基于事件触发的分布式优化算法

杨 涛[1] 徐 磊[1] 易新蕾[2] 张圣军[3] 陈蕊娟[4] 李渝哲[1]

摘 要 本文研究了一类分布式优化问题, 其目标是通过局部信息交换使由局部成本函数之和构成的全局成本函数最小. 针对无向连通图, 我们提出了两种基于比例积分策略的分布式优化算法. 在局部成本函数可微且凸的条件下, 证明了所提算法渐近收敛到全局最小值点. 更进一步, 在局部成本函数具有局部 Lipschitz 梯度和全局成本函数关于全局最小值点是有限强凸的条件下, 证明了所提算法的指数收敛性. 此外, 为了避免智能体之间的连续通信和减少通信负担, 将所提的两种分布式优化算法与事件触发通信相结合, 提出了两种基于事件触发的分布式优化算法. 证明了提出的事件触发优化算法不存在 Zeno 行为, 并且在相应条件下保持了与连续通信下分布式优化算法一样的收敛性. 最后, 通过数值仿真验证了上述理论结果.

关键词 分布式优化, 事件触发通信, Zeno 行为, 比例积分算法

引用格式 杨涛, 徐磊, 易新蕾, 张圣军, 陈蕊娟, 李渝哲. 基于事件触发的分布式优化算法. 自动化学报, 2022, **48**(1): 133−143

DOI 10.16383/j.aas.c200838

Event-triggered Distributed Optimization Algorithms

YANG Tao[1] XU Lei[1] YI Xin-Lei[2] ZHANG Sheng-Jun[3] CHEN Rui-Juan[4] LI Yu-Zhe[1]

Abstract This paper studies a class of distributed optimization problems, whose objective is to minimize the global cost function formed by a sum of local cost functions through local information exchanges. For undirected connected graphs, we propose two distributed optimization algorithms based on the proportional-integral feedback mechanism. Under the condition that the local cost functions are differentiable and convex, it is proved that the proposed algorithms asymptotically converge to a global minimum. For the case that the local cost functions have local Lipschitz gradient and the global cost function is strongly convex with respect to the global minimum, the exponential convergences of the two distributed optimization algorithms are established. In addition, in order to avoid continuous communication between agents and reduce communication burden, by integrating the two proposed distributed optimization algorithms with event-triggered communications, two event-triggered based distributed optimization algorithms are developed. It is shown that the two proposed event-triggered optimization algorithms are free of Zeno behavior. Moreover, the two proposed event-triggered based distributed optimization algorithms maintain the same convergence properties as the distributed optimization algorithms with continuous communications under the corresponding conditions. Finally, the above theoretical results are verified by numerical simulations.

Key words Distributed optimization, event-triggered communications, Zeno behavior, proportional-integral algorithm

Citation Yang Tao, Xu Lei, Yi Xin-Lei, Zhang Sheng-Jun, Chen Rui-Juan, Li Yu-Zhe. Event-triggered distributed optimization algorithms. *Acta Automatica Sinica*, 2022, **48**(1): 133−143

在多智能体系统中, 每个智能体 (节点) 都具有一个局部成本函数, 分布式优化的目标是使由局部成本函数之和所构成的全局成本函数最小. 分布式优化的研究由来已久, 至少可以追溯到[1−2]. 近年来, 由于其在电力系统、机器学习和传感器网络等领域的广泛应用, 这一研究重新引起了关注. 研究者设计了多种分布式优化算法, 详见综述文章[3−10], 大致可分为离散时间算法和连续时间算法.

收稿日期 2020-10-10 录用日期 2021-01-26
Manuscript received October 10, 2020; accepted January 26, 2021
国家自然科学基金委重大项目 (61991400, 61991403, 61991404, 61890924) 资助
Supported by Major Program of National Natural Science Foundation of China (61991400, 61991403, 61991404, 61890924)
本文责任编委 贺威
Recommended by Associate Editor HE Wei

1. 东北大学流程工业综合自动化国家重点实验室 沈阳 110819 中国 2. 瑞典皇家理工学院电气工程与计算机科学学院决策与控制系统系 斯德哥尔摩 10044 瑞典 3. 北德克萨斯大学电气工程系 德克萨斯州 丹顿 76203 美国 4. 华中科技大学人工智能与自动化学院 武汉 430074 中国
1. State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China 2. The Division of Decision and Control Systems, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm 10044, Sweden 3. The Department of Electrical Engineering, University of North Texas, Denton, TX 76203, USA 4. School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430074, China

