廣東工業大學

QG 中期考核详细报告书

题　　目　文献《 *Differentially private average consensus: Obstructions, trade-offs, and optimal algorithm design* 》 的仿真复现

学　　院　*计算机学院*

专　　业　*计算机类*

年级班别　*2024 级 1 班*

学　　号　*3124004052*

学生姓名　*陈英锐*

2025 年 4 月 5 日

目录

一、引言

1.1 目的

阅读一篇关于多智能体平均共识相关的文献《Differentially private average consensus: Obstructions, trade-offs, and optimal algorithm design》，熟悉论文的基本框架和基础算法，并且完成模拟实验 1、2、4 的仿真复现。

1.2 研究背景

社交网络、电网、智能交通等现代社会网络系统，系统内的各个智能体在一些应用领域要传递、交换信息，以实现达成系统的一致性，文献研究在隐私保护的情况下以实现多智能体系统的平均共识。

1.3 研究贡献

研究了在差分隐私保护下，一组智能体如何达成对其局部变量平均值的共识。隐私保护旨在防止外部对手和群组内其他成员获取智能体的初始状态。

证明了不可能结果：若算法具有差分隐私性，则无法保证智能体状态收敛到初始值的平均值；设计了线性拉普拉斯共识算法；在期望上实现平均共识，满足差分隐私性，并刻画了准确性和收敛速率；优化了设计参数：通过一次性拉普拉斯噪声扰动，最小化算法收敛点的方差，得到最优解。

二、正文

2.1 理论学习

下面总结的都是在后续进行模拟实验时需要用到的重要公式，并没有涵盖论文中所有公式及其推导。

2.1.1 系统状态方程

这是整篇文章讨论的基石，后续文章的讨论都要从这几条公式出发，得到最后的结论，下面先给出公式，然后对公式进行解释和分析：

$$\theta(k+1) = \theta(k) - hLx(k) + S\eta(k)$$

$$x(k) = \theta(k) + \eta(k)$$

$$\eta_i(k) \sim Lap\big(b_i(k)\big)$$

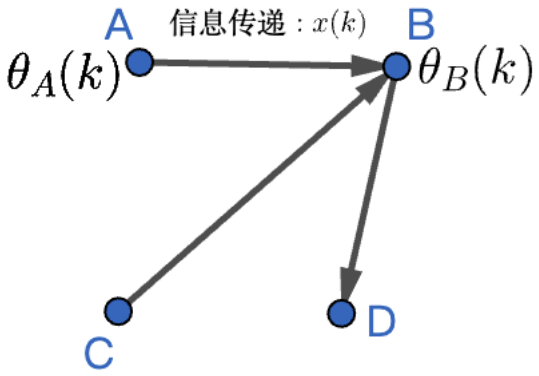
$$b_i(k) = c_i q_i^k$$

我们借助表格更加直观的展示公式的各个元素释义：

代数式	释义
$\boldsymbol{\theta(k)}$	智能体在 k 次模拟以后的状态

$\boldsymbol{x(k)}$	智能体在 k 次模拟时传递的信息
$\boldsymbol{\eta(k)}$	在第 k 次模拟生成的噪声参数
$\boldsymbol{Lap(x)}$	拉普拉斯分布，默认第一个位置参数为 0，尺度参数为 x

考虑一个简单的多智能体系统，将各个参数在图中表示出来，我们就可以很直观的理解各个代数式所代表的含义，其中，各个智能体之间信息传递参数 x 服从拉普拉斯分布，并且尺度参数为 b。



2.1.2 渐进收敛性

这是文献中的一个重要命题，文章后续各种命题的讨论与研究都建立在这个命题的基础之上：

$$\text{所有智能体的状态几乎必然收敛到}\theta_\infty = \mathrm{Ave}(\theta_0) + \sum_{i=1}^{n}\frac{s_i}{n}\sum_{j=0}^{\infty}\eta_i(j)$$

2.1.3 差分隐私概念

这是一个重要的概念，因为算法的设计必须要满足差分隐私要求，最后在数据的准确性和隐私性之间得到权衡：

$$\textit{差分隐私参数}\epsilon_i = \delta\frac{q_i}{c_i(q_i - |s_i - 1|)}$$

这条式子关联了很多参数，也是后续模拟实验的关键公式，在后续计算的时候，需要用到参数 c，我们可以利用这个公式求解参数。

2.1.4 收敛状态方差

这条公式对应收敛状态时的方差，可以判断系统最后的收敛性：

$$var\{\theta_\infty\} = \frac{2}{n^2}\sum_{i=1}^{n}\frac{s_i^2 c_i^2}{1-q_i^2}$$

