学习算法的 MSE 值以不同的速率下降, 并最终趋于平缓以达到稳定状态, 即所有方案都会收敛, 但会收敛到不同的稳态值。这种现象证实了引理 2.1 和命题 2.1 中的结果, 即信道加性白噪声 AWGN 不会影响联邦学习算法的收敛性, 但会影响联邦学习算法收敛的稳态值。

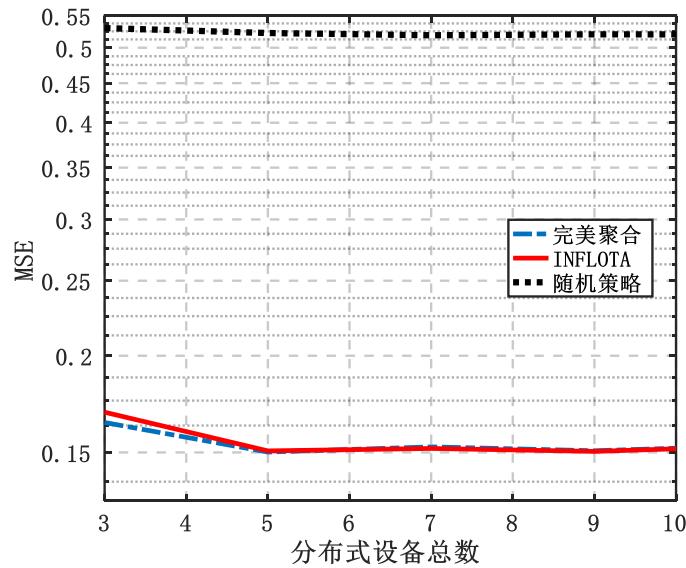


图 2.4 MSE 随分布式设备总数变化

Figure 2.4 MSE as the number of distributed devices varies

图 2.4 显示了 MSE 如何随分布式设备总数  $U$  变化。一般来说，所有考虑的联邦学习方案的 MSE 性能随着  $U$  的增加而降低。这是因为分布式设备总数的增加导致可用于联邦学习训练的数据量整体增加，从而提高了估计模型参数的准确性。此外，随着分布式设备数量的增加，无线传输对全局联邦学习模型精度的影响开始减弱。这是因为当  $U$  超过某个水平时，数据样本可能已经足以进行准确的训练。

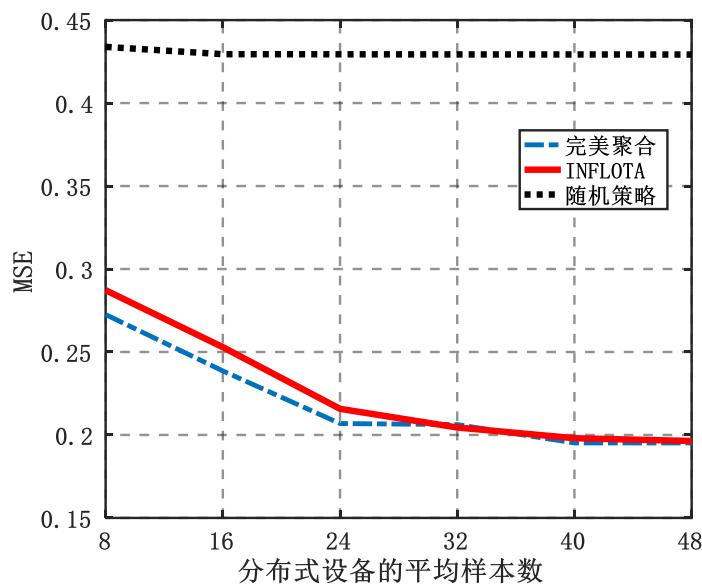


图 2.5 MSE 随每个设备的平均样本数变化

Figure 2.5 MSE as the number of data samples per worker varies

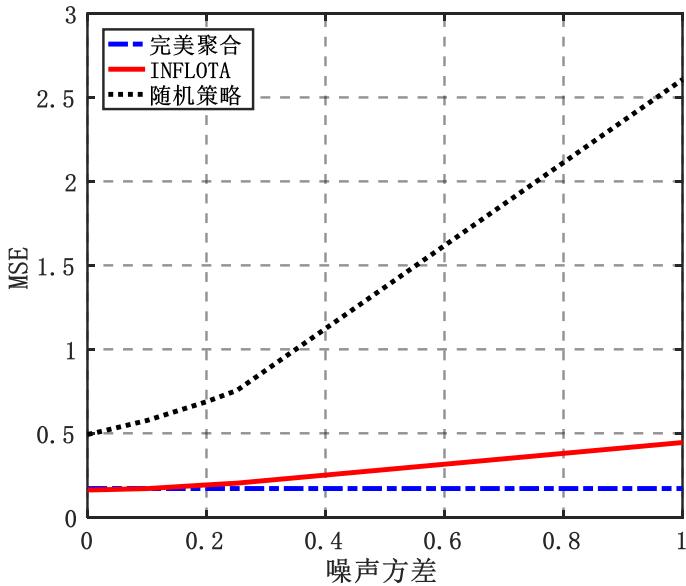


图 2.6 MSE 随噪声方差变化

Figure 2.6 MSE as the noise variance varies

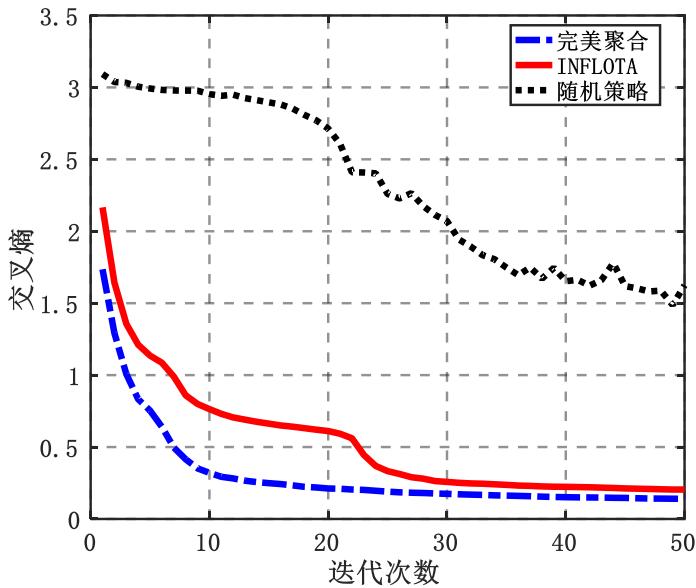


图 2.7 交叉熵随迭代次数变化

Figure 2.7 Cross entropy as the number of iterations varies.

在图 2.5 中，本章节展示了 MSE 如何随着每个分布式设备的平均样本数  $\bar{K} = K/U$  变化而变化。每个分布式设备的数据样本数围绕平均数波动，即设置  $K_i = \text{round}(\text{uniform}[\bar{K} - 5, \bar{K} + 5])$ 。随着  $\bar{K}$  的增加，所有考虑的联邦学习算法都有更多的数据样本可用于训练，因此在图 2.5 中所有考虑的联邦学习算法的 MSE 都会降低。然而，随着每个分布式设备的平均数据样本量不断增加，MSE 降低速度减慢并最终饱和。这是因为随着每个分布式设备的数据样本量的不断增加，数据样

本足以训练联邦学习模型。

图 2.6 展示了参数服务器接收到的 AWGN 如何影响 MSE 的结果。可以看到，随着噪声方差的增加，所有考虑的联邦学习算法的 MSE 值都会增加，除了完美聚合下不考虑噪声的理想情况。当噪声方差较小时（例如，小于  $10^{-1}$ ），它对联邦学习算法的性能影响不大。

### 2.4.3 图像识别实验性能分析

为了评估本章所提出的 INFLOTA 算法在具有真实数据的实际应用场景中的性能，本节在 MNIST 数据集<sup>[99]</sup>上训练了一个多层感知器 (Multi-layer Perceptron, MLP) 神经网络，其包含一个 784 个神经元的输入层，一个 64 个神经元的隐藏层，一个 10 个神经元的以 softmax 函数为输出函数的输出层。仿真实验采用交叉熵作为损失函数，采用线性整流函数 (Rectified Linear Unit, ReLU) 作为激活函数。MLP 神经网络中的参数总数为 50890。学习速率设置为 0.1。在 MNIST 数据集中，有 60000 个训练样本和 10000 个测试样本。仿真实验分别随机抽取 500-1000 个训练样本，并分发给 20 名分布式设备作为他们的本地数据集。然后用 10000 个测试样本对三个训练好的联邦学习模型进行测试。本节分别在图 2.7 和图 2.8 中提供交叉熵和测试精度随着迭代轮次变化的结果。由于 MNIST 数据集是为手写数字识别而设计的，因此测试准确度呈现的是识别准确度。该仿真实验方案设计面向非凸优化的场景。如图所示，本章提出的 INFLOTA 算法优于随机策略方案，并可以达到与理想情况下的完美聚合方案相当的性能。

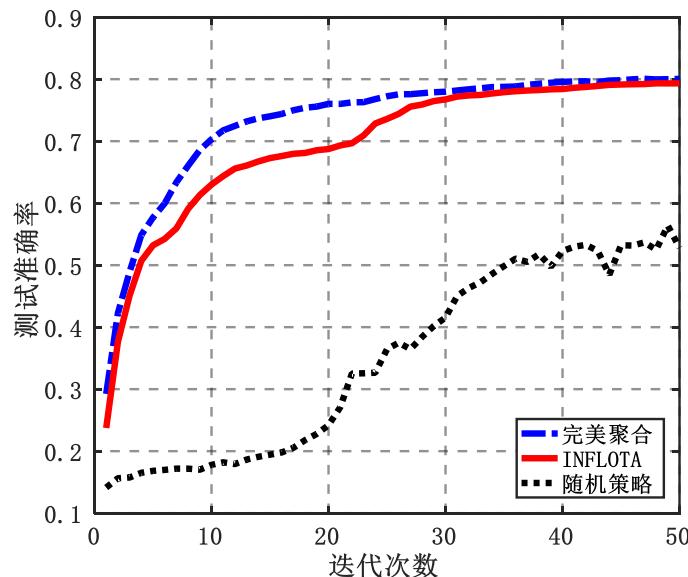


图 2.8 测试准确率随迭代次数变化

Figure 2.8 Test accuracy as the number of iterations varies

## 2.5 本章小结

本章研究了 FLOA 中的传算联合优化问题，其中在有限通信资源的约束下考虑了分布式设备选择和发射功率控制。在 **GD** 或 **SGD** 实现的凸和非凸损失函数情况下，本章分别推导了 FLOA 算法的预期收敛速率的封闭式表达式，其可以量化模拟聚合通信情况下资源受限的无线通信对联邦学习性能的影响。通过对预期收敛性能的分析，本章提出了考虑分布式设备选择和功率控制的通信和学习联合优化方案，其可以减轻无线通信对联邦学习算法收敛性和学习准确率性能的影响。更重要的是，本章所提出的联合优化框架适用于凸和非凸损失函数情况下使用 **GD** 或 **SGD** 来实现的联邦学习的场景。仿真结果表明，本章在收敛性理论分析的指导下所提出的联合优化方案在减轻无线通信对联邦学习的影响方面是有效的。本章研究内容发表在 IEEE ICC2022 会议论文集中，以及 IEEE TWC 期刊上。

### 3 基于 FLOA 传算联合优化框架的高效性方案研究

利用无线多址信道的波形叠加特性，模拟聚合可以在传输中自动直接获得联邦学习所需的平均更新，这推动了 FLOA 的繁荣。尽管当前存在一些面向 FLOA 的研究工作，但一些基本问题仍未得到解答，阻碍了实现高效通信和高性能的 FLOA。首先，联邦学习与模拟聚合通信之间的量化关系尚不清楚<sup>[100,101]</sup>。参与联邦学习的分布式设备数量的简单最大化与学习无关，因此不一定是最优的，因为这将计算和通信的优化进行了解耦，例如文献[31,32]中的工作。其次，为了便于功率控制，大多数现有工作都是基于一个强假设开展的，即从本地设备传输的信号，即本地梯度，可以归一化为具有零均值和单位方差的变量<sup>[32,59,61,62]</sup>。然而，联邦学习中的梯度统计在训练迭代和特征维度上都是不同的，并且是先验未知的<sup>[60]</sup>。因此，在不事先了解参数服务器处的局部梯度的情况下设计最优功率控制是不可行的，特别是对于 FLOA 中的非编码线性模拟调制。第三，在 FLOA 中引入了稀疏化作为局部梯度的有损压缩的手段<sup>[28,29]</sup>，这可能会引入聚合错误，但是这些聚合错误对联邦学习的影响尚不清楚，更不用说如何减轻它们的副作用。

为了解决上述问题，本章通过引入 1 比特压缩感知 (Compressive Sensing, CS) 以实现高效的 FLOA，并在该方案下进一步提出分布式设备选择和功率控制优化策略。据了解，本章是第一个将 1 比特压缩感知<sup>[102-104]</sup>引入 FLOA 以提高通信效率的工作，其中局部梯度的维度和量化位数都可以显著地减少。此外，由于 1 比特量化，本章的功率控制变得可行，因为它取决于已知幅度的量化值，而不依赖于任何先验知识或梯度统计或特定分布的假设。更重要的是，在 1 比特压缩感知技术下，本章的工作对联邦学习和模拟聚合之间的关系提供了基本解释，以实现计算和通信的联合优化。本章的主要贡献概述如下：

- (1) 本章提出一种利用 1 比特压缩感知和模拟聚合 (**One Bit CS Analog Aggregation, OBCSAA**) 技术来实现高效联邦学习的方案。在所提 OBCSAA 中，本章设计一套压缩、模拟聚合传输以及信号重构的解决方案，以实现高效通信的联邦学习。
- (2) 本章为所提出的 OBCSAA 的预期收敛速率推导出一个封闭形式的表达式。此封闭形式的表达式量化由稀疏化、降维、量化、模拟聚合传输、信号重构和 AWGN 引起的聚合误差导致的性能权衡，其为设计模拟无线传输系统提供新的视角。
- (3) 以收敛性理论结果为指导，本章制定计算和通信的联合优化问题，以优化分布式设备选择和功率控制。鉴于允许的峰值发射功率和可用带宽的实际限制，该优化问题旨在减轻聚合错误以提高学习准确率。为了解决这个非

凸的联合优化问题，本章分别针对小型网络和大型网络场景提出两种不同的解决方案：基于枚举的方法和基于交替方向乘子法（Alternating Direction Method of Multipliers, ADMM）的方法。

本章在解决基于 MNIST 数据集的图像分类问题上评估所提出的 OBCSAA。仿真结果表明，本章所提出的 OBCSAA 可实现与现有工作[62]，以及通过在无差错无线信道上实现完美聚合的联邦学习的理想情况相当的性能，并大大提高通信效率。

值得注意的是，与数字通信下的联邦学习不同，FLOA 的优化设计由于模拟聚合而面临着大大降低的自由度，并且尚未在已有文献[28,29,32,59,61,62]中得到很好的探索。与文献[32,59,61]中将调度设备的比例作为学习指标（分离了通信和计算过程）相比，本章的学习指标是关于压缩感知和通信因素的学习收敛性，其提供通信与计算之间的密切关系。不同于文献[32,59,61,62]中假设局部更新必须遵循具有零均值和单位方差的独立同分布（Independent and Identically Distributed, IID），本章的工作采用 1 比特压缩感知技术，其使得所提方案能够实现即使没有文献[32,59,61,62]中所需的任何梯度统计信息，也可以对单个分布式设备进行功率控制。与文献[28,29]相比，本章的工作不仅在降维后应用 1 比特量化，而且还提供基于 1 比特压缩感知技术的 FLOA 的收敛性分析，从而实现计算和通信的联合优化。通过信号重建，本章在参数服务器处恢复稀疏的原始局部梯度的叠加值进行训练，这与使用量化的局部梯度的现有工作[29]不同。简而言之，本章的工作是梯度稀疏化、降维、量化和信号重建的整体集成，以实现高效的 FLOA，并在收敛性的基础上进一步设计分布式设备选择和功率缩放策略。

本章具体内容安排如下：基于 1 比特压缩感知技术 FLOA 的系统模型在 3.1 节中介绍。3.2 节中推导预期收敛速率的封闭式表达式，以量化聚合误差对联邦学习的影响。3.3 节研究通信和联邦学习的联合优化问题，以优化分布式设备选择和功率控制。作为研究工作的延伸，3.4 节讨论所提方案在实施随机梯度下降（SGD）法的情况下收敛性和算法设计。性能仿真结果在 3.5 节中展示。最后，在 3.6 节中提供关于本章的总结。

### 3.1 系统模型

如图 3.1 所示，本章考虑一个由单个参数服务器和  $U$  个分布式设备组成的无线联邦学习系统。利用 1 比特压缩感知技术的无线模拟聚合传输，参数服务器和所有分布式设备协作训练一个共享学习模型。

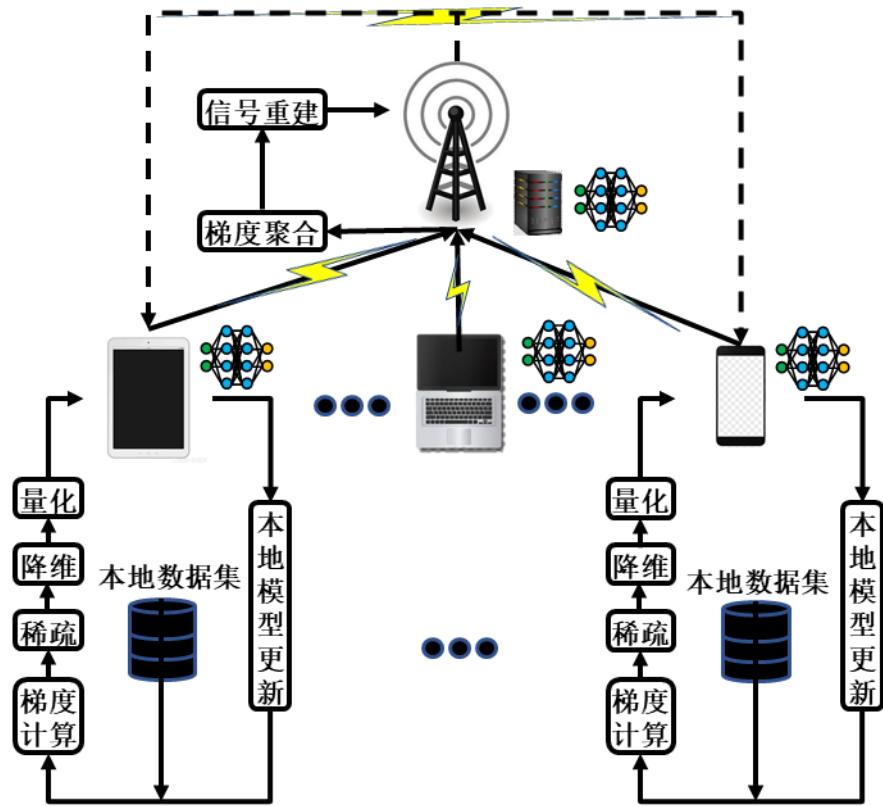


图 3.1 基于 1 比特压缩感知的模拟聚合联邦通信模型框图

Figure 3.1 An analog-aggregated federated learning model based on 1-bit compressed sensing

### 3.1.1 联邦学习模型

所有训练数据集的并集记为  $\mathfrak{D} = \bigcup_i \mathfrak{D}_i$ , 其中  $\mathfrak{D}_i = \{\mathbf{x}_{i,k}, \mathbf{y}_{i,k}\}_{k=1}^{K_i}$  是在第  $i$  个分布式设备的本地数据集, 其中  $K_i = |\mathfrak{D}_i|$  是样本数量,  $i = 1 \dots U$ 。在本地数据集  $\mathfrak{D}_i$  中, 第  $k$  个数据输入样本及其标签分别表示为  $\mathbf{x}_{i,k}$  和  $\mathbf{y}_{i,k}$ ,  $k = 1, 2, \dots, K_i$ 。联邦学习的训练过程的目标是最小化由  $D$  维参数  $\mathbf{w} = [w^1, \dots, w^D] \in \mathcal{R}^D$  参数化的全局共享学习模型的全局损失函数  $F(\mathbf{w}; \mathfrak{D})$ , 即,

$$\text{P3.1: } \mathbf{w}^* = \arg \min_{\mathbf{w} \in \mathcal{R}^D} F(\mathbf{w}; \mathfrak{D}), \quad (3-1)$$

其中,  $F(\mathbf{w}; \mathfrak{D}) = \frac{1}{K} \sum_{i=1}^U \sum_{k=1}^{K_i} f(\mathbf{w}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})$  是  $K = \sum_{i=1}^U K_i$  个样本损失函数之和, 由学习模型定义。注意, 问题 P3.1 的解  $\mathbf{w}^*$  可以是凸优化情况下的全局最优解, 非凸优化情况下的局部最优解或鞍点。

为了避免将原始本地数据集直接上传到参数服务器处进行集中训练, 问题 P3.1 中的学习过程通过迭代梯度平均算法<sup>[85,105]</sup>以分布式的形式进行。具体来说, 在每次迭代  $t$  过程中, 一个一阶算法, 例如梯度下降 (GD) 法, 并行应用于分布式

设备以最小化本地损失函数<sup>3</sup>

$$F_i(\mathbf{w}_i; \mathfrak{D}_i) = \frac{1}{K_i} \sum_{k=1}^{K_i} f(\mathbf{w}_i; \mathbf{x}_{i,k}, \mathbf{y}_{i,k}), \quad i = 1, \dots, U, \quad (3-2)$$

其中， $\mathbf{w}_i = [w_i^1, \dots, w_i^D] \in \mathcal{R}^D$  是本地模型参数。每个分布式设备根据自己的本地数据集从接收到的全局学习模型更新其本地梯度如下

$$(本地梯度更新) \quad \mathbf{g}_i = \frac{1}{K_i} \sum_{k=1}^{K_i} \nabla f(\mathbf{w}_i; \mathbf{x}_{i,k}, \mathbf{y}_{i,k}), \quad i = 1, \dots, U, \quad (3-3)$$

其中， $\nabla f(\mathbf{w}_i; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})$  是样本损失函数  $f(\mathbf{w}_i; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})$  相对于  $\mathbf{w}_i$  的梯度。

然后分布式设备将其本地梯度更新值发送到参数服务器处聚合为全局梯度更新如下

$$(全局梯度更新) \quad \mathbf{g} = \frac{1}{K} \sum_{i=1}^U K_i \mathbf{g}_i, \quad (3-4)$$

并且全局梯度  $\mathbf{g}$  被发送回分布式设备，然后将其用于全局共享模型更新为

$$(全局模型更新) \quad \mathbf{w} = \mathbf{w} - \alpha \mathbf{g}, \quad (3-5)$$

其中， $\alpha$  为学习速率。

联邦学习迭代地实现公式 (3-3)、公式 (3-4) 和 公式 (3-5)<sup>4</sup>，直到它收敛或达到最大迭代次数。

### 3.1.2 1 比特压缩感知和模拟聚合传输过程

当联邦学习应用于大规模网络和高维模型参数的场景中时，模型更新参数在参数服务器和分布式设备之间的传输会消耗大量通信资源并导致训练延迟。同时，由于实际应用中无线通信所带来的传输功率和带宽限制，使用数字通信的以单独的方式逐一传输和重建所有梯度向量中的元素是非常消耗通信资源的。因此，为了减少传输开销并减少通信时间，本章提出在 FLOA 中应用 1 比特压缩感知<sup>[102-104]</sup>。这是出于两个原因，一个是大规模网络学习问题中涉及的梯度通常是可压缩的，即

---

<sup>3</sup> 本章节采用基本梯度下降法 GD 作为一种范例，其可以通过在每个分布式设备处使用小批量进行训练扩展到 SGD，如第 3.5 节所示。注意，与 GD 相比，SGD 以更多迭代为代价来获得计算效率的提高，但传输消耗则更多。

<sup>4</sup> 实际实现中可能涉及在通信前使用小批量样本的梯度更新以及多轮本地迭代。但是，它们不会影响本文中的概念开展。

梯度向量中只有少数具有重要值的元素<sup>[105-107]</sup>。另一个是联邦学习通常运行在基于平均值的分布式学习机制中。

在本章的研究工作中，梯度稀疏化和压缩感知的压缩特性可以降低传输梯度向量的维数。同时，模拟聚合传输使所有分布式设备能够同时使用相同的时频资源将其更新传输到参数服务器。此外，1 比特量化不仅可以最大限度地减少量化开销，而且还可以规避对已知局部梯度分布的不切实际要求。接下来本章节详细说明所提出的 1 比特压缩感知在 FLOA 中应用过程。

- 1) **稀疏：**在第  $t$  次迭代传输之前，所有分布式设备将其更新的本地梯度  $\mathbf{g}_{i,t}$  的  $\kappa$  个元素之外的所有元素设置为 0，从而实现  $\kappa$  级稀疏表示为

$$\tilde{\mathbf{g}}_{i,t} = \text{sparse}_\kappa(\mathbf{g}_{i,t}), \quad (3-6)$$

其中， $\text{sparse}_\kappa(\cdot)$  表示向量的稀疏化操作，其使得  $\tilde{\mathbf{g}}_{i,t}$  的长度为  $D$ ，稀疏阶为  $\kappa$ 。在本章中，所涉及的稀疏以 top- $\kappa$  稀疏化策略为例，即保留具有最大的  $\kappa$  个幅度的元素，而将其它的元素设置为 0。

- 2) **降维：**为了传输其稀疏的局部梯度向量的非零元素，分布式设备需要将非零元素的索引和具体值分别传输到参数服务器，这会导致额外的数据传输。为了避免这种开销，本章使用与文献[28,29]中类似的方法，即所有分布式设备使用相同的测量矩阵  $\Phi \in \mathbb{R}^{S \times D}$  ( $S \ll D$ )，该测量矩阵是一个随机高斯矩阵。请注意，在经过 top- $\kappa$  稀疏化之后，稀疏梯度的特定的  $\kappa$  个非零索引在不同分布式设备处是不同的<sup>5</sup>。这导致叠加的梯度信号的稀疏级  $\bar{\kappa}$  增加，即  $\bar{\kappa} > \kappa$ 。为了可靠地重建压缩梯度，压缩感知过程需要满足受限等距属性（Restricted Isometry Property, RIP）条件，即  $\kappa U \leq S \ll D$  且测量矩阵  $\Phi$  中的每个元素服从独立同分布的高斯分布  $\mathcal{N}(0, \sigma_{sp}^2)$ ，其中  $\kappa U$  是聚合的稀疏梯度的稀疏性的上限，即  $\kappa U > \bar{\kappa}$ 。此外，测量矩阵  $\Phi$  在传输之前需要在分布式设备和参数服务器之间共享。
- 3) **量化：**接下来，对  $\Phi \tilde{\mathbf{g}}_{i,t}$  应用 1 比特量化，这样得到的每个分布式设备的压缩局部梯度  $\mathcal{C}(\mathbf{g}_{i,t})$  由下式给出

$$\begin{aligned} \mathcal{C}(\mathbf{g}_{i,t}) &= \text{sign}(\Phi \text{sparse}_\kappa(\mathbf{g}_{i,t})) \\ &= \text{sign}(\Phi \tilde{\mathbf{g}}_{i,t}), \quad i = 1, \dots, U, \end{aligned} \quad (3-7)$$

其中  $\mathcal{C}(\cdot)$  表示整体 1 比特压缩感知操作，包括 top- $\kappa$  稀疏化、降维和 1 比特

<sup>5</sup> 当分布式设备的本地数据集是独立同分布的时，不同分布式设备的稀疏化梯度的  $\kappa$  个非零索引值将会出现较多的重复。

量化。此外，令  $\mathcal{C}(\mathbf{g}_{i,t}) = [c_{i,t}^1, \dots, c_{i,t}^s, \dots, c_{i,t}^S]^T$ ，其中  $c_{i,t}^s = \pm 1$  是由于 1 比特量化。

- 4) **模拟聚合传输：**在公式 (3-7) 中收集压缩测量后，所有分布式设备在功率控制下以模拟传输的方式将其本地压缩测量值  $\mathcal{C}(\mathbf{g}_{i,t})$  传输到参数服务器，并在参数服务器上通过无线方式聚合以实现公式 (3-4) 中的全局梯度更新步骤。具体地，每个分布式设备  $i$  处的本地  $\mathcal{C}(\mathbf{g}_{i,t})$  乘以功率控制因子  $p_{i,t}$ ，然后发送到参数服务器处。当所有参与用户同步传输时，在参数服务器处接收到的信号向量由下式给出

$$\mathbf{y}_t = \sum_{i=1}^U h_{i,t} p_{i,t} \mathcal{C}(\mathbf{g}_{i,t}) + \mathbf{z}_t, \quad (3-8)$$

其中  $\mathbf{z}_t \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$  是 AWGN 向量， $h_{i,t}$  表示在第  $t$  次迭代时第  $i$  个分布式设备和参数服务器之间的信道系数<sup>6</sup>。

功率控制策略包括设备选择和发射功率缩放。令  $\beta_{i,t}$  表示分布式设备选择的指示因子，即  $\beta_{i,t} = 1$  表示第  $t$  次迭代时第  $i$  个分布式设备被选中执行联邦学习算法，否则  $\beta_{i,t} = 0$  表示其未被选中。为了实现公式 (3-4) 中的平均梯度步骤，第  $t$  次迭代时，在参数服务器处的感兴趣信号向量由下式给出

$$\mathbf{y}_t^{desired} = \frac{\sum_{i=1}^U K_i \beta_{i,t} \mathcal{C}(\mathbf{g}_{i,t})}{\sum_{i=1}^U K_i \beta_{i,t}}. \quad (3-9)$$

为了获得上述公式 (3-9) 中的感兴趣的信号向量，本章将功率控制因子  $p_{i,t}$  设计为

$$p_{i,t} = \frac{\beta_{i,t} K_i b_t}{h_{i,t}}, \quad (3-10)$$

其中， $b_t$  是功率缩放因子。通过这种功率缩放，第  $i$  个分布式设备的发射功率满足功率限制  $P_i^{\text{Max}}$  为

$$|p_{i,t} c_{i,t}^s|^2 = \left( \frac{\beta_{i,t} K_i b_t}{h_{i,t}} c_{i,t}^s \right)^2 = \frac{\beta_{i,t}^2 K_i^2 b_t^2}{h_{i,t}^2} \leq P_i^{\text{Max}}, \quad (3-11)$$

其中，由于  $c_{i,t}^s = \pm 1$ ， $c_{i,t}^s$  被消去<sup>7</sup>。

在应用功率控制  $p_{i,t}$  并将公式 (3-10) 代入公式 (3-8) 后，公式 (3-8)

<sup>6</sup> 本章考虑块衰落信道，其中信道状态信息 (CSI) 在联邦学习的每次迭代中保持不变，但可能在一次迭代之间独立变化。此外，本章假设 CSI 在参数服务器和分布式设备处都是完全已知的，因此可以在分布式设备传输梯度更新之前补偿信道相位偏移。

<sup>7</sup> 从公式 (3-11) 中看出，功率控制与特定的局部梯度无关，这使得在没有任何梯度或梯度统计的先验知识的情况下能够优化功率控制。这是本章将 1 比特压缩感知通过无线的方式整合到 FLOA 中的动机之一。

中的接收信号向量可以重写为

$$\mathbf{y}_t = \sum_{i=1}^U K_i b_t \beta_{i,t} \mathcal{C}(\mathbf{g}_{i,t}) + \mathbf{z}_t. \quad (3-12)$$

在接收到上式 (3-12) 中的信号  $\mathbf{y}_t$  后，参数服务器通过后处理操作将感兴趣的信号向量估计为<sup>8</sup>

$$\begin{aligned} \hat{\mathbf{y}}_t^{desired} &= \left( \sum_{i=1}^U K_i \beta_{i,t} b_t \right)^{-1} \mathbf{y}_t \\ &= \left( \sum_{i=1}^U K_i \beta_{i,t} \right)^{-1} \sum_{i=1}^U K_i \beta_{i,t} \mathcal{C}(\mathbf{g}_{i,t}) + \left( \sum_{i=1}^U K_i \beta_{i,t} b_t \right)^{-1} \mathbf{z}_t \\ &= \mathbf{y}_t^{desired} + \frac{\mathbf{z}_t}{\sum_{i=1}^U K_i \beta_{i,t} b_t}, \end{aligned} \quad (3-13)$$

其中， $(\sum_{i=1}^U K_i \beta_{i,t} b_t)^{-1}$  是后处理因子。

- 5) 信号重构：**从公式 (3-13) 得到  $\hat{\mathbf{y}}_t^{desired}$  后，参数服务器需要使用压缩感知重建算法  $\mathcal{C}^{-1}(\cdot)$  来估计聚合后的全局梯度，即  $\hat{\mathbf{g}}_t = \mathcal{C}^{-1}(\hat{\mathbf{y}}_t^{desired})$ 。重构算法  $\mathcal{C}^{-1}(\cdot)$  存在许多选项可用，例如二进制迭代硬阈值 (Binary Iteration Hard Thresholding, BIHT) 算法<sup>[103]</sup>、定点延续算法<sup>[108]</sup>、基础追踪算法<sup>[109]</sup>和其他一些贪婪匹配追踪算法<sup>[110]</sup>。

然后参数服务器将估计的全局梯度  $\hat{\mathbf{g}}_t$  广播给所有分布式设备，以更新共享模型参数，如下所示

$$\mathbf{w}_{i,t+1} = \mathbf{w}_{i,t} - \alpha \hat{\mathbf{g}}_t, \quad i = 1, 2, \dots, U. \quad (3-14)$$

比较公式 (3-14) 和公式 (3-5)，由于模拟聚合传输、top- $\kappa$ 稀疏化、压缩感知降维和 1 比特量化，聚合误差可能会在基于 1 比特压缩感知的 FLOA 中引入。

## 3.2 压缩和通信对 FLOA 的定量化描述

本节通过分析收敛性来研究模拟聚合传输和 1 比特压缩感知对 FLOA 的影响。

### 3.2.1 基本假设

为了便于收敛性分析，本小节对损失函数和梯度向量做出以下优化问题研究

<sup>8</sup> 值得注意的是，这种信道反转方法可以通过在策略中采用截断信道值来改进，这会使得 FLOA 在深度衰落信道上的学习性能更好<sup>[61]</sup>。

中常见的标准假设<sup>[91,92,96,111]</sup>。

**假设 3.1 (Lipschitz 连续性, 平滑) :** 损失函数  $F(\mathbf{w})$  的梯度  $\nabla F(\mathbf{w})$  为  $L$ -Lipschitz 连续平滑的, 即,

$$\|\nabla F(\mathbf{w}_{t+1}) - \nabla F(\mathbf{w}_t)\| \leq L\|\mathbf{w}_{t+1} - \mathbf{w}_t\|, \quad (3-15)$$

其中,  $L$  表示连续可微函数  $F(\cdot)$  的非负 Lipschitz 常数<sup>[111]</sup>。

**假设 3.2 (二阶连续可微) :** 损失函数  $F(\mathbf{w})$  是二次连续可微和  $L$ -平滑的。因此,  $F(\mathbf{w})$  的 Hessian 矩阵的特征值以下式为界<sup>[111]</sup>

$$\nabla^2 F(\mathbf{w}_t) \preceq L\mathbf{I}. \quad (3-16)$$

**假设 3.3 (样本梯度有界) :** 分布式设备的本地样本梯度值受限于它们的全局梯度值, 即<sup>[91,92]</sup>

$$\|\nabla f(\mathbf{w}_t)\|^2 \leq \rho_1 + \rho_2 \|\nabla F(\mathbf{w}_t)\|^2, \quad (3-17)$$

其中,  $\rho_1 \geq 0$ ,  $0 \leq \rho_2 < 1$ 。

**假设 3.4 (本地梯度有界) :** 本地梯度更新值受限于<sup>[96]</sup>

$$\|\mathbf{g}_{i,t}\|^2 \leq G^2, \forall i, t, \quad (3-18)$$

其中,  $G$  是一个非负常数。

### 3.2.2 收敛性分析

本小节首先分析公式 (3-14) 中恢复的平均梯度与公式 (3-5) 中的理想平均梯度之间的总误差, 包括由稀疏化、量化、AWGN 和重建算法引起的误差。基于上述假设 3.4, 本小节推导出以下引理 3.1 来描述总误差。