现有的大多数离散时间算法均基于一致性算法和分布式次梯度下降 (Distributed gradient descent, DGD) 算法[11−15]. 尽管 DGD 算法可以处理非光滑凸函数的分布式优化问题, 并在通讯延迟、丢包等多个方向上进行扩展以处理更为实际的情况, 但由于使用了衰减步长, 因此收敛速度较慢. 在步长固定的情况下, 虽然 DGD 算法收敛速度快, 但只能收敛到最小值点的邻域[16−17]. 最近的研究集中在利用历史信息来设计具有固定步长的加速算法. 具体而言, 文献 [18−19] 中提出的算法是基于比例积分 (Proportional integral, PI) 控制策略, 文献 [20−25] 中提出的算法是基于分布式不精确梯度算法和分布式动态平均梯度跟踪技术[26]. 现有的连续时间算法可以分为两类: 第一类是文献 [27−29] 中提出的基于梯度的算法, 这类算法本质上都是基于 PI 控制策略, 其中每个智能体使用一个辅助状态(积分反馈) 来校正由不同局部梯度引起的误差; 第二类算法使用二阶 Hessian 信息, 例如文献 [30−32].

为了避免连续通信和减少通信负担, 事件触发通信和控制的思想最初是针对单个系统[33−35] 提出的. 后来这种思想被应用到分布式一致性问题[36−42]. 近年来, 研究者提出了基于事件触发通信的分布式优化算法[29, 43−50]. 文献 [29] 提出了一种不存在 Zeno 行为[51] 的事件触发算法, 即在有限时间内不会触发无限多次事件, 并针对无向连通图, 在局部成本函数强凸以及梯度局部 Lipschitz 且可微的条件下, 证明了算法指数收敛到最小值点的邻域. 受文献 [30] 提出的零梯度和 (Zero-gradient-sum, ZGS) 算法的启发, 文献 [44] 提出周期性的事件触发机制; 文献 [45] 则设计了基于动态事件触发的 ZGS 算法. 针对无向连通图或权平衡强连通的有向图, 在局部成本函数强凸且具有局部 Lipschitz Hessians 的条件下, 证明了算法的指数收敛性.

本文提出了两种基于比例积分策略的分布式优化算法, 并证明了算法的收敛性. 在此基础上, 为了减少通信负担, 我们提出了基于事件触发的分布式优化算法, 并证明了提出的基于事件触发的优化算法不存在 Zeno 行为, 且保持了与连续通信下分布式优化算法相同的收敛性. 文献 [29] 提出的事件触发算法只有在局部成本函数强凸且具有局部 Lipschitz 梯度的条件下收敛到全局最小值点的一个邻域, 而我们提出的算法在局部成本函数可微且凸的条件下, 即可精确地指数收敛到唯一的全局最小值点. 与文献 [46] 中提出的算法相比, 我们提出的算法更简单, 因为在执行文献 [46] 中的算法时需要一些特殊设计的增益参数. 与文献 [44−45] 中提出的基于二阶 Hessian 信息的事件触发 ZGS 算法相比, 我们提出的算法是基于一阶梯度的, 易于实现. 此外, ZGS 算法需要特殊的初始化, 而我们提出的算法允许任意初始化.

本文其余部分安排如下. 第 1 节介绍图论的基础知识. 第 2 节介绍本文所考虑的分布式优化问题. 第 3 节提出两种基于 PI 控制策略的分布式优化算法, 并分析所提算法的收敛性. 第 4 节提出基于事件触发通信机制的分布式优化算法并分析其收敛性. 第 5 节利用数值仿真验证理论结果. 第 6 节总结本文的主要结果并介绍未来的研究方向.

1 基础概念

在这一部分, 我们介绍图论的一些基本知识[52]. 考虑一个包含 N 个智能体的无向图 $\mathcal{G}=(\mathcal{V},\mathcal{E},\mathcal{A})$, 其中 $\mathcal{V}=\{1,2,\cdots,N\}$ 表示智能体的集合, $\mathcal{E}=\{(i,j): i,j\in\mathcal{V}, i\neq j\}$ 表示边的集合, $\mathcal{A}=[a_{ij}]\in\mathbf{R}^{N\times N}$ 表示加权邻接矩阵, 其中, 当 $(i,j)\in\mathcal{E}$ 时, $a_{ij}>0$; 当 $(i,j)\notin\mathcal{E}$ 时, $a_{ij}=0$. 在本文中, 我们还假设图中没有自环边, 即对于所有的 $i\in\mathcal{V}$, $a_{ii}=0$. 智能体 i 的邻居集合定义为 $\mathcal{N}_i=\{j\in\mathcal{V}|a_{ij}>0\}$. 在无向图中从节点 i_1 到节点 i_k 的路径是指存在节点序列 $i_1,\cdots,i_k$, 使得 $(i_j,i_{j+1})\in\mathcal{E}$, $j=1,\cdots,k-1$.