2.2 算法实现

2.2.1 Fig1-收敛点在系统状态平均值附近的损失函数等价函数

首先需要明确我们本次实验的目标，即绘制等价函数的图像，那么我们就要明确函数的重要组成部分：函数参数和函数表达式，这些并没有在文章的命题或结论中出现，我们最后发现它出现在一个命题的证明过程中，最终确定了函数的表达式：

$$\phi(\alpha, s) = \frac{s^2(\alpha + (1-\alpha)|s-1|)^2}{\alpha^2(1-|s-1|)^2[1-(\alpha+(1-\alpha)|s-1|)^2]}$$

于是我们可以得到关键代码，最后通过 python 的 plt 库绘制就能得到理想结果：

```
def phi(alpha, s):
    a = s ** 2 * (alpha + (1 - alpha) * abs(s - 1)) ** 2
    b = (alpha ** 2 * (1 - abs(s - 1)) ** 2) * (1 - (alpha + (1 - alpha) * abs(s - 1)) ** 2)
    if b == 0:
        return np.nan
    return a / b
```

2.2.2 Fig3-收敛点的方差和收敛时间

明确我们的目标，我们需要计算方差和收敛时间对应变量 s 的关系，在论文中找到对应的公式，那么实际上，有关方差的重要公式已经在 2.1.4 中总结出来了，那么收敛时间的计算，需要我们在模拟计算的循环中实现，题目给出了收敛的阀值，在模拟过程中，当误差小于收敛阀值时，就可以认为最后的状态收敛了，记录下此时的收敛时间（循环次数）

这是有关 fig3 模拟的关键代码，它是参数 s 单次模拟的函数，后面我们只要对 s 进行遍历，反复调用函数并记录数值就能得到最终结果：

```
def simulation_single(theta0, c, A, B, q):
    theta = theta0.copy()
    convergence = False
    q_powered = 1
    for k in range(max_iterations):
        b = c * q_powered
        eta = np.random.laplace(scale=b, size=n)
        next_theta = A @ theta + B @ eta
        # if np.max(np.abs(next_theta - theta)) < threshold:
        if np.linalg.norm(next_theta - theta) < threshold:
            convergence = True; theta = next_theta
            break
```

```
        theta = next_theta
        q_powered *= q
    return (k if convergence else max_iterations), theta
```

代码中的部分参数还没有解释，这些都是笔者在编写代码时所发现的问题，我们需要回到论文当中寻找目标参数，而这些参数都在 2.1 节中进行了归纳总结，关于 2.1.3 中差分隐私的参数也在这里得到了应用。

2.2.3 Fig4-隐私和准确性的权衡关系

实际上图一的实现较为简单，我们对隐私差分参数进行扫描，最后计算误差取绝对值就能得到二者的关系图，fig4 有一点值得注意，它的实验条件有别于前面的模拟实验，参数 q=0，那么我们就要修改参数 c 的计算公式（实际上得到了简化）：

$$c = \frac{\delta}{\epsilon}$$

对于单次模拟，代码类似于 fig3，我们只需要修改参数 c 的计算代码，并且略去含有 q 的式子，就可以完成单次模拟，同样的，方差的理论值也随之要修改：

$$var(\theta) = \frac{2\delta^2}{n \cdot \epsilon^2}$$

对应的代码也相对简单，这里只展示部分重要过程代码：

```
for eps in tqdm(epsilons):
    final_avgs = []
    for _ in range(num_runs):
        final_avg = one_shot_consensus(theta0, L, epsilon=eps, max_iter=max_iter)
        final_avgs.append(final_avg)

    errors.extend([abs(avg - true_avg) for avg in final_avgs])
    variances.append(np.var(final_avgs))
    theoretical_var.append(2 * delta**2 / (n * eps**2))
```