**引理 3.1** 基于 1 比特压缩感知的 FLOA 中, 第  $t$  次迭代的总误差  $\mathbf{e}_t = \hat{\mathbf{g}}_t - \mathbf{g}_t$  的上界表述如下

$$\begin{aligned} \mathbb{E}\|\mathbf{e}_t\|^2 &= \mathbb{E}(\|\hat{\mathbf{g}}_t - \mathbf{g}_t\|^2) \leq 2 \sum_{i=1}^U \beta_{i,t} \frac{D-\kappa}{D} G^2 \\ &\quad + 2C^2 \left( 1 + (1+\delta) \frac{D-\kappa}{SD} G^2 + \frac{\sigma^2}{\left(\sum_{i=1}^U K_i \beta_{i,t} b_t\right)^2} \right), \end{aligned} \quad (3-19)$$

其中,  $0 < \delta < 1$  是 RIP 条件下的常数,  $C = \frac{2\varpi}{1-\rho}$ ,  $\varpi = \frac{2\sqrt{1+\delta}}{\sqrt{1-\delta}}$  和  $\varrho = \frac{\sqrt{2}\delta}{1-\delta}$ 。

证明: 参见附录 A.5。 ■

**引理 3.1** 说明较大的  $\kappa$  会导致较小的误差, 这也同时表明应用稀疏化来提高通信高效性是以牺牲准确性为代价的。此外, 由于压缩较少, 较大的  $S$  会导致较小的降维误差。

接下来，本小节提出基于 1 比特压缩感知技术的 FLOA 预期收敛速率的主要定理，如下面定理 3.1 所示。

**定理 3.1** 给定功率缩放因子  $b_t$ 、分布式设备选择向量  $\beta_{i,t}$  和学习速率  $\alpha = \frac{1}{L}$ ，在第  $T$  迭代时，基于 1 比特压缩感知技术的 FLOA 有以下收敛速率

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \| \nabla F(\mathbf{w}_{t-1}) \|^2 &\leq \frac{2L}{T(1 - 2\rho_2(U + K))} \mathbb{E}[F(\mathbf{w}_0) - F(\mathbf{w}^*)] \\ &+ \frac{2L}{T(1 - 2\rho_2(U + K))} \sum_{t=1}^T B_t, \end{aligned} \quad (3-20)$$

其中，

$$\begin{aligned} B_t &= \frac{\rho_1(U + K)}{LK} \sum_{i=1}^U K_i(1 - \beta_{i,t}) + 2 \sum_{i=1}^U \beta_{i,t} \frac{D - \kappa}{LD} G^2 \\ &+ \frac{2C^2}{L} \left( 1 + (1 + \delta) \frac{D - \kappa}{SD} G^2 + \frac{\sigma^2}{\left( \sum_{i=1}^U K_i \beta_{i,t} b_t \right)^2} \right), \end{aligned} \quad (3-21)$$

$\mathbf{w}^*$  是问题 P3.1 的可行解。

证明：参见附录 A.6。 ■

在定理 3.1 中，梯度范数的期望被用作收敛性的指标<sup>[88,114,115]</sup>。也就是说，如果满足以下条件，联邦学习算法将获得  $\tau$  水平的次优解：

$$\frac{1}{T} \sum_{t=1}^T \| \nabla F(\mathbf{w}_{t-1}) \|^2 \leq \tau. \quad (3-22)$$

上式 (3-22) 保证了算法收敛到一个固定点。如果目标函数  $F(\mathbf{w})$  是非凸的，则联邦学习可能会收敛到局部最小值或鞍点。

根据定理 3.1，可以得到

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \| \nabla F(\mathbf{w}_{t-1}) \|^2 &\leq \frac{2L}{T(1 - 2\rho_2(U + K))} \mathbb{E}[F(\mathbf{w}_0) - F(\mathbf{w}^*)] \\ &+ \frac{2L}{T(1 - 2\rho_2(U + K))} \sum_{t=1}^T B_t \xrightarrow{T \rightarrow \infty} \frac{2L}{T(1 - 2\rho_2(U + K))} \sum_{t=1}^T B_t. \end{aligned} \quad (3-23)$$

上式 (3-23) 给出了收敛时的误差底限。很显然，最小化这个误差底限可以提高联邦学习的收敛性能。利用这一封闭式的理论结果，本章接下来研究通信和计算的联合优化问题。

如定理 3.1 所示，参与的分布式设备数量的最大化不一定是实现最佳收敛性能的最优解。这是因为，参与的分布式设备数量越多，学习误差就越大。因此，有必要正确选择分布式设备，而不是简单地最大化参与的分布式设备的数量。

如果不考虑稀疏化，同时分布式设备传输的信号可以在参数服务器处完美恢复，则有  $\kappa = D$  成立，使得公式 (3-21) 的右侧第二项减小为 0。当更多的分布式设备参与联邦学习过程时，对于分布式设备  $i$ ，分布式设备选择向量中的更多元素变为  $\beta_{i,t} = 1$ ，这导致公式 (3-21) 中的  $B_t$  更小，进而公式 (3-23) 中的误差底限更低。因此，最大限度地增加参与联邦学习的分布式设备的数量可以降低误差底限来提高联邦学习性能。然而，参与的分布式设备的选择需要满足它们个人的最大功率约束。因此，需要适当地设定功率缩放因子  $b_t$ 。

从公式 (3-21) 和公式 (3-23) 可以看出，较大的  $b_t$  会降低 AWGN 对联邦学习的负面影响。但是， $b_t$  越大，可以参与联邦学习的分布式设备的数量就越少。因此，需要联合优化  $b_t$  和  $\beta_{i,t}$ 。

显然，稀疏比越大（即  $\kappa$  越大），收敛性能就越好。此外，误差底限随着压缩维度大小  $S$  的增加而降低。然而，通信效率随着  $\kappa$  和  $S$  的增加而降低。因此，在实际应用中需要在学习性能和通信效率之间取得平衡，然后根据实际需求来设置  $\kappa$  和  $S$  这两个变量的值。

在分布式数据集为非独立同分布的情况下，假设 3.3 中的因子  $\rho_1$  会变得更大。如公式 (3-23) 所示，较大的  $\rho_1$  会导致较大的误差底限，从而导致联邦学习的性能变差。

### 3.3 通信与学习联合优化方法

在 3.3 节的收敛性分析中，误差底限被指出是左右联邦学习性能的关键性因素。根据定理 3.1，本节制定一个联合优化问题，以最小化公式 (3-23) 中基于 1 比特压缩感知技术的 FLOA 的误差底限。在解决此优化问题时，本章首先通过离散规划来获得最优解，然后针对大规模无线网络应用场景提出基于 ADMM 的可扩展的次优解方案。

#### 3.3.1 联合优化问题建模

在实施基于 1 比特压缩感知技术的 FLOA 时，公式 (3-23) 中的误差底限在迭代中累积，导致  $F(\mathbf{w}_{t-1})$  和  $F(\mathbf{w}^*)$  之间的性能差距。因此，本章节设计一个实时策略来在每次迭代中最小化这个差距，这相当于在公式 (3-11) 中的最大发射功率限制的约束下迭代地最小化  $B_t$ 。

在每次迭代  $t$ ，参数服务器旨在确定功率缩放因子  $b_t$  和分布式设备调度指示因子  $\beta_t = [\beta_{1,t}, \beta_{2,t}, \dots, \beta_{U,t}]$  以最小化  $B_t$ ，对于给定的与 1 比特压缩感知相关的因子（即  $C$ ， $\kappa$  和  $S$ ）。丢弃不相关的项，最小化  $B_t$  等价于最小化

$$\begin{aligned}
R_t = & \frac{\rho_1(U+K) \sum_{i=1}^U K_i(1-\beta_{i,t})}{K} + 2C^2 \left( \sum_{i=1}^U K_i \beta_{i,t} b_t \right)^{-2} \sigma^2 \\
& + 2 \sum_{i=1}^U \beta_{i,t} \frac{D-\kappa}{D} G^2.
\end{aligned} \tag{3-24}$$

因此，联合优化问题被表述为

$$\begin{aligned}
\mathbf{P3.2:} \quad & \min_{b_t, \beta_t} R_t \\
\text{s.t.} \quad & \frac{\beta_{i,t}^2 K_i^2 b_t^2}{h_{i,t}^2} \leq P_i^{\max}, \quad \beta_{i,t} \in \{0, 1\}, i = 1, 2, \dots, U.
\end{aligned} \tag{3-25}$$

为了在参数服务器上实现解决上述问题的优化算法，在联邦学习实施之前，本地分布式设备应将本地样本的数量和每个分布式设备的最大功率发送到参数服务器。这些数字是固定的，在迭代过程中不会改变，故而只需要一次通信过程。

### 3.3.2 基于遍历算法的最优解

显然，优化问题**P3.2**是一个混合整数规划 (MIP)问题，是非凸的，并且由于实值功率缩放因子 $b_t$ 和二进制调度指示因子 $\beta_t$ 的耦合而使得它的求解具有挑战性。注意到，一旦给出 $\beta_t$ ，优化问题**P3.2**就简化为一个凸优化问题，其中最优功率缩放因子 $b_t$ 可以使用现成的优化算法来求解，例如，内点法<sup>[116]</sup>。因此，一种直接的方法是枚举 $\beta_t$ 的所有 $2^U$ 个可能性并最终输出产生最低目标函数值的可能性。这种基于枚举的方法在下面**算法 3.1** 中进行了总结。

表 3.1 基于枚举的最优解算法

Table 3.1 Optimal solution via the enumeration-based method

---

#### 算法 3-1 通过枚举法的最优解

---

**输入:**

系统参数 $\{P_i^{\max}, h_{i,t}, K_i\}_{i=1}^U, \Phi, G, \kappa$ 。

**输出:**

最优解 $\{b_t^*, \beta_t^*\}$ 。

1. **Repeat**

2. 从 $\beta_t$ 的可能性中选择一种；
  3. 给定 $\beta_t$ ，解决凸优化问题**P3.2**得到可行解 $b_t$ ；
  4. 在新的解 $\{b_t, \beta_t\}$ 下，如果目标函数值低于当前的最优值，则更新最优值 $\{b_t^*, \beta_t^*\}$ ；
  5. **Until** 遍历所有可能的 $\beta_t$
  6. **Return**  $\{b_t^*, \beta_t^*\}$
-

基于枚举的方法适用于少数分布式设备参与的场景，例如， $U \leq 10$ 。然而，随着 $U$ 的增加，它很快在计算上变得不可行。

### 3.3.3 基于 ADMM 算法的次优解

上一小节中提出的基于枚举的方法实现起来很简单，因为计算只涉及基本的函数评估。然而，具有大量分布式设备参与下的大规模网络使其容易受到高计算复杂性的影响。为了解决这个问题，本小节提出一种基于 ADMM 的算法来联合优化分布式设备选择和功率控制。正如后文将展示的，所提出的基于 ADMM 的方法具有关于 $U$ 的线性扩展的计算复杂度。

本节的主要思想是将组合优化问题**P3.2**分解为 $U$ 个并行的更小的整数规划问题。尽管如此，由于耦合变量 $\{b_t, \beta_t\}$ 和公式（3-25）中的约束，传统的分解技术，例如对偶分解，不能直接应用于**P3.2**中。为了消除这些耦合因素，本章节首先引入一个辅助向量 $\mathbf{r}_t = [r_{1,t}, r_{2,t}, \dots, r_{U,t}]$ ，并将两个辅助函数定义为

$$Q_1(\mathbf{r}_t) = 2C^2 \left( \sum_{i=1}^U K_i r_{i,t} \right)^{-2} \sigma^2, \quad (3-26)$$

和

$$Q_2(\boldsymbol{\beta}_t) = \frac{\rho_1(U+K) \sum_{i=1}^U K_i (1 - \beta_{i,t})}{K} + 2 \sum_{i=1}^U \beta_{i,t} \frac{D - \kappa}{D} G^2. \quad (3-27)$$

然后本章节引入另一个辅助向量 $\mathbf{q}_t = [q_{1,t}, q_{2,t}, \dots, q_{U,t}]$ ，并重新公式化优化问题**P3.2**为如下优化问题**P3.3**。

$$\begin{aligned} \mathbf{P3.3}: \min_{b_t, \{r_{i,t}, q_{i,t}, \beta_{i,t}\}_{i=1}^U} \quad & Q_1(\mathbf{r}_t) + Q_2(\boldsymbol{\beta}_t) + \sum_{i=1}^U Q_{3,i}(r_{i,t}) \\ \text{s.t.} \quad & r_{i,t} = \beta_{i,t} q_{i,t}, \quad i = 1, 2, \dots, U, \\ & q_{i,t} = b_t, \quad i = 1, 2, \dots, U, \\ & r_{i,t} > 0, b_t > 0, \quad i = 1, 2, \dots, U, \\ & \beta_{i,t} \in \{0, 1\}, \quad i = 1, 2, \dots, U, \end{aligned} \quad (3-28)$$

其中，

$$Q_{3,i}(r_{i,t}) = \begin{cases} 0, & r_{i,t} \in \left\{ r_{i,t} \mid \left| \frac{K_i r_{i,t}}{h_{i,t}} \right|^2 \leq P_i^{\max} \right\}, \\ \infty, & \text{otherwise.} \end{cases} \quad (3-29)$$

这里，公式（3-28）中的限定条件 $r_{i,t} = \beta_{i,t} q_{i,t}$ 和 $q_{i,t} = b_t$ ,  $i = 1, 2, \dots, U$ 是为了解耦 $\beta_{i,t}$ 和 $b_t$ ，从而使得优化问题**P3.2**与优化问题**P3.3**等价。

通过引入乘子,  $\xi_{i,t} \geq 0$  和  $\varsigma_{i,t} \geq 0$  到公式 (3-28) 中的限定条件里, 优化问题 P3.3 的部分增广拉格朗日可以被写成

$$\begin{aligned} & \mathcal{L}(b_t, \beta_t, \mathbf{r}_t, \mathbf{q}_t, \boldsymbol{\xi}_t, \boldsymbol{\varsigma}_t) \\ &= Q_1(\mathbf{r}_t) + Q_2(\beta_t) + \sum_{i=1}^U Q_{3,i}(r_{i,t}) \\ &+ \sum_{i=1}^U \xi_{i,t}(r_{i,t} - \beta_{i,t} q_{i,t}) + \frac{c}{2} \sum_{i=1}^U (r_{i,t} - \beta_{i,t} q_{i,t})^2 \\ &+ \sum_{i=1}^U \varsigma_{i,t}(q_{i,t} - b_t) + \frac{c}{2} \sum_{i=1}^U (q_{i,t} - b_t)^2, \end{aligned} \quad (3-30)$$

其中,  $\boldsymbol{\xi}_t = [\xi_{1,t}, \xi_{2,t}, \dots, \xi_{U,t}]$ ,  $\boldsymbol{\varsigma}_t = [\varsigma_{1,t}, \varsigma_{2,t}, \dots, \varsigma_{U,t}]$  以及  $c > 0$  为固定步长。对应的对偶问题可以被表述为

$$\begin{aligned} \text{P3.4: } & \max_{\{\xi_{i,t}, \varsigma_{i,t}\}_{i=1}^U} \mathcal{M}(\boldsymbol{\xi}_t, \boldsymbol{\varsigma}_t) \\ \text{s.t. } & \xi_{i,t} \geq 0, \varsigma_{i,t} \geq 0, i = 1, 2, \dots, U, \end{aligned} \quad (3-31)$$

其中,  $\mathcal{M}(\boldsymbol{\xi}_t, \boldsymbol{\varsigma}_t)$  为对偶函数, 表示如下

$$\begin{aligned} \mathcal{M}(\boldsymbol{\xi}_t, \boldsymbol{\varsigma}_t) = & \min_{b_t, \{r_{i,t}, q_{i,t}, \beta_{i,t}\}_{i=1}^U} \mathcal{L}(b_t, \mathbf{r}_t, \mathbf{q}_t, \boldsymbol{\beta}_t) \\ \text{s.t. } & r_{i,t} > 0, b_t > 0, q_{i,t} > 0, \beta_{i,t} \in \{0, 1\}, i = 1, 2, \dots, U. \end{aligned} \quad (3-32)$$

ADMM 技术<sup>[117]</sup>通过迭代更新  $\{\mathbf{r}_t, b_t\}$ ,  $\{\mathbf{q}_t, \boldsymbol{\beta}_t\}$  和  $\{\boldsymbol{\xi}_t, \boldsymbol{\varsigma}_t\}$  来解决对偶问题 P3.4。本章将第  $l$  次迭代的值表示为  $\{\mathbf{r}_t^{(l)}, b_t^{(l)}\}$ ,  $\{\mathbf{q}_t^{(l)}, \boldsymbol{\beta}_t^{(l)}\}$  和  $\{\boldsymbol{\xi}_t^{(l)}, \boldsymbol{\varsigma}_t^{(l)}\}$ 。然后, 上述变量在第  $(l+1)$  次迭代中按如下顺序更新:

- 1) 步骤一: 给定  $\{\mathbf{q}_t^{(l)}, \boldsymbol{\beta}_t^{(l)}\}$  和  $\{\boldsymbol{\xi}_t^{(l)}, \boldsymbol{\varsigma}_t^{(l)}\}$ , 本章首先最小化关于  $\{\mathbf{r}_t, b_t\}$  的目标函数  $\mathcal{L}$ , 其中,

$$\{\mathbf{r}_t^{(l+1)}, b_t^{(l+1)}\} = \arg \min_{\mathbf{r}_t, b_t} \mathcal{L}(\mathbf{r}_t, b_t). \quad (3-33)$$

注意公式 (3-33) 是一个严格的凸问题, 可以很容易地求解以获得最优解。另一方面, 该优化问题也可以分解为  $U+2$  并行的凸的子问题。具体而言, 让  $Q_1(r_t) = 2C^2(r_t)^{-2}\sigma^2$  代替公式 (3-26) 中的  $Q_1$ , 并且  $r_t = \sum_{i=1}^U K_i r_{i,t}$  为一个额外的约束。然后公式 (3-33) 可以通过求解  $U+2$  个并行的子问题来求解 (由于有  $U+2$  个优化变量, 即  $r_t, r_{1,t}, \dots, r_{U,t}, b_t$ )。由于求解这些  $U+2$  子问题的复杂度不随  $U$  而变化, 因此步骤一的总体计算复杂性为  $\mathcal{O}(U)$ 。

- 2) 步骤二: 给定  $\{\mathbf{r}_t^{(l+1)}, b_t^{(l+1)}\}$  和  $\{\boldsymbol{\xi}_t^{(l)}, \boldsymbol{\varsigma}_t^{(l)}\}$ , 然后最小化关于  $\{\mathbf{q}_t, \boldsymbol{\beta}_t\}$  的目标函数  $\mathcal{L}$ , 其中,

$$\{\mathbf{q}_t^{\{l+1\}}, \boldsymbol{\beta}_t^{\{l+1\}}\} = \arg \min_{\mathbf{q}_t, \boldsymbol{\beta}_t} \mathcal{L}(\mathbf{q}_t, \boldsymbol{\beta}_t). \quad (3-34)$$

上述优化问题可以分解为  $U$  个并行的子优化问题。在每个子优化问题（例如，第  $i$  个子优化问题）中，分别考虑  $\beta_{i,t} = 0$  和  $\beta_{i,t} = 1$ ，第  $i$  个子问题可以表示为

$$\{q_{i,t}\}^{\{l+1\}} = \begin{cases} \arg \min_{q_{i,t}} \mathcal{L}(q_{i,t}, 0), & \beta_{i,t} = 0, \\ \arg \min_{q_{i,t}} \mathcal{L}(q_{i,t}, 1), & \beta_{i,t} = 1, \end{cases} \quad (3-35)$$

其中，

$$\begin{aligned} \mathcal{L}(q_{i,t}, 0) = & \frac{\rho_1(U+K)K_i}{K} + \{\xi_{i,t}\}^{\{l\}} \{r_{i,t}\}^{\{l+1\}} \\ & + \frac{c}{2} \left( \{r_{i,t}\}^{\{l+1\}} \right)^2 + \varsigma_{i,t} \left( q_{i,t} - \{b_t\}^{\{l+1\}} \right) \\ & + \frac{c}{2} \left( q_{i,t} - \{b_t\}^{\{l+1\}} \right)^2, \end{aligned} \quad (3-36)$$

和

$$\begin{aligned} \mathcal{L}(q_{i,t}, 1) = & (1+\delta) \frac{D-\kappa}{D} G^2 + \frac{c}{2} \left( q_{i,t} - \{b_t\}^{\{l+1\}} \right)^2 \\ & + \{\xi_{i,t}\}^{\{l\}} \left( \{r_{i,t}\}^{\{l+1\}} - q_{i,t} \right) \\ & + \varsigma_{i,t} \left( q_{i,t} - \{b_t\}^{\{l+1\}} \right) \\ & + \frac{c}{2} \left( \{r_{i,t}\}^{\{l+1\}} - q_{i,t} \right)^2. \end{aligned} \quad (3-37)$$

对于  $\beta_{i,t} = 0$  和  $\beta_{i,t} = 1$ ，公式 (3-35) 解决的是一个严格凸的问题，很容易获得最优解。因此，可以简单地在  $\beta_{i,t} = 0$  和  $\beta_{i,t} = 1$  之间进行选择，从而确定在公式 (3-35) 中产生较小的目标函数值的  $\beta_{i,t}$  作为  $\{\beta_{i,t}\}^{\{l+1\}}$ ，以及  $\{q_{i,t}\}^{\{l+1\}}$  对应的最优解。在求解了  $U$  个并行的子优化问题后，公式 (3-34) 的最优解为  $\{\mathbf{q}_t^{\{l+1\}}, \boldsymbol{\beta}_t^{\{l+1\}}\}$ 。请注意，求解公式 (3-34) 中每个子问题的复杂度与  $U$  成比例，因此步骤二的整体计算复杂度为  $\mathcal{O}(U)$ 。

- 3) **步骤三：**最后，给定  $\{\mathbf{r}_t^{\{l+1\}}, b_t^{\{l+1\}}\}$  和  $\{\mathbf{q}_t^{\{l+1\}}, \boldsymbol{\beta}_t^{\{l+1\}}\}$ ，最大化与  $\{\xi_t, \varsigma_t\}$  相关的目标函数  $\mathcal{L}$ 。该优化通过如下更新乘子来实现

$$\{\xi_{i,t}\}^{\{l+1\}} = \{\xi_{i,t}\}^{\{l\}} + c \left( \{r_{i,t}\}^{\{l+1\}} - \{\beta_{i,t}\}^{\{l+1\}} \{q_{i,t}\}^{\{l+1\}} \right), \quad i = 1, \dots, U, \quad (3-38)$$

$$\{\varsigma_{i,t}\}^{\{l+1\}} = \{\varsigma_{i,t}\}^{\{l\}} + c \left( \{q_{i,t}\}^{\{l+1\}} - \{b_t\}^{\{l+1\}} \right), \quad i = 1, \dots, U. \quad (3-39)$$

显然，步骤三的计算复杂度也是  $\mathcal{O}(U)$ 。

本章所提出的 ADMM 方法迭代地执行上述步骤一到步骤三直到满足指定的停止条件。通常，停止标准由两个阈值指定<sup>[117]</sup>：绝对容差（例如，

$\sum_{i=1}^U |\{q_{i,t}\}^{\{l+1\}} - \{b_t\}^{\{l+1\}}|$  和相对容差（例如， $|\{b_t\}^{\{l+1\}} - \{b_t\}^{\{l\}}|$ ）。基于 ADMM 的方法求解优化问题 P3.3 的伪代码总结在 **算法 3.2** 中。

表 3.2 基于 ADMM 的次优解算法

Table 3.2 ADMM-based suboptimal solution

**算法 3.2** 基于 ADMM 的次优解**输入：**系统参数  $\{P_i^{\text{Max}}, h_{i,t}, K_i\}_{i=1}^U, \Phi, G, \kappa$ 。**输出：**最优解  $\{b_t^*, \beta_t^*\}$ 。**1. Repeat**

2. 通过求解公式 (3-33) 更新  $\{\mathbf{r}_t^{\{l+1\}}, b_t^{\{l+1\}}\}$ ;
3. 通过求解公式 (3-34) 更新  $\{\mathbf{q}_t^{\{l+1\}}, \beta_t^{\{l+1\}}\}$ ;
4. 通过求解公式 (3-38) 和公式 (3-39) 更新  $\{\xi_t^{\{l+1\}}, \varsigma_t^{\{l+1\}}\}$ ;
5. **Until** 满足收敛阈值或达到最大迭代次数
6. **Return**  $\{b_t^*, \beta_t^*\}$

本章所提出的 **算法 3.2** 的收敛性可以保证，因为对偶问题 P3.4 是凸优化问题。它的收敛对步长  $c$  不敏感<sup>[118]</sup>。由于非凸问题的潜在对偶差距，**算法 3.2** 可能不会完全收敛到优化问题 P3.3 的原始最优解。因此，对偶问题 P3.4 通过 **算法 3.2** 得到最优解  $\{b_t^*, \beta_t^*\}$  事实上是原始优化问题 P3.3 的近似解。

在本章节中，复杂度的计算是关于分布式设备的数量的，即复杂度如何随分布式设备的数量增加而变化。因此，本章节推导出 ADMM 迭代（包括 3 个步骤）的计算复杂度为  $\mathcal{O}(U)$ ，因为这三个步骤的最高复杂度为  $\mathcal{O}(U)$ 。这种复杂度  $\mathcal{O}(U)$  在于  $U$  线性缩放，即具有线性复杂度，而不是在基于枚举的方法中的指数复杂度  $\mathcal{O}(2^U)$ 。

### 3.4 SGD 算法下的拓展方案研究

本章将重点介绍上文所提理论分析和性能优化在 SGD 情形下的收敛行为和算法设计原理，即将上文描述方法应用在使用小批量 SGD 算法更新联邦学习模型的情形下。在这里，在应用具有恒定小批量大小为  $K_b$  的小批量 SGD 的算法下，本章提供基于 1 比特压缩感知的 FLOA 的预期收敛速度，而结果通过设置  $K_b = 1$  可直接应用于标准的 SGD 算法。

### 3.4.1 收敛性分析

在应用小批量 SGD 算法的情况下，基于 1 比特压缩感知的 FLOA 的收敛行为总结为如下定理 3.2 所示。

**定理 3.2** 给定功率缩放因子  $b_t$ ，分布式设备选择向量  $\beta_{i,t}$  和学习速率  $\alpha = \frac{1}{L}$ ，在恒定小批量大小为  $K_b$  的小批量 SGD 的算法下，基于 1 比特压缩感知的 FLOA 在  $T$  次迭代中有如下收敛性能

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \| \nabla F(\mathbf{w}_{t-1}) \|^2 &\leq \frac{2LK^2}{T(K^2 - 2\rho_2(\vartheta + 2UK^2))} \mathbb{E}[F(\mathbf{w}_0) - F(\mathbf{w}^*)] \\ &+ \frac{2LK^2}{T(K^2 - 2\rho_2(\vartheta + 2UK^2))} \sum_{t=1}^T B_t^{\text{sgd}}, \end{aligned} \quad (3-40)$$

其中，

$$B_t^{\text{sgd}} = \frac{\rho_1 \left( \vartheta + 2UK^2 \left( \sum_{i=1}^U \beta_{i,t} \right)^{-1} \right)}{LK^2} + \frac{2}{L} \left( \frac{C^2}{S} \varepsilon_t + \sum_{i=1}^U \beta_{i,t} \frac{D - \kappa}{D} G^2 \right), \quad (3-41)$$

且有  $\vartheta = 2U^2K_b^2 + K - 4KUK_b - UK_b$ 。

证明：参见附录 A.7 ■

### 3.4.2 性能优化

根据定理 3.2，可以最小化  $B_t^{\text{sgd}}$  以减少每次迭代时由于 1 比特压缩感知和模拟聚合传输而引入的聚合误差导致的 FLOA 与最优学习性能之间的差距，这相当于最小化

$$R_t^{\text{sgd}} = U \rho_1 \left( \sum_{i=1}^U \beta_{i,t} \right)^{-1} + C^2 \left( \sum_{i=1}^U K_b \beta_{i,t} b_t \right)^{-2} \sigma^2 + \sum_{i=1}^U \beta_{i,t} \frac{D - \kappa}{D} G^2. \quad (3-42)$$

对于与 1 比特压缩感知相关的因子（即  $C$ ,  $\kappa$  和  $S$ ）的给定值，SGD 情况下的联合优化问题因此被表述为

$$\begin{aligned} \mathbf{P3.5:} \quad &\min_{b_t, \beta_t} R_t^{\text{sgd}} \\ \text{s.t.} \quad &\frac{\beta_{i,t}^2 K_b^2 b_t^2}{h_{i,t}^2} \leq P_i^{\text{Max}}, \quad \beta_{i,t} \in \{0, 1\}, i = 1, 2, \dots, U. \end{aligned} \quad (3-43)$$

为了解决上述优化问题 P3.5，依然还可以应用章节 3.4 中描述的基于枚举的方法和基于 ADMM 的方法。对于基于 ADMM 的方法，定义的两个辅助函数分别重新表示为

$$Q_1^{\text{sgd}}(\mathbf{r}_t) = U \rho_1 \left( \sum_{i=1}^U r_{i,t} \right)^{-1} + C^2 (\sum_{i=1}^U K_i r_{i,t})^{-2} \sigma^2, \quad (3-44)$$

$$Q_2^{\text{sgd}}(\boldsymbol{\beta}_t) = \sum_{i=1}^U \beta_{i,t} \frac{D - \kappa}{D} G^2. \quad (3-45)$$

进而，其余的求解过程可以很容易以类似于第 3.4.3 节中的全批量 GD 情况的方式进行设计，本章在此省略。

从公式 (3-42) 中的  $R_t^{\text{sgd}}$  可以看出批量大小  $K_b$  越大， $R_t^{\text{sgd}}$  越小，即基于 1 比特压缩感知的 FLOA 具有更好的学习表现。

## 3.5 性能仿真与分析

本节对本章所提出的基于 1 比特压缩感知的 FLOA 方案 (OBCSAA) 用于图像分类任务的性能进行评估。仿真实验均在 MNIST 数据集上完成。

### 3.5.1 系统设定

在本小节的仿真实验中，除非另有说明，否则系统参数设置如下。本仿真实验设定无线网络下的联邦学习系统有  $U = 10$  个分布式设备，并且将它们的最大峰值功率设置为  $P_i^{\text{Max}} = 10 + i \text{ mW}$ ,  $i \in [1, U]$ 。分布式设备和参数服务器之间的无线信道建模为独立同分布的瑞利衰落，通过从不同的  $i$  和  $t$  下的正态分布  $\mathcal{N}(0, 1)$  生成相应的  $h_{i,t}$ 。不失一般性，AWGN 在参数服务器处的方差设置为  $\sigma^2 = 10^{-4} \text{ mW}$ 。仿真实验执行  $\text{top-}\kappa = 1000$  的稀疏化，稀疏的局部梯度  $\mathcal{C}(\mathbf{g}_i)$  的压缩维度设置为  $S = 2000$ 。测量矩阵  $\Phi$  的元素由正态分布  $\mathcal{N}(0, 1/S)$  随机生成。此外，文献[103]中的 BIHT 算法被选作为参数服务器处信号的重构算法。

本章所使用的 MNIST 数据集考虑手写数字识别的学习任务，该数据集由 10 个类别组成，范围从数字“0”到“9”。在 MNIST 数据集中，共有 60000 个标记的训练数据样本和 10000 个测试样本可用于训练学习模型。在本章的实验中，训练了一个具有 784 个神经元输入层、一个具有 64 个神经元的隐藏层和一个具有 10 个神经元的以 softmax 为输出函数的输出层的 MLP 神经网络。仿真实验采用交叉熵作为损失函数，采用 ReLU 作为激活函数。MLP 神经网络中的参数总数为  $D = 50890$ 。学习速率  $\alpha$  设置为 0.1。实验随机为每个分布式设备选择 3000 个不同的训练样本，并将它们作为不同的本地数据集分发给所有分布式设备，即对于任何  $i \in [1, U]$ ，

$$K_i = \bar{K} = 3000.$$

为了进行比较，本章在仿真对比实验中使用了一个基准方案，其局部梯度更新的传输始终可靠且无错误以实现完美聚合，即忽略无线信道的影响。这个基准方案是一个理想情况，被命名为完美聚合方案。此外，本节将本章所提出的 OBCSAA 方案与文献[62]中现有的基于模拟聚合的方案（被命名为 OBDA）进行了比较，该方案采用了 signSGD<sup>[119]</sup>的思想。对于性能评估，本章提供了不同参数设置下训练损失和测试准确度在不同通信轮次下的结果，如下文所示。

### 3.5.2 仿真结果与分析

在图 3.2 中，本章节首先通过评估 MLP 的训练损失和测试精度来探索不同稀疏化算子对所提出的 OBCSAA 的影响。为了满足 RIP 条件， $S$  设置为 10000。可以观察到，在一定程度的稀疏性的情况下，例如， $\kappa = 1000$ ，其中稀疏率是  $1000/50890$ ，本章所提出的 OBCSAA 可以提供联邦学习任务所需的性能（接近 OBDA 和完美聚合）。随着  $\kappa$  的增加，当所有联邦学习算法收敛时，训练损失减少，测试准确率增加。这是因为  $\kappa$  越大，每轮通信丢失的梯度更新信息越少。

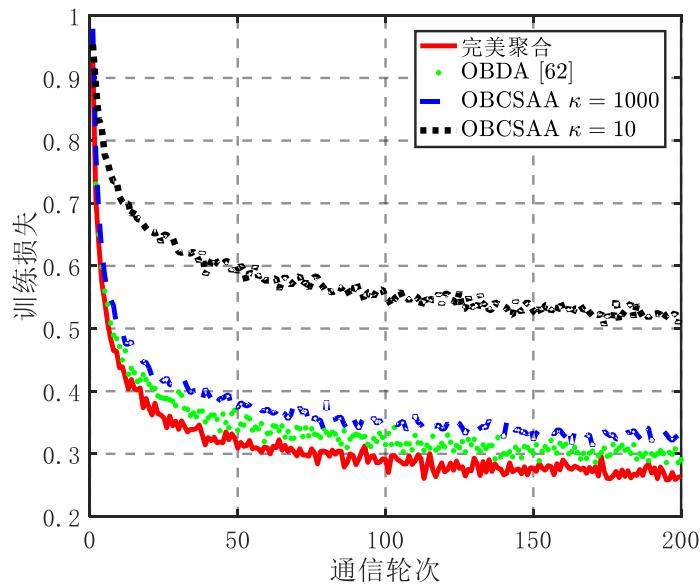


图 3.2 (a) 不同稀疏化算子下的训练损失

Figure 3.2(a) Training loss under different sparsification operators

图 3.3 展示了在  $\kappa = 1000$  的情况下减小压缩维度大小  $S$  对本章所提出的 OBCSAA 性能的影响，其中学习性能随着  $S$  的增加而提高。当  $S$  足够大时，性能几乎不会再提高。这是因为  $S$  越大，越有利于信号重构。当  $S$  足够大时，重构算法的

性能达到最优。事实上， $S$ 越大，需要的通信资源就越多。因此，在联邦学习性能和通信效率之间存在折衷。与采用数字通信的传统无压缩联邦学习相比，本章所提出的 OBCSAA 在  $S = 5000$  和  $\kappa = 1000$  情况下仅占用一个通道和  $\frac{5000}{50890}$  传输时间，而性能仅比 OBDA 和完美聚合方案低 5% 至 10%。这些结果表明，本章所提出的 OBCSAA 在适当的压缩参数下可以大大降低通信开销和传输延迟，同时可以确保可观的联邦学习性能。