对于无向加权图 $\mathcal{G}$, 加权 Laplacian 矩阵 $L=[L_{ij}]\in\mathbf{R}^{N\times N}$ 的定义是 $L_{ii}=\sum_{j=1}^{N}a_{ij}$, 对于 $i\neq j$, 有 $L_{ij}=-a_{ij}$, 因此, Laplacian 矩阵的行和为零. 如果无向加权图 $\mathcal{G}$ 是连通的, 则 Laplacian 矩阵 L 有唯一的 0 特征值, 其对应的右特征向量为 1, 其他所有特征值均大于零.

符号说明: 给定一个矩阵 A, A^{T} 表示其转置矩阵. 对称矩阵 A 是半正 (负) 定的当且仅当其所有特征值均为非负 (非正) 时. 给定两个对称矩阵 M, N, $M\leq N$ 意味着 $M-N$ 是半负定的. 记号 $A\otimes B$ 表示矩阵 A 和 B 之间的 Kronecker 积. $\rho(\cdot)$ 代表矩阵的谱半径, $\rho_2(\cdot)$ 表示非负矩阵的最小正特征值. I_n 表示维数为 $n\times n$ 的单位矩阵. 1_n 表示 n 维的列向量, 其每个元素都为 1. $\|\cdot\|$ 表示向量的欧几里德范数或矩阵的诱导 2 范数. 给定一个向量 $[a_1,\cdots,a_N]^{\mathrm{T}}\in\mathbf{R}^N$, $\mathrm{diag}\{a_1,\cdots,a_N\}$ 是第 i 个对角线元素为 a_i 的对角矩阵. 对于列向量 $x_1,x_2,\cdots,x_N$, 那么由其组成的堆栈列向量用 $[x_1,x_2,\cdots,x_N]$ 表示.

2 问题描述

考虑由 N 个智能体组成的网络, 每个智能体都有一个局部凸函数 $f_i:\mathbf{R}^n\to\mathbf{R}$, $i\in\mathcal{V}$. 所有智能体共同协作以找到一个最小值点 x^*, 使全局成本函

数 $f(x)=\sum_{i=1}^{N} f_i(x)$ 最小, 即:

$$\min_{x\in\mathbf{R}^n} f(x)=\sum_{i=1}^{N} f_i(x) \tag{1}$$

智能体之间的通信用无向加权图 $\mathcal{G}=(\mathcal{V},\mathcal{E},\mathcal{A})$ 来描述, 其中 $\mathcal{V}=\{1,2,\cdots,N\}$ 是智能体的集合, $\mathcal{E}\subseteq\mathcal{V}\times\mathcal{V}$ 是边的集合, $\mathcal{A}$ 是加权邻接矩阵.

如引言部分所述, 为了避免智能体之间的连续通信, 研究者设计了一些分布式事件触发算法. 然而, 大多数现有的算法要么需要特殊的初始化, 要么只收敛到全局最小值点的一个邻域, 这些启发了本文的研究. 更具体地说, 我们的目标是设计任意初始化的事件触发算法, 并精确地收敛到全局最小值点.

我们对局部和全局成本函数作出以下假设:

假设 1. 对每个 $i\in\mathcal{V}$, 局部成本函数 $f_i(x)$ 是连续可微凸函数. 此外, $\sum_{i=1}^{N} f_i(x)$ 的全局最小值是有界的.

假设 2. 全局成本函数 $f(x)=\sum_{i=1}^{N} f_i(x)$ 关于全局最小值点 x^* 是 m_f-(有限) 强凸的, 即存在常数 $m_f>0$, 使得 $\sum_{i=1}^{N}(\nabla f_i(x)-\nabla f_i(x^*))^{\mathrm{T}}(x-x^*)\geq m_f\|x-x^*\|^2$, 对任意的 $x\in\mathbf{R}^n$ 成立.

假设 3. 对于每个 $i\in\mathcal{V}$, 局部成本函数 $f_i(x)$ 具有局部 Lipschitz 梯度, 即对任意紧集 $D\subseteq\mathbf{R}^n$, 存在常数 $M_i(D)>0$, 使得

$$\|\nabla f_i(x)-\nabla f_i(y)\|\leq M_i(D)\|x-y\|,\quad \forall x,y\in D$$

其中, $M_i(D)$ 称为函数 $f_i(x)$ 在紧集 D 上的 Lipschitz 常数.