2.3 结果分析与展示

2.3.1 Fig1

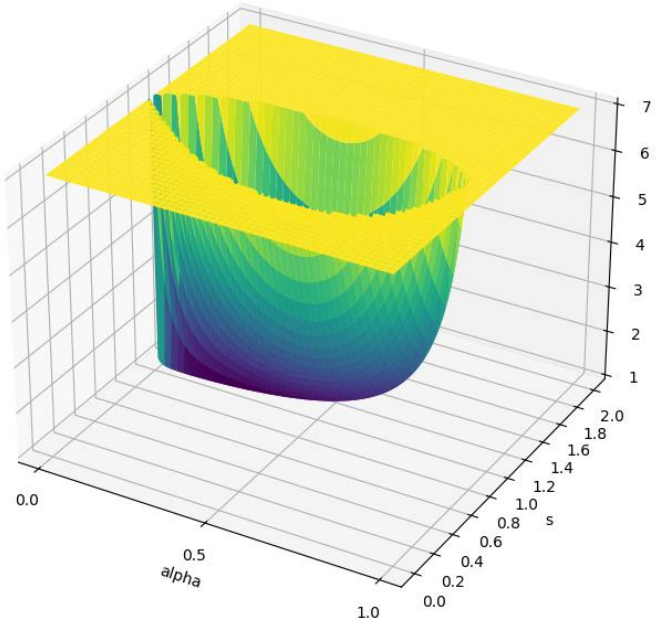
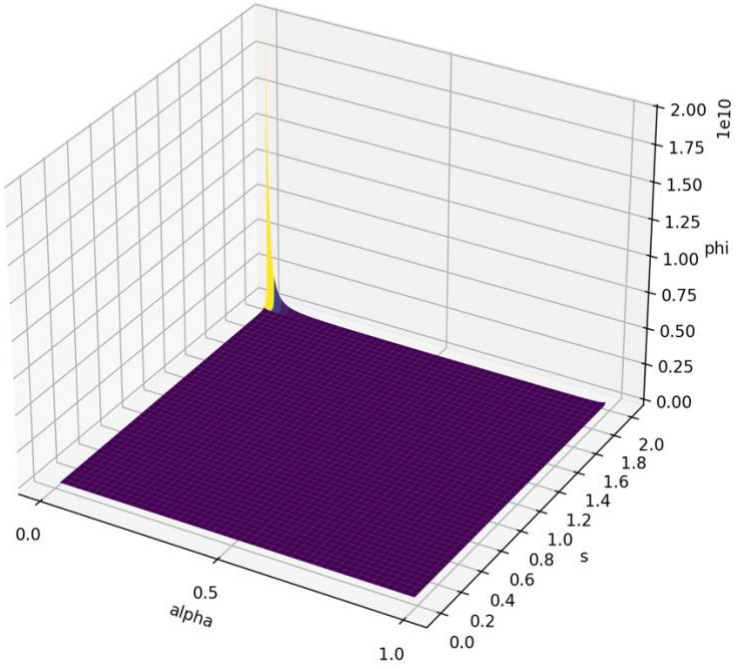
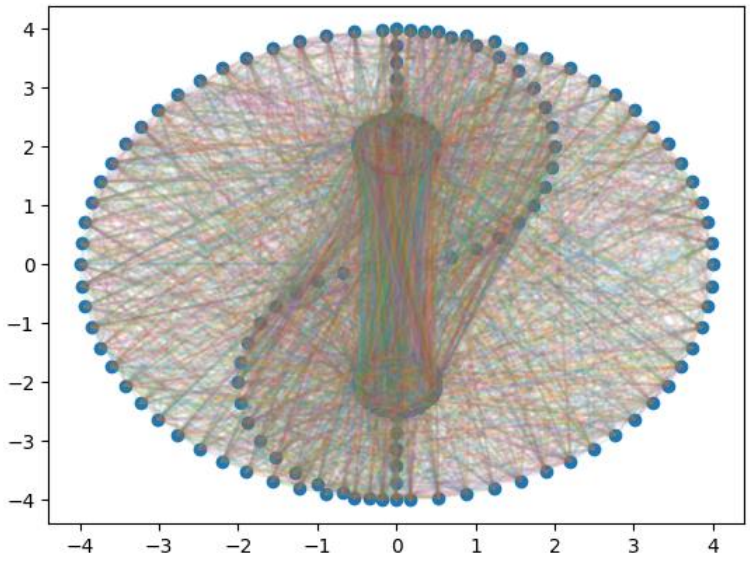


fig1 展示的是智能体的损失函数的局部函数图像，两个变量分别是 alpha 和 s，值得注意一点的是，为了让图像方便可视化，我们要将 y 轴的坐标限制在 7 以内，在进行模拟实验的时候，如果没有添加限制坐标的代码将会导致我们很难观察到图像的下界，其余地方的函数值达到了 1e10 的量级，如下图所示：