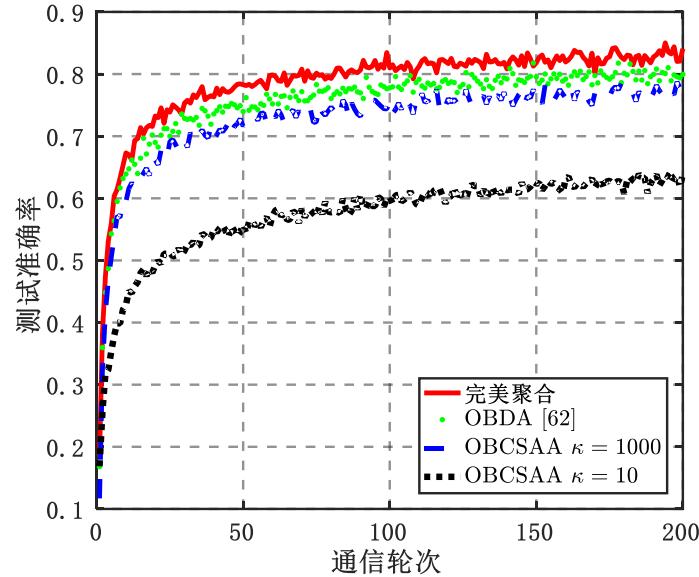


图 3.2 (b) 不同稀疏化算子下的测试准确率

Figure 3.2(b) Test accuracy under different sparsification operators

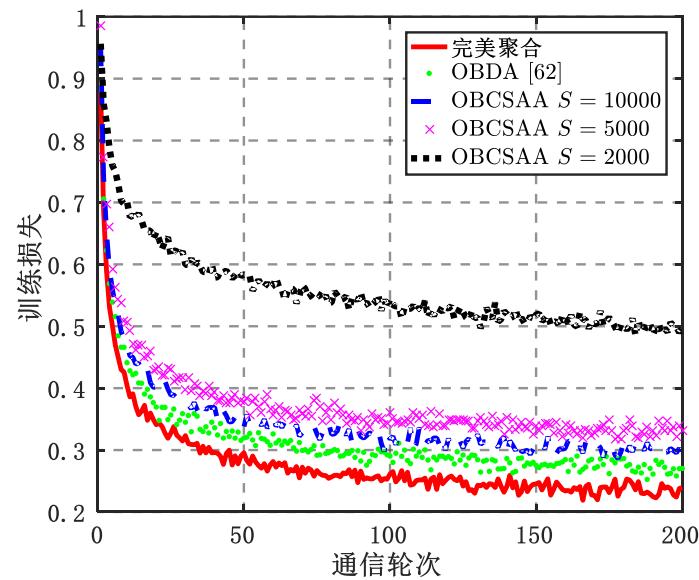


图 3.3 (a) 不同压缩后的维度大小下的训练损失

Figure 3.3(a) Training loss under different reduced dimension size  $S$

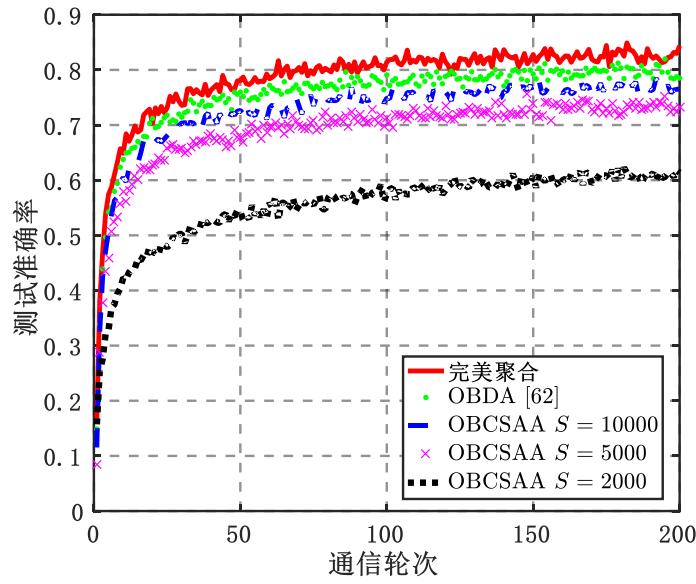


图 3.3 (b) 不同稀疏化算子下的测试准确率

Figure 3.3 (b) Test accuracy under different reduced dimension size  $S$ 

在不同分布式设备总数 $U$ 下, 所提出的 OBCSAA 在基于枚举的方法和 ADMM 的方法下的性能在图 3.4 中进行了比较, 其中基于枚举的方法与基于 ADMM 的方法相比具有更好的性能。该结果准确地证明了本章所提出的联合优化方案的有效性, 其可以减轻聚合误差对 FLOA 的影响。此外, 从图 3.4 中也可以看到, 当分布式设备总数 $U$ 更大时, 性能更高。这是因为分布式设备数量的增加导致可用于联邦学习算法的数据量增加, 并且可以选择更多具有高信道增益的分布式设备。

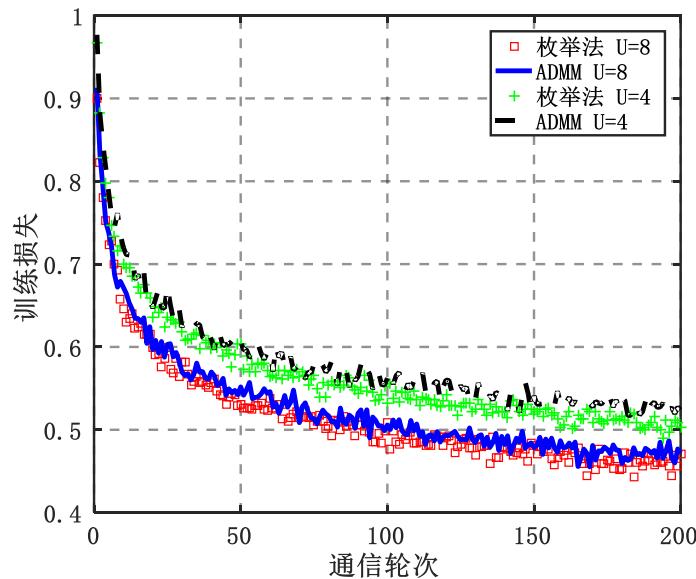


图 3.4 (a) 不同联合优化算法在不同设备总数下的训练损失

Figure 3.4 (a) Training loss under different joint optimization solving methods with different total number of the distributed devices

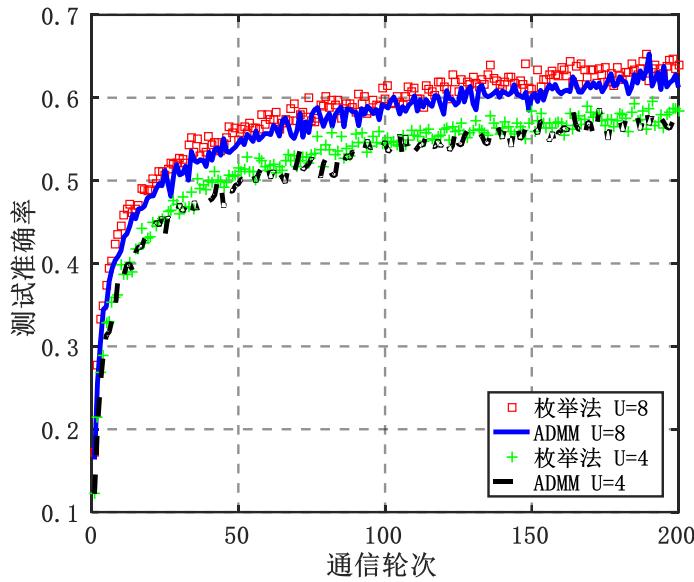


图 3.4 (b) 不同联合优化算法在不同设备总数下的测试准确率

Figure 3.4 (b) Test accuracy under different joint optimization solving methods with different total number of the distributed devices

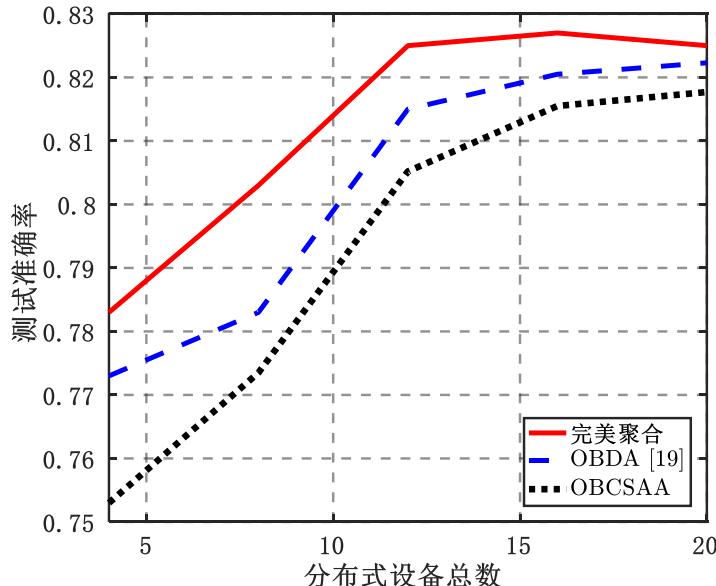


图 3.5 不同联邦学习方案在不同设备总数下的测试准确率

Figure 3.5 Test accuracy under different schemes for federated learning with different total number of the distributed devices

在图 3.5 中，本章评估所提出的 OBCSAA 在实施 SGD 算法的情况下与对比方案相比的测试准确性。图 3.5 提供了测试准确性随分布式设备总数  $U$  变化的结果，其中实验批量大小被设置为  $K_b = 64$ 。从图 3.5 中可以看出，随着分布式设备总数的增加，所有方案的性能都在提高。但是随着分布式设备的总数  $U$  继续增加，所有方

案的性能的改进最终都会缩小并趋于平缓。这是因为当 $U$ 超过某个水平时，数据样本就足以进行准确的训练。

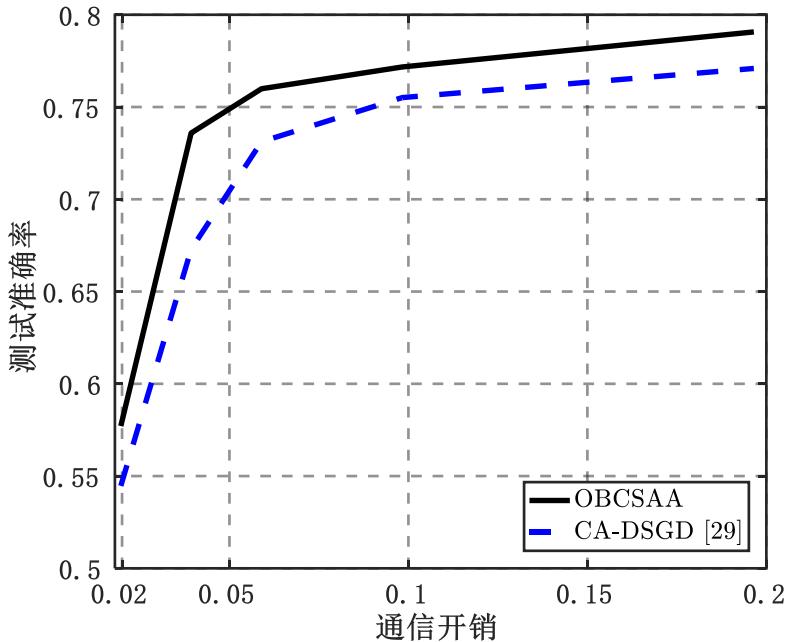


图 3.6 学习准确率和通信开销之间的权衡

Figure 3.6 The tradeoff between learning accuracy and communication cost

在图 3.6 中，本章节进一步比较了所提出的 OBCSAA 与现有模拟压缩方案(称为 CA-DSGD<sup>[29]</sup>)的通信效率和学习准确率，其反映了通信开销和学习准确率之间的权衡。通信开销被定义为  $\frac{S}{D}$ ，其中  $S$  越大表明通信开销越大。根据图 3.6，为了获得所需的学习准确率，本章所提出的 OBCSAA 消耗的通信开销低于 CA-DSGD。换言之，给定相同的通信开销，本章所提出的 OBCSAA 可以实现比 CA-DSGD 更高的学习准确率。由于功率控制和设备选择的联合优化，本章所提出的 OBCSAA 在学习和通信之间实现了比 CA-DSGD 更好的权衡，由于 CA-DSGD 仅采用固定功率分配且没有设备选择。

### 3.6 本章小结

本章研究了一种基于 1 比特压缩感知技术和模拟聚合传输的高效通信联邦学习方案（OBCSAA）。首先本章为 OBCSAA 方案在给定分布式设备和功率控制的情况下预期收敛速率推导了一个封闭形式的表达式。该理论结果揭示了由于稀疏化、降维压缩、量化、信号重构和信道噪声引起的聚合误差，收敛性能和通信效

率之间的权衡。在这一发现的指导下，本章构建了一个通信和学习的联合优化问题，以减轻聚合误差对联邦学习的影响，从而实现最佳的分布式设备选择和功率控制。为了解决这个具有挑战性的非凸优化问题，本章提出了一种基于枚举的方法和一种基于 ADMM 的方法，分别可以获得小规模网络的最优解和大规模网络的次优解。最后，通过仿真实验结果表明，本章所提出的 OBCSAA 方案可以极大地提高通信效率，同时确保所需的学习性能。本章研究内容发表在 IEEE ICC2021 会议论文集中，以及 IEEE TWC 期刊上。



## 4 基于 FLOA 传算联合优化框架的鲁棒性方案研究

联邦学习利用分布式数据的同时保护了数据的隐私，为边缘智能提供了一个新范式<sup>[122,123]</sup>。FLOA 凭借其通信高效的模型更新聚合，吸引了多个领域越来越多的研究兴趣，例如功率控制<sup>[59,60,100,101]</sup>，设备调度<sup>[31,61,101]</sup>，梯度压缩<sup>[28,29,62,124,125]</sup>，波束赋形设计<sup>[30,32,63]</sup>和学习速率优化<sup>[64]</sup>。这些研究从通信、优化和机器学习等多个角度推进了 FLOA 的发展。

除了比传统联邦学习具有更高的通信效率外，FLOA 还增强了数据隐私，这要归功于它固有的对单个局部梯度的不可访问性，从而防止了潜在的模型反转攻击，例如梯度的深度泄漏<sup>[65]</sup>。虽然 FLOA 关闭了梯度深度泄漏的大门，但它也为攻击者提供了执行拜占庭攻击的机会。事实上，即使是单一的拜占庭式故障也可能摧毁联邦学习。现存文献针对传统联邦学习中的拜占庭鲁棒性进行了大量研究<sup>[66-69]</sup>，其中大部分使用了基于筛选的方法，例如利用几何中值<sup>[70-73]</sup>，坐标中位数<sup>[67]</sup>，坐标修剪均值<sup>[73]</sup>，Krum/Multi-Krum<sup>[74]</sup>，Bulyan<sup>[75,76]</sup>，Zeno/Zeno++<sup>[77,78]</sup>等等<sup>[66]</sup>。这些筛选方法的基本思想是在聚合局部梯度时排除异常值。它们的应用均取决于局部梯度的各个值的知识，但是由于 FLOA 中所有局部梯度在空中模拟叠加，分布式设备的个体更新值无法访问。因此，为传统联邦学习设计的现有拜占庭鲁棒方法不能应用于 FLOA 中，这促使本文设计一种为 FLOA 定制的新拜占庭鲁棒性方法。

据了解，除本文外目前还没有关于 FLOA 中针对拜占庭攻击的研究文献，当然也没有任何针对 FLOA 的反攻击措施的设计。本章节旨在深入了解拜占庭攻击如何影响 FLOA，然后提供相应的防御策略。本章节针对 FLOA 的鲁棒性方案的研究的主要贡献有三方面，简要罗列如下：

- (1) 鉴于大多数现有的研究工作都在 FLOA 中采用信道反转（Channel Inversion, CI）功率控制（或其变体）<sup>[28,61,62,64,100,101,124-127]</sup>来传输本地更新，本章首先从理论上证明衰落信道下的 CI 方法可以达到接近理想无错误情况的性能，这可以解释为什么该策略被广泛用于克服 FLOA 中的传输错误。同时，本章的分析表明，CI 对拜占庭攻击的防御能力非常有限。因此，本章提出一种新的鲁棒传输策略来对抗拜占庭攻击，称为最大努力投票（Best Effort Voting, BEV）功率控制策略，其中分布式设备以最大功率传输其本地梯度更新。
- (2) 为了研究拜占庭攻击对 FLOA 的影响，本章推导拜占庭攻击者为了实现最强攻击应使用的传输策略，即如何伪造梯度和设计传输功率可以最大程度地阻止 FLOA 的收敛。由于这是最强的攻击，因此评估其在各种传输策略下对 FLOA 的影响是有意义的，这反过来又有助于阐明

这些策略各自的鲁棒性水平。

- (3) 为了证明本章所提出的 BEV 策略在最强攻击下与现有的 CI 策略相比的有效性，本章分别提供 BEV 策略和现有 CI 策略的收敛性分析。本章的理论结果证明，在最强的拜占庭攻击下，BEV 策略在收敛行为和学习准确性方面优于 CI 策略。

此外，本章还使用 MNIST 数据集在图像分类问题上测试所提出的方法。仿真结果表明，在没有拜占庭攻击的情况下，BEV 策略的学习性能略逊于 CI 策略，而 BEV 策略在抵抗拜占庭攻击的鲁棒性方面明显优于 CI 策略。因此，本章的理论分析和仿真实验结果表明，在实际应用中由于无法确知是否存在拜占庭攻击，BEV 策略整体上优于 CI 策略。

本章的其余部分内容安排如下。FLOA 的系统模型在 4.1 节中介绍，本章节提供两种功率控制策略，即 CI 和 BEV 策略。在 4.2 节，在这两种功率控制策略下的 FLOA 的预期收敛速率的封闭式表达式被推导出来，用于比较不同功率控制策略的性能，其中本章还推导出拜占庭攻击者的所能实现的最强攻击的具体实施措施。第 4.3 节中提供性能仿真分析结果来验证所提方案的性能。最后，第 4.4 节总结本章的研究内容并给出结论。

## 4.1 系统模型

### 4.1.1 联邦学习模型

如图 4.1 所示，本章考虑一个具有一个参数服务器和  $U$  个分布式设备的无线网络分布式计算模型。每个分布式设备存储有  $K$  个数据点<sup>[67-69]</sup>，这些数据点是从大型数据集  $\mathcal{D}$  中抽取的独立且同分布 (Independent and Identically Distributed, i.i.d.) 的样本<sup>9</sup>。令  $(\mathbf{x}_{i,k}, \mathbf{y}_{i,k})$  表示为第  $i$  个分布式设备的第  $k$  个样本数据。令  $f(\mathbf{w}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})$  表示与每个样本数据点  $(\mathbf{x}_{i,k}, \mathbf{y}_{i,k})$  相关的损失函数，其中  $\mathbf{w} = [w^1, \dots, w^D]$  表示维度大小为  $D$  的训练模型的参数。对应的总体损失函数记为  $F(\mathbf{w}) := \mathbb{E}_{\mathcal{D}}[f(\mathbf{w}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})]$ 。协作训练的目标是，参数服务器和分布式设备通过最小化协作学习模型的总体损失函数来获得最优的学习模型参数向量  $\mathbf{w}$ ，如下列优化问题所示

$$\text{P4.1: } \mathbf{w}^* = \arg \min_{\mathbf{w}} F(\mathbf{w}). \quad (4-1)$$

---

<sup>9</sup> 非独立同分布的 (non-i.i.d.) 数据集下的拜占庭鲁棒性问题涉及较多，留待以后研究。

全局目标函数  $F(\mathbf{w})$  的最小化通常通过随机梯度下降 (SGD) 算法进行。在参数服务器处，第  $t$  迭代中的模型参数  $\mathbf{w}_t$  更新规则如下

$$( \text{全局模型更新} ) \quad \mathbf{w}_t = \mathbf{w}_{t-1} - \alpha \frac{\sum_{i=1}^U \mathbf{g}_{i,t}}{U}, \quad (4-2)$$

其中， $\alpha$  是学习速率， $\mathbf{g}_{i,t} = \nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})$  表示在第  $i$  个分布式设备处使用其随机选择的样本数据（例如第  $k$  个样本数据）来计算的局部梯度。为了让参数服务器从分布式设备那里获取公式 (4-2) 中的局部梯度之和，需要制定一些通信和聚合方案。

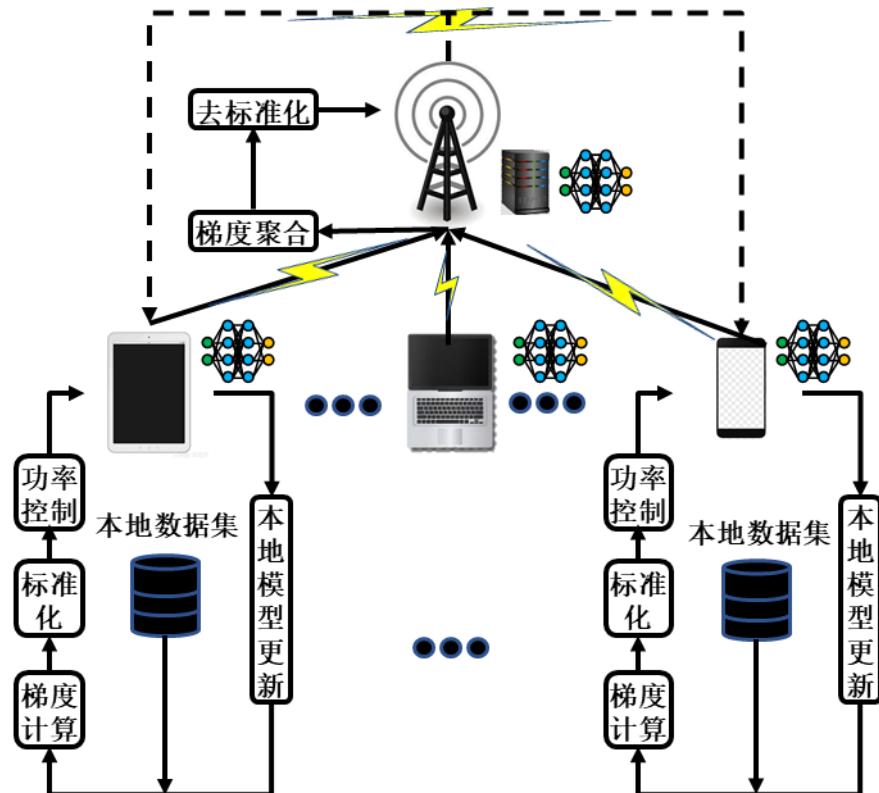


图 4.1 不同功率控制策略下的 FLOA 模型

Figure 4.1 A federated learning model based on analog aggregation communication under different power control strategies

假设  $U$  个分布式设备中有  $N$  个是拜占庭攻击者，剩余  $M = U - N$  分布式设备是正常的。但是，拜占庭攻击者不需要遵循给正常设备设定的协议，可以向参数服务器发送任意消息。更坏的情况是，这些攻击者可能对学习系统和算法有完整的了解，并且可以相互勾结。此外，参数服务器和分布式设备之间的通信不可避免地会引入信道噪声，而拜占庭攻击者也可以利用这个机会来破坏 FLOA。接下来，本章将展示不同的预定义模拟聚合传输协议导致存在拜占庭攻击时的 FLOA 方案的不同性能。

#### 4.1.2 模拟聚合传输过程

在 FLOA 中，为了利用空中计算技术进行低延迟梯度聚合，所有的分布式设备需要对分布式局部梯度进行幅度调制以进行模拟传输，并同时通过相同的时频信道资源传输到参数服务器。假设分布式设备之间通过同步信道实现符号级同步<sup>[61]</sup>。为了便于功率控制设计，传输的符号记为  $\tilde{\mathbf{g}}_{i,t} = [\tilde{g}_{i,t}^1, \dots, \tilde{g}_{i,t}^d, \dots, \tilde{g}_{i,t}^D]$ ，其是通过标准化的，梯度向量中的所有元素的均值为 0，方差为 1，即  $\mathbb{E}[(\tilde{g}_{i,t}^d)^2] = 1, \forall i, t$ 。通过这种方式，在参数服务器处可以设计功率控制策略，而无需了解特定的传输符号。请注意，标准化因子对于所有局部梯度都是一致的，因此可以在参数服务器处接收到来自分布式设备的信号后进行反转变换。

由于梯度的统计数据可能会随着迭代而改变，因此在所有通信轮次中都需要执行局部梯度的标准化过程。具体来说，在每一轮通信开始时，每个分布式设备需估计其本地学习到的梯度的均值和方差，分别表示为  $\bar{g}_{i,t} = \frac{1}{D} \sum_{d=1}^D g_{i,t}^d$  和  $\epsilon_{i,t}^2 = \frac{1}{D} \sum_{d=1}^D (g_{i,t}^d - \bar{g}_{i,t})^2$ 。然后将局部估计的均值和方差传输到参数服务器通过平均的形式进行全局梯度统计估计。给定接收到的  $\bar{g}_{i,t}$  和  $\epsilon_{i,t}^2$ ，参数服务器对所有局部估计进行平均，得到梯度均值和方差的全局估计分别表示为  $\bar{g}_t = \frac{1}{U} \sum_{i=1}^U \bar{g}_{i,t}$  和  $\epsilon_t^2 = \frac{1}{U} \sum_{i=1}^U \epsilon_{i,t}^2$ 。然后估计的全局标准化因子  $\bar{g}_t$  和  $\epsilon_t^2$  被广播回本地分布式设备并用于发射信号标准化。

在接收到全局标准化因子  $\bar{g}_t$  和  $\epsilon_t^2$  后，每个本地分布式设备执行发射信号标准化如下：

$$\tilde{\mathbf{g}}_{i,t} = \frac{\mathbf{g}_{i,t} - \bar{g}_t \mathbf{1}}{\epsilon_t}, \quad (4-3)$$

其中， $\mathbf{1}$  是一个维数等于  $\mathbf{g}_{i,t}$  的全一向量。

考虑到在每个通信轮次中仅传输两个符号 ( $\bar{g}_{i,t}$  和  $\epsilon_{i,t}^2$ )，参数服务器一一收集各个个体局部估计的均值和方差。本章假设这种标准化所需的通信是无噪声的，不会引入错误。请注意，拜占庭攻击者知道设计的标准化方法，他们会发送其局部梯度的真实均值和方差，以避免在标准化阶段暴露自己。否则，攻击者可能很容易被参数服务器检测到，然后被参数服务器过滤掉，因为普通分布式设备和拜占庭攻击者拥有 i.i.d. 数据集。

标准化后，所有本地分布式设备将其标准化的局部梯度  $\tilde{\mathbf{g}}_{i,t}$  以一定的发射功率  $p_{i,t}$  传输到参数服务器 ( $p_{i,t}$  的设定将在本节后面讨论)。每个本地分布式设备的传输受传输功率约束如下：

$$\begin{aligned}\mathbb{E}[\|p_{i,t}\tilde{\mathbf{g}}_{i,t}\|^2] &= \mathbb{E}\left[p_{i,t}^2 \sum_{d=1}^D (\tilde{g}_{i,t}^d)^2\right] = p_{i,t}^2 \sum_{d=1}^D \mathbb{E}[(\tilde{g}_{i,t}^d)^2] \\ &= Dp_{i,t}^2 \leq p_i^{\max}, \quad \forall i.\end{aligned}\tag{4-4}$$

因此，功率约束归结为  $p_{i,t}^2 \leq \frac{p_i^{\max}}{D}$ 。

另一方面，拜占庭攻击者可以将  $\hat{\mathbf{g}}_{n,t}$  的任何值作为他们的假的梯度更新报告给参数服务器，从而扭曲联邦学习过程。第  $n$  个拜占庭攻击者的发射功率  $\hat{p}_{n,t}$  也应满足其最大发射功率，即

$$\mathbb{E}[\|\hat{p}_{n,t}\hat{\mathbf{g}}_{n,t}\|^2] \leq p_n^{\max}, \quad \forall n.\tag{4-5}$$

考虑块衰落信道，其中无线信道在联邦学习中的每次迭代中保持不变，但可以独立地随迭代改变。本章将一次迭代的持续时间定义为一个时间块，以  $t$  为索引。在第  $t$  次迭代中，在参数服务器处接收到的信号由下式给出

$$\mathbf{y}_t = \sum_{m=1}^M p_{m,t} |h_{m,t}| \tilde{\mathbf{g}}_{m,t} + \sum_{n=1}^N \hat{p}_{n,t} |h_{n,t}| \hat{\mathbf{g}}_{n,t} + \mathbf{z}_t,\tag{4-6}$$

其中，上式 (4-6) 右侧第一项、第二项和第三项分别对应于普通分布式设备、拜占庭攻击者和信道噪声。具体地， $|h_{i,t}|$  是在第  $t$  次迭代中从第  $i$  个分布式设备到参数服务器的信道增益，而  $\mathbf{z}_t \sim \mathcal{N}(0, z^2 \mathbf{I})$  是独立于梯度更新的加性高斯白噪声 (AWGN)。信道遵循独立的瑞利衰落，即  $h_{i,t} \sim \mathcal{CN}(0, \sigma_i^2)$ 。在这项研究工作中，本章假设这些信道状态信息 (CSI) 在分布式设备和参数服务器处是完美已知的。有了完美的 CSI，新到的相位偏移在本地分布式设备传输其梯度更新之前得到补偿。

在从本地分布式设备接收到公式 (4-6) 中的信号  $\mathbf{y}_t$  后，参数服务器执行去标准化操作，通过公式 (4-3) 中反转的标准化来获得估计的聚合梯度如下：

$$\begin{aligned}\tilde{\mathbf{g}}_t &= \epsilon_t \mathbf{y}_t + \left( \sum_{i=1}^U p_{i,t} |h_{i,t}| \right) \bar{g}_t \mathbf{1} \\ &= \epsilon_t \left( \sum_{m=1}^M p_{m,t} |h_{m,t}| \tilde{\mathbf{g}}_{m,t} + \sum_{n=1}^N \hat{p}_{n,t} |h_{n,t}| \hat{\mathbf{g}}_{n,t} + \mathbf{z}_t \right) + \left( \sum_{i=1}^U p_{i,t} |h_{i,t}| \right) \bar{g}_t \mathbf{1} \\ &= \epsilon_t \left( \sum_{m=1}^M p_{m,t} |h_{m,t}| \frac{\mathbf{g}_{m,t} - \bar{g}_t \mathbf{1}}{\epsilon_t} + \sum_{n=1}^N \hat{p}_{n,t} |h_{n,t}| \hat{\mathbf{g}}_{n,t} + \mathbf{z}_t \right) \\ &\quad + \left( \sum_{i=1}^U p_{i,t} |h_{i,t}| \right) \bar{g}_t \mathbf{1} \\ &= \sum_{m=1}^M p_{m,t} |h_{m,t}| \mathbf{g}_{m,t} + \epsilon_t \sum_{n=1}^N \hat{p}_{n,t} |h_{n,t}| \hat{\mathbf{g}}_{n,t} \\ &\quad + \left( \sum_{n=1}^N p_{n,t} |h_{n,t}| \right) \bar{g}_t \mathbf{1} + \epsilon_t \mathbf{z}_t,\end{aligned}\tag{4-7}$$

其中，上式（4-7）中右侧第一项对应于来自正常分布式设备的聚合梯度，第二项加上第三项表示拜占庭攻击者对梯度更新的恶意贡献，最后一项则来自信道加性高斯白噪声。

通过使用估计的聚合梯度，全局模型参数在第 $t$ 次迭代时更新如下

$$\mathbf{w}_t = \mathbf{w}_{t-1} - \alpha \tilde{\mathbf{g}}_t. \quad (4-8)$$

接下来，本章讨论普通分布式设备可以采用的两种用于 $p_{i,t}$ 设计的发射功率分配方案：现有的信道反转（CI）传输策略<sup>[61,62]</sup>和本章提出的最大努力投票（BEV）传输策略。

#### 4.2.2.1 信道反转传输方案

给定完美已知的 CSI，在 CI 方案<sup>[61,62]</sup>中，通过信道反转功率控制策略，以便以相同的幅度接收不同本地分布式设备传输的梯度参数，这可以使得分布式梯度参数在参数服务器处进行幅度对齐之后聚合。第 $i$ 个分布式设备的传输功率表示为  $p_{i,t}^2 = \frac{b_t^2}{|h_{i,t}|^2}, \forall i$ ，其中， $b_t^2 = \min\{\frac{P_i^{\max}}{D} |h_{i,t}|^2, i = 1, 2, \dots, U\}$  是用于满足公式（4-4）中的功率约束的功率缩放比例因子。

很明显， $b_t$  满足

$$\mathbb{E}[b_t^2] \geq P_0^{\max} \mathbb{E}[\min\{|h_{i,t}|^2, i = 1, 2, \dots, U\}], \quad (4-9)$$

其 中  $P_0^{\max} = \min\{\frac{P_i^{\max}}{D}, i = 1, 2, \dots, U\}$ 。因此可以设置  $b_t^2 = P_0^{\max} \mathbb{E}[\min\{|h_{i,t}|^2, i = 1, 2, \dots, U\}]$  用于功率分配。由于信道系数为瑞利分布  $h_{i,t} \sim \mathcal{CN}(0, \sigma_i^2)$ ，故而， $|h_{i,t}|^2$  服从均值为  $\frac{1}{\lambda_i} = 2\sigma_i^2$  指数分布。因此，有  $\mathbb{E}[\min\{|h_{i,t}|^2, i = 1, 2, \dots, U\}] = \frac{1}{\sum_{i=1}^U \lambda_i} \doteq \lambda$ 。

于是，为了在实践中实现信道反转功率控制方案，第 $i$ 个分布式设备的发射功率设置为

$$p_{i,t} = \frac{b_0}{|h_{i,t}|}, \quad \forall i, \quad (4-10)$$

其中， $b_0$  被设置为  $b_0^2 \doteq b_t^2 = P_0^{\max} \lambda$ 。

#### 4.2.2.2 最大努力投票传输方案

为了对抗智能拜占庭攻击者，本章的核心思想是让正常的本地分布式设备尽最大努力对抗潜在拜占庭攻击的影响，使 FLOA 收敛到正确的方向，因此被命名

为最大努力投票 (BEV) 方案。在 BEV 方案中，普通本地分布式设备通过使用与他们的 CSI 无关的最大传输功率来传输他们的本地梯度更新。在 BEV 方案中，第  $i$  个分布式设备的发射功率由下式给出

$$p_{i,t} = \sqrt{\frac{p_i^{\max}}{D}}, \quad \forall i. \quad (4-11)$$

本章将在下面讨论上述两种不同的功率分配方案对拜占庭攻击者具有的不同鲁棒性。

## 4.2 收敛性对比与分析

本节比较上述两种功率分配方案 CI 和 BEV 的收敛性能。首先证明存在一种最强攻击使得拜占庭攻击者可以最大限度地阻止 FLOA 的收敛。然后在这种情况下，本节分别推导出两种传输方案应用时的 FLOA 收敛速率。

### 4.2.1 基本假设

为了便于收敛分析，本章节对损失函数和局部梯度估计做了几个标准假设。请注意，本章节的理论推导并未假设损失函数具有凸性。因此，本章节的方法适用于广泛使用的深度神经网络（Deep Neural Network，DNN）或卷积神经网络（Convolutional Neural Network，CNN）学习模型。

**假设 4.1** （损失函数 Lipschitz 连续平滑） 损失函数  $F$  是 Lipschitz 连续平滑的，即

$$F(\mathbf{w}_t) \leq F(\mathbf{w}_{t-1}) + \mathbf{g}_t^T (\mathbf{w}_t - \mathbf{w}_{t-1}) + \frac{L}{2} \|\mathbf{w}_t - \mathbf{w}_{t-1}\|^2, \quad (4-12)$$

其中， $L$  是一个正常数，称为函数  $F(\cdot)$  的 Lipschitz 常数<sup>[128]</sup>。

**假设 4.2** （局部梯度无偏估计和误差有界） 随机局部梯度估计是具有方差界限的全局梯度的独立且无偏估计<sup>[62,119]</sup>，即

$$\mathbb{E}(\mathbf{g}_{i,t}) = \mathbf{g}_t, \quad \forall i, t, \quad (4-13)$$

$$\mathbb{E}(\|\mathbf{g}_{i,t} - \mathbf{g}_t\|^2) \leq \delta^2, \quad \forall i, t, \quad (4-14)$$