在假设 1 的条件下, 分布式优化问题 (1) 的全局最小值点 x^* 可能不唯一. 但是, 如果假设 2 成立, 则很容易证明全局最小值点 x^* 是唯一的. 与大多数文献中局部成本函数是强凸的假设相比, 这种假设的限制较小. 详细讨论, 请参见文献 [18, 46], 假设 3 在现有文献中也被广泛使用.

3 基于 PI 的分布式优化算法

针对问题 (1), 我们提出两种基于 PI 反馈策略的分布式优化算法, 其中 $x_i(t)\in\mathbf{R}^n$ 表示第 i 个节点在时刻 t 对全局最小值点 x^* 的一个估计, 积分项 $q_i(t)\in\mathbf{R}^n$ 是用来校正第 i 个节点由于固定步长所产生的误差. 第一种算法如下:

$$\dot{x}_i(t)=-\sum_{j=1}^{N} L_{ij}x_j(t)-\sum_{j=1}^{N} L_{ij}q_j(t)-\nabla f_i(x_i(t)),\ \forall x_i(0) \tag{2a}$$

$$\dot{q}_i(t)=\sum_{j=1}^{N} L_{ij}x_j(t),\ \forall q_i(0) \tag{2b}$$

算法 (2) 的收敛性如下:

定理 1. 假设无向图是连通的, 并且假设 1 成立. 如果每个智能体 $i\in\mathcal{V}$ 运行分布式优化算法 (2), 则有:

1) 每个 $x_i(t),\ i\in\mathcal{V}$, 渐近收敛到全局最小值点 x^*;

2) 如果假设 2 ~ 3 也成立, 则每个 $x_i(t),\ i\in\mathcal{V}$, 以不小于 $\frac{\epsilon_2}{\epsilon_3}$ 的速率指数收敛到唯一的全局最小值点 x^*, 其中 ϵ_2 和 ϵ_3 是两个正常数, 并在证明中给出.

证明. 见附录 B.

注 1. 算法 (2) 与文献 [27−28] 中所提出的分布式优化算法相类似. 但是, 文献 [27−28] 只给出了在凸成本函数下, 算法渐近收敛的结果. 而定理 1 给出了在全局成本函数对最小值点有限强凸的附加条件下, 算法指数收敛的结果. 针对有向图为权平衡强连通的情形, 当局部成本函数是可微的凸函数并且具有全局 Lipschitz 梯度时, 文献 [28] 中的定理 5.4 给出了算法渐近收敛的证明. 这里我们考虑的是无向连通图, 在全局成本函数对全局最小值点是有限强凸的附加条件下, 定理 1 给出了算法的指数收敛性.

下面, 介绍第二种分布式优化算法:

$$\dot{x}_i(t)=-\sum_{j=1}^{N} L_{ij}x_j(t)-q_i(t)-\nabla f_i(x_i(t)),\ \forall x_i(0) \tag{3a}$$

$$\dot{q}_i(t)=\sum_{j=1}^{N} L_{ij}x_j(t),\ \sum_{i=1}^{N} q_i(0)=0 \tag{3b}$$

算法 (3) 的收敛性如下:

定理 2. 假设无向图是连通的, 并且假设 1 成立. 如果每个智能体 $i\in\mathcal{V}$ 运行分布式优化算法 (3), 则有:

1) 每个 $x_i(t),\ i\in\mathcal{V}$, 渐近收敛到全局最小值点 x^*;

2) 如果假设 2 ~ 3 也成立, 则每个 $x_i(t),\ i\in\mathcal{V}$, 指数收敛到唯一的全局最小值点 x^*.

证明. 该定理的证明与定理 1 的证明相类似, 这里不再赘述.

注 2. 算法 (3) 与文献 [29] 中所提出的算法相类似. 但是, 文献 [29] 考虑的是局部成本函数强凸的情况, 而本文中只要求全局成本函数关于全局最小值点是有限强凸的, 是较之更一般的情况.

注 3. 算法 (2) 中 $x_i(0)$ 和 $q_i(0)$ 均可以任意选择的, 而在算法 (3) 中, 虽然 $x_i(0)$ 可以任意选择, 但要求 $\sum_{i=1}^{N} q_i(0)=0$, 因此算法 (2) 对初始条件 $q_i(0)$ 是更为鲁棒的. 然而, 与算法 (3) 相比, 式 (2b) 需要额外通信 q_j, 因此比算法 (3) 需要更多的通信开销.

4 基于事件触发的分布式 PI 优化算法

为了避免智能体之间的连续通信和减少通信负担, 将第 4 节所提出的分布式 PI 算法与事件触发通信相结合, 提出了两种分布式事件触发算法并给出其收敛性分析. 首