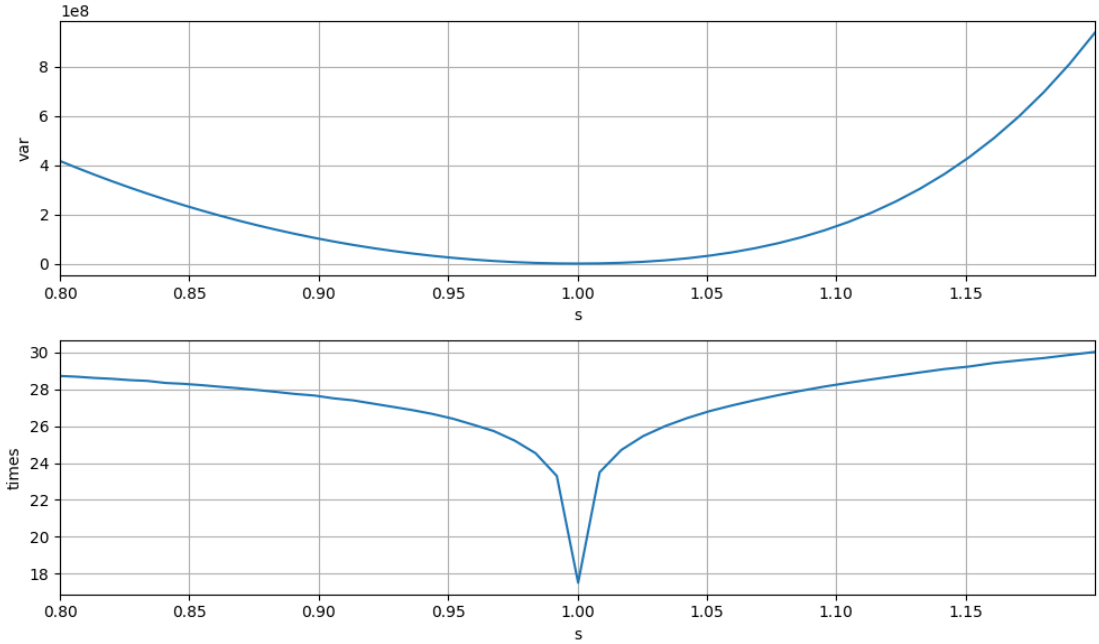


2.3.2 Fig2



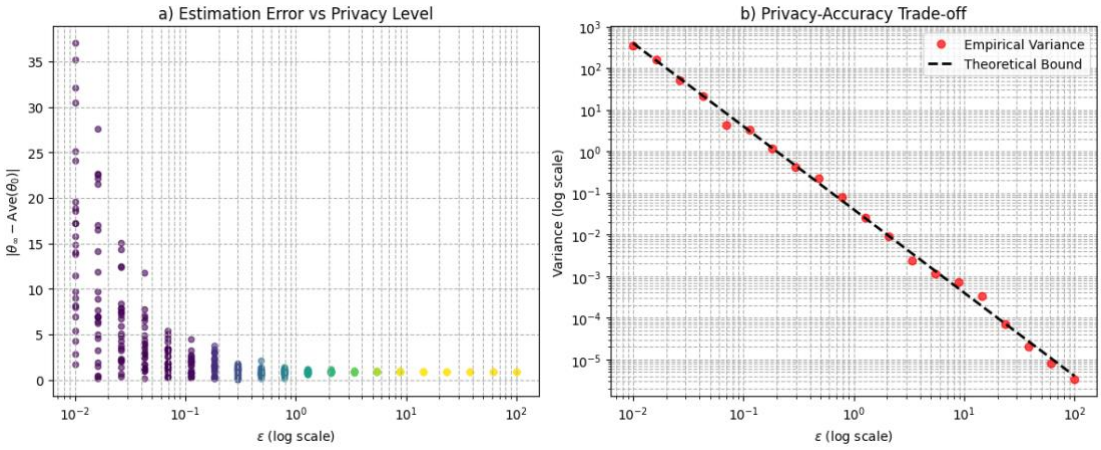
这张图片是对数据集的可视化实现，代码读取 npy 文件，利用散点图的方式绘制出每个点，在根据文章所提到的每个边的概率是两个节点的伯努利概率之和随机生成的边，再利用折线图绘制存在的边得到的图像。