其中，本章在这项工作中考虑标准 SGD，即批量大小为 1。如果应用批量大小为  $K_b$  的批量 SGD，则方差以  $\frac{\delta^2}{K_b}$  为界。

**假设 4.3** （标准化系数的无偏性和有界性） 标准化因子  $\bar{g}_t$  和  $\epsilon_t^2$  是全局梯度的无偏估计，其中方差有界如下<sup>[61]</sup>

$$\mathbb{E}[\bar{g}_t] = \frac{\sum_{d=1}^D g_t^d}{D}, \quad \forall t, \quad (4-15)$$

$$\epsilon_t \leq \epsilon, \quad \forall t. \quad (4-16)$$

### 4.2.2 最强拜占庭攻击

虽然拜占庭攻击者可以发送任意信号，但存在一种拜占庭攻击者可以实现的最强攻击，以阻止 FLOA 的收敛。直观地说，拜占庭攻击者希望沿与正常本地分布式设备梯度更新方向相反的方向影响参数服务器处的全局聚合梯度。为此，拜占庭攻击者会以最大发射功率  $\hat{p}_{n,t}$  向参数服务器发送  $\hat{\mathbf{g}}_{n,t} = -\mathbf{g}_{n,t}$ 。具体而言，给定全局模型参数  $\mathbf{w}_{t-1}$ ，拜占庭攻击者使用自己的本地数据计算自己的梯度  $\mathbf{g}_{n,t}$ 。此外，发射功率  $\hat{p}_{n,t}$  满足最大功率约束，即  $\mathbb{E}[\|\hat{p}_{n,t}\hat{\mathbf{g}}_{n,t}\|^2] = p_n^{\max}$ 。这是 FLOA 可能遇到的最坏情况。本文在以下定理 4.1 中理论上证明，这是拜占庭攻击者可以施加的最强攻击，以阻止 FLOA 的收敛。

**定理 4.1** 在拜占庭攻击者存在的情况下，采用 SGD 的 FLOA 系统，最强的攻击可以被执行为

$$\hat{\mathbf{g}}_{n,t} = -\mathbf{g}_{n,t}, \quad (4-17)$$

$$\hat{p}_{n,t} = \sqrt{\frac{p_n^{\max}}{(\bar{g}_t^2 + \epsilon_t^2)D}}. \quad (4-18)$$

**证明：**参见附录 A8。 ■

由于上述最强攻击已被证明是 FLOA 可以经历的最坏情况，接下来本章节将通过收敛性分析来评估不同传输方案的防御效率。

本章节采用众所周知的策略，将梯度范数与预期改进联系起来，以显示非凸优化的收敛性<sup>[62,88,119]</sup>，即

$$\min_{0,1,\dots,T} \mathbb{E}[\|\mathbf{g}_t\|^2] \leq \mathbb{E} \left[ \sum_{t=1}^T \frac{1}{T} \|\mathbf{g}_t\|^2 \right] \leq \mathcal{O}\left(\frac{1}{T^q}\right), \quad (4-19)$$

其中， $q > 0$  是迭代总数  $T$  的阶数。正如所看到的，如果公式 (4-19) 成立，那么随着  $T$  增加到无穷大，梯度的范数期望会收敛到 0，这意味着联邦学习算法渐近收敛。收敛速度取决于阶值  $q$ ，是接下来要评估的关键参数。

### 4.2.3 信道反转方案下的收敛性分析

在每个本地分布式的 CSI 可获得的情况下，CI 功率控制可以按照公式 (4-10) 中执行。于是，CI 传输方案在最强攻击下的最终收敛速率推导如下面定理 4.2 所示。

**定理 4.2** 对于具有基于 SGD 的模型更新、针对普通分布式设备的采用基于 CI 的功率控制以及  $N$  个拜占庭攻击者采取如公式 (4-17) 和公式 (4-18) 所示的最强攻击的 FLOA 系统，学习算法的收敛速率由下式给出

$$\begin{aligned} \mathbb{E}\left[\sum_{t=1}^T \frac{1}{T} \|\mathbf{g}_t\|^2\right] &\leq \frac{1}{\sqrt{T}} \left( \frac{2L\Omega_{CI}}{\omega_{CI}^2 \bar{\alpha}} (F(\mathbf{w}_0) - F(\mathbf{w}^*)) \right. \\ &\quad \left. + \bar{\alpha} \left( \delta^2 + \frac{1}{\Omega_{CI}} \epsilon^2 z^2 \right) \right), \end{aligned} \quad (4-20)$$

其中，

$$\omega_{CI} = Mb_0 - \sum_{n=1}^N \sqrt{\frac{\pi \sigma_n^2 p_n^{\max}}{2D}}, \quad (4-21)$$

$$\Omega_{CI} = (U + N) \left( Ub_0^2 + \sum_{n=1}^N \frac{2\sigma_n^2 p_n^{\max}}{D} \right), \quad (4-22)$$

且有  $\bar{\alpha} = \frac{L\Omega_{CI}\sqrt{T}}{\omega_{CI}^2}$  为一正常数，满足  $\bar{\alpha} < 2\sqrt{T}$ ，同时  $b_0$  按照公式 (4-10) 初始化。如果  $\frac{\alpha^2 L}{2} \Omega_{CI} - \alpha \omega_{CI} < 0$  成立，则算法的收敛性可以保证，这对参数  $\alpha, L, b_0, \sigma_n, p_n^{\max}, M, N, D$  施加了约束。

**证明：** 参见附录 A9。 ■

对于较小的学习速率，上述定理 4.2 的渐近收敛速率由  $O(\frac{\Omega_{CI}}{\omega_{CI}^2 \sqrt{T}})$  主导。另外，收敛条件由  $\frac{\alpha^2 L}{2} \Omega_{CI} - \alpha \omega_{CI} < 0$  给出，其证明也在附录 A9 中提供。这个收敛条件以  $\alpha < \frac{2\omega_{CI}}{L\Omega_{CI}}$  的形式对学习速率施加了一个上限。此外，当学习速率设置得足够小时， $\alpha^2$  接近于 0，联邦学习算法在  $\omega_{CI} > 0$  的简化条件下收敛。从这个收敛条件可以看出，即使是一个拜占庭攻击者也可以破坏 FLOA 收敛性。因为如果这个攻击者的发射功率非常大或者它的信道增益非常大，例如，如果  $p_n^{\max}$  或  $\sigma_n^2$  对于任何  $n$  的都很大，就很难保证  $\omega_{CI} > 0$ 。

对于所有本地分布式设备具有相同最大功率（即  $p_i^{\max} = p^{\max}, \forall i$ ）和独立且同分布的信道（即  $\sigma_i = \sigma, \forall i$ ）的情况，有收敛条件  $\omega_{CI} = (\frac{M}{\sqrt{U}} - \sqrt{\frac{N^2 p^{\max}}{4}}) \sqrt{\frac{2p^{\max} \sigma^2}{D}} > 0$ 。因此，本章节得出结论，在这种特殊情况下，攻击者的数量在不超过  $\frac{U}{1 + \sqrt{\pi U}}$ ，才可以使 CI 方案能够抵御拜占庭攻击。

当没有拜占庭攻击者时，即  $N = 0$  时，本章节有以下引理 4.1 来说明此情况下的联邦学习算法的收敛性。

**引理 4.1** 针对采用基于 SGD 的模型更新的 FLOA 系统，普通本地分布式设备采用 CI 功率控制策略，在没有拜占庭攻击者的情况下，学习算法的收敛速率由下式给出

$$\begin{aligned} \mathbb{E}\left[\sum_{t=1}^T \frac{1}{T} \|\mathbf{g}_t\|^2\right] &\leq \frac{1}{\sqrt{T}} \left( \frac{2L}{\bar{\alpha}} (F(\mathbf{w}_0) - F(\mathbf{w}^*)) \right. \\ &\quad \left. + \bar{\alpha} \left( \delta^2 + \frac{1}{U^2 b_0^2} \epsilon^2 z^2 \right) \right), \end{aligned} \quad (4-23)$$

其中,  $\alpha = \frac{1}{LUb_0\sqrt{T}}\bar{\alpha}$ 。

**证明:** 当  $N = 0$  时, 有  $\omega_{CI}^2 = \Omega_{CI}$ 。然后设置  $\alpha = \frac{\omega_{CI}}{L\Omega_{CI}\sqrt{T}}\bar{\alpha} = \frac{1}{LUb_0\sqrt{T}}\bar{\alpha}$ , 将  $\alpha$ 、 $\omega_{CI}$  和  $\Omega_{CI}$  代入公式 (4-20), 则引理 4.1 得证。 ■

从公式 (4-23) 可以看出, 在没有拜占庭攻击者的 CI 功率控制的情况下, 得到最快的渐近收敛速率为  $O(\frac{1}{\sqrt{T}})$ , 这与不考虑无线信道和噪声的影响的无差错 (Error-free, EF) 下的联邦学习情况相同。

#### 4.2.4 最大努力投票方案下的收敛性分析

对于在最强攻击下的本章所提出的 BEV 传输方案, 得到的收敛速度推导如下面定理 4.3 所示。

**定理 4.3** 对于具有基于 SGD 的模型更新、针对普通分布式设备的采用基于 BEV 的功率控制以及  $N$  个拜占庭攻击者采取如公式 (4-17) 和公式 (4-18) 所示的最强攻击的 FLOA 系统, 学习算法的收敛速率由下式给出

$$\begin{aligned} \mathbb{E}\left[\sum_{t=1}^T \frac{1}{T} \|\mathbf{g}_t\|^2\right] &\leq \frac{1}{\sqrt{T}} \left( \frac{2L\Omega_{BEV}}{\bar{\alpha}\omega_{BEV}^2} (F(\mathbf{w}_0) - F(\mathbf{w}^*)) \right. \\ &\quad \left. + \bar{\alpha} \left( \delta^2 + \frac{1}{\Omega_{BEV}} \epsilon^2 z^2 \right) \right), \end{aligned} \quad (4-24)$$

其中,

$$\omega_{BEV} = \sum_{i=1}^M \sqrt{\frac{p_i^{\max} \pi}{2D}} \sigma_i - \sum_{n=1}^N \sqrt{\frac{p_n^{\max} \pi}{2D}} \sigma_n, \quad (4-25)$$

$$\Omega_{BEV} = (U + N) \sum_{i=1}^U \frac{2\sigma_i^2 p_i^{\max}}{D}, \quad (4-26)$$

且有  $\bar{\alpha} = \frac{L\Omega_{BEV}\sqrt{T}}{\omega_{BEV}}$  为正常数, 满足  $\bar{\alpha} < 2\sqrt{T}$ 。如果  $\frac{\alpha^2 L}{2} \Omega_{BEV} - \alpha \omega_{BEV} < 0$  成立, 则算法的收敛性可以保证, 这对参数  $\alpha, L, \sigma_i, p_i^{\max}, M, N, D$  施加了约束。

**证明:** 参见附录 A10。 ■

收敛条件  $\frac{\alpha^2 L}{2} \Omega_{BEV} - \alpha \omega_{BEV} < 0$  的证明见附录 A10。这个条件以  $\alpha < \frac{2\omega_{BEV}}{L\Omega_{BEV}}$  的形式对学习速率施加了一个上限。此外, 当学习速率设置得足够小时,  $\alpha^2$  接近于 0, 联邦学习在  $\omega_{BEV} > 0$  的简化条件下收敛。如果所有攻击者和普通分布式设备都是

同构的（即它们有相同的最大发射功率以及独立同分布的信道），本章所提出的 BEV 可以在  $N \leq \frac{U}{2}$  时防御拜占庭攻击。由于  $\frac{U}{2} \geq \frac{U}{1+\sqrt{\pi}U}$ ，本章所提出的 BEV 方案可以防御比 CI 更多的拜占庭攻击者。

对于较小的学习速率，如果 CI 方案和本章所提出的 BEV 方案都能收敛，则渐近收敛速率由  $O(\frac{\Omega}{\omega^2\sqrt{T}})$  主导。两方案中的主导项  $O(\frac{\Omega_{CI}}{\omega_{CI}^2\sqrt{T}})$  和  $O(\frac{\Omega_{BEV}}{\omega_{BEV}^22\sqrt{T}})$  取决于具体参数。对于较大的学习速率，如果 CI 方案和 BEV 方案都能收敛，则渐近收敛速度由  $O(\frac{1}{\Omega\sqrt{T}})$  支配。由于  $\Omega_{BEV} > \Omega_{CI}$ ，BEV 方案的收敛速度快于 CI 方案。

当没有拜占庭攻击者时，即  $N = 0$  时，有  $\omega_{BEV}^2 \leq \Omega_{BEV}$ 。考虑到较小的学习速率，本章所提出的 BEV 的渐近收敛速度以  $O(\frac{\Omega_{BEV}}{\omega_{BEV}^2\sqrt{T}})$  为主，比 CI 方案和以及不考虑信道噪声的完美聚合方案性能都差。

### 4.3 性能仿真与分析

为了评估本章所提出的 BEV 方案对拜占庭攻击的鲁棒性，本节提供其在图像分类任务中的仿真实验结果，同时也对比 CI 方案的性能。

#### 4.3.1 系统参数设定

除非另有说明，否则仿真实验参数设定如下。FLOA 系统中有  $U = 10$  个分布式设备。分布式设备和参数服务器之间的无线信道建模为独立同分布的瑞利衰落，即在不同的  $i$  和  $t$  下以复高斯分布  $\mathcal{CN}(0, 1)$  生成信道系数  $h_{i,t}$ 。本地分布式设备的平均接收 SNR 设定为  $\frac{P_i^{\max}}{Dz^2} = 10 \text{ dB}$ <sup>[62]</sup>。

本章仍使用 MNIST 数据集完成手写数字识别的学习任务，该数据集由 10 个类别组成，范围从数字“0”到“9”。在 MNIST 数据集中，共有 60000 个标记的训练数据样本和 10000 个测试样本。在本章的实验中，训练了一个具有 784 个神经元的输入层、一个 64 个神经元的隐藏层和一个 10 个神经元的 softmax 为输出函数的输出层的多层感知器 (MLP)。采用 ReLU 作为激活函数，交叉熵作为损失函数。MLP 中的参数总数为  $D = 50890$ 。本章为每个分布式设备随机选择 3000 个不同的训练样本作为他们的本地数据集，即对于任何  $i \in [1, U]$ ， $K_i = \bar{K} = 3000$ 。

在不同的攻击下本章评估所提出的 BEV 方案，包括：1) 没有任何攻击行为的情况，2) 只有一个远离参数服务器的攻击者，其被看作是弱攻击者，3) 只有一个靠近参数服务器的攻击者，其被看作是强攻击者，4) 随机选择几个攻击者。本章选择两个基准进行方案对比：1) CI 方案和 2) 理想无错误 (EF) 下的完美聚合情况的 FLOA，其不考虑无线信道和噪声的影响。

### 4.3.1 无攻击场景下性能仿真与分析

本小节将无错误情况设置为在参数服务器处完美聚合局部梯度的基准情况，即设置信道系数  $h_{i,t} = 1$  并设置 AWGN 为  $\mathbf{z}_t = 0$ 。在图 4.2 中，本小节比较了没有拜占庭攻击的 BEV，CI 和 EF 的性能。考虑到  $\alpha < \frac{\omega}{L\Omega}$ ，本小节设置学习速率  $\alpha$  为它的缩放版本  $\hat{\alpha} = \frac{\bar{\alpha}}{L\sqrt{T}} = \frac{\Omega}{\omega}\alpha = 0.1$ ，其中  $\hat{\alpha}$  表示  $\alpha$  的调节因子。正如从图 5.2 中看到的，CI 方案的性能几乎与 EF 相同。但是，与 CI 和 EF 相比，BEV 的性能损失约为 2%。该结果与本章在定理 4.3 中的理论分析一致，已在前文中进行了讨论。也就是说，当且仅当不存在拜占庭攻击者时，CI 的收敛速度比本章所提出的 BEV 方案快一点。然而，实际学习应用通常在可能的对抗环境中运行的，恶意攻击者的情况不得不考虑。

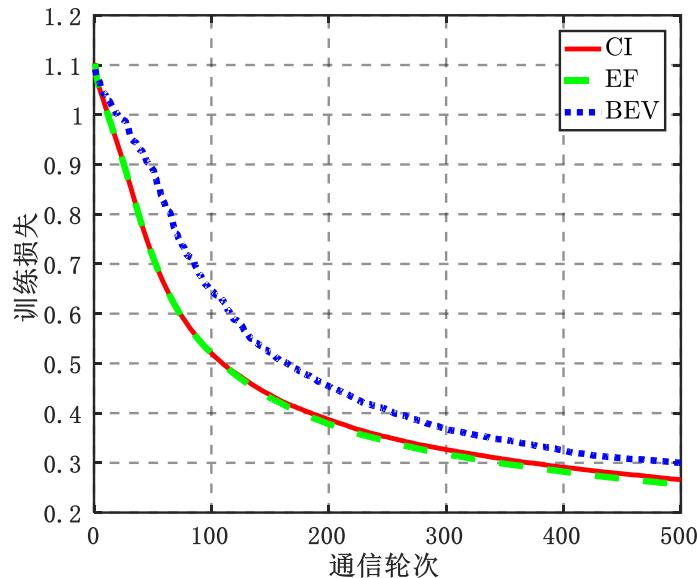


图 4.2 (a) 没有攻击者的情况下训练损失

Figure 4.2 (a) Training loss without attackers

### 4.3.2 存在单个信道增益较弱的攻击者的性能仿真与分析

在图 4.3 中，本小节比较了 BEV 方案和 CI 方案在单个拜占庭攻击者下的性能。假设攻击者在所有分布式设备中拥有最低的信道增益。它仍然采用最强的攻击策略来破坏 FLOA。由于对 FLOA 的拜占庭攻击相对较弱，如果选择适当的学习速率  $\hat{\alpha} = \frac{\bar{\alpha}}{L\sqrt{T}} = \frac{\Omega}{\omega}\alpha$ ，BEV 方案和 CI 方案都可以收敛。另一方面，当学习速率没有被正确选择时，例如，当  $\hat{\alpha} = 2$  时，在图 4.3 中，BEV 方案可以收敛但 CI 方案失

败。当 $\hat{\alpha} = 1$ 时，BEV 方案和 CI 方案都能收敛，但 BEV 方案的收敛速度比 CI 方案快。这是因为对于较大的学习率，渐近收敛率由 $O(\frac{1}{\Omega\sqrt{T}})$ 支配且 $\Omega_{BEV} > \Omega_{CI}$ 。当 $\hat{\alpha} = 0.1$ 时，BEV 方案的性能比 CI 方案稍微弱。当可以保证收敛时，大的学习速率可实现快速的收敛速度，因此在实践中大的学习速率更有优势。在较大的学习速率下，例如 $\hat{\alpha} = 1$ ，本章所提的 BEV 方案比 CI 方案效果更好。

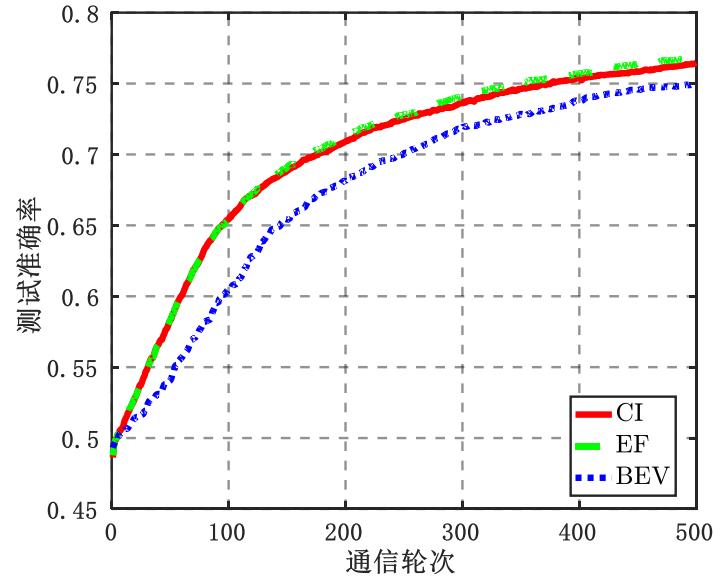


图 4.2 (b) 没有攻击者的情况下测试准确率

Figure 4.2(b) Test accuracy without attackers

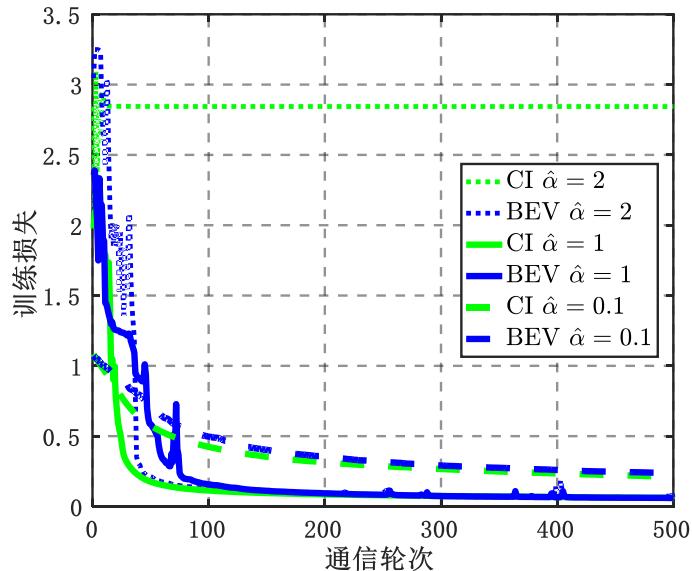


图 4.3 (a) 存在单个信道增益较弱的攻击者的训练损失

Figure 4.3 (a) Training loss under single Byzantine attacker whose channel gain is the lowest

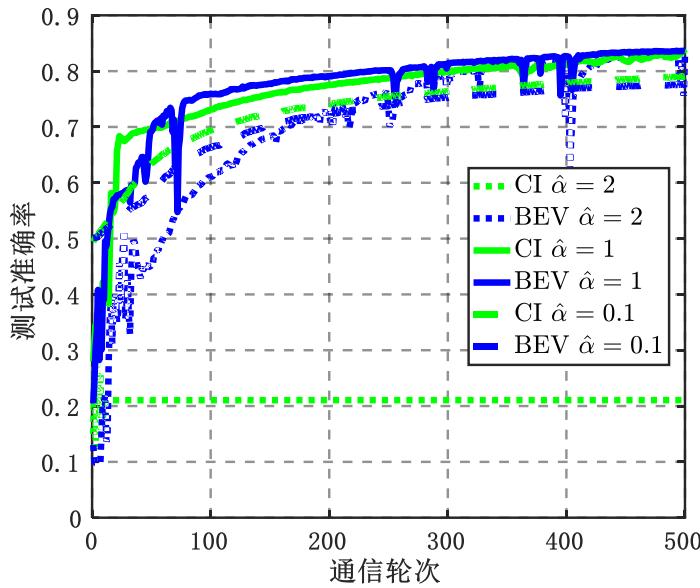


图 4.3 (b) 存在单个信道增益较弱的攻击者的测试准确率

Figure 4.3 (b) Test accuracy under single Byzantine attacker whose channel gain is the lowest

### 4.3.3 存在单个信道增益较强的攻击者的性能仿真与分析

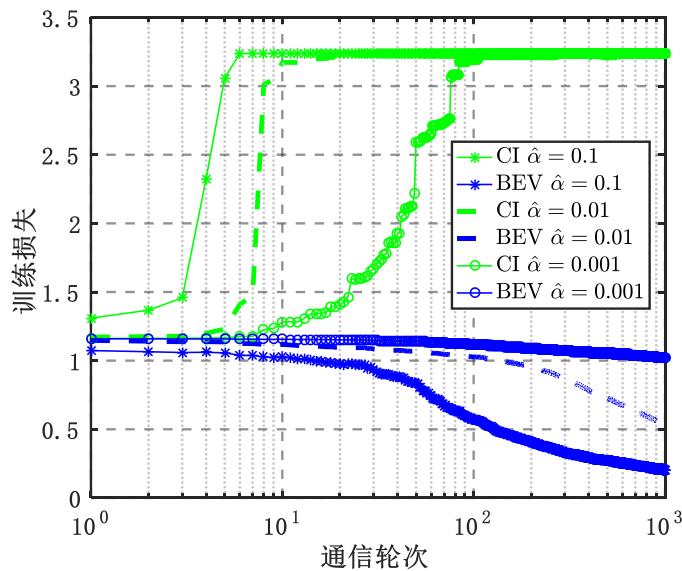


图 4.4 (a) 存在单个信道增益较强的攻击者的训练损失

Figure 4.4 (a) Training loss under single Byzantine attacker whose channel gain is the highest

在图 4.4 中，本小节比较了 BEV 方案与 CI 方案在单个拜占庭攻击者下的性能，该攻击者的信道增益在所有分布式设备中最高。因此，这可以被看作是强的攻击。在这种强攻击的情况下，本小节比较了 BEV 方案和 CI 方案的性能。由于收敛条件  $\omega_{CI} > 0$  很难保证，从图 4.4 可以看出，CI 方案不能收敛或收敛到失败的情况。

随着 $\hat{\alpha}$ 的减小，CI 方案在前几次迭代中可以向正确的方向收敛，但它仍然无法防御几次迭代后的攻击。另一方面，BEV 方案却仍然可以收敛。因此，在存在强攻击的情况下，BEV 方案是比 CI 方案更好的选择。此外，收敛速度随着 $\hat{\alpha}$ 的降低而降低。这意味着在保证收敛的情况下，应该选择更大的学习速率。

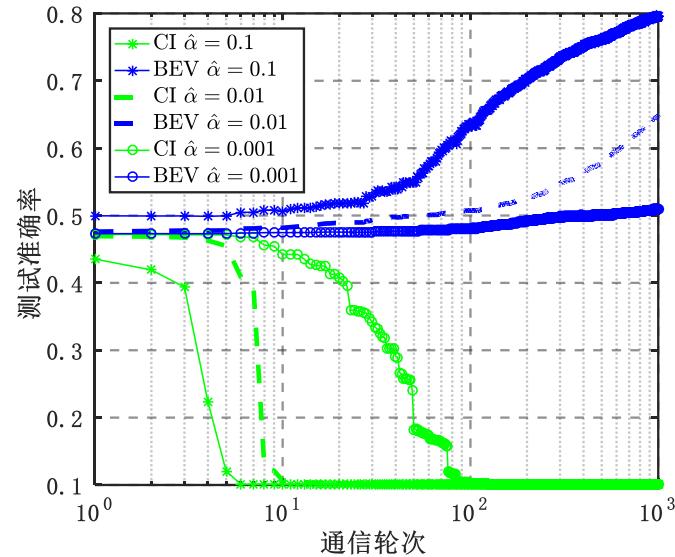


图 4.4 (b) 存在单个信道增益较强的攻击者的测试准确率

Figure 4.4 (b) Test accuracy under single Byzantine attacker whose channel gain is the highest

#### 4.3.4 存在多个随机选择的攻击者的性能仿真与分析

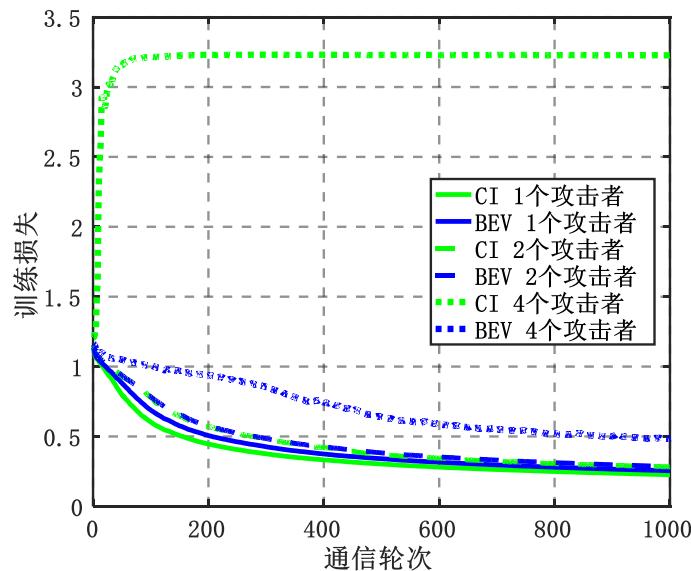


图 4.5 (a) 不同数量的攻击者下的训练损失

Figure 4.5 (a) Training loss under the different number of Byzantine attackers

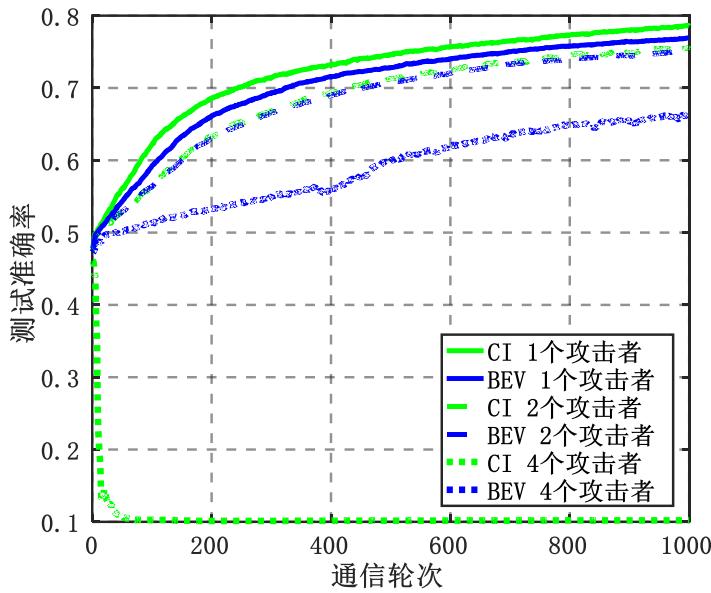


图 4.5 (b) 不同数量的攻击者下的测试准确率

Figure 4.5 (b) Test accuracy under the different number of Byzantine attackers

在图 4.5 中, 本小节比较了不同拜占庭攻击者数量下的 BEV 方案与 CI 方案的性能。可以看出, 当拜占庭攻击者数量小于 4 时, BEV 方案和 CI 方案都可以收敛, 但收敛速度随着拜占庭攻击者数量的增加而降低。当拜占庭攻击者数量为 4 时, 即  $N > \frac{U}{1+\sqrt{\pi}U}$ , CI 方案不能收敛到正确的方向, 而 BEV 方案仍然会收敛到正确的方向但以较慢的速度收敛。这些结果与前文理论结果一致。

#### 4.4 本章小结

本章研究了 FLOA 对拜占庭攻击的鲁棒性, 对不同传输方案下的学习算法的收敛性能进行了理论分析。本章的分析结果揭示了拜占庭攻击者可以施加的最强攻击。针对该攻击模式, 本章提出了一种新的传输功率控制方案, 即 BEV 方案。此外, 本章通过理论分析和仿真实验对比现存的 CI 方案和 BEV 方案在各种对抗环境下的收敛行为, 说明了所提 BEV 方案的有效性。具体而言, 在没有任何拜占庭攻击者的情况下, CI 方案的性能与理想的无错误情况下的联邦学习的性能相当, 而 BEV 方案的性能损失约为 2%。在弱的拜占庭攻击中, 对于较大的学习率, CI 方案和 BEV 方案都可以收敛, 而 BEV 方案的收敛速度比 CI 方案快。如果存在强拜占庭攻击者, 则无法保证 CI 方案的收敛, 但 BEV 方案仍然可以收敛。在实践中, 由于无法确定潜在攻击的强度, BEV 方案是对抗拜占庭攻击的更好选择, 因为它在各种攻击情况下都表现良好。本章研究内容发表在 IEEE ICC2021 会议论文集中, 以及 IEEE IoT-J 期刊上。

## 5 总结与展望

随着物联网的飞速发展，海量智能设备产生了大量的可用信息在网络边缘。这些设备和数据信息结合机器学习技术促进了人工智能在物联网中的发展，构建出以边缘智能为动力的智慧物联网，深层次地改变了人们的生活方式。作为边缘智能的关键技术，联邦学习以其可以充分利用分布式设备的数据和算力的同时又保护了数据隐私的优势，受到了广泛的关注。为了推动联邦学习的发展，使其在智慧物联网中更好地发挥作用，本文研究了联邦学习中的通信高效性、鲁棒性以及安全隐私保护问题，针对不同目标的理论优化模型进行了研究并提出了可行方法，实现了强鲁棒、高效通信、隐私增强的边缘智慧传算协同一体化解决方案，为智慧物联网的发展做出了从理论到具体实践的技术支撑。下面本章将首先对本文的研究内容得出的一些具体成果做出总结并突出本文的贡献，然后指出未来可以进一步研究的方向。

### 5.1 本文研究总结

本论文考虑海量设备、高维复杂的神经网络模型、潜在恶意用户攻击行为等场景，以增强大规模边缘智能网络的计算学习性能为驱动，提炼无线通信物理层链路属性、边缘智能网络规模、边缘节点特征等，针对模型参数高效传输、计算学习性能优化、鲁棒与隐私增强三方面问题，构建了“传输—计算”协同的边缘智能优化方法，分析了传算协同性能并形成定量化指标，研究了相应的模型架构、联合优化方法、高效性提升策略和鲁棒性增强方案，形成了完备的理论、模型、算法与策略，实现了无线网络资源和计算学习资源的一体化管理体系。具体而言，本文通过以上的研究，得出以下主要成果和结论：

- (1) **推导了模拟聚合通信对 FLOA 影响的定量化描述表达式：**针对海量设备场景下造成的通信瓶颈，本文采用的空中计算技术允许所有参与联邦学习的分布式设备利用相同的时频资源同时传输它们的本地模型更新信息，大大节省了通信带宽和通信时延。此外，虽然联邦学习不共享原始数据，但分布式设备与中心服务器之间传递的模型参数仍会遭遇模型反演攻击，存在泄露隐私的风险。借助于空中计算技术，分布式设备传输的模型更新参数在空中叠加，使得恶意用户无法获得单个分布式设备的具体更新信息，进而保护了分布式设备的数据隐私。由于空中计算技术采用模拟聚合通信模式，一些传输误差会对联邦学习造成负面影响。因此，本文通过提炼无线链路特征参数，推导了 FLOA 在不同假设下的收敛性闭式表达式，量化了模拟聚合无

**线通信对 FLOA 的影响。**针对不同学习网络的损失函数（如凸损失函数、非凸损失函数等）和不同计算学习算法（如标准梯度下降法、随机梯度下降法等），本文从学习算法收敛性的角度全面分析了无线传输对计算学习的影响，为后续提出传算联合优化方法的构建奠定了坚实的理论基础。

- (2) 提出了一个基于收敛性分析的 FLOA 的传算联合优化基础框架：当前关于基于无线通信的联邦学习的研究大多将通信和计算两个过程分开，独立进行优化，鲜有将两个过程一起考虑研究通信资源和学习资源的联合优化的方案。为了填补这个研究空白，基于本文关于无线链路对学习性能的定量化描述，针对不同学习网络的损失函数和不同计算学习算法，本文建立了传算联合优化问题，并提出了传算联合优化方法，形成了一个可扩展的传算联合优化模型与解决方案框架。该框架在节省通信带宽，降低通信时延，同时保护用户数据隐私的基础上，可以缓解通信链路对联邦学习性能的负面影响，提高学习准确率。通过理论分析和仿真实验，本文证实了所提方案的有效性。
- (3) 提出了基于 1 比特压缩感知的高效性 FLOA 方案：针对边缘智能应用中所使用的高维神经网络模型，本文研究了复杂模型更新参数的高效性通信问题。本文将 1 比特压缩感知技术引入到 FLOA 传算联合优化框架中，构建了一个集稀疏化、降维、量化、模拟聚合传输以及信号重建为一体的联邦学习高效方案。此方案除了空中计算技术可以节省通信带宽资源外，由于使用 1 比特压缩感知技术减少了所要传输的本地更新信息的信息量，进一步降低了通信负载与通信时延，实现了快速高效的联邦学习。通过推导联邦学习算法的收敛性闭式表达式，本文揭示了由于稀疏化、降维、量化、信号重构和信道噪声引起的聚合误差对学习性能的影响。此收敛性分析的理论结果，不仅为通信效率和学习性能之间的权衡做出了量化分析，而且还可以指导参数设计与优化。在该理论分析结果的指导下，本文构建了一个通信和学习的联合优化问题，并分别提出了可以获得小规模网络的最优解和大规模网络的次优解的两种解决方法。此联合优化方案实现了最佳的分布式设备选择和功率控制，减轻了聚合误差对联邦学习的影响。最后，通过仿真实验结果表明了，本文所提出的方案可以极大地提高通信效率，同时确保所需的学习性能。
- (4) 提出了基于功率控制的 FLOA 的鲁棒性方案：本文采用的 FLOA 方案虽然阻止了由模型反演攻击导致的数据隐私泄露，但是却使得参与学习的恶意用户更容易发动拜占庭攻击破坏学习任务。针对智慧物联网中可能存在恶意用户实行拜占庭攻击的破坏行为，本文从物理层信号处理的角度提出了

FLOA 鲁棒性方案。首先，本文讨论了拜占庭攻击者的攻击方式，从理论上证明了拜占庭攻击影响联邦学习的效果，并推导出了拜占庭攻击者的攻击存在一种最强的攻击方式。由于不同功率控制策略对拜占庭攻击的抵御能力不同，本文进一步分析了各种功率控制策略下，拜占庭攻击者的最强攻击效果。通过推导出的联邦学习算法收敛性的闭式表达式，本文从理论上证明了拜占庭攻击下学习算法的性能边界。最后，根据对拜占庭攻击者的攻击特点分析，提出了一种对抗拜占庭攻击的新型功率控制策略，通过算法收敛性分析，比较了所提出的功率控制策略和现存策略的性能差异。通过理论分析和仿真实验表明，本文所提出的对抗策略可以有效抑制拜占庭攻击对学习算法带来的负面影响。

## 5.2 下一步研究展望

本论文针对边缘智能中的学习与通信的联合优化问题进行了深入研究，取得了一些成果。根据本文所获研究经验，结合智慧物联网发展态势，在此对未来边缘智能研究领域的一些热点问题、基础理论和关键技术做出以下展望：

- (1) 针对复杂场景的边缘智能研究：
  - 1) 针对分布式设备的异构性：在大规模的智慧物联网环境中，设备种类各异，很难保证每种设备具有相同的计算能力。而本文所涉及的空中计算技术很大程度上依赖于通信上的同步。如果有的设备计算速度慢，就会产生通信上的等待延迟。一些计算能力差的设备会成为整个联邦学习计划的累赘。但是，这些计算能力差的设备的本地数据可能是高质量的。因此，针对分布式设备的异构性，从公平性的角度来设计分布式设备的选择方案，同步机制以及聚合规则是一个未来亟需解决的问题。
  - 2) 针对非独立同分布的分布式数据：现存的关于联邦学习的研究通常假设分布式数据是独立同分布的，这样使得聚合的模型参数是中心式机器学习的模型更新参数的无偏估计。但是在大规模智慧网络中，分布式设备的本地数据往往是非独立同分布的。非独立同分布的数据将导致联邦学习算法的性能下降。因此，如何设计新的计算学习算法来克服由非独立同分布的数据导致的性能下降问题，且考虑传输时的不同设备权重比例设计传算联合优化方案是未来研究工作的一个重点。
  - 3) 针对无中心的网络拓扑场景：本文研究的联邦学习与无线通信的联合优化问题是在有中心的网络拓扑场景下的，而在大规模智慧物联网中，由于网络的复杂性，有时很难找到一个可靠的中心服务器来进行有中心的

联邦学习。完全分布式的联邦学习在算法上与有中心的联邦学习有很大区别，并且通信方式也将改变。此场景下的传算联合优化方案研究也将助力智慧物联网的发展。

- 4) **针对动态环境：**本文研究的传算联合优化方案均是在静态网络环境中进行的。对于实时物联网应用，每个分布式设备通常在本地接收连续的流数据，同时所有分布式设备协作以动态形式来训练一个通用的学习模型。由于每个分布式设备只能访问和处理当前数据而不是整个数据集，因此传统的确定性优化方法变得不适用。另一方面，给定流数据，学习问题的数据相关目标函数若随随时间变化，即动态目标，也会加大训练难度。为了实现以在线方式处理到达的流数据和动态的目标函数，新的动态优化方案亟待研究以支撑未来智慧物联网的实时边缘智能应用。
- (2) **针对半实物实验平台的边缘智能研究：**绝大多数现存的联邦学习的研究仅处在理论研究阶段，其中的实验部分是采用虚拟仿真形式进行，从而说明理论研究结果的有效性。然而在实际中以真实硬件设备进行的合作学习任务的性能可能和理论上的性能有很大出入。如何搭建以真实硬件设备为主体的半实物边缘智能原型系统，用以测试理论研究结果，是支撑智慧物联网的理论研究走向实用的必经之路。

## 参考文献

- [1] Knud Lasse Lueth, State of the IoT 2020 [R/OL], IoT Analytics, 2020.11.
- [2] 中国信息通信研究院, 物联网白皮书[R/OL], 2020.12.
- [3] Abhishek Mukherjee, Bill Rojas, Hugh Ujhazy, Business Models for the Long-Term Storage of Internet of Things Use Case Data[R/OL], IDC, 2020.7.
- [4] Lim W Y B, Luong N C, Hoang D T, et al. Federated learning in mobile edge networks: A comprehensive survey[J]. IEEE Communications Surveys & Tutorials, 2020, 22(3): 2031-2063.
- [5] 乔德文, 郭松涛, 何静, 朱永东, 边缘智能: 研究进展及挑战[J], 无线电通信技术, 2021.12-05.
- [6] S. Deng, H. Zhao, W. Fang, J. Yin, S. Dustdar and A. Y. Zomaya, Edge Intelligence: The Confluence of Edge Computing and Artificial Intelligence[J], IEEE Internet of Things Journal, vol. 7, no. 8, pp. 7457-7469, 2020.8.
- [7] J. Park, S. Samarakoon, M. Bennis, and M. Debbah, Wireless network intelligence at the edge[J], Proceedings of the IEEE, vol. 107, no. 11, pp. 2204–2239, 2019.
- [8] H. B. McMahan, E. Moore, D. Ramage, and B. A. Y. Arcas, Federated learning of deep networks using model averaging [J], 2016. [Online]. Available: arxiv.org/pdf/1602.05629v1.
- [9] M. Chiang and T. Zhang, Fog and IoT: An overview of research opportunities[J], IEEE Internet of Things Journal, vol. 3, no. 6, pp. 854–864, 2016.
- [10] 杨强. AI 与数据隐私保护: 联邦学习的破解之道[J]. 信息安全研究, 2019, 5(11): 961-965.
- [11] J. Konecny, H. B. McMahan, D. Ramage, and P. Richtárik, Federated optimization: Distributed machine learning for on-device intelligence[J], arXiv preprint arXiv:1610.02527, 2016.
- [12] M. I. Jordan, J. D. Lee, and Y. Yang, Communication-efficient distributed statistical inference[J], Journal of the American Statistical Association, vol. 114, no. 526, pp. 668–681, 2019.
- [13] Y. Lin, S. Han, H. Mao, Y. Wang, and B. Dally, Deep gradient compression: Reducing the communication bandwidth for distributed training[C], International Conference on Learning Representations, 2018.
- [14] A. F. Aji and K. Heafield, Sparse communication for distributed gradient descent[J], arXiv preprint arXiv:1704.05021, 2017.
- [15] Y. Liu, K. Yuan, G. Wu, Z. Tian, and Q. Ling, Decentralized dynamic ADMM with quantized and censored communications[C], 2019 53rd Asilomar Conference on Signals, Systems, and Computers. IEEE, 2019, pp. 1496–1500.
- [16] F. Seide, H. Fu, J. Droppo, G. Li, and D. Yu, 1-bit stochastic gradient descent and its application to data-parallel distributed training of speech DNNs[C], Fifteenth Annual Conference of the International Speech Communication Association, 2014.
- [17] D. Alistarh, D. Grubic, J. Li, R. Tomioka, and M. Vojnovic, QSGD: Communication-

- efficient SGD via gradient quantization and encoding[J], Advances in Neural Information Processing Systems, 2017, pp. 1709–1720.
- [18] Y. Liu, W. Xu, G. Wu, Z. Tian, and Q. Ling, Communication-censored ADMM for decentralized consensus optimization[J], IEEE Transactions on Signal Processing, vol. 67, no. 10, pp. 2565–2579, 2019.
  - [19] P. Xu, Z. Tian, Z. Zhang, and Y. Wang, Coke: Communication-censored kernel learning via random features[C], 2019 IEEE Data Science Workshop (DSW), 2019, pp. 32–36.
  - [20] T. Chen, G. Giannakis, T. Sun, and W. Yin, LAG: Lazily aggregated gradient for communication-efficient distributed learning[J], Advances in Neural Information Processing Systems, 2018, pp. 5050–5060.
  - [21] P. Xu, Z. Tian, and Y. Wang, An energy-efficient distributed average consensus scheme via infrequent communication[C], 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2018, pp. 648–652.
  - [22] P. Xu, Y. Wang, X. Chen, and T. Zhi, Coke: Communication-censored kernel learning for decentralized non-parametric learning[J], arXiv preprint arXiv:2001.10133, 2020.
  - [23] B. Nazer and M. Gastpar, Computation over multiple-access channels[J], IEEE Transactions on Information Theory, vol. 53, no. 10, pp. 3498–3516, 2007.
  - [24] F. Ang, L. Chen, N. Zhao, Y. Chen, and F. R. Yu, Robust design for massive CSI acquisition in analog function computation networks[J], IEEE Transactions on Vehicular Technology, vol. 68, no. 3, pp. 2361–2373, 2019.
  - [25] L. Chen, N. Zhao, Y. Chen, F. R. Yu, and G. Wei, Over-the-air computation for IoT networks: Computing multiple functions with antenna arrays[J], IEEE Internet of Things Journal, vol. 5, no. 6, pp. 5296–5306, 2018.
  - [26] G. Zhu and K. Huang, MIMO over-the-air computation for high-mobility multimodal sensing[J], IEEE Internet of Things Journal, vol. 6, no. 4, pp. 6089–6103, 2018.
  - [27] M. Goldenbaum, H. Boche, and S. Stanczak, Harnessing interference for analog function computation in wireless sensor networks[J], IEEE Transactions on Signal Processing, vol. 61, no. 20, pp. 4893–4906, 2013.
  - [28] M. M. Amiri and D. Gündüz, Machine learning at the wireless edge: Distributed stochastic gradient descent over-the-air[J], IEEE Transactions on Signal Processing, vol. 68, pp. 2155–2169, 2020.
  - [29] M. M. Amiri and D. Gündüz, Federated learning over wireless fading channels[J], IEEE Transactions on Wireless Communications, 2020.
  - [30] M. M. Amiri, T. M. Duman, and D. Gündüz, Collaborative machine learning at the wireless edge with blind transmitters[J], arXiv preprint arXiv:1907.03909, 2019.
  - [31] Y. Sun, S. Zhou, and D. Gündüz, Energy-aware analog aggregation for federated learning with redundant data[J], arXiv preprint arXiv:1911.00188, 2019.
  - [32] K. Yang, T. Jiang, Y. Shi, and Z. Ding, Federated learning via over-the-air computation[J], IEEE Transactions on Wireless Communications, vol. 19, no. 3, pp. 2022–2035, 2020.

- [33] T. Li, A. K. Sahu, A. Talwalkar and V. Smith, Federated Learning: Challenges, Methods, and Future Directions[J], IEEE Signal Processing Magazine, vol. 37, no. 3, pp. 50-60, May 2020.
- [34] J. Chen, X. Pan, R. Monga, S. Bengio, and R. Jozefowicz, Revisiting distributed synchronous SGD[J], arXiv preprint arXiv:1604.00981, 2016.
- [35] R. Tandon, Q. Lei, A. G. Dimakis, and N. Karampatziakis, Gradient coding: Avoiding stragglers in distributed learning[C], International Conference on Machine Learning, 2017, pp. 3368–3376.
- [36] N. Raviv, I. Tamo, R. Tandon, and A. G. Dimakis, Gradient coding from cyclic MDS codes and expander graphs[J], arXiv preprint arXiv:1707.03858, 2017.
- [37] M. Ye and E. Abbe, Communication-computation efficient gradient coding[J], arXiv preprint arXiv:1802.03475, 2018.
- [38] M. Kamp, L. Adilova, J. Sicking, F. Hüger, P. Schlicht, T. Wirtz, and S. Wrobel, Efficient decentralized deep learning by dynamic model averaging[C], Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, 2018, pp. 393–409.
- [39] X. Wang, Y. Han, C. Wang, Q. Zhao, X. Chen and M. Chen, In-Edge AI: Intelligentizing Mobile Edge Computing, Caching and Communication by Federated Learning[J], IEEE Network, vol. 33, no. 5, pp. 156-165, Sept.-Oct. 2019.
- [40] H. H. Yang, Z. Liu, T. Q. S. Quek and H. V. Poor, Scheduling Policies for Federated Learning in Wireless Networks[J], IEEE Transactions on Communications, vol. 68, no. 1, pp. 317-333, Jan. 2020.
- [41] Q. Zeng, Y. Du, K. K. Leung, and K. Huang, Energy-efficient radio resource allocation for federated edge learning[J], arXiv preprint arXiv:1907.06040, 2019.
- [42] S. Wang, T. Tuor, T. Salonidis, K. K. Leung, C. Makaya, T. He, and K. Chan, Adaptive federated learning in resource constrained edge computing systems[J], IEEE Journal on Selected Areas in Communications, vol. 37, no. 6, pp. 1205–1221, 2019.
- [43] N. H. Tran, W. Bao, A. Zomaya, N. M. NH, and C. S. Hong, Federated learning over wireless networks: Optimization model design and analysis[C], IEEE INFOCOM 2019-IEEE Conference on Computer Communications. IEEE, 2019, pp. 1387–1395.
- [44] M. Chen, Z. Yang, W. Saad, C. Yin, H. V. Poor, and S. Cui, A joint learning and communications framework for federated learning over wireless networks[J], arXiv preprint arXiv:1909.07972, 2019.
- [45] M. Gastpar, Uncoded Transmission Is Exactly Optimal for a Simple Gaussian Sensor Network, IEEE Transactions on Information Theory[J], vol. 54, no. 11, pp. 5247-5251, Nov. 2008.
- [46] A. B. Wagner, S. Tavildar and P. Viswanath, Rate Region of the Quadratic Gaussian Two-Encoder Source-Coding Problem[J], IEEE Transactions on Information Theory, vol. 54, no. 5, pp. 1938-1961, May 2008.
- [47] R. Soundararajan and S. Vishwanath, Communicating Linear Functions of Correlated Gaussian Sources Over a MAC[J], IEEE Transactions on Information

- Theory, vol. 58, no. 3, pp. 1853-1860, March 2012.
- [48] J. -J. Xiao, S. Cui, Z. -Q. Luo and A. J. Goldsmith, Linear Coherent Decentralized Estimation[J], IEEE Transactions on Signal Processing, vol. 56, no. 2, pp. 757-770, Feb. 2008.
- [49] M. Goldenbaum and S. Stanczak, On the Channel Estimation Effort for Analog Computation over Wireless Multiple-Access Channels[J], IEEE Wireless Communications Letters, vol. 3, no. 3, pp. 261-264, June 2014.
- [50] C. Wang, A. S. Leong and S. Dey, Distortion Outage Minimization and Diversity Order Analysis for Coherent Multiaccess[J], IEEE Transactions on Signal Processing, vol. 59, no. 12, pp. 6144-6159, Dec. 2011.
- [51] M. Goldenbaum, S. Stańczak and H. Boche, On achievable rates for analog computing real-valued functions over the wireless channel[C], 2015 IEEE International Conference on Communications (ICC), 2015, pp. 4036-4041.
- [52] O. Abari, H. Rahul, and D. Katabi, Over-the-air function computation in sensor networks[J], CoRR, vol. abs/1612.02307, 2016.
- [53] M. Goldenbaum and S. Stanczak, Robust Analog Function Computation via Wireless Multiple-Access Channels[J], IEEE Transactions on Communications, vol. 61, no. 9, pp. 3863-3877, September 2013.
- [54] O. Abari, H. Rahul, D. Katabi and M. Pant, AirShare: Distributed coherent transmission made seamless[C], 2015 IEEE Conference on Computer Communications (INFOCOM), 2015, pp. 1742-1750.
- [55] X. Li, G. Zhu, Y. Gong, and K. Huang, Wirelessly powered data aggregation for IoT via over-the-air functional computation: Beamforming and power control[J], [Online]. Available: <https://arxiv.org/pdf/1808.04616.pdf>, 2018.
- [56] D. Wen, G. Zhu and K. Huang, Reduced-Dimension Design of MIMO Over-the-Air Computing for Data Aggregation in Clustered IoT Networks[J], IEEE Transactions on Wireless Communications, vol. 18, no. 11, pp. 5255-5268, Nov. 2019.
- [57] B. Nazer and M. Gastpar, Compute-and-Forward: Harnessing Interference Through Structured Codes[J], IEEE Transactions on Information Theory, vol. 57, no. 10, pp. 6463-6486, Oct. 2011.
- [58] A. Sakzad, J. Harshan and E. Viterbo, Integer-Forcing MIMO Linear Receivers Based on Lattice Reduction[J], IEEE Transactions on Wireless Communications, vol. 12, no. 10, pp. 4905-4915, October 2013.
- [59] X. Cao, G. Zhu, J. Xu and K. Huang, Optimal Power Control for Over-the-Air Computation[C], 2019 IEEE Global Communications Conference (GLOBECOM), 2019, pp. 1-6.
- [60] N. Zhang and M. Tao, Gradient Statistics Aware Power Control for Over-the-Air Federated Learning[J], IEEE Transactions on Wireless Communications, vol. 20, no. 8, pp. 5115-5128, Aug. 2021.
- [61] G. Zhu, Y. Wang and K. Huang, Broadband Analog Aggregation for Low-Latency Federated Edge Learning[J], IEEE Transactions on Wireless Communications, vol. 19, no. 1, pp. 491-506, Jan. 2020.

- [62] G. Zhu, Y. Du, D. Gündüz and K. Huang, One-Bit Over-the-Air Aggregation for Communication-Efficient Federated Edge Learning: Design and Convergence Analysis[J], *IEEE Transactions on Wireless Communications*, vol. 20, no. 3, pp. 2120-2135, March 2021.
- [63] S. Wang, Y. Hong, R. Wang, Q. Hao, Y.-C. Wu, and D. W. K. Ng, Edge Federated Learning via Unit-modulus Over-the-air Computation (extended version) [J], arXiv preprint arXiv:2101.12051, 2021.
- [64] C. Xu, S. Liu, Z. Yang, Y. Huang, and K.-K. Wong, Learning rate optimization for federated learning exploiting over-the-air computation[J], arXiv preprint arXiv:2102.02946, 2021.
- [65] L. Zhu and S. Han, Deep leakage from gradients in Federated learning[J]. Springer, 2020, pp. 17–31.
- [66] Z. Yang, A. Gang and W. U. Bajwa, Adversary-Resilient Distributed and Decentralized Statistical Inference and Machine Learning: An Overview of Recent Advances Under the Byzantine Threat Model[J], *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 146-159, May 2020.
- [67] D. Yin, Y. Chen, R. Kannan, and P. Bartlett, Byzantine-robust distributed learning: Towards optimal statistical rates[C], International Conference on Machine Learning. PMLR, 2018, pp. 5650–5659.
- [68] Y. Dong, J. Cheng, M. J. Hossain and V. C. M. Leung, Secure Distributed On-Device Learning Networks with Byzantine Adversaries[J], *IEEE Network*, vol. 33, no. 6, pp. 180-187, Nov.-Dec. 2019.
- [69] G. Damaskinos, E. M. El Mhamdi, R. Guerraoui, A. H. A. Guirguis, and S. L. A. Rouault, Aggregathor: Byzantine machine learning via robust gradient aggregation[C], The Conference on Systems and Machine Learning (SysML), 2019, no. CONF, 2019.
- [70] S. Minsker, Geometric Median and Robust Estimation in Banach Spaces[J], *Bernoulli*, vol. 21, no. 4, pp. 2308–2335, 2015.
- [71] Z. Wu, Q. Ling, T. Chen and G. B. Giannakis, Federated Variance-Reduced Stochastic Gradient Descent With Robustness to Byzantine Attacks, [J] *IEEE Transactions on Signal Processing*, vol. 68, pp. 4583-4596, 2020.
- [72] Y. Chen, L. Su, and J. Xu, Distributed Statistical Machine Learning in Adversarial Settings: Byzantine Gradient Descent[C], Proceedings of the ACM on Measurement and Analysis of Computing Systems, vol. 1, no. 2, pp. 1–25, 2017.
- [73] S. Huang, Y. Zhou, T. Wang and Y. Shi, Byzantine-Resilient Federated Machine Learning via Over-the-Air Computation[C], 2021 IEEE International Conference on Communications Workshops (ICC Workshops), 2021, pp. 1-6.
- [74] P. Blanchard, E. M. El Mhamdi, R. Guerraoui, and J. Stainer, Machine Learning with Adversaries: Byzantine Tolerant Gradient Descent[C], The 31st International Conference on Neural Information Processing Systems, 2017, pp. 118–128.
- [75] R. Guerraoui, S. Rouault et al., The Hidden Vulnerability of Distributed Learning in Byzantium[C], International Conference on Machine Learning. PMLR, 2018, pp.

- 3521–3530.
- [76] E.-M. El-Mhamdi and R. Guerraoui, Fast and Secure Distributed Learning in High Dimension[J], arXiv e-prints, pp. arXiv1905, 2019.
- [77] C. Xie, O. Koyejo, and I. Gupta, Zeno: Byzantine-suspicious Stochastic Gradient Descent[J], arXiv preprint arXiv:1805.10032, vol. 24, 2018.
- [78] C. Xie, O. Koyejo, and I. Gupta, Zeno++: Robust Asynchronous SGD with Arbitrary Number of byzantine workers[J], arXiv preprint arXiv:1903.07020, 2019.
- [79] K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition[J], arXiv preprint arXiv:1409.1556, 2014.
- [80] Ruder S. An overview of gradient descent optimization algorithms[J]. arXiv preprint arXiv:1609.04747, 2016.
- [81] Bottou L. Large-scale machine learning with stochastic gradient descent[M]//Proceedings of COMPSTAT'2010. Physica-Verlag HD, 2010: 177-186.
- [82] R. C. Buck, approximate complexity and functional representation[J], J. Math. Anal. Appl., vol. 70, no. 1, pp. 280–298, Jul. 1979.
- [83] P. Andres-Maldonado, P. Ameigeiras, J. Prados-Garzon, J. Navarro-Ortiz, and J. M. Lopez-Soler, Narrowband iot data transmission procedures for massive machine-type communications[J], Ieee Network, vol. 31, no. 6, pp. 8–15, 2017.
- [84] A. KhannaandS. Kaur, Internet of things (iot), applications and challenges: a comprehensive review[J], Wireless Personal Communications, vol. 114, no. 2, pp. 1687–1762, 2020.
- [85] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, Communication-efficient learning of deep networks from decentralized data[J], in Artificial intelligence and statistics. PMLR, 2017, pp. 1273–1282.
- [86] G. Zhu, D. Liu, Y. Du, C. You, J. Zhang, and K. Huang, Toward an intelligent edge: wireless communication meets machine learning[J], IEEE Communications Magazine, vol. 58, no. 1, pp. 19–25, 2020.
- [87] T. T. Vu, D. T. Ngo, N. H. Tran, H. Q. Ngo, M. N. Dao, and R. H. Middleton, Cell-free massive mimo for wireless federated learning[J], IEEE Transactions on Wireless Communications, vol. 19, no. 10, pp. 6377–6392, 2020.
- [88] J. Wang and G. Joshi, Cooperative sgd: A unified framework for the design and analysis of local-update sgd algorithms[J], Journal of Machine Learning Research, vol. 22, no. 213, pp. 1–50, 2021.
- [89] O. Shamir, N. Srebro, and T. Zhang, Communication-efficient distributed optimization using an approximate newton-type method[C], in International conference on machine learning, 2014, pp. 1000–1008.
- [90] S. Magnússon, H. Shokri-Ghadikolaei, and N. Li, On maintaining linear convergence of distributed learning and optimization under limited communication[J], IEEE Transactions on Signal Processing, 2020.
- [91] D. P. Bertsekas, J. N. Tsitsiklis, and J. Tsitsiklis, Neuro-Dynamic Programming[J]. Athena Scientific, 1996.
- [92] M. P. Friedlander and M. Schmidt, Hybrid deterministic-stochastic methods for data

- fitting[J], SIAM Journal on Scientific Computing, vol. 34, no. 3, pp. A1380–A1405, 2012.
- [93] D. Alistarh, T. Hoefler, M. Johansson, N. Konstantinov, S. Khirirat, and C. Renggli, The convergence of sparsified gradient methods[J], in Advances in Neural Information Processing Systems, 2018, pp. 5973–5983.
- [94] K. Yuan, Q. Ling, and W. Yin, On the convergence of decentralized gradient descent[J], SIAM Journal on Optimization, vol. 26, no. 3, pp. 1835–1854, 2016.
- [95] L. Bottou, F. E. Curtis, and J. Nocedal, Optimization methods for large-scale machine learning[J], Siam Review, vol. 60, no. 2, pp. 223–311, 2018.
- [96] S. U. Stich, J.-B. Cordonnier, and M. Jaggi, Sparsified sgd with memory[J], in Advances in Neural Information Processing Systems, 2018, pp. 4447–4458.
- [97] H. Tang, C. Yu, X. Lian, T. Zhang, and J. Liu, Doublesqueeze: Parallel stochastic gradient descent with double-pass error-compensated compression[C], in International Conference on Machine Learning. PMLR, 2019, pp. 6155–6165.
- [98] D. Alistarh, D. Grubic, J. Li, R. Tomioka, and M. Vojnovic, Qsgd: Communication-efficient sgd via gradient quantization and encoding[J], in Advances in Neural Information Processing Systems, 2017, pp. 1709–1720.
- [99] LeCun Y. The MNIST database of handwritten digits[J]. <http://yann. lecun. com/exdb/mnist/>, 1998.
- [100] X. Fan, Y. Wang, Y. Huo, and Z. Tian, Joint optimization of communications and federated learning over the air[J], IEEE Transactions on Wireless Communications, pp. 1–1, 2022.
- [101] X. Fan, Y. Wang, Y. Huo, and Z. Tian, Joint optimization for federated learning over the air[C], in 2022 IEEE International Conference on Communications (ICC 2022). IEEE, 2022, pp. 1–6.
- [102] P. T. Boufounos and R. G. Baraniuk, 1-bit compressive sensing[C], in 2008 42nd Annual Conference on Information Sciences and Systems. IEEE, 2008, pp. 16–21
- [103] L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors[J], IEEE Transactions on Information Theory, vol. 59, no. 4, pp. 2082–2102, 2013.
- [104] D.-Q. Dai, L. Shen, Y. Xu, and N. Zhang, Noisy 1-bit compressive sensing: models and algorithms[J], Applied and Computational Harmonic Analysis, vol. 40, no. 1, pp. 1–32, 2016.
- [105] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, Federated learning: Strategies for improving communication efficiency[J], arXiv preprint arXiv:1610.05492, 2016.
- [106] N. Ström, Sparse connection and pruning in large dynamic artificial neural networks[C], in Fifth European Conference on Speech Communication and Technology. Citeseer, 1997.
- [107] N. Ström, Scalable distributed dnn training using commodity gpu cloud computing[C], in Sixteenth Annual Conference of the International Speech Communication Association, 2015.

- [108] E. T. Hale, W. Yin, and Y. Zhang, A fixed-point continuation method for  $\ell_1$ -regularized minimization with applications to compressed sensing[J], CAAM TR07-07, Rice University, vol. 43, p. 44, 2007.
- [109] A. Moshtaghpour, L. Jacques, V. Cambareri, K. Degraux, and C. De Vleeschouwer, Consistent basis pursuit for signal and matrix estimates in quantized compressed sensing[J], IEEE signal processing letters, vol. 23, no. 1, pp. 25–29, 2015.
- [110] D. L. Donoho, Compressed sensing[J], IEEE Transactions on information theory, vol. 52, no. 4, pp. 1289–1306, 2006.
- [111] S. Bubeck, Convex optimization: Algorithms and complexity[J], Foundations and Trends in Machine Learning, vol. 8, no.3-4, pp. 231–357, 2015.
- [112] D. P. Bertsekas, J. N. Tsitsiklis, and J. Tsitsiklis, Neuro-Dynamic Programming[J]. Athena Scientific, 1996.
- [113] M. P. Friedlander and M. Schmidt, Hybrid deterministic-stochastic methods for data fitting[J], SIAM Journal on Scientific Computing, vol. 34, no. 3, pp. A1380–A1405, 2012.
- [114] X. Lian, Y. Huang, Y. Li, and J. Liu, Asynchronous parallel stochastic gradient for nonconvex optimization[J], Advances in Neural Information Processing Systems, vol. 28, pp. 2737–2745, 2015.
- [115] J. Zeng and W. Yin, On nonconvex decentralized gradient descent[J], IEEE Transactions on signal processing, vol. 66, no. 11, pp. 2834–2848, 2018.
- [116] S. Boyd, S. P. Boyd, and L. Vandenberghe, Convex optimization[M]. Cambridge university press, 2004.
- [117] S. Boyd, N. Parikh, and E. Chu, Distributed optimization and statistical learning via the alternating direction method of multipliers[J]. Now Publishers Inc, 2011.
- [118] E. Ghadimi, A. Teixeira, I. Shames, and M. Johansson, Optimal parameter selection for the alternating direction method of multipliers (admm): quadratic problems[J], IEEE Transactions on Automatic Control, vol. 60, no. 3, pp. 644–658, 2014.
- [119] J. Bernstein, Y.-X. Wang, K. Azizzadenesheli, and A. Anandkumar, signsgd: Compressed optimisation for non-convex problems[C], in International Conference on Machine Learning. PMLR, 2018, pp. 560–569.
- [120] E. J. Candes et al., The restricted isometry property and its implications for compressed sensing[J], Comptes rendus mathematique, vol. 346, no. 9-10, pp. 589–592, 2008.
- [121] E. J. Candes, J. K. Romberg, and T. Tao, Stable signal recovery from incomplete and inaccurate measurements[J], Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences, vol. 59, no. 8, pp. 1207–1223, 2006.
- [122] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, A survey on mobile edge computing: The communication perspective[J], IEEE Communications Surveys & Tutorials, vol. 19, no. 4, pp. 2322–2358, 2017.
- [123] Q. Yang, Y. Liu, T. Chen, and Y. Tong, Federated machine learning: Concept and applications[J], ACM Transactions on Intelligent Systems and Technology (TIST),

- vol. 10, no. 2, pp. 1–19, 2019.
- [124] X. Fan, Y. Wang, Y. Huo, and Z. Tian, Communication-efficient federated learning through 1-bit compressive sensing and analog aggregation[C], in 2021 IEEE International Conference on Communications Workshops (ICC Workshops). IEEE, 2021, pp. 1–6.
  - [125] X. Fan, Y. Wang, Y. Huo, and Z. Tian, 1-bit compressive sensing for efficient federated learning over the air[J], arXiv preprint arXiv:2103.16055, 2021.
  - [126] S. Xia, J. Zhu, Y. Yang, Y. Zhou, Y. Shi, and W. Chen, Fast convergence algorithm for analog federated learning[C], in ICC 2021-IEEE International Conference on Communications. IEEE, 2021, pp. 1–6.
  - [127] D. Liu and O. Simeone, Privacy for free: Wireless federated learning via uncoded transmission with adaptive power control[J], IEEE Journal on Selected Areas in Communications, vol. 39, no. 1, pp. 170–185, 2020.
  - [128] Y. Nesterov, Introductory lectures on convex programming volume i: Basic course[J], Lecture notes, vol. 3, no. 4, p. 5, 1998.