2.3.3 Fig3



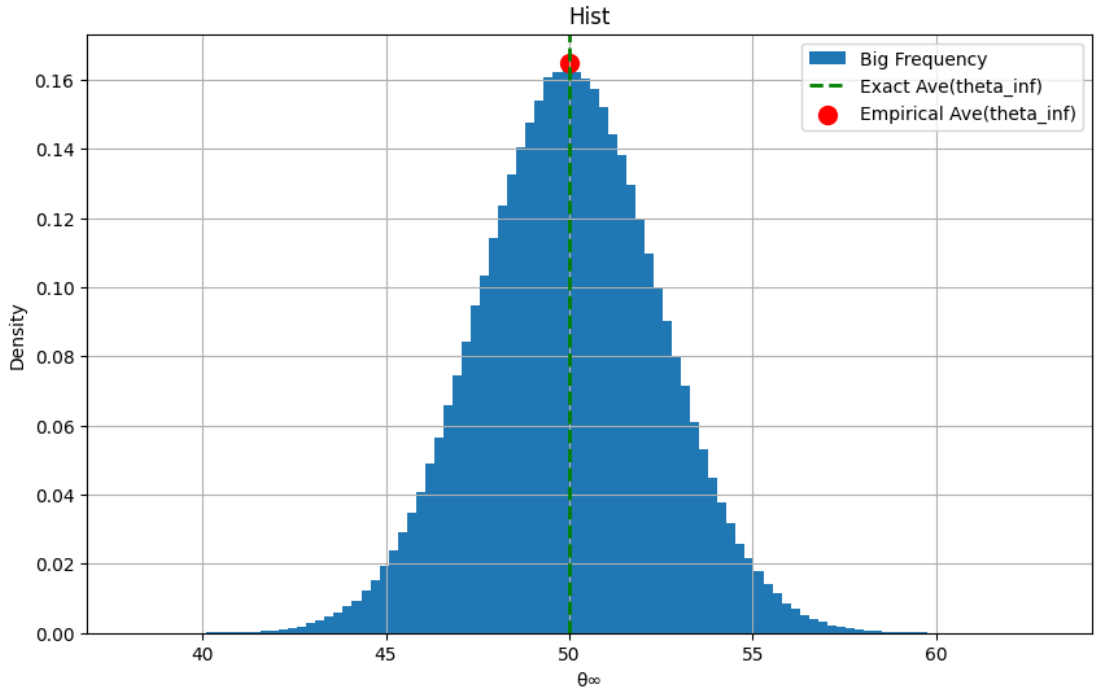
这是关于多智能体系统收敛方差和收敛时间随 s 变化的关系图，方差图像的凹凸性与文章所给出模拟实验结果似乎不一致，但是二者都能得出的结论是：当 s=1 时，收敛经验方差和收敛时间最小，而 s 参数对应噪声，图像告诉我们如何选择最优的噪声参数。

2.3.4 Fig4



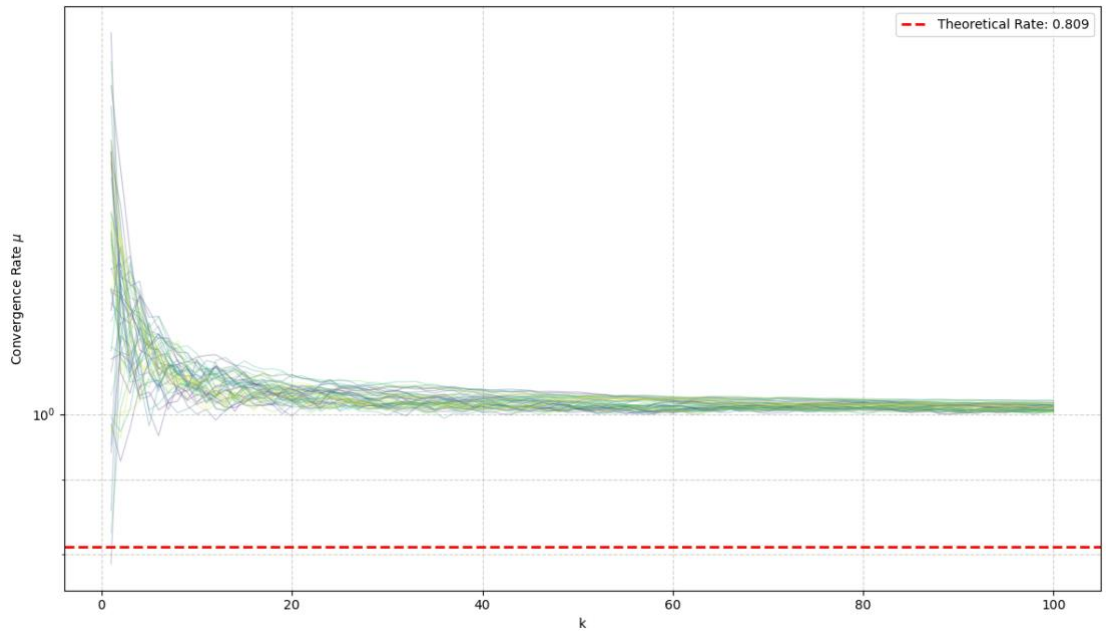
这张图片展示的是如何权衡隐私与数据准确性，左图展示了在 25 种不同噪声条件下最终的误差，我们可以看到它是呈递减趋势的；右图展示的是 100 次噪声下的样本方差和理论值的对照，根据模拟实验得出的图像，我们可以发现这些散点稳定的落在理论值直线附近，证明了理论推导的正确性。