## 附录 A

### A1. 定理 2.1 的证明

**证明：**定理 2.1 考虑凸损失函数下的全批量 GD 方法。本证明过程首先写出假设 2.1 所隐含的不等式，如下所示

$$F(\mathbf{w}_t) \leq F(\mathbf{w}_{t-1}) + (\mathbf{w}_t - \mathbf{w}_{t-1})^T \nabla F(\mathbf{w}_{t-1}) + \frac{L}{2} \|\mathbf{w}_t - \mathbf{w}_{t-1}\|^2. \quad (\text{A-1})$$

使用标准的全批量 GD 方法，第  $i$  个分布式设备在第  $t$  次迭代时更新其本地联邦学习模型参数  $\mathbf{w}_{i,t}$  如下

$$\mathbf{w}_{i,t} = \mathbf{w}_{t-1} - \frac{\alpha}{K_i} \sum_{k=1}^{K_i} \nabla f(\mathbf{w}_{t-1}, \mathbf{x}_{i,k}, \mathbf{y}_{i,k}), \quad i = 1, 2, \dots, U. \quad (\text{A-2})$$

将公式 (A-2) 代入到公式 (2-9) 中，可得

$$\begin{aligned} \mathbf{w}_t &= \mathbf{w}_{t-1} + \left( \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot -1} \odot \mathbf{z}_t \\ &\quad - \alpha \left( \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \right)^{\odot -1} \odot \sum_{i=1}^U \sum_{k=1}^{K_i} \boldsymbol{\beta}_{i,t} \odot \nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, y_{i,k}) \\ &= \mathbf{w}_{t-1} - \alpha (\nabla F(\mathbf{w}_{t-1}) - \mathbf{o}), \end{aligned} \quad (\text{A-3})$$

其中，

$$\begin{aligned} \mathbf{o} &= \nabla F(\mathbf{w}_{t-1}) + \left( \alpha \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot -1} \odot \mathbf{z}_t \\ &\quad - \left( \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \right)^{\odot -1} \odot \sum_{i=1}^U \sum_{k=1}^{K_i} \boldsymbol{\beta}_{i,t} \odot \nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, y_{i,k}). \end{aligned} \quad (\text{A-4})$$

给定学习速率  $\alpha = \frac{1}{L}$ （此为不失一般性的简单表达式的特殊设置），公式 (A-1) 中全局损失函数的期望  $\mathbb{E}[F(\mathbf{w}_t)]$  可以表示为

$$\begin{aligned} \mathbb{E}[F(\mathbf{w}_t)] &\leq \mathbb{E} \left[ F(\mathbf{w}_{t-1}) - \alpha (\nabla F(\mathbf{w}_{t-1}) - \mathbf{o})^T \nabla F(\mathbf{w}_{t-1}) + \frac{L\alpha^2}{2} \|\nabla F(\mathbf{w}_{t-1}) - \mathbf{o}\|^2 \right] \\ &\stackrel{(a)}{=} \mathbb{E}[F(\mathbf{w}_{t-1})] - \frac{1}{2L} \|\nabla F(\mathbf{w}_{t-1})\|^2 + \frac{1}{2L} \mathbb{E}[\|\mathbf{o}\|^2], \end{aligned} \quad (\text{A-5})$$

其中步骤(a)的推导源于以下事实

$$\frac{L\alpha^2}{2} \|\nabla F(\mathbf{w}_{t-1}) - \mathbf{o}\|^2 = \frac{1}{2L} \|\nabla F(\mathbf{w}_{t-1})\|^2 - \frac{1}{L} \mathbf{o}^T \nabla F(\mathbf{w}_{t-1}) + \frac{1}{2L} \|\mathbf{o}\|^2. \quad (\text{A-6})$$

公式 (A-6) 右侧的最后一项中的  $\mathbb{E}[\|\mathbf{o}\|^2]$  可以被推导如下

$$\begin{aligned}
\mathbb{E}[\|\mathbf{o}\|^2] &= \mathbb{E}\left[\left\|\nabla F(\mathbf{w}_{t-1}) + \left(\alpha \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t\right)^{\odot -1} \odot \mathbf{z}_t - \right.\right. \\
&\quad \left.\left(\sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t}\right)^{\odot -1} \odot \sum_{i=1}^U \sum_{k=1}^{K_i} \boldsymbol{\beta}_{i,t} \odot \nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})\right\|^2\Big] \\
&= \mathbb{E}\left[\left\|\frac{\sum_{i=1}^U \sum_{k=1}^{K_i} \nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})}{K} - \right.\right. \\
&\quad \left.\left(\sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t}\right)^{\odot -1} \odot \sum_{i=1}^U \sum_{k=1}^{K_i} \boldsymbol{\beta}_{i,t} \odot \nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k}) \right. \\
&\quad \left. + \left(\alpha \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t\right)^{\odot -1} \odot \mathbf{z}_t\right\|^2\Big] \\
&= \mathbb{E}\left[\left\|\left(\alpha \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t\right)^{\odot -1} \odot \mathbf{z}_t + \sum_{i=1}^U \left(\frac{1}{K} - \right.\right. \\
&\quad \left.\left.\boldsymbol{\beta}_{i,t} \odot \left(\sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t}\right)^{\odot -1}\right) \odot \sum_{k=1}^{K_i} \nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})\right\|^2\Big], \quad (\text{A-7})
\end{aligned}$$

其中 $\mathbf{1}$ 是长度为 $D$ 的全 $1$ 向量，其维度与 $\boldsymbol{\beta}_{i,t}$ 的维度相同。

利用范数的三角不等式 $\|\mathbf{X} + \mathbf{Y}\| \leq \|\mathbf{X}\| + \|\mathbf{Y}\|$ ，范数的子乘性质 $\|\mathbf{XY}\| \leq \|\mathbf{X}\|\|\mathbf{Y}\|$ 和Jensen不等式，公式(A-7)可以被进一步推导如下

$$\begin{aligned}
\mathbb{E}[\|\mathbf{o}\|^2] &\leq \mathbb{E}\left[\left\|\sum_{i=1}^U \left(\frac{1}{K} - \boldsymbol{\beta}_{i,t} \odot \left(\sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t}\right)^{\odot -1}\right) \odot \sum_{k=1}^{K_i} \nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})\right\|^2\right] \\
&\quad + \mathbb{E}\left[\left\|\left(\alpha \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t\right)^{\odot -1} \odot \mathbf{z}_t\right\|^2\right] \\
&\leq \mathbb{E}\left[\left\|\left(\alpha \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t\right)^{\odot -1} \odot \mathbf{z}_t\right\|^2\right] + \mathbb{E}\left[K \sum_{i=1}^U \left\|\frac{1}{K} - \right.\right. \\
&\quad \left.\left.\boldsymbol{\beta}_{i,t} \odot \left(\sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t}\right)^{\odot -1}\right\|^2 \sum_{k=1}^{K_i} \|\nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})\|^2\right] \\
&\leq \left\|\left(\sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t\right)^{\odot -1}\right\|^2 \sigma^2 L^2 + K \sum_{i=1}^U \left\|\frac{1}{K} - \right.\right. \\
&\quad \left.\left.\boldsymbol{\beta}_{i,t} \odot \left(\sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t}\right)^{\odot -1}\right\|^2 \sum_{k=1}^{K_i} \|\nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})\|^2. \quad (\text{A-8})\right.
\end{aligned}$$

应用假设2.3中的公式(2-12)到公式(A-8)中，可以得到

$$\begin{aligned}
\mathbb{E}[\|\mathbf{o}\|^2] &\leq K \sum_{i=1}^U \left\| \frac{\mathbf{1}}{K} - \boldsymbol{\beta}_{i,t} \odot \left( \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \right)^{\odot-1} \right\|^2 K_i(\rho_1 \\
&\quad + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) + \left\| \left( \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot-1} \right\|^2 \sigma^2 L^2 \\
&= K \sum_{i=1}^U \sum_{d=1}^D \left( \frac{1}{K} - \frac{\beta_{i,t}^d}{\sum_{i=1}^U K_i \beta_{i,t}^d} \right)^2 K_i(\rho_1 \\
&\quad + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) + \left\| \left( \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot-1} \right\|^2 \sigma^2 L^2 \\
&= K \sum_{i=1}^U \sum_{d=1}^D \left( \frac{1}{K^2} - \frac{2}{K} \frac{\beta_{i,t}^d}{\sum_{i=1}^U K_i \beta_{i,t}^d} \right. \\
&\quad \left. + \frac{(\beta_{i,t}^d)^2}{(\sum_{i=1}^U K_i \beta_{i,t}^d)^2} \right) K_i(\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) \\
&\quad + \left\| \left( \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot-1} \right\|^2 \sigma^2 L^2 \\
&= \sum_{d=1}^D \left( \frac{K}{\sum_{i=1}^U K_i \beta_{i,t}^d} - 1 \right) (\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) \\
&\quad + \left\| \left( \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot-1} \right\|^2 \sigma^2 L^2. \tag{A-9}
\end{aligned}$$

将公式 (A-9) 代入公式 (A-5) 中, 可得

$$\begin{aligned}
\mathbb{E}[F(\mathbf{w}_t)] &\leq \mathbb{E}[F(\mathbf{w}_{t-1})] - \frac{1}{2L} \|\nabla F(\mathbf{w}_{t-1})\|^2 + \\
&\quad \frac{1}{2L} \left( \sum_{d=1}^D \left( \frac{K}{\sum_{i=1}^U K_i \beta_{i,t}^d} - 1 \right) (\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) \right. \\
&\quad \left. + \left\| \left( \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot-1} \right\|^2 \sigma^2 L^2 \right). \tag{A-10}
\end{aligned}$$

从公式 (A-10) 的两边减去  $\mathbb{E}[F(\mathbf{w}^*)]$ , 可得

$$\begin{aligned}
\mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}^*)] &\leq \mathbb{E}[F(\mathbf{w}_{t-1}) - F(\mathbf{w}^*)] \\
&\quad + \frac{1}{2L} \left( \sum_{d=1}^D \left( \frac{K}{\sum_{i=1}^U K_i \beta_{i,t}^d} - 1 \right) (\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) \right. \\
&\quad \left. + \left\| \left( \sum_{i=1}^U K_i \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot-1} \right\|^2 \sigma^2 L^2 \right) - \frac{1}{2L} \|\nabla F(\mathbf{w}_{t-1})\|^2. \tag{A-11}
\end{aligned}$$

为了最小化公式 (2-11) 的两边, 可得

$$\begin{aligned} \min_{\mathbf{w}_t} F(\mathbf{w}_t) &\geq \min_{\mathbf{w}_t} (F(\mathbf{w}_{t-1}) + (\mathbf{w}_t - \mathbf{w}_{t-1})^T \nabla F(\mathbf{w}_{t-1}) \\ &\quad + \frac{\mu}{2} \|\mathbf{w}_t - \mathbf{w}_{t-1}\|^2). \end{aligned} \quad (\text{A-12})$$

公式 (A-12) 左边的最小化是通过  $\mathbf{w}_t = \mathbf{w}^*$  实现的, 而右边的最小化是通过  $\mathbf{w}_t = \mathbf{w}_{t-1} - \frac{1}{\mu} \nabla F(\mathbf{w}_{t-1})$  实现的。因此, 可得

$$F(\mathbf{w}^*) \geq F(\mathbf{w}_{t-1}) - \frac{1}{2\mu} \|\nabla F(\mathbf{w}_{t-1})\|^2. \quad (\text{A-13})$$

进而可得

$$\|\nabla F(\mathbf{w}_{t-1})\|^2 \geq 2\mu(F(\mathbf{w}_{t-1}) - F(\mathbf{w}^*)). \quad (\text{A-14})$$

将公式 (A-14) 代入公式 (A-11) 可得

$$\begin{aligned} \mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}^*)] &\leq (1 - \frac{\mu}{L}) \mathbb{E}[F(\mathbf{w}_{t-1}) - F(\mathbf{w}^*)] \\ &\quad + \frac{1}{2L} \left( \sum_{d=1}^D \left( \frac{K}{\sum_{i=1}^U K_i \beta_{i,t}^d} - 1 \right) (\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) \right. \\ &\quad \left. + \left\| \left( \sum_{i=1}^U K_i \beta_{i,t} \odot \mathbf{b}_t \right)^{\odot -1} \right\|^2 \sigma^2 L^2 \right) \\ &= (1 - \frac{\mu}{L}) \mathbb{E}[F(\mathbf{w}_{t-1}) - F(\mathbf{w}^*)] + \frac{\rho_2}{2L} \left( \sum_{d=1}^D \left( \frac{K}{\sum_{i=1}^U K_i \beta_{i,t}^d} \right. \right. \\ &\quad \left. \left. - 1 \right) \right) \|\nabla F(\mathbf{w}_{t-1})\|^2 + \frac{\rho_1}{2L} \sum_{d=1}^D \left( \frac{K}{\sum_{i=1}^U K_i \beta_{i,t}^d} - 1 \right) \\ &\quad + \left\| \left( \sum_{i=1}^U K_i \beta_{i,t} \odot \mathbf{b}_t \right)^{\odot -1} \right\|^2 \frac{L\sigma^2}{2}. \end{aligned} \quad (\text{A-15})$$

接下来, 与推导公式 (A-14) 的方式相同, 最小化公式 (A-1) 的两边, 可以得到

$$\|\nabla F(\mathbf{w}_{t-1})\|^2 \leq 2L(F(\mathbf{w}_{t-1}) - F(\mathbf{w}^*)). \quad (\text{A-16})$$

将公式 (A-16) 代入公式 (A-15), 可以得到

$$\begin{aligned} \mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}^*)] &\leq \left( 1 - \frac{\mu}{L} + \rho_2 \sum_{d=1}^D \left( \frac{K}{\sum_{i=1}^U K_i \beta_{i,t}^d} - 1 \right) \right) \mathbb{E}[F(\mathbf{w}_{t-1}) - F(\mathbf{w}^*)] \\ &\quad + \frac{\rho_1}{2L} \sum_{d=1}^D \left( \frac{K}{\sum_{i=1}^U K_i \beta_{i,t}^d} - 1 \right) \\ &\quad + \left\| \left( \sum_{i=1}^U K_i \beta_{i,t} \odot \mathbf{b}_t \right)^{\odot -1} \right\|^2 \frac{L\sigma^2}{2} \\ &= B_t + A_t \mathbb{E}[F(\mathbf{w}_{t-1}) - F(\mathbf{w}^*)], \end{aligned} \quad (\text{A-17})$$

其中，

$$A_t = 1 - \frac{\mu}{L} + \rho_2 \sum_{d=1}^D \left( \frac{K}{\sum_{i=1}^U K_i \beta_{i,t}^d} - 1 \right), \quad (\text{A-18})$$

$$\begin{aligned} B_t = & \frac{\rho_1}{2L} \sum_{d=1}^D \left( \frac{K}{\sum_{i=1}^U K_i \beta_{i,t}^d} - 1 \right) \\ & + \left\| \left( \sum_{i=1}^U K_i \beta_{i,t} \odot \mathbf{b}_t \right)^{\odot -1} \right\|^2 \frac{L\sigma^2}{2}. \end{aligned} \quad (\text{A-19})$$

证毕。 ■

## A2. 定理 2.2 的证明

**证明：**定理 2.2 考虑非凸优化问题的全批量 GD 方法。此定理 2.2 的证明遵循定理 2.1 的证明，直到公式 (A-10)。根据公式 (A-10)，可以得到

$$\begin{aligned} \mathbb{E}[F(\mathbf{w}_t)] &\leq \mathbb{E}[F(\mathbf{w}_{t-1})] + \frac{\rho_1}{2L} \sum_{d=1}^D \left( \frac{K}{\sum_{i=1}^U K_i \beta_{i,t}^d} - 1 \right) \\ &\quad - \frac{1}{2L} \left( 1 - \rho_2 \sum_{d=1}^D \left( \frac{K}{\sum_{i=1}^U K_i \beta_{i,t}^d} - 1 \right) \right) \|\nabla F(\mathbf{w}_{t-1})\|^2 \\ &\quad + \left\| \left( \sum_{i=1}^U K_i \beta_{i,t} \odot \mathbf{b}_t \right)^{\odot -1} \right\|^2 \frac{L\sigma^2}{2} \\ &= \mathbb{E}[F(\mathbf{w}_{t-1})] - \frac{2 - A_t - \frac{\mu}{L}}{2L} \|\nabla F(\mathbf{w}_{t-1})\|^2 + B_t. \end{aligned} \quad (\text{A-20})$$

将上式 (A-20) 从  $t = 1$  到  $T$  加起来，可以得到

$$\mathbb{E}[F(\mathbf{w}_t)] - \mathbb{E}[F(\mathbf{w}_0)] \leq - \sum_{t=1}^T \frac{2 - A_t - \frac{\mu}{L}}{2L} \|\nabla F(\mathbf{w}_{t-1})\|^2 + \sum_{t=1}^T B_t. \quad (\text{A-21})$$

由上式 (A-21) 可得

$$\begin{aligned} \sum_{t=1}^T \frac{2 - A_t - \frac{\mu}{L}}{2L} \|\nabla F(\mathbf{w}_{t-1})\|^2 &\leq \mathbb{E}[F(\mathbf{w}_0)] - \mathbb{E}[F(\mathbf{w}_t)] + \sum_{t=1}^T B_t \\ &\leq \mathbb{E}[F(\mathbf{w}_0)] - \mathbb{E}[F(\mathbf{w}^*)] + \sum_{t=1}^T B_t. \end{aligned} \quad (\text{A-22})$$

根据命题 2.1 中的结论，可以得到

$$0 \leq \frac{1 - \rho_2 D \left( \frac{K}{K_{min}} - 1 \right)}{2L} \leq \frac{2 - A_t - \frac{\mu}{L}}{2L} \leq \frac{1}{2L}, \quad \forall t. \quad (\text{A-23})$$

将公式 (A-23) 代入公式 (A-22) 可以得到

$$\begin{aligned}
& \frac{1}{T} \sum_{t=1}^T \frac{1 - \rho_2 D(\frac{K}{K_{min}} - 1)}{2L} \|\nabla F(\mathbf{w}_{t-1})\|^2 \\
& \leq \frac{1}{T} \sum_{t=1}^T \frac{2 - A_t - \frac{\mu}{L}}{2L} \|\nabla F(\mathbf{w}_{t-1})\|^2 \\
& \leq \frac{1}{T} (\mathbb{E}[F(\mathbf{w}_0)] - \mathbb{E}[F(\mathbf{w}^*)]) + \frac{1}{T} \sum_{t=1}^T B_t.
\end{aligned} \tag{A-24}$$

综上，可以得到定理 2.2 中的结论如下

$$\begin{aligned}
\frac{1}{T} \sum_{t=1}^T \|\nabla F(\mathbf{w}_{t-1})\|^2 & \leq \frac{2L \sum_{t=1}^T B_t}{T(1 - \rho_2 D(\frac{K}{K_{min}} - 1))} \\
& + \frac{2L}{T(1 - \rho_2 D(\frac{K}{K_{min}} - 1))} \mathbb{E}[F(\mathbf{w}_0) - F(\mathbf{w}^*)].
\end{aligned} \tag{A-25}$$

由此即证。 ■

### A3. 定理 2.3 的证明

**证明：**采用批量大小  $K_b$  为小批量 SGD，第  $i$  个分布式设备的本地参数在第  $t$  次迭代时更新为

$$\mathbf{w}_{i,t} = \mathbf{w}_{t-1} - \alpha \mathbb{E}_{\mathcal{D}_i} \left[ \frac{\sum_{k=1}^{K_b} \nabla f(\mathbf{w}_{t-1}, \mathbf{x}_{i,k}, \mathbf{y}_{i,k})}{K_b} \right], \tag{A-26}$$

其中， $\mathbb{E}_{\mathcal{D}_i}[\cdot]$  是期望值，表示第  $i$  个分布式设备从其本地数据集  $\mathcal{D}_i$  中随机选择  $K_b$  个样本来计算局部梯度。

将公式 (A-26) 代入公式 (2-9) 可以得到平均梯度的估计值如下

$$\begin{aligned}
\mathbf{w}_t & = \mathbf{w}_{t-1} - \alpha \left( \sum_{i=1}^U K_b \boldsymbol{\beta}_{i,t} \right)^{\odot -1} \\
& \odot \sum_{i=1}^U \left( K_b \boldsymbol{\beta}_{i,t} \odot \mathbb{E}_{\mathcal{D}_i} \left[ \frac{\sum_{k=1}^{K_b} \nabla f(\mathbf{w}_{t-1}, \mathbf{x}_{i,k}, \mathbf{y}_{i,k})}{K_b} \right] \right) \\
& + \left( \sum_{i=1}^U K_b \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot -1} \odot \mathbf{z}_t \\
& = \mathbf{w}_{t-1} - \alpha (\nabla F(\mathbf{w}_{t-1}) - \mathbf{o}),
\end{aligned} \tag{A-27}$$

其中，

$$\begin{aligned}
 \mathbf{o} = & \nabla F(\mathbf{w}_{t-1}) - \left( \sum_{i=1}^U K_b \boldsymbol{\beta}_{i,t} \right)^{\odot-1} \\
 & \odot \sum_{i=1}^U \left( \boldsymbol{\beta}_{i,t} \odot \mathbb{E}_{\mathcal{D}_i} \left[ \sum_{k=1}^{K_b} \nabla f(\mathbf{w}_{t-1}, \mathbf{x}_{i,k}, \mathbf{y}_{i,k}) \right] \right) \\
 & + \left( \alpha \sum_{i=1}^U K_b \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot-1} \odot \mathbf{z}_t. \tag{A-28}
 \end{aligned}$$

令  $\mathcal{N}_{i,t}$  表示在第  $t$  次迭代中第  $i$  个分布式设备没有选择用于计算梯度的样本集，则  $\mathbb{E}[\|\mathbf{o}\|^2]$  可以推导如下

$$\begin{aligned}
 \mathbb{E}[\|\mathbf{o}\|^2] = & \mathbb{E} \left[ \left\| \frac{\sum_{i=1}^U \sum_{k=1}^{K_i} \nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})}{K} \right. \right. \\
 & + \left( \alpha \sum_{i=1}^U K_b \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot-1} \odot \mathbf{z}_t - \left( \sum_{i=1}^U K_b \boldsymbol{\beta}_{i,t} \right)^{\odot-1} \\
 & \left. \left. \odot \sum_{i=1}^U \left( \boldsymbol{\beta}_{i,t} \odot \mathbb{E}_{\mathcal{D}_i} \left[ \sum_{k=1}^{K_b} \nabla f(\mathbf{w}_{t-1}, \mathbf{x}_{i,k}, \mathbf{y}_{i,k}) \right] \right) \right\|^2 \right] \\
 = & \mathbb{E} \left[ \left\| \sum_{i=1}^U \left( \frac{1}{K} - \boldsymbol{\beta}_{i,t} \odot \left( \sum_{i=1}^U K_b \boldsymbol{\beta}_{i,t} \right)^{\odot-1} \right) \right. \right. \\
 & \odot \mathbb{E}_{\overline{\mathcal{N}}_{i,t}} \left[ \sum_{k \in \overline{\mathcal{N}}_{i,t}} \nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k}) \right] \\
 & + \frac{\sum_{i=1}^U \mathbb{E}[\sum_{k \in \mathcal{N}_{i,t}} \nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})]}{K} \\
 & \left. \left. + \left( \alpha \sum_{i=1}^U K_b \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot-1} \odot \mathbf{z}_t \right\|^2 \right] \\
 \leq & \left\| \left( \sum_{i=1}^U K_b \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot-1} \right\|^2 \sigma^2 L^2 + \left( \sum_{i=1}^U K_b \right) \sum_{i=1}^U \left\| \frac{1}{K} - \boldsymbol{\beta}_{i,t} \right. \\
 & \left. \odot \left( \sum_{i=1}^U K_b \boldsymbol{\beta}_{i,t} \right)^{\odot-1} \right\|^2 \mathbb{E}_{\overline{\mathcal{N}}_{i,t}} \left[ \sum_{k \in \overline{\mathcal{N}}_{i,t}} \|\nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})\|^2 \right] \\
 & + \frac{\|\sum_{i=1}^U \mathbb{E}[\sum_{k \in \mathcal{N}_{i,t}} \nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})]\|^2}{K^2}. \tag{A-29}
 \end{aligned}$$

由假设 2.3 可得

$$\begin{aligned}
\mathbb{E}[\|\mathbf{o}\|^2] &\leq \left( \sum_{i=1}^U K_b \right) \sum_{d=1}^D \left( \frac{\left( \sum_{i=1}^U K_b \right) - 2K}{K^2} \right. \\
&\quad \left. + \frac{1}{\sum_{i=1}^U K_b \beta_{i,t}^d} \right) (\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) + \\
&\quad \frac{(\sum_{i=1}^U (K_i - K_b))^2}{K^2} (\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) \\
&\quad + \left\| \left( \sum_{i=1}^U K_b \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot-1} \right\|^2 \sigma^2 L^2. \tag{A-30}
\end{aligned}$$

将公式 (A-30) 代入公式 (A-5) 可得

$$\begin{aligned}
\mathbb{E}[F(\mathbf{w}_t)] &\leq \frac{1}{2L} \left( \left\| \left( \sum_{i=1}^U K_b \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot-1} \right\|^2 \sigma^2 L^2 \right. \\
&\quad + \left( \sum_{d=1}^D \left( \frac{\left( \sum_{i=1}^U K_b \right)^2 - 2K \left( \sum_{i=1}^U K_b \right)}{K^2} + \frac{\left( \sum_{i=1}^U K_b \right)}{\sum_{i=1}^U K_b \beta_{i,t}^d} \right. \right. \\
&\quad \left. \left. + \frac{(\sum_{i=1}^U (K_i - K_b))^2}{K^2} \right) \right) (\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) \\
&\quad \left. + \mathbb{E}[F(\mathbf{w}_{t-1})] - \frac{1}{2L} \|\nabla F(\mathbf{w}_{t-1})\|^2 \right). \tag{A-31}
\end{aligned}$$

从公式 (A-31) 两边均减去  $\mathbb{E}[F(\mathbf{w}^*)]$ ，并利用公式 (A-14) 和公式 (A-16)，可以得出

$$\mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}^*)] \leq B_t^{SGD} + A_t^{SGD} \mathbb{E}[F(\mathbf{w}_{t-1}) - F(\mathbf{w}^*)], \tag{A-32}$$

其中，

$$\begin{aligned}
A_t^{SGD} &= 1 - \frac{\mu}{L} + \rho_2 \left( \sum_{d=1}^D \left( \frac{(\sum_{i=1}^U K_b)^2 - 2K(\sum_{i=1}^U K_b)}{K^2} \right. \right. \\
&\quad \left. \left. + \frac{(\sum_{i=1}^U K_b)}{\sum_{i=1}^U K_b \beta_{i,t}^d} \right) + \frac{(\sum_{i=1}^U (K_i - K_b))^2}{K^2} \right), \tag{A-33}
\end{aligned}$$

$$\begin{aligned}
B_t^{SGD} &= \frac{\rho_1}{2L} \left( \sum_{d=1}^D \left( \frac{(\sum_{i=1}^U K_b)^2 - 2K(\sum_{i=1}^U K_b)}{K^2} \right. \right. \\
&\quad \left. \left. + \frac{(\sum_{i=1}^U K_b)}{\sum_{i=1}^U K_b \beta_{i,t}^d} \right) + \frac{(\sum_{i=1}^U (K_i - K_b))^2}{K^2} \right) \\
&\quad + \left\| \left( \sum_{i=1}^U K_b \boldsymbol{\beta}_{i,t} \odot \mathbf{b}_t \right)^{\odot-1} \right\|^2 \frac{L\sigma^2}{2}. \tag{A-34}
\end{aligned}$$

递归地应用公式 (A-32)，可以得到

$$\begin{aligned} \mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}^*)] &\leq \sum_{i=1}^{t-1} \prod_{j=1}^i A_{t+1-j}^{SGD} B_{t-i}^{SGD} + B_t^{SGD} \\ &+ \prod_{j=1}^t A_j^{SGD} \mathbb{E}[F(\mathbf{w}_0) - F(\mathbf{w}^*)]. \end{aligned} \quad (\text{A-35})$$

**定理 2.1** 由此得证。 ■

#### A4. 定理 2.4 的证明

**证明：**从公式 (2-36)、公式 (2-37) 和公式 (2-38) 可以看出，为了最小化第  $t$  次迭代中的目标损失函数  $R_t$ ，应该最大化所选分布式设备的数量和功率缩放因子。因此，选定的分布式设备应以最大功率发送其参数。为了在参数服务器处达到如公式 (2-5) 中所期望的参数聚合，每个分布式设备需要使用相同的发射功率缩放因子  $b_t$ 。事实上，参数  $b_t$  一旦确定，分布式设备的选择参数  $\beta_t$  也相应确定。因此，参数  $b_t$  才是一个真正需要优化的参数。根据公式 (2-36)、公式 (2-37) 和公式 (2-38)，较大的  $b_t$  会导致较小的  $R_t$ 。另一方面，公式 (2-42) 暗示较大的  $b_t$  会导致可以被选中的分布式设备减少，这反而又会导致  $R_t$  的增加。所以，参数  $b_t$  的优化是一个平衡性问题。

重写公式 (2-39) 中的限定条件并将  $|w_{i,t}|$  替换为  $(|w_{t-1}| + \eta)$ ，可以得到第  $i$  个分布式设备的最大可接受的功率缩放因子  $b_{i,t}^{\max}$  如下

$$b_{i,t}^{\max} = \left\lfloor \frac{\sqrt{P_i^{\max}} h_{i,t}}{K_i(|w_{t-1}| + \eta)} \right\rfloor. \quad (\text{A-36})$$

因此，参数  $b_t$  应该从  $\{b_{i,t}^{\max}\}_{i=1}^U$  中选择。一旦确定了参数  $b_t$ ，就可以通过验证发射功率是否满足公式 (2-42) 中的条件来确定  $\beta_t$ 。由此可得到优化问题 P2 的缩减解空间为

$$\begin{aligned} \mathcal{S} = \left\{ \left\{ \left( b_t^{(k)}, \beta_{i,t}^{(k)} \right) \right\}_{k=1}^U \middle| b_t^{(k)} = b_{k,t}^{\max}, \right. \\ \left. \beta_t^{(k)} \left( b_t^{(k)} \right) = \left[ \beta_{1,t}^{(k)}, \dots, \beta_{U,t}^{(k)} \right], k = 1, \dots, U \right\}, \end{aligned} \quad (\text{A-37})$$

其中，

$$\beta_{U,t}^{(k)} = H \left( P_U^{\max} - \left\lceil \frac{K_U b_t^{(k)} (|w_{t-1}| + \eta)}{h_{U,t}} \right\rceil \right), \quad (\text{A-38})$$

Heaviside 阶跃函数为

$$H(x) = \begin{cases} 1, & x > 0, \\ 0, & x \leq 0. \end{cases} \quad (\text{A-39})$$

定理 2.4 证毕。 ■

## A5. 引理 3.1 的证明

**证明：**在假设 3.4 成立的情况下，由于稀疏化导致的误差  $\mathbf{e}_{i,t}^s \in \mathbb{R}^D, \forall i, t$  满足下列不等式

$$\mathbb{E}\|\mathbf{e}_{i,t}^s\|^2 = \mathbb{E}\|\tilde{\mathbf{g}}_{i,t} - \mathbf{g}_{i,t}\|^2 \leq \frac{D-\kappa}{D}G^2. \quad (\text{A-40})$$

由于测量矩阵  $\Phi$  满足 RIP 条件<sup>[120]</sup>，则

$$(1-\delta)\|\mathbf{x}\|^2 \leq \|\Phi\mathbf{x}\|^2 \leq (1+\delta)\|\mathbf{x}\|^2, \quad (\text{A-41})$$

其中  $\mathbf{x}$  是一个  $k$ -稀疏的向量。由此，量化误差  $\mathbf{e}_{i,t}^q \in \mathbb{R}^S$  可以被推导如下

$$\begin{aligned} \mathbb{E}\|\mathbf{e}_{i,t}^q\|^2 &= \mathbb{E}\|\text{sign}(\Phi\tilde{\mathbf{g}}_{i,t}) - \Phi\tilde{\mathbf{g}}_{i,t}\|^2 \\ &\leq \mathbb{E}(\|\text{sign}(\Phi\tilde{\mathbf{g}}_{i,t})\|^2 + \|\Phi\tilde{\mathbf{g}}_{i,t}\|^2) \\ &\leq S + (1+\delta)\frac{D-\kappa}{D}G^2. \end{aligned} \quad (\text{A-42})$$

当参数服务器通过公式 (3-13) 中获得  $\hat{\mathbf{y}}_t^{desired}$  时，它会在存在范数限制的测量误差  $\mathbf{e}_t^r$  的情况下重构信号  $\hat{\mathbf{g}}_t$ 。在参数服务器处，通过求解下列优化问题可以实现稳健的信号重建<sup>[102]</sup>

$$\hat{\mathbf{g}}_t = \arg \min_{\tilde{\mathbf{g}}_t} \|\tilde{\mathbf{g}}_t\|_1 \quad \text{s.t. } \|\hat{\mathbf{y}}_t^{desired} - \Phi\tilde{\mathbf{g}}_t\|^2 \leq \varepsilon_t, \quad (\text{A-43})$$

其中  $\varepsilon_t$  是范数限制边界，由下式给出

$$\begin{aligned} \mathbb{E}\|\hat{\mathbf{y}}_t^{desired} - \Phi\tilde{\mathbf{g}}_t\|^2 &= \mathbb{E}\left\|\hat{\mathbf{y}}_t^{desired} - \frac{\sum_{i=1}^U K_i \beta_{i,t} (\Phi\tilde{\mathbf{g}}_{i,t})}{\sum_{i=1}^U K_i \beta_{i,t}}\right\|^2 \\ &= \mathbb{E}\left\|\frac{\sum_{i=1}^U K_i \beta_{i,t} \mathbf{e}_{i,t}^q}{\sum_{i=1}^U K_i \beta_{i,t}} + \frac{\mathbf{z}_t}{\sum_{i=1}^U K_i \beta_{i,t} b_t}\right\|^2 \\ &= \mathbb{E}\left\|\mathbf{e}_{1,t}^q + \frac{\mathbf{z}_t}{\sum_{i=1}^U K_i \beta_{i,t} b_t}\right\|^2 \\ &\leq \mathbb{E}\|\mathbf{e}_{1,t}^q\|^2 + \mathbb{E}\left\|\frac{\mathbf{z}_t}{\sum_{i=1}^U K_i \beta_{i,t} b_t}\right\|^2 \\ &\leq S + (1+\delta)\frac{D-\kappa}{D}G^2 + \frac{S\sigma^2}{\left(\sum_{i=1}^U K_i \beta_{i,t} b_t\right)^2} \doteq \varepsilon_t. \end{aligned} \quad (\text{A-44})$$

在这种情况下，重构误差范数上界可以表示为

$$\|\hat{\mathbf{g}}_t - \tilde{\mathbf{g}}_t\|^2 \leq \frac{C^2}{S} \varepsilon_t, \quad (\text{A-45})$$

其中  $C$  是常数，其取值取决于测量矩阵  $\Phi$  的属性，但与信号本身无关<sup>[121]</sup>。根据文献[120]中的定理 1.2，如果  $\Phi$  满足  $\delta \leq \sqrt{2} - 1$ ，则  $C$  可以由下式给出

$$C = \frac{2\varpi}{1-\varrho}, \quad (\text{A-46})$$

其中， $\varpi = \frac{2\sqrt{1+\delta}}{\sqrt{1-\delta}}$ ,  $\varrho = \frac{\sqrt{2}\delta}{1-\delta}$ 。

注意到公式 (A-45) 中的  $\tilde{\mathbf{g}}_t$  是经过分布式设备选择后，参数服务器处所需的稀疏全局梯度。因此，基于 1 比特压缩感知的 FLOA 中第  $t$  次迭代后的总误差的上界由下列不等式给出

$$\begin{aligned} \mathbb{E}\|\mathbf{e}_t\|^2 &= \mathbb{E}(\|\hat{\mathbf{g}}_t - \mathbf{g}_t\|^2) = \mathbb{E}(\|\hat{\mathbf{g}}_t - (\tilde{\mathbf{g}}_t + \mathbf{e}_t^s)\|^2) \\ &\leq \mathbb{E}(\|\hat{\mathbf{g}}_t - \tilde{\mathbf{g}}_t\| + \|\mathbf{e}_t^s\|)^2 \leq \mathbb{E}(2\|\hat{\mathbf{g}}_t - \tilde{\mathbf{g}}_t\|^2 + 2\|\mathbf{e}_t^s\|^2) \\ &\leq 2\frac{C^2}{S}\varepsilon_t + 2\sum_{i=1}^U \beta_{i,t} \frac{D-\kappa}{D} G^2 \\ &= 2C^2 \left( 1 + (1+\delta) \frac{D-\kappa}{SD} G^2 + \frac{\sigma^2}{\left(\sum_{i=1}^U K_i \beta_{i,t} b_t\right)^2} \right) + 2\sum_{i=1}^U \beta_{i,t} \frac{D-\kappa}{D} G^2, \quad (\text{A-47}) \end{aligned}$$

其中， $\mathbf{e}_t^s = \sum_{i=1}^U \beta_{i,t} \mathbf{e}_{i,t}^s$ 。

引理 3.1 得证。 ■

## A6. 定理 3.1 的证明

证明：为了证明定理 3.1，首先将  $F(\mathbf{w}_t)$  重写为其二阶泰勒展开式的表达式，由下式给出

$$\begin{aligned} F(\mathbf{w}_t) &= F(\mathbf{w}_{t-1}) + (\mathbf{w}_t - \mathbf{w}_{t-1})^T \nabla F(\mathbf{w}_{t-1}) \\ &\quad + \frac{1}{2} (\mathbf{w}_t - \mathbf{w}_{t-1})^T \nabla^2 F(\mathbf{w}_{t-1}) (\mathbf{w}_t - \mathbf{w}_{t-1}) \\ &\stackrel{(a)}{\leq} F(\mathbf{w}_{t-1}) + (\mathbf{w}_t - \mathbf{w}_{t-1})^T \nabla F(\mathbf{w}_{t-1}) + \frac{L}{2} \|\mathbf{w}_t - \mathbf{w}_{t-1}\|^2, \quad (\text{A-48}) \end{aligned}$$

其中，步骤(a)应用了假设 3.2。

通过求解公式 (A-45) 中的优化问题，从接收信号中恢复所需的  $\hat{\mathbf{g}}_t$  后，全局模型更新为

$$\begin{aligned} \mathbf{w}_t &= \mathbf{w}_{t-1} - \alpha \hat{\mathbf{g}}_t \\ &= \mathbf{w}_{t-1} - \alpha (\nabla F(\mathbf{w}_{t-1}) - \mathbf{o}), \end{aligned} \quad (\text{A-49})$$

其中，

$$\mathbf{o} = \nabla F(\mathbf{w}_{t-1}) - \hat{\mathbf{g}}_t. \quad (\text{A-50})$$

给定学习速率  $\alpha = \frac{1}{L}$  (一个特殊的设置，用于更简单的表达而不失一般性)，

那么来自公式 (A-48) 的  $\mathbb{E}[F(\mathbf{w}_t)]$  的期望优化函数可以表示为

$$\begin{aligned} \mathbb{E}[F(\mathbf{w}_t)] &\leq \mathbb{E}\left[F(\mathbf{w}_{t-1}) - \alpha(\nabla F(\mathbf{w}_{t-1}) - \mathbf{o})^T \nabla F(\mathbf{w}_{t-1})\right. \\ &\quad \left.+ \frac{L\alpha^2}{2} \|\nabla F(\mathbf{w}_{t-1}) - \mathbf{o}\|^2\right] \\ &\stackrel{(b)}{=} \mathbb{E}[F(\mathbf{w}_{t-1})] - \frac{1}{2L} \|\nabla F(\mathbf{w}_{t-1})\|^2 + \frac{1}{2L} \mathbb{E}[\|\mathbf{o}\|^2], \end{aligned} \quad (\text{A-51})$$

其中，上式中的步骤(b)的推导是基于以下事实

$$\frac{L\alpha^2}{2} \|\nabla F(\mathbf{w}_{t-1}) - \mathbf{o}\|^2 = \frac{1}{2L} \|\nabla F(\mathbf{w}_{t-1})\|^2 - \frac{1}{L} \mathbf{o}^T \nabla F(\mathbf{w}_{t-1}) + \frac{1}{2L} \|\mathbf{o}\|^2. \quad (\text{A-52})$$

根据公式 (A-47)，可得  $\|\mathbf{e}_t\|^2 \leq \frac{2C^2}{S} \varepsilon_t + 2 \sum_{i=1}^U \beta_{i,t} \frac{D-\kappa}{D} G^2$ 。进而可以推导出  $\mathbb{E}[\|\mathbf{o}\|^2]$  如下

$$\begin{aligned} \mathbb{E}[\|\mathbf{o}\|^2] &= \mathbb{E}[\|\nabla F(\mathbf{w}_{t-1}) - \hat{\mathbf{g}}_t\|] = \mathbb{E}[\|\nabla F(\mathbf{w}_{t-1}) - \mathbf{g}_t - \mathbf{e}_t\|] \\ &= \mathbb{E}\left[\left\|\frac{\sum_{i=1}^U \sum_{k=1}^{K_i} \nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})}{K} \right.\right. \\ &\quad \left.\left.- (\sum_{i=1}^U K_i \beta_{i,t})^{-1} \sum_{i=1}^U \sum_{k=1}^{K_i} \beta_{i,t} \nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k}) - \mathbf{e}_t\right\|^2\right] \\ &\leq \mathbb{E}\left[\left\|\left(\frac{1}{K} - \frac{\beta_{i,t}}{\sum_{i=1}^U K_i \beta_{i,t}}\right) \sum_{k=1}^{K_i} \nabla f(\mathbf{w}_{t-1}, \mathbf{x}_{i,k}, \mathbf{y}_{i,k}) - \mathbf{e}_t\right\|^2\right]. \end{aligned} \quad (\text{A-53})$$

应用范数的三角不等式： $\|\mathbf{X} + \mathbf{Y}\| \leq \|\mathbf{X}\| + \|\mathbf{Y}\|$ ，范数的相容性质： $\|\mathbf{XY}\| \leq \|\mathbf{X}\|\|\mathbf{Y}\|$ ，以及 Jensen 不等式： $(\sum_{i=1}^n a_i)^2 \leq n \sum_{i=1}^n a_i^2$ ，可以进一步推导公式 (A-53) 如下

$$\begin{aligned} \mathbb{E}[\|\mathbf{o}\|^2] &\leq \mathbb{E}\left[2\left\|\left(\frac{1}{K} - \frac{\beta_{i,t}}{\sum_{i=1}^U K_i \beta_{i,t}}\right) \sum_{k=1}^{K_i} \nabla f(\mathbf{w}_{t-1}, \mathbf{x}_{i,k}, \mathbf{y}_{i,k})\right\|^2\right] \\ &\quad + \mathbb{E}[2\|\mathbf{e}_{t-1}\|^2] \\ &\leq \mathbb{E}\left[2(U+K) \sum_{i=1}^U \left(\frac{1}{K} - \frac{\beta_{i,t}}{\sum_{i=1}^U K_i \beta_{i,t}}\right)^2 \sum_{k=1}^{K_i} \|\nabla f(\mathbf{w}_{t-1}, \mathbf{x}_{i,k}, \mathbf{y}_{i,k})\|^2\right] \\ &\quad + 2\mathbb{E}[\|\mathbf{e}_{t-1}\|^2] \\ &\leq \frac{4C^2}{S} \varepsilon_t + 4 \sum_{i=1}^U \beta_{i,t} \frac{D-\kappa}{D} G^2 \\ &\quad + 2(U+K) \sum_{i=1}^U \left(\frac{1}{K} - \frac{\beta_{i,t}}{\sum_{i=1}^U K_i \beta_{i,t}}\right)^2 \sum_{k=1}^{K_i} \|\nabla f(\mathbf{w}_{t-1}, \mathbf{x}_{i,k}, \mathbf{y}_{i,k})\|^2. \end{aligned} \quad (\text{A-54})$$

将假设 3.3 中的公式 (3-17) 应用于公式 (A-54)，可以进一步推导出以下的结果

$$\begin{aligned}
\mathbb{E}[\|\mathbf{o}\|^2] &\leq 2(U+K) \sum_{i=1}^U \left( \frac{1}{K} - \frac{\beta_{i,t}}{\sum_{i=1}^U K_i \beta_{i,t}} \right)^2 K_i (\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) \\
&\quad + \frac{4C^2}{S} \varepsilon_t + 4 \sum_{i=1}^U \beta_{i,t} \frac{D-\kappa}{D} G^2 \\
&= 2(U+K) \left( \frac{1}{\sum_{i=1}^U K_i \beta_{i,t}} - \frac{1}{K} \right) (\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) \\
&\quad + \frac{4C^2}{S} \varepsilon_t + \sum_{i=1}^U 4\beta_{i,t} \frac{D-\kappa}{D} G^2 \\
&\leq \frac{2(U+K) \sum_{i=1}^U K_i (1 - \beta_{i,t})}{K} (\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) \\
&\quad + \frac{4C^2}{S} \varepsilon_t + 4 \sum_{i=1}^U \beta_{i,t} \frac{D-\kappa}{D} G^2.
\end{aligned} \tag{A-55}$$

将公式 (A-55) 代入 (A-51) 中，可得

$$\begin{aligned}
\mathbb{E}[F(\mathbf{w}_t)] &\leq \frac{2(U+K) \sum_{i=1}^U K_i (1 - \beta_{i,t})}{2LK} (\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) \\
&\quad + \frac{4C^2 \varepsilon_t}{2LS} + 4 \sum_{i=1}^U \beta_{i,t} \frac{D-\kappa}{2LD} G^2 + \mathbb{E}[F(\mathbf{w}_{t-1})] - \frac{1}{2L} \|\nabla F(\mathbf{w}_{t-1})\|^2 \\
&= \mathbb{E}[F(\mathbf{w}_{t-1})] + \frac{\rho_1 (U+K) \sum_{i=1}^U K_i (1 - \beta_{i,t})}{LK} \\
&\quad + \left( \frac{\rho_2 (U+K) \sum_{i=1}^U K_i (1 - \beta_{i,t})}{LK} - \frac{1}{2L} \right) \|\nabla F(\mathbf{w}_{t-1})\|^2 \\
&\quad + \frac{2}{L} \left( \frac{C^2}{S} \varepsilon_t + \sum_{i=1}^U \beta_{i,t} \frac{D-\kappa}{D} G^2 \right).
\end{aligned} \tag{A-56}$$

将上述不等式从  $t = 1$  到  $t = T$  加起来，可以得到

$$\mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}_0)] \leq - \sum_{t=1}^T A_t \|\nabla F(\mathbf{w}_{t-1})\|^2 + \sum_{t=1}^T B_t, \tag{A-57}$$

其中，

$$A_t = \frac{1}{2L} - \frac{\rho_2 (U+K) \sum_{i=1}^U K_i (1 - \beta_{i,t})}{LK}, \tag{A-58}$$

$$B_t = \frac{\rho_1 (U+K) \sum_{i=1}^U K_i (1 - \beta_{i,t})}{LK} + \frac{2}{L} \left( \frac{C^2}{S} \varepsilon_t + \sum_{i=1}^U \beta_{i,t} \frac{D-\kappa}{D} G^2 \right). \tag{A-59}$$

不等式 (A-57) 可以被重新描述为

$$\begin{aligned}
\sum_{t=1}^T A_t \|\nabla F(\mathbf{w}_{t-1})\|^2 &\leq \mathbb{E}[F(\mathbf{w}_0) - F(\mathbf{w}_t)] + \sum_{t=1}^T B_t \\
&\leq \mathbb{E}[F(\mathbf{w}_0) - F(\mathbf{w}^*)] + \sum_{t=1}^T B_t. \tag{A-60}
\end{aligned}$$

由  $\frac{1}{2L} - \frac{\rho_2(U+K)}{L} \leq A_t \leq \frac{1}{2L}$ , 可以得到

$$\begin{aligned}
\frac{1}{T} \sum_{t=1}^T \left( \frac{1}{2L} - \frac{\rho_2(U+K)}{L} \right) \|\nabla F(\mathbf{w}_{t-1})\|^2 &\leq \frac{1}{T} \sum_{t=1}^T A_t \|\nabla F(\mathbf{w}_{t-1})\|^2 \\
&\leq \frac{1}{T} \mathbb{E}[F(\mathbf{w}_0) - F(\mathbf{w}^*)] + \frac{1}{T} \sum_{t=1}^T B_t. \tag{A-61}
\end{aligned}$$

综上可得,

$$\begin{aligned}
\frac{1}{T} \sum_{t=1}^T \|\nabla F(\mathbf{w}_{t-1})\|^2 &\leq \frac{2L}{T(1-2\rho_2(U+K))} \mathbb{E}[F(\mathbf{w}_0) - F(\mathbf{w}^*)] \\
&\quad + \frac{2L}{T(1-2\rho_2(U+K))} \sum_{t=1}^T B_t. \tag{A-62}
\end{aligned}$$

定理 3.1 证毕。 ■

## A7. 定理 3.2 的证明

证明: 对于在基于 1 比特压缩感知的 FLOA 中应用 SGD 的方法, 第*i*个分布式设备的本地梯度在第*t*次迭代时的计算为

$$\mathbf{g}_{i,t} = \mathbb{E}_{\mathfrak{D}_i} \left[ \frac{\sum_{k=1}^{K_b} \nabla f(\mathbf{w}_{t-1}, \mathbf{x}_{i,k}, \mathbf{y}_{i,k})}{K_b} \right], \quad i = 1, 2, \dots, U, \tag{A-63}$$

其中,  $\mathbb{E}_{\mathfrak{D}_i}[\cdot]$  为取期望值操作, 表示第*i*个分布式设备从其本地数据集  $\mathfrak{D}_i$  中随机选择  $K_b$  个样本来计算本地梯度。

给定公式 (A-63), 原本公式 (3-9) 中面向 GD 情况下第*t*次迭代中参数服务器 处的期望信号向量现在表示为

$$\mathbf{y}_t^{desired} = \frac{\sum_{i=1}^U K_b \beta_{i,t} \mathcal{C}(\mathbf{g}_{i,t})}{\sum_{i=1}^U K_b \beta_{i,t}} = \frac{\sum_{i=1}^U \beta_{i,t} \mathcal{C}(\mathbf{g}_{i,t})}{\sum_{i=1}^U \beta_{i,t}}. \tag{A-64}$$

因此, 通过公式 (3-13) 中的后处理操作来估计感兴趣的信号向量被重描述为

$$\begin{aligned}
\hat{\mathbf{y}}_t^{desired} &= \left( \sum_{i=1}^U K_b \beta_{i,t} b_t \right)^{-1} \mathbf{y}_t \\
&= \left( \sum_{i=1}^U K_b \beta_{i,t} \right)^{-1} \sum_{i=1}^U K_b \beta_{i,t} \mathcal{C}(\mathbf{g}_{i,t}) + \left( \sum_{i=1}^U K_b \beta_{i,t} b_t \right)^{-1} \mathbf{z}_t \\
&= \mathbf{y}_t^{desired} + \frac{\mathbf{z}_t}{\sum_{i=1}^U K_b \beta_{i,t} b_t}, \tag{A-65}
\end{aligned}$$

其中,  $(\sum_{i=1}^U K_b \beta_{i,t} b_t)^{-1}$  用作后处理因子。

于是, 得到在公式 (A-43) 中定义的范数边界  $\varepsilon_t^{\text{sgd}}$  为

$$\begin{aligned} \varepsilon_t^{\text{sgd}} &\doteq \mathbb{E} \|\hat{\mathbf{y}}_t^{\text{desired}} - \Phi \tilde{\mathbf{g}}_t\|^2 = \mathbb{E} \left\| \hat{\mathbf{y}}_t^{\text{desired}} - \frac{\sum_{i=1}^U K_b \beta_{i,t} (\Phi \tilde{\mathbf{g}}_{i,t})}{\sum_{i=1}^U K_b \beta_{i,t}} \right\|^2 \\ &= \mathbb{E} \left\| \frac{\sum_{i=1}^U \beta_{i,t} \mathbf{e}_{i,t}^q}{\sum_{i=1}^U \beta_{i,t}} + \frac{\mathbf{z}_t}{\sum_{i=1}^U K_b \beta_{i,t} b_t} \right\|^2 = \mathbb{E} \left\| \mathbf{e}_{1,t}^q + \frac{\mathbf{z}_t}{\sum_{i=1}^U K_b \beta_{i,t} b_t} \right\|^2 \\ &\leq \mathbb{E} \|\mathbf{e}_{1,t}^q\|^2 + \mathbb{E} \left\| \frac{\mathbf{z}_t}{\sum_{i=1}^U K_b \beta_{i,t} b_t} \right\|^2 \\ &\leq S + (1 + \delta) \frac{D - \kappa}{D} G^2 + \frac{S \sigma^2}{\left( \sum_{i=1}^U K_b \beta_{i,t} b_t \right)^2}. \end{aligned} \quad (\text{A-66})$$

因此, 经过第  $t$  次迭代产生的总误差为

$$\begin{aligned} \mathbb{E} \|\mathbf{e}_t\|^2 &= \mathbb{E} (\|\hat{\mathbf{g}}_t - \mathbf{g}_t\|^2) = \mathbb{E} (\|\hat{\mathbf{g}}_t - (\tilde{\mathbf{g}}_t + \mathbf{e}_t^s)\|^2) \\ &\leq \mathbb{E} (2 \|\hat{\mathbf{g}}_t - \tilde{\mathbf{g}}_t\|^2 + 2 \|\mathbf{e}_t^s\|^2) \leq \frac{2C^2}{S} \varepsilon_t^{\text{sgd}} + 2 \sum_{i=1}^U \beta_{i,t} \frac{D - \kappa}{D} G^2 \\ &= 2C^2 \left( 1 + (1 + \delta) \frac{D - \kappa}{SD} G^2 + \frac{\sigma^2}{\left( \sum_{i=1}^U K_b \beta_{i,t} b_t \right)^2} \right) + 2 \sum_{i=1}^U \beta_{i,t} \frac{D - \kappa}{D} G^2. \end{aligned} \quad (\text{A-67})$$

令  $\mathcal{N}_{i,t}$  表示在第  $t$  次迭代中第  $i$  个分布式设备没有选择的样本集, 则有

$$\begin{aligned} \mathbb{E} [\|\mathbf{o}\|^2] &= \mathbb{E} [\|\nabla F(\mathbf{w}_{t-1}) - \hat{\mathbf{g}}_t\|] = \mathbb{E} [\|\nabla F(\mathbf{w}_{t-1}) - \mathbf{g}_t - \mathbf{e}_t\|] \\ &= \mathbb{E} \left[ \left\| \frac{\sum_{i=1}^U \sum_{k=1}^{K_i} \nabla f(\mathbf{w}_{t-1}; \mathbf{x}_{i,k}, \mathbf{y}_{i,k})}{K} \right. \right. \\ &\quad \left. \left. - \left( \sum_{i=1}^U K_b \beta_{i,t} \right)^{-1} \sum_{i=1}^U \beta_{i,t} \mathbb{E}_{\mathfrak{D}_i} \left[ \sum_{k=1}^{K_b} \nabla f(\mathbf{w}_{t-1}, \mathbf{x}_{i,k}, \mathbf{y}_{i,k}) \right] - \mathbf{e}_t \right\|^2 \right] \\ &\leq \mathbb{E} \left[ \left\| \sum_{i=1}^U \left( \frac{1}{K} - \frac{\beta_{i,t}}{\sum_{i=1}^U K_i \beta_{i,t}} \right) \mathbb{E}_{\mathcal{N}_{i,t}} \left[ \sum_{k \in \mathcal{N}_{i,t}} \nabla f(\mathbf{w}_{t-1}, \mathbf{x}_{i,k}, \mathbf{y}_{i,k}) \right] \right. \right. \\ &\quad \left. \left. + \frac{\sum_{i=1}^U \mathbb{E} [\sum_{k \in \mathcal{N}_{i,t}} \nabla f(\mathbf{w}_{t-1}, \mathbf{x}_{i,k}, \mathbf{y}_{i,k})]}{K} - \mathbf{e}_t \right\|^2 \right] \\ &\leq \mathbb{E} \left[ 2 \left\| \sum_{i=1}^U \left( \frac{1}{K} - \frac{\beta_{i,t}}{\sum_{i=1}^U K_i \beta_{i,t}} \right) \mathbb{E}_{\mathcal{N}_{i,t}} \left[ \sum_{k \in \mathcal{N}_{i,t}} \nabla f(\mathbf{w}_{t-1}, \mathbf{x}_{i,k}, \mathbf{y}_{i,k}) \right] \right. \right. \\ &\quad \left. \left. + \frac{\sum_{i=1}^U \mathbb{E} [\sum_{k \in \mathcal{N}_{i,t}} \nabla f(\mathbf{w}_{t-1}, \mathbf{x}_{i,k}, \mathbf{y}_{i,k})]}{K} \right\|^2 \right] + 2 \|\mathbf{e}_t\|^2, \\ &\leq \left( 4 \sum_{i=1}^U K_b \right) \sum_{i=1}^U \left( \frac{1}{K} - \beta_{i,t} \left( \sum_{i=1}^U K_b \beta_{i,t} \right)^{-1} \right)^2 \mathbb{E}_{\mathcal{N}_{i,t}} \left[ \sum_{k \in \mathcal{N}_{i,t}} \|\nabla f(\mathbf{w}_{t-1}, \mathbf{x}_{i,k}, \mathbf{y}_{i,k})\|^2 \right] \\ &\quad + 2 \frac{\|\sum_{i=1}^U \mathbb{E} [\sum_{k \in \mathcal{N}_{i,t}} \nabla f(\mathbf{w}_{t-1}, \mathbf{x}_{i,k}, \mathbf{y}_{i,k})]\|^2}{K^2} + 2 \|\mathbf{e}_t\|^2. \end{aligned} \quad (\text{A-68})$$

将假设 3.3 中的公式 (3-17) 应用到公式 (A-68) 中, 可以进一步推导出以下结果

$$\begin{aligned}
\mathbb{E}[\|\mathbf{o}\|^2] &\leq 4UK_b \sum_{i=1}^U \left( \frac{1}{K} - \beta_{i,t} \left( \sum_{i=1}^U K_b \beta_{i,t} \right)^{-1} \right)^2 K_b (\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) \\
&+ 2 \frac{\sum_{i=1}^U (K_i - K_b)(\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2)}{K^2} + 2\|\mathbf{e}_t\|^2 \\
&\leq 4UK_b \sum_{i=1}^U \left( \frac{K_b}{K^2} - 2K_b \frac{1}{K} \beta_{i,t} \left( \sum_{i=1}^U K_b \beta_{i,t} \right)^{-1} \right. \\
&+ K_b \beta_{i,t} \left( \sum_{i=1}^U K_b \beta_{i,t} \right)^{-2} \left. \right) (\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) \\
&+ 2 \frac{\sum_{i=1}^U (K_i - K_b)(\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2)}{K^2} + 2\|\mathbf{e}_t\|^2 \\
&\leq 4UK_b \left( \frac{UK_b}{K^2} - 2 \frac{1}{K} + \left( \sum_{i=1}^U K_b \beta_{i,t} \right)^{-1} \right) (\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) \\
&+ 2 \frac{(K - UK_b)(\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2)}{K^2} + \frac{4C^2}{S} \varepsilon_t^{\text{sgd}} + 4 \sum_{i=1}^U \beta_{i,t} \frac{D - \kappa}{D} G^2. \quad (\text{A-69})
\end{aligned}$$

将公式 (A-69) 代入公式 (A-51) 中可得

$$\begin{aligned}
\mathbb{E}[F(\mathbf{w}_t)] &\leq \frac{\vartheta + 2UK^2 \left( \sum_{i=1}^U \beta_{i,t} \right)^{-1}}{LK^2} (\rho_1 + \rho_2 \|\nabla F(\mathbf{w}_{t-1})\|^2) \\
&+ \frac{4C^2 \varepsilon_t^{\text{sgd}}}{2LS} + 4 \sum_{i=1}^U \beta_{i,t} \frac{D - \kappa}{2LD} G^2 + \mathbb{E}[F(\mathbf{w}_{t-1})] - \frac{1}{2L} \|\nabla F(\mathbf{w}_{t-1})\|^2 \\
&= \mathbb{E}[F(\mathbf{w}_{t-1})] + \frac{\rho_1 \left( \vartheta + 2UK^2 \left( \sum_{i=1}^U \beta_{i,t} \right)^{-1} \right)}{LK^2} \\
&+ \left( \frac{\rho_2 \left( \vartheta + 2UK^2 \left( \sum_{i=1}^U \beta_{i,t} \right)^{-1} \right)}{LK^2} - \frac{1}{2L} \right) \|\nabla F(\mathbf{w}_{t-1})\|^2 \\
&+ \frac{2}{L} \left( \frac{C^2}{S} \varepsilon_t + \sum_{i=1}^U \beta_{i,t} \frac{D - \kappa}{D} G^2 \right). \quad (\text{A-70})
\end{aligned}$$

其中,  $\vartheta = 2U^2 K_b^2 + K - 4KUK_b - UK_b$ 。

将上述不等式从  $t = 1$  到  $T$  加起来, 可以得到

$$\mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}_0)] \leq - \sum_{t=1}^T A_t^{\text{sgd}} \|\nabla F(\mathbf{w}_{t-1})\|^2 + \sum_{t=1}^T B_t^{\text{sgd}}, \quad (\text{A-71})$$

其中,

$$A_t^{\text{sgd}} = \frac{1}{2L} - \frac{\rho_2 (\vartheta + 2UK^2 (\sum_{i=1}^U \beta_{i,t})^{-1})}{LK^2}, \quad (\text{A-72})$$

$$B_t^{\text{sgd}} = \frac{\rho_1 (\vartheta + 2UK^2 (\sum_{i=1}^U \beta_{i,t})^{-1})}{LK^2} + \frac{2}{L} \left( \frac{C^2}{S} \varepsilon_t + \sum_{i=1}^U \beta_{i,t} \frac{D-\kappa}{D} G^2 \right). \quad (\text{A-73})$$

不等式 (A-73) 可以被重新描述为

$$\begin{aligned} \sum_{t=1}^T A_t^{\text{sgd}} \| \nabla F(\mathbf{w}_{t-1}) \|^2 &\leq \mathbb{E}[F(\mathbf{w}_0) - F(\mathbf{w}_t)] + \sum_{t=1}^T B_t^{\text{sgd}} \\ &\leq \mathbb{E}[F(\mathbf{w}_0) - F(\mathbf{w}^*)] + \sum_{t=1}^T B_t^{\text{sgd}}. \end{aligned} \quad (\text{A-74})$$

由  $\frac{1}{2L} - \frac{\rho_2(\vartheta+2UK^2)}{LK^2} \leq A_t^{\text{sgd}} \leq \frac{1}{2L}$ , 可以得到

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \left( \frac{1}{2L} - \frac{\rho_2(\vartheta+2UK^2)}{LK^2} \right) \| \nabla F(\mathbf{w}_{t-1}) \|^2 &\leq \frac{1}{T} \sum_{t=1}^T A_t^{\text{sgd}} \| \nabla F(\mathbf{w}_{t-1}) \|^2 \\ &\leq \frac{1}{T} \mathbb{E}[F(\mathbf{w}_0) - F(\mathbf{w}^*)] + \frac{1}{T} \sum_{t=1}^T B_t^{\text{sgd}}. \end{aligned} \quad (\text{A-75})$$

综上可得,

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \| \nabla F(\mathbf{w}_{t-1}) \|^2 &\leq \frac{2LK^2}{T(K^2 - 2\rho_2(\vartheta + 2UK^2))} \mathbb{E}[F(\mathbf{w}_0) - F(\mathbf{w}^*)] \\ &\quad + \frac{2LK^2}{T(K^2 - 2\rho_2(\vartheta + 2UK^2))} \sum_{t=1}^T B_t^{\text{sgd}}. \end{aligned} \quad (\text{A-76})$$

**定理 3.2** 证毕。 ■

## A8. 定理 4.1 的证明

**证明:** 给定公式 (4-7) 中全局梯度的估计, 可得模型参数的更新规则如下

$$\begin{aligned} \mathbf{w}_t &= \mathbf{w}_{t-1} - \alpha \tilde{\mathbf{g}}_t \\ &= \mathbf{w}_{t-1} - \alpha \left( \sum_{m=1}^M p_{m,t} |h_{m,t}| \mathbf{g}_{m,t} + \epsilon_t \sum_{n=1}^N \hat{p}_{n,t} |h_{n,t}| \hat{\mathbf{g}}_{n,t} \right. \\ &\quad \left. + \sum_{n=1}^N p_{n,t} |h_{n,t}| \bar{g}_t \mathbf{1} + \epsilon_t \mathbf{z}_t \right). \end{aligned} \quad (\text{A-77})$$

将上述公式 (A-77) 代入公式 (4-12) 可得

$$\begin{aligned}
F(\mathbf{w}_t) &\leq F(\mathbf{w}_{t-1}) + \mathbf{g}_t^T(\mathbf{w}_t - \mathbf{w}_{t-1}) + \frac{L}{2}\|\mathbf{w}_t - \mathbf{w}_{t-1}\|^2 \\
&= F(\mathbf{w}_{t-1}) - \alpha \mathbf{g}_t^T \left( \sum_{m=1}^M p_{m,t} |h_{m,t}| \mathbf{g}_{m,t} \right. \\
&\quad \left. + \epsilon_t \sum_{n=1}^N \hat{p}_{n,t} |h_{n,t}| \hat{\mathbf{g}}_{n,t} + \sum_{n=1}^N p_{n,t} |h_{n,t}| \bar{g}_t \mathbf{1} + \epsilon_t \mathbf{z}_t \right) \\
&\quad + \frac{\alpha^2 L}{2} \left\| \sum_{m=1}^M p_{m,t} |h_{m,t}| \mathbf{g}_{m,t} + \epsilon_t \sum_{n=1}^N \hat{p}_{n,t} |h_{n,t}| \hat{\mathbf{g}}_{n,t} \right. \\
&\quad \left. + \sum_{n=1}^N p_{n,t} |h_{n,t}| \bar{g}_t \mathbf{1} + \epsilon_t \mathbf{z}_t \right\|^2. \tag{A-78}
\end{aligned}$$

重写上述不等式 (A-78) 并两边分别取期望, 可得

$$\begin{aligned}
\mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}_{t-1})] &\leq -\alpha \mathbf{g}_t^T \left( \sum_{m=1}^M p_{m,t} |h_{m,t}| \mathbf{g}_t \right. \\
&\quad \left. + \sum_{n=1}^N \hat{p}_{n,t} |h_{n,t}| \hat{\mathbf{g}}_{n,t} + \sum_{n=1}^N p_{n,t} |h_{n,t}| \mathbb{E}[\bar{g}_t] \mathbf{1} \right) \\
&\quad + \frac{\alpha^2 L}{2} \mathbb{E} \left[ \left\| \sum_{m=1}^M p_{m,t} |h_{m,t}| \mathbf{g}_{m,t} + \epsilon_t \sum_{n=1}^N \hat{p}_{n,t} |h_{n,t}| \hat{\mathbf{g}}_{n,t} \right. \right. \\
&\quad \left. \left. + \sum_{n=1}^N p_{n,t} |h_{n,t}| \bar{g}_t \mathbf{1} + \epsilon_t \mathbf{z}_t \right\|^2 \right]. \tag{A-79}
\end{aligned}$$

由于  $\mathbf{g}_t^T \mathbb{E}[\bar{g}_t] \mathbf{1} = \mathbf{g}_t^T \frac{\sum_{d=1}^D g_t^d}{D} \mathbf{1} = \frac{(\sum_{d=1}^D g_t^d)^2}{D} \geq 0$ , 可得

$$\begin{aligned}
\mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}_{t-1})] &\leq \\
&\quad -\alpha \left( \sum_{i=1}^M p_{i,t} |h_{i,t}| \|\mathbf{g}_t\|^2 + \sum_{n=1}^N \hat{p}_{n,t} |h_{n,t}| \mathbf{g}_t^T \hat{\mathbf{g}}_{n,t} \right) \\
&\quad + \frac{\alpha^2 L}{2} \mathbb{E} \left[ \left\| \sum_{m=1}^M p_{m,t} |h_{m,t}| \mathbf{g}_{m,t} + \epsilon_t \sum_{n=1}^N \hat{p}_{n,t} |h_{n,t}| \hat{\mathbf{g}}_{n,t} \right. \right. \\
&\quad \left. \left. + \sum_{n=1}^N p_{n,t} |h_{n,t}| \bar{g}_t \mathbf{1} + \epsilon_t \mathbf{z}_t \right\|^2 \right]. \tag{A-80}
\end{aligned}$$

如果  $\mathbb{E}(F(\mathbf{w}_t) - F(\mathbf{w}_{t-1})) \leq 0$ , 目标函数单调递减, 则联邦学习在期望上将收敛。从上述公式 (A-80) 可以看出, 如果将学习速率设置得足够小, 则公式 (A-80) 右侧的第二项会减小至忽略不计, 于是联邦学习算法的收敛性可以保证, 只要满足

下列条件

$$\sum_{i=1}^M p_{i,t} |h_{i,t}| \|\mathbf{g}_t\|^2 + \sum_{n=1}^N \hat{p}_{n,t} |h_{n,t}| \mathbf{g}_t^T \hat{\mathbf{g}}_{n,t} > 0. \quad (\text{A-81})$$

为了破坏公式(A-81)中的收敛条件,  $N$ 个拜占庭攻击者将寻求使  $\mathbf{g}_t^T \hat{\mathbf{g}}_{n,t} < 0$ ,  $\forall n$ 。显然, 对它们来说最好的办法就是以最大的功率发送  $\hat{\mathbf{g}}_{n,t} = -\mathbf{g}_{n,t}$ , 从而使  $\mathbb{E}[\mathbf{g}_t^T \hat{\mathbf{g}}_{n,t}] = -\|\mathbf{g}_t\|^2 < 0$ 。

给定公式(A-82)中的功率约束条件, 可得

$$\begin{aligned} \mathbb{E}[\|\hat{p}_{n,t} \hat{\mathbf{g}}_{n,t}\|^2] &= \hat{p}_{n,t}^2 \sum_{d=1}^D \mathbb{E}[(g_{n,t}^d)^2] \\ &= \hat{p}_{n,t}^2 D(\epsilon_t^2 + \bar{g}_t^2) \leq p_n^{\max}. \end{aligned} \quad (\text{A-82})$$

因此, 作为一种最强的攻击, 拜占庭攻击者应该以最大功率  $\hat{p}_{n,t} = \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}}$  发送  $\hat{\mathbf{g}}_{n,t} = -\mathbf{g}_{n,t}$ 。

至此定理 4.1 证毕。 ■

## A9. 定理 4.2 的证明

**证明:** 给定公式(4-7)中的全局梯度估计, 公式(4-10)中的CI功率分配策略, 以及定理 4.1 中的最强攻击模式, 可得模型参数的更新规则如下

$$\begin{aligned} \mathbf{w}_t &= \mathbf{w}_{t-1} - \alpha \tilde{\mathbf{g}}_t \\ &= \mathbf{w}_{t-1} - \alpha \left( \sum_{m=1}^M p_{m,t} |h_{m,t}| \mathbf{g}_{m,t} + \epsilon_t \sum_{n=1}^N \hat{p}_{n,t} |h_{n,t}| \hat{\mathbf{g}}_{n,t} \right. \\ &\quad \left. + \sum_{n=1}^N p_{n,t} |h_{n,t}| \bar{g}_t \mathbf{1} + \epsilon_t \mathbf{z}_t \right) \\ &= \mathbf{w}_{t-1} - \alpha \left( \sum_{m=1}^M b_0 \mathbf{g}_{m,t} - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \mathbf{g}_{n,t} \right. \\ &\quad \left. + \sum_{n=1}^N b_0 \bar{g}_t \mathbf{1} + \epsilon_t \mathbf{z}_t \right). \end{aligned} \quad (\text{A-83})$$

将上式(A-83)代入公式(4-12)可得

$$\begin{aligned}
F(\mathbf{w}_t) &\leq F(\mathbf{w}_{t-1}) + \mathbf{g}_t^T(\mathbf{w}_t - \mathbf{w}_{t-1}) + \frac{L}{2}\|\mathbf{w}_t - \mathbf{w}_{t-1}\|^2 \\
&= F(\mathbf{w}_{t-1}) - \alpha \mathbf{g}_t^T \left( \sum_{m=1}^M b_0 \mathbf{g}_{m,t} \right. \\
&\quad \left. - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \mathbf{g}_{n,t} + \sum_{n=1}^N b_0 \bar{g}_t \mathbf{1} + \epsilon_t \mathbf{z}_t \right) \\
&\quad + \frac{\alpha^2 L}{2} \left\| \sum_{m=1}^M b_0 \mathbf{g}_{m,t} - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \mathbf{g}_{n,t} + \sum_{n=1}^N b_0 \bar{g}_t \mathbf{1} + \epsilon_t \mathbf{z}_t \right\|^2. \quad (\text{A-84})
\end{aligned}$$

重写上述不等式 (A-84) 并两边取期望, 可以得到

$$\begin{aligned}
\mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}_{t-1})] &\leq -\alpha \mathbf{g}_t^T \left( \sum_{m=1}^M b_0 \mathbf{g}_t \right. \\
&\quad \left. - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} \mathbb{E}[|h_{n,t}|] \mathbf{g}_t + \sum_{n=1}^N b_0 \mathbb{E}[\bar{g}_t] \mathbf{1} \right) \\
&\quad + \frac{\alpha^2 L}{2} \mathbb{E} \left[ \left\| \sum_{m=1}^M b_0 \mathbf{g}_{m,t} + \sum_{n=1}^N b_0 \bar{g}_t \mathbf{1} \right. \right. \\
&\quad \left. \left. - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \mathbf{g}_{n,t} + \epsilon_t \mathbf{z}_t \right\|^2 \right], \quad (\text{A-85})
\end{aligned}$$

其中  $\mathbb{E}[|h_{i,t}|] = \sigma_i \sqrt{\frac{\pi}{2}}$ , 因为  $|h_{i,t}|$  服从瑞利分布。

由于  $\mathbf{g}_t^T \mathbb{E}[\bar{g}_t] \mathbf{1} = \mathbf{g}_t^T \frac{\sum_{d=1}^D g_t^d}{D} \mathbf{1} = \frac{(\sum_{d=1}^D g_t^d)^2}{D} \geq 0$ , 可得

$$\begin{aligned}
\mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}_{t-1})] &\leq -\alpha (Mb_0 \\
&\quad - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} \sigma_n \sqrt{\frac{\pi}{2}}) \|\mathbf{g}_t\|^2 \\
&\quad + \frac{\alpha^2 L}{2} \mathbb{E} \left[ \left\| \sum_{m=1}^M b_0 \mathbf{g}_{m,t} - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \mathbf{g}_{n,t} \right. \right. \\
&\quad \left. \left. + \sum_{n=1}^N b_0 \bar{g}_t \mathbf{1} + \epsilon_t \mathbf{z}_t \right\|^2 \right] \\
&\leq -\alpha \left( Mb_0 - \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D}} \sigma_n \sqrt{\frac{\pi}{2}} \right) \|\mathbf{g}_t\|^2 \\
&\quad + \frac{\alpha^2 L}{2} \mathbb{E} \left[ \left\| \sum_{m=1}^M b_0 \mathbf{g}_{m,t} - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \mathbf{g}_{n,t} \right. \right. \\
&\quad \left. \left. + \sum_{n=1}^N b_0 \bar{g}_t \mathbf{1} + \epsilon_t \mathbf{z}_t \right\|^2 \right]. \quad (\text{A-86})
\end{aligned}$$