2.3.5 Fig5



这张图片展示的是收敛点的统计分布，发现它与二项分布的图像非常相似，图中的绿色虚线代表理论均值，图中的红点代表实验均值，在图中我们也可以发现，红点正好落在绿色的虚线附近，这也表明了最后收敛状态下的期望值就是初始状态的平均值，符合理论的预期。

2.3.6 Fig6



这张图片展示的是收敛速率与 k 值的关系图，fig6 的模拟实验与原文的模拟实验结果偏差较大！最后的收敛速率不能趋近于理论值，编写的代码仍然有待改善！但是我们我可以从图中得到的结论是：收敛速率最终还是会趋向于某一个稳定的值（文中推导的理论值）。

2.4 学习思考与指引

笔者认为这是初学者在能力受限的情况下首次阅读论文最重要的一小节，论文的研究内容可能对初学者知识方面起不到重要帮助，但是对于初学者在能力突破、对未来论文探索上有很大的帮助！

【障碍一——翻译问题】第一次阅读论文遇到的障碍就是**语言障碍以及翻译不准确的障碍**，对于部分专业名词以及数学公式，仍需要根据翻译版并且**对照原文**（部分翻译甚至修改了公式的角标！这是一个很严重的问题），这样才能保证公式的准确性；部分翻译不连贯，专业名词翻译偏差，这也需要**对照原文**，才能够连贯的理解全文；

【障碍二——难以理解的专业名词】专业名词的不理解，学习者能力受限的情况如何获取文章的重要思想，实现文章内容的部分复现；笔者认为首先可以阅读论文的 **Abstract 和 Conclusion 部分**，这是论文的最终要实现的目标以及文章最后讨论的结果；还有**文章反复出现的命题和结论**，以本篇为例，状态方程 12～14 就是这篇文章讨论的基石，那么类似于这样的命题就要重点关注了

【障碍三——无从下手的公式推导】还有文章大篇幅的理论证明，这对于首次阅读的笔者是一个极大挑战！但是部分公式我们可以对其进行简化，简化后分析其逻辑过程，比如一个大串的公式我们可以把它改写成一个具有很多参数函数 f(a,b,c...)，这样就省去了很多复杂的

数学计算符号，而且还能知道背后推理的大致逻辑，类似于上一点，中间部分重要的过程，作者都会有标记，这些重要过程可以慢慢理解，**初学者可以更多的侧重于结论的运用！**

【障碍四——繁杂的公式】在论文结尾，笔者在编写图形复现的代码时，发现有很多参数，有印象但是不知道其具体出处，导致编写代码的时候发懵；解决方案：**归纳整理重要的公式**，如上文，罗列出了一些反复使用的公式还有各种参数的出处，这样在遇到某些参数的时候就可以在代码中很好的定义了！

在进行模拟实验之后，重新对各种参数有了自己的认知，现在回头看论文，就对论文大体思路有所了解（虽然还是对其中部分的推导过程不够了解）。

三、附录

3.1 参考文献

[1]Nozari, E., Tallapragada, P., & Cortés, J. (2017). Differentially private average consensus: Obstructions, trade-offs, and optimal algorithm design. Automatica, 83, 251-264.