使用范数的三角不等式和 Jensen 不等式，可得

$$\begin{aligned}
& \mathbb{E} \left[ \left\| \sum_{m=1}^M b_0 \mathbf{g}_{m,t} - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \mathbf{g}_{n,t} \right. \right. \\
& \quad \left. \left. + \sum_{n=1}^N b_0 \bar{g}_t \mathbf{1} + \epsilon_t \mathbf{z}_t \right\|^2 \right] \\
& = \mathbb{E} \left[ \left\| \sum_{m=1}^M b_0 \mathbf{g}_{m,t} - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \mathbf{g}_{n,t} \right. \right. \\
& \quad \left. \left. + \sum_{n=1}^N b_0 \bar{g}_t \mathbf{1} \right\|^2 \right] + \mathbb{E}[\|\epsilon_t \mathbf{z}_t\|^2] \\
& \leq \mathbb{E} \left[ \left( \sum_{m=1}^M \|b_0 \mathbf{g}_{m,t}\| + \sum_{n=1}^N \left\| \epsilon_t \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \mathbf{g}_{n,t} \right\| \right. \right. \\
& \quad \left. \left. + \sum_{n=1}^N \|b_0 \bar{g}_t \mathbf{1}\| \right)^2 \right] + \epsilon^2 z^2 \\
& \leq \mathbb{E} \left[ (U+N) \left( \sum_{m=1}^M b_0^2 \|\mathbf{g}_{m,t}\|^2 + \sum_{n=1}^N b_0^2 \|\bar{g}_t \mathbf{1}\|^2 \right. \right. \\
& \quad \left. \left. + \sum_{n=1}^N \frac{\epsilon_t^2 p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)} |h_{n,t}|^2 \|\mathbf{g}_{n,t}\|^2 \right) \right] + \epsilon^2 z^2 \\
& = (U+N) \left( \sum_{m=1}^M b_0^2 \mathbb{E}[\|\mathbf{g}_{m,t}\|^2] + \sum_{n=1}^N b_0^2 D \left( \frac{\sum_{d=1}^D g_t^d}{D} \right)^2 \right. \\
& \quad \left. + \sum_{n=1}^N \frac{\epsilon_t^2 p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)} \mathbb{E}[|h_{n,t}|^2] \mathbb{E}[\|\mathbf{g}_{n,t}\|^2] \right) + \epsilon^2 z^2 \\
& \leq (U+N) \left( \sum_{m=1}^M b_0^2 (\|\mathbf{g}_t\|^2 + \delta^2) + \sum_{n=1}^N b_0^2 \|\mathbf{g}_t\|^2 \right. \\
& \quad \left. + \sum_{n=1}^N \frac{\epsilon_t^2 p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)} 2\sigma_n^2 (\|\mathbf{g}_t\|^2 + \delta^2) \right) + \epsilon^2 z^2 \\
& \leq (U+N) \left( \left( Ub_0^2 + \sum_{n=1}^N \frac{2\sigma_n^2 p_n^{\max}}{D} \right) \|\mathbf{g}_t\|^2 \right. \\
& \quad \left. + \left( Mb_0^2 + \sum_{n=1}^N \frac{2\sigma_n^2 p_n^{\max}}{D} \right) \delta^2 \right) + \epsilon^2 z^2. \tag{A-87}
\end{aligned}$$

将上述公式 (A-87) 代入公式 (A-86) 可得

$$\begin{aligned}
\mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}_{t-1})] &\leq -\alpha \left( Mb_0 - \sum_{n=1}^N \sqrt{\frac{\pi \sigma_n^2 p_n^{\max}}{2D}} \right) \|\mathbf{g}_t\|^2 \\
&+ \frac{\alpha^2 L}{2} \left( (U + N) \left( \left( Ub_0^2 + \sum_{n=1}^N \frac{2\sigma_n^2 p_n^{\max}}{D} \right) \|\mathbf{g}_t\|^2 \right. \right. \\
&\left. \left. + \left( Mb_0^2 + \sum_{n=1}^N \frac{2\sigma_n^2 p_n^{\max}}{D} \right) \delta^2 \right) + \epsilon^2 z^2 \right) \\
&\leq \left( \frac{\alpha^2 L}{2} \Omega_{CI} - \alpha \omega_{CI} \right) \|\mathbf{g}_t\|^2 + \frac{\alpha^2 L}{2} (\Omega_{CI} \delta^2 + \epsilon^2 z^2),
\end{aligned} \tag{A-88}$$

其中，

$$\omega_{CI} = Mb_0 - \sum_{n=1}^N \sqrt{\frac{\pi \sigma_n^2 p_n^{\max}}{2D}}, \tag{A-89}$$

$$\Omega_{CI} = (U + N) \left( Ub_0^2 + \sum_{n=1}^N \frac{2\sigma_n^2 p_n^{\max}}{D} \right). \tag{A-90}$$

如果  $\mathbb{E}(F(\mathbf{w}_t) - F(\mathbf{w}_{t-1})) \leq 0$ ，目标函数单调递减，则联邦学习算法收敛将在期望上收敛。因此，为了确保收敛，有以下收敛条件需要满足

$$\frac{\alpha^2 L}{2} \Omega_{CI} - \alpha \omega_{CI} < 0. \tag{A-90}$$

现在将期望的随机性扩展到迭代轨迹中，并在迭代上执行放缩求和，可得

$$\begin{aligned}
F(\mathbf{w}_0) - F(\mathbf{w}^*) &\geq F(\mathbf{w}_0) - \mathbb{E}[F(\mathbf{w}_T)] \\
&= \mathbb{E} \left[ \sum_{t=1}^T (F(\mathbf{w}_{t-1}) - F(\mathbf{w}_t)) \right] \\
&\geq \mathbb{E} \left[ \sum_{t=1}^T \left( \left( \alpha \omega_{CI} - \frac{\alpha^2 L}{2} \Omega_{CI} \right) \|\mathbf{g}_t\|^2 \right. \right. \\
&\quad \left. \left. - \frac{\alpha^2 L}{2} (\Omega_{CI} \delta^2 + \epsilon^2 z^2) \right) \right]. \tag{A-91}
\end{aligned}$$

上式 (A-91) 可以被重新组织以得到

$$\begin{aligned}
&\mathbb{E} \left[ \sum_{t=1}^T \left( \left( \alpha \omega_{CI} - \frac{\alpha^2 L}{2} \Omega_{CI} \right) \|\mathbf{g}_t\|^2 \right) \right] \\
&\leq F(\mathbf{w}_0) - F(\mathbf{w}^*) + \frac{\alpha^2 L}{2} T (\Omega_{CI} \delta^2 + \epsilon^2 z^2).
\end{aligned} \tag{A-92}$$

如果联邦学习算法收敛，则条件 (A-90) 成立，有  $\alpha \omega_{CI} - \frac{\alpha^2 L}{2} \Omega_{CI} > 0$ 。于是可以得到

$$\begin{aligned} \mathbb{E} \left[ \sum_{t=1}^T \frac{1}{T} \|\mathbf{g}_t\|^2 \right] &\leq \frac{1}{T(\alpha\omega_{CI} - \frac{\alpha^2 L}{2}\Omega_{CI})} (F(\mathbf{w}_0) - F(\mathbf{w}^*)) \\ &\quad + \frac{\alpha^2 L}{2} T(\Omega_{CI}\delta^2 + \epsilon^2 z^2). \end{aligned} \quad (\text{A-93})$$

令  $\alpha = \frac{\omega_{CI}}{L\Omega_{CI}\sqrt{T}}\bar{\alpha}$ , 其中  $\bar{\alpha} < 2\sqrt{T}$  是一个正常数, 则可得

$$\begin{aligned} \mathbb{E} \left[ \sum_{t=1}^T \frac{1}{T} \|\mathbf{g}_t\|^2 \right] &\leq \frac{1}{T \left( \bar{\alpha} \frac{\omega_{CI}^2}{L\Omega_{CI}\sqrt{T}} - \frac{\bar{\alpha}^2 \omega_{CI}^2}{2LT\Omega_{CI}} \right)} (F(\mathbf{w}_0) \\ &\quad - F(\mathbf{w}^*) + \frac{\bar{\alpha}^2 \omega_{CI}^2}{2L\Omega_{CI}} \left( \delta^2 + \frac{1}{\Omega_{CI}} \epsilon^2 z^2 \right)) \\ &= \frac{1}{T \left( \frac{\bar{\alpha}}{\sqrt{T}} - \frac{\bar{\alpha}^2}{2T} \right)} \left( \frac{L\Omega_{CI}}{\omega_{CI}^2} (F(\mathbf{w}_0) - F(\mathbf{w}^*)) \right. \\ &\quad \left. + \frac{\bar{\alpha}^2}{2} \left( \delta^2 + \frac{1}{\Omega_{CI}} \epsilon^2 z^2 \right) \right) \\ &\leq \frac{1}{T \frac{\bar{\alpha}}{2\sqrt{T}}} \left( \frac{L\Omega_{CI}}{\omega_{CI}^2} (F(\mathbf{w}_0) - F(\mathbf{w}^*)) \right. \\ &\quad \left. + \frac{\bar{\alpha}^2}{2} \left( \delta^2 + \frac{1}{\Omega_{CI}} \epsilon^2 z^2 \right) \right) \\ &= \frac{1}{\sqrt{T}\bar{\alpha}} \left( \frac{2L\Omega_{CI}}{\omega_{CI}^2} (F(\mathbf{w}_0) - F(\mathbf{w}^*)) \right. \\ &\quad \left. + \bar{\alpha}^2 \left( \delta^2 + \frac{1}{\Omega_{CI}} \epsilon^2 z^2 \right) \right). \end{aligned} \quad (\text{A-94})$$

定理 4.2 由此得证。 ■

## A10. 定理 4.3 的证明

**证明:** 给定公式 (4-7) 中的全局梯度估计, 公式 (4-12) 中的 BEV 功率分配策略, 以及定理 4.1 中的最强攻击模式, 可得模型参数的更新规则如下

$$\begin{aligned} \mathbf{w}_t &= \mathbf{w}_{t-1} - \alpha \tilde{\mathbf{g}}_t \\ &= \mathbf{w}_{t-1} - \alpha \left( \sum_{m=1}^M p_{m,t} |h_{m,t}| \mathbf{g}_{m,t} + \epsilon_t \sum_{n=1}^N \hat{p}_{n,t} |h_{n,t}| \hat{\mathbf{g}}_{n,t} \right. \\ &\quad \left. + \sum_{n=1}^N p_{n,t} |h_{n,t}| \bar{g}_t \mathbf{1} + \epsilon_t \mathbf{z}_t \right) \\ &= \mathbf{w}_{t-1} - \alpha \left( \sum_{m=1}^M \sqrt{\frac{p_m^{\max}}{D}} |h_{m,t}| \mathbf{g}_{m,t} + \epsilon_t \mathbf{z}_t \right. \\ &\quad \left. - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \mathbf{g}_{n,t} + \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D}} |h_{n,t}| \bar{g}_t \mathbf{1} \right). \end{aligned} \quad (\text{A-95})$$

将上式 (A-95) 代入公式 (4-12) 可得

$$\begin{aligned}
F(\mathbf{w}_t) &\leq F(\mathbf{w}_{t-1}) + \mathbf{g}_t^T(\mathbf{w}_t - \mathbf{w}_{t-1}) + \frac{L}{2}\|\mathbf{w}_t - \mathbf{w}_{t-1}\|^2 \\
&= F(\mathbf{w}_{t-1}) - \alpha \mathbf{g}_t^T \left( \sum_{m=1}^M \sqrt{\frac{p_m^{\max}}{D}} |h_{m,t}| \mathbf{g}_{m,t} \right. \\
&\quad \left. - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \mathbf{g}_{n,t} + \epsilon_t \mathbf{z}_t \right. \\
&\quad \left. + \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D}} |h_{n,t}| \bar{g}_t \mathbf{1} \right) + \frac{\alpha^2 L}{2} \left\| \sum_{m=1}^M \sqrt{\frac{p_m^{\max}}{D}} |h_{m,t}| \mathbf{g}_{m,t} \right. \\
&\quad \left. - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \mathbf{g}_{n,t} + \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D}} |h_{n,t}| \bar{g}_t \mathbf{1} + \epsilon_t \mathbf{z}_t \right\|^2. \tag{A-96}
\end{aligned}$$

重新组织上述不等式 (A-84) 并两边取期望，可以得到

$$\begin{aligned}
\mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}_{t-1})] &\leq -\alpha \mathbf{g}_t^T \left( \sum_{m=1}^M \sqrt{\frac{p_m^{\max}}{D}} \mathbb{E}[|h_{m,t}| \mathbf{g}_{m,t}] \right. \\
&\quad \left. - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} \mathbb{E}[|h_{n,t}| \mathbf{g}_{n,t}] \right. \\
&\quad \left. + \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D}} \mathbb{E}[|h_{n,t}| \bar{g}_t \mathbf{1}] + \mathbb{E}[\epsilon_t \mathbf{z}_t] \right) \\
&\quad + \frac{\alpha^2 L}{2} \mathbb{E} \left[ \left\| \sum_{m=1}^M \sqrt{\frac{p_m^{\max}}{D}} |h_{m,t}| \mathbf{g}_{m,t} + \epsilon_t \mathbf{z}_t \right. \right. \\
&\quad \left. \left. - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \mathbf{g}_{n,t} + \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D}} |h_{n,t}| \bar{g}_t \mathbf{1} \right\|^2 \right] \\
&\leq -\alpha \left( \sum_{i=1}^M \sqrt{\frac{p_i^{\max}}{D}} \sigma_i \sqrt{\frac{\pi}{2}} - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} \sigma_n \sqrt{\frac{\pi}{2}} \right) \|\mathbf{g}_t\|^2 \\
&\quad + \frac{\alpha^2 L}{2} \mathbb{E} \left[ \left\| \sum_{m=1}^M \sqrt{\frac{p_m^{\max}}{D}} |h_{m,t}| \mathbf{g}_{m,t} + \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D}} |h_{n,t}| \bar{g}_t \mathbf{1} \right. \right. \\
&\quad \left. \left. - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \mathbf{g}_{n,t} + \epsilon_t \mathbf{z}_t \right\|^2 \right] \\
&\leq -\alpha \sqrt{\frac{\pi}{2}} \left( \sum_{i=1}^M \sqrt{\frac{p_i^{\max}}{D}} \sigma_i - \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D}} \sigma_n \right) \|\mathbf{g}_t\|^2 \\
&\quad + \frac{\alpha^2 L}{2} \mathbb{E} \left[ \left\| \sum_{m=1}^M \sqrt{\frac{p_m^{\max}}{D}} |h_{m,t}| \mathbf{g}_{m,t} + \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D}} |h_{n,t}| \bar{g}_t \mathbf{1} \right. \right. \\
&\quad \left. \left. - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \mathbf{g}_{n,t} + \epsilon_t \mathbf{z}_t \right\|^2 \right]. \tag{A-97}
\end{aligned}$$

使用范数的三角不等式和 Jensen 不等式，可得

$$\begin{aligned}
& \mathbb{E} \left[ \left\| \sum_{m=1}^M \sqrt{\frac{p_m^{\max}}{D}} |h_{m,t}| \mathbf{g}_{m,t} + \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D}} |h_{n,t}| \bar{g}_t \mathbf{1} \right. \right. \\
& \quad \left. \left. - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \mathbf{g}_{n,t} + \epsilon_t \mathbf{z}_t \right\|^2 \right] \\
& = \mathbb{E} \left[ \left\| \sum_{m=1}^M \sqrt{\frac{p_m^{\max}}{D}} |h_{m,t}| \mathbf{g}_{m,t} + \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D}} |h_{n,t}| \bar{g}_t \mathbf{1} \right. \right. \\
& \quad \left. \left. - \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \mathbf{g}_{n,t} \right\|^2 \right] + \mathbb{E}[\|\epsilon_t \mathbf{z}_t\|^2] \\
& \leq \mathbb{E} \left[ \left( \sum_{m=1}^M \sqrt{\frac{p_m^{\max}}{D}} |h_{m,t}| \|\mathbf{g}_{m,t}\| + \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D}} |h_{n,t}| \|\bar{g}_t \mathbf{1}\| \right. \right. \\
& \quad \left. \left. + \epsilon_t \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)}} |h_{n,t}| \|\mathbf{g}_{n,t}\| \right)^2 \right] + \epsilon^2 z^2 \\
& \leq \mathbb{E} \left[ (U+N) \left( \sum_{m=1}^M \frac{p_m^{\max}}{D} |h_{m,t}|^2 \|\mathbf{g}_{m,t}\|^2 \right. \right. \\
& \quad \left. \left. + \sum_{n=1}^N \frac{\epsilon_t^2 p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)} |h_{n,t}|^2 \|\mathbf{g}_{n,t}\|^2 + \sum_{n=1}^N \frac{p_n^{\max}}{D} |h_{n,t}|^2 \|\mathbf{g}_t\|^2 \right) \right] + \epsilon^2 z^2 \\
& \leq (U+N) \left( \sum_{i=1}^U \frac{p_i^{\max}}{D} 2\sigma_i^2 \|\mathbf{g}_t\|^2 + \sum_{m=1}^M \frac{p_m^{\max}}{D} 2\sigma_m^2 \delta^2 \right. \\
& \quad \left. + \sum_{n=1}^N \frac{\epsilon_t^2 p_n^{\max}}{D(\epsilon_t^2 + \bar{g}_t^2)} 2\sigma_n^2 \delta^2 \right) + \epsilon^2 z^2 \\
& \leq (U+N) \left( \sum_{i=1}^U \frac{p_i^{\max}}{D} 2\sigma_i^2 \|\mathbf{g}_t\|^2 + \sum_{i=1}^U \frac{p_i^{\max}}{D} 2\sigma_i^2 \delta^2 \right) + \epsilon^2 z^2. \tag{A-98}
\end{aligned}$$

将上述公式 (A-98) 代入公式 (A-97) 可得

$$\begin{aligned}
& \mathbb{E}[F(\mathbf{w}_t) - F(\mathbf{w}_{t-1})] \\
& \leq -\alpha \sqrt{\frac{\pi}{2}} \left( \sum_{i=1}^M \sqrt{\frac{p_i^{\max}}{D}} \sigma_i - \sum_{n=1}^N \sqrt{\frac{p_n^{\max}}{D}} \sigma_n \right) \|\mathbf{g}_t\|^2 \\
& \quad + \frac{\alpha^2 L}{2} \left( (U+N) \left( \sum_{i=1}^U \frac{p_i^{\max}}{D} 2\sigma_i^2 \|\mathbf{g}_t\|^2 \right. \right. \\
& \quad \left. \left. + \sum_{i=1}^U \frac{p_i^{\max}}{D} 2\sigma_i^2 \delta^2 \right) + \epsilon^2 z^2 \right) \\
& = \left( \frac{\alpha^2 L}{2} \Omega_{BEV} - \alpha \omega_{BEV} \right) \|\mathbf{g}_t\|^2 + \frac{\alpha^2 L}{2} (\Omega_{BEV} \delta^2 + \epsilon^2 z^2), \tag{A-99}
\end{aligned}$$

其中，

$$\omega_{BEV} = \sum_{i=1}^M \sqrt{\frac{p_i^{\max}\pi}{2D}}\sigma_i - \sum_{n=1}^N \sqrt{\frac{p_n^{\max}\pi}{2D}}\sigma_n, \quad (\text{A-100})$$

$$\Omega_{BEV} = (U + N) \sum_{i=1}^U \frac{2\sigma_i^2 p_i^{\max}}{D}. \quad (\text{A-101})$$

如果  $\mathbb{E}(F(\mathbf{w}_t) - F(\mathbf{w}_{t-1})) \leq 0$ ，目标函数单调递减，则联邦学习算法收敛将在期望上收敛。因此，为了确保收敛，有以下收敛条件需要满足

$$\frac{\alpha^2 L}{2} \Omega_{BEV} - \alpha \omega_{BEV} < 0. \quad (\text{A-102})$$

现在将期望的随机性扩展到迭代轨迹中，并在迭代上执行放缩求和，则可以得到

$$\begin{aligned} F(\mathbf{w}_0) - F(\mathbf{w}^*) &\geq F(\mathbf{w}_0) - \mathbb{E}[F(\mathbf{w}_T)] \\ &= \mathbb{E} \left[ \sum_{t=1}^T (F(\mathbf{w}_{t-1}) - F(\mathbf{w}_t)) \right] \\ &\geq \mathbb{E} \left[ \sum_{t=1}^T \left( \left( \alpha \omega_{BEV} - \frac{\alpha^2 L}{2} \Omega_{BEV} \right) \|\mathbf{g}_t\|^2 \right. \right. \\ &\quad \left. \left. - \frac{\alpha^2 L}{2} (\Omega_{BEV} \delta^2 + \epsilon^2 z^2) \right) \right]. \end{aligned} \quad (\text{A-103})$$

上式 (A-103) 可以被重新组织以得到

$$\begin{aligned} &\mathbb{E} \left[ \sum_{t=1}^T \left( \alpha \omega_{BEV} - \frac{\alpha^2 L}{2} \Omega_{BEV} \right) \|\mathbf{g}_t\|^2 \right] \\ &\leq F(\mathbf{w}_0) - F(\mathbf{w}^*) + \frac{\alpha^2 L}{2} \sum_{t=1}^T (\Omega_{BEV} \delta^2 + \epsilon^2 z^2). \end{aligned} \quad (\text{A-104})$$

如果联邦学习算法收敛，则条件 (A-102) 成立，有  $\alpha \omega_{BEV} - \frac{\alpha^2 L}{2} \Omega_{BEV} > 0$ 。于是可以得到

$$\mathbb{E} \left[ \sum_{t=1}^T \frac{1}{T} \|\mathbf{g}_t\|^2 \right] \leq \frac{F(\mathbf{w}_0) - F(\mathbf{w}^*) + \frac{\alpha^2 L}{2} \sum_{t=1}^T (\Omega_{BEV} \delta^2 + \epsilon^2 z^2)}{T(\alpha \omega_{BEV} - \frac{\alpha^2 L}{2} \Omega_{BEV})}. \quad (\text{A-105})$$

令  $\alpha = \frac{\omega_{BEV}}{L \Omega_{BEV} \sqrt{T}} \bar{\alpha}$ ，其中  $\bar{\alpha} < 2\sqrt{T}$  是一个正常数，则可得

$$\begin{aligned}
& \mathbb{E} \left[ \sum_{t=1}^T \frac{1}{T} \|\mathbf{g}_t\|^2 \right] \\
& \leq \frac{F(\mathbf{w}_0) - F(\mathbf{w}^*) + \frac{\bar{\alpha}^2 \omega_{BEV}^2}{2L\Omega_{BEV}} \left( \delta^2 + \frac{1}{\Omega_{BEV}} \epsilon^2 z^2 \right)}{T \left( \bar{\alpha} \frac{\omega_{BEV}^2}{L\Omega_{BEV}\sqrt{T}} - \frac{\bar{\alpha}^2 \omega_{BEV}^2}{2LT\Omega_{BEV}} \right)} \\
& = \frac{\frac{L\Omega_{BEV}}{\omega_{BEV}^2} (F(\mathbf{w}_0) - F(\mathbf{w}^*)) + \frac{\bar{\alpha}^2}{2} \left( \delta^2 + \frac{1}{\Omega_{BEV}} \epsilon^2 z^2 \right)}{T \left( \frac{\bar{\alpha}}{\sqrt{T}} - \frac{\bar{\alpha}^2}{2T} \right)} \\
& \leq \frac{\frac{L\Omega_{BEV}}{\omega_{BEV}^2} (F(\mathbf{w}_0) - F(\mathbf{w}^*)) + \frac{\bar{\alpha}^2}{2} \left( \delta^2 + \frac{1}{\Omega_{BEV}} \epsilon^2 z^2 \right)}{T \frac{\bar{\alpha}}{2\sqrt{T}}} \\
& = \frac{\frac{2L\Omega_{BEV}}{\bar{\alpha}\omega_{BEV}^2} (F(\mathbf{w}_0) - F(\mathbf{w}^*)) + \bar{\alpha} \left( \delta^2 + \frac{1}{\Omega_{BEV}} \epsilon^2 z^2 \right)}{\sqrt{T}}. \tag{A-106}
\end{aligned}$$

由此，定理 4.3 证毕。 ■



## 作者简历及攻读博士学位期间取得的研究成果

### 一、作者简历

范新，2012年9月录取至北京交通大学理学院理科试验班（思源班）专业。

2014年9月，通过理科试验班自由选专业政策转入北京交通大学电子信息工程学院通信工程（理科试验班）专业。

2016年7月，获得北京交通大学通信工程（理科试验班）专业学士学位。同年9月，考入北京交通大学电子信息工程学院电子与通信工程专业攻读硕士学位。

2018年7月，硕士研究生毕业并获得硕士学位。同年9月考入北京交通大学电子信息工程学院信息与通信工程专业攻读博士学位至今。

攻读博士学位期间，获留学基金委资助，于2020年1月至2022年8月由国家公派赴美国乔治梅森大学（George Mason University）电气与计算机工程系联合培养。

### 二、已发表/录用文章

- [1]. **Xin Fan**, Yue Wang, Yan Huo, Zhi Tian. Joint Optimization of Communications and Federated Learning Over the Air[J]. IEEE Transactions on Wireless Communications, vol. 21, no. 6, pp. 4434-4449, June 2022. (SCI期刊; 2022年影响因子 8.346; 中科院一区; JCR一区; 北交大 A+; WOS:000809406400064; 对应学位论文第二章)
- [2]. **Xin Fan**, Yue Wang, Yan Huo, Zhi Tian. 1-Bit Compressive Sensing for Efficient Federated Learning Over the Air[J]. IEEE Transactions on Wireless Communications, oct, 2022, early access. (SCI期刊; 2022年影响因子 8.346; 中科院一区; JCR一区; 北交大 A+; 对应学位论文第三章)
- [3]. **Xin Fan**, Yue Wang, Yan Huo, Zhi Tian. BEV-SGD: Best Effort Voting SGD Against Byzantine Attacks for Analog-Aggregation-Based Federated Learning Over the Air[J]. IEEE Internet of Things Journal, vol. 9, no. 19, pp. 18946-18959, Oct, 2022. (SCI期刊; 2022年影响因子 10.238; 中科院一区; JCR一区; 北交大 A+; WOS:000857705300062, 对应学位论文第四章)
- [4]. Yan Huo, **Xin Fan**, Liran Ma, Xiuzhen Cheng, Zhi Tian, Dechang Chen. Secure communications in tiered 5G wireless networks with cooperative jamming[J]. IEEE Transactions on Wireless Communications, 2019, 18(6): 3265-3280. (SCI期刊; 2022年影响因子 8.346; 中科院一区; JCR一区; 北交大 A+; WOS:000471120800027)
- [5]. **Xin Fan**, Yue Wang, Yan Huo, Zhi Tian. Joint Optimization for Federated Learning Over the Air [C] //2022 IEEE International Conference on Communications (IEEE ICC 2022). (EI/ISTP 国际会议; CCF C类; 北交大 A类; 对应学位论文第二章)
- [6]. **Xin Fan**, Yue Wang, Yan Huo, Zhi Tian. Communication-efficient Federated Learning Through 1-Bit Compressive Sensing and Analog Aggregation[C] //2021 IEEE International Conference on Communications (IEEE ICC 2021). (EI/ISTP 国际会议; CCF C类; 北交大 A类; WOS:000848412200333; 对应学位论文第三章)
- [7]. **Xin Fan**, Yue Wang, Yan Huo, Zhi Tian. Best Effort Voting Power Control for Byzantine-resilient Federated Learning Over the Air[C] //2022 IEEE International Conference on Communications (IEEE ICC 2022). (EI/ ISTP 国际会议; CCF C类; 北交大 A类;

**WOS:000848467200135, 对应学位论文第四章)**

- [8]. **Xin Fan**, Yan Huo. Cooperative secure transmission against collusive eavesdroppers in Internet of Things[J]. International Journal of Distributed Sensor Networks, 2020, 16(6). (SCI 期刊, 2022 年影响因子 1.938, 中科院四区; JCR 三区; WOS:000542331100001)
- [9]. **Xin Fan**, Yue Wang, Guangkai Li, Yan Huo, Zhi Tian. Hybrid Uplink-Downlink NOMA for Secure Coordinated Multi-Point Networks[C] //2021 IEEE International Conference on Communications (IEEE ICC 2021). (EI/ISTP 国际会议; CCF C 类; 北交大 A 类; WOS:000848412200120)
- [10].**Xin Fan**, Yan Huo. Security Analysis of Cooperative Jamming in Internet of Things with Multiple Eavesdroppers[C]//2019 IEEE Global Communications Conference (GLOBECOM). IEEE, 2019: 1-6. (EI/ISTP 国际会议; CCF C 类; 北交大 A 类; WOS:000552238604125)
- [11].**Xin Fan**, Yan Huo. Blockchain Based Dynamic Spectrum Access of Non-Real-Time Data in Cyber-Physical-Social Systems[J]. IEEE Access, 2020, 8: 64486-64498. (SCI 期刊; 2022 年影响因子 3.476; 中科院三区; JCR 二区; WOS:000530832200155)
- [12].Yihui Zhang, **Xin Fan**, Yan Huo, Yanfei Lu, An Improved Off-grid Channel Estimation Scheme for Lens Antenna Arrays in mmWave Systems [J]. Procedia Computer Science, 2020, 174: 555-560.. (EI 检索; 检索号 20220111431171)
- [13].Yan Huo, Mi Xu, **Xin Fan** and Tao Jing. A novel secure relay selection strategy for energy-harvesting-enabled Internet of things[J]. EURASIP Journal on Wireless Communications and Networking, 2018, 2018(1): 264. (SCI 期刊; 2022 年影响因子 2.559; 中科院四区; JCR 三区; WOS:000450503600003)

### 三、已提交论文

- [1]. **Xin Fan**, Yue Wang, Yan Huo, Zhi Tian. Robust Distributed Swarm Learning for Intelligent IoT [C], submitted to IEEE ICC 2023.
- [2]. **Xin Fan**, Yue Wang, Yan Huo, Zhi Tian. Efficient Distributed Swarm Learning for Edge Computing [C], submitted to IEEE ICC 2023.
- [3]. **Xin Fan**, Yue Wang, Yan Huo, Zhi Tian. CB-DSL: Communication-efficient and Byzantine-robust Distributed Swarm Learning on Non-i.i.d. Data[J]. arXiv preprint arXiv:2208.05578, 2022. Submitted to IEEE Transactions on Cognitive Communications and Networking, under review.
- [4]. **Xin Fan**, Yan Huo. An Overview of Low latency for Wireless Communications: An Evolutionary Perspective[J] arXiv preprint arXiv:2107.03484, 2021.
- [5]. **Xin Fan**, Yue Wang, Weishan Zhang, Yan Huo, Zhi Tian. GANFed: GAN-based Federated Learning with Non-IID Datasets[J]. Submitted to IEEE wireless communication letters, under review.
- [6]. Yue Wang, Zhi Tian, **Xin Fan**, Yan Huo, Cameron Nowzari, Kai Zeng, Distributed Swarm Learning for Internet of Things at the Edge: Where Artificial Intelligence Meets Biological Intelligence[J], Submitted to IEEE Wireless Communications, under review.

#### 四、已授权发明专利

- [1]. 范新, 霍炎, 荆涛. 基于区块链的频谱接入和管理方法, ZL 2019 1 0197190.3
- [2]. 荆涛, 安茜雯, 范新, 霍炎. 一种基于原子范数最小化的三维信道参数估计方法, ZL 2019 1 0483315.9

#### 五、参与主要科研项目

- [1]. 主持中央高校基本科研业务费专项资金资助项目: NOMA 系统中联合安全传输方案研究 (项目批准号: 2019YJS010)
- [2]. 作为项目书第二参与人参与国家自然科学基金面上项目: 面向下一代无线网络的多维度物理层安全绿色通信系统研究 (项目批准号: 61871023)
- [3]. 作为项目书第二参与人参与北京市自然科学基金面上项目: 面向智慧物联的协作认知动态安全传输系统研究与设计 (项目批准号: 4202054)
- [4]. 作为项目书第一参与人参与铁科研横向项目: 升降式站台门系统障碍物探测算法及电机同步性研究 (项目批准号: W19L00460)
- [5]. 作为主研人员之一参与国家自然科学基金重点项目: 面向未来移动通信的低功耗安全智慧物联网系统关键技术研究 (项目批准号: 61931001)
- [6]. 作为主研人员之一参与美国自然科学基金项目: Communication-efficient and Robust Learning from Distributed Data (项目批准号: NSF-CCF-1939553)
- [7]. 作为主研人员之一参与美国自然科学基金项目: Distributed Swarm Learning for Internet of Things at the Edge (项目批准号: NSF- ECCS- 2231209)

#### 六、获得主要奖励与荣誉

- [1] 优秀共产党员, 2019 年
- [2] 三好研究生, 2019 年
- [3] 国家公派奖学金, 2019 年
- [4] 交控科技奖学金, 2021 年
- [5] 学业奖学金, 2018-2021 年



## 答辩决议书

论文针对高效鲁棒联邦学习的传算联合优化方法展开研究，选题具有重要的理论意义和应用前景。

论文主要工作和创新点如下：

1. 揭示了基于空中计算技术的联邦学习（Federated Learning Over the Air, FLOA）系统中无线通信对联邦学习收敛性和准确性的影响规律。推导了 FLOA 在不同假设下的预期收敛速率的闭式表达式，分析了分布式设备选择、信道噪声、功率缩放因子等通信参数对联邦学习算法收敛性和准确性的影响，定量刻画了模拟聚合通信与联邦学习的关系。
2. 提出了基于计算收敛性的 FLOA 传算联合优化框架。建立了学习、设备选择和功率控制的联合优化问题，联合优化了无线网络资源和计算学习资源，降低了通信误差对学习算法的影响，提升了学习算法的准确性。
3. 提出了基于 1 比特压缩感知的高效性 FLOA 方案，构建了集稀疏化、降维、量化、模拟聚合传输和信号重建为一体的高效 FLOA 方案，推导了所提方案收敛速率的闭式表达式，实现了设备选择和功率控制的优化，降低了聚合错误，进一步提升了 FLOA 的高效性。
4. 提出了基于功率控制的鲁棒性 FLOA 方案，提出了最大努力投票功率控制方案来对抗拜占庭攻击，推导了所提方案和现存方案在攻击者存在的情况下收敛速率闭式表达式，所提方案提升了 FLOA 的鲁棒性。

论文工作表明作者已掌握了坚实宽广的基础理论和本学科系统深入的专门知识，具有独立从事科研工作的能力。答辩过程中讲述清楚，回答问题正确。论文结构清晰，图文规范，是一篇优秀的博士学位论文。

答辩委员会经无记名投票，一致同意通过范新博士学位论文答辩，并建议授予工学博士学位。

答辩委员会主席：



2022 年 12 月 5 日



## 独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作和取得的研究成果，除了文中特别加以标注和致谢之处外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得北京交通大学或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名：  签字日期： 2022 年 12 月 8 日



# 学位论文数据集

表 1.1: 数据集页

关键词*	密级*	中图分类号	UDC	论文资助				
联合优化；联邦学习；空中计算技术；高效性；鲁棒性；计算收敛性	公开							
学位授予单位名称*		学位授予单位代码*	学位类别*	学位级别*				
北京交通大学		10004	工学	博士				
论文题名*		并列题名		论文语种*				
面向高效鲁棒联邦学习的传算联合优化方法研究				中文				
作者姓名*	范新		学号*	18111012				
培养单位名称*		培养单位代码*	培养单位地址	邮编				
北京交通大学		10004	北京市海淀区西直门外上园村 3 号	100044				
学科专业*		研究方向*	学制*	学位授予年*				
信息与通信工程		无线通信与机器学习	4 年					
论文提交日期*								
导师姓名*	霍炎		职称*	教授				
评阅人	答辩委员会主席*		答辩委员会成员					
艾渤 王莉 赵雄文 赵友平 徐少毅	艾渤（北京交通大学）		王莉（北京邮电大学） 赵雄文（华北电力大学） 赵友平（北京交通大学） 徐少毅（北京交通大学）					
电子版论文提交格式 文本( ) 图像( ) 视频( ) 音频( ) 多媒体( ) 其他( )								
推荐格式: application/msword; application/pdf								
电子版论文出版(发布)者	电子版论文出版(发布)地		权限声明					
论文总页数*	125							
共 33 项，其中带*为必填数据，为 21 项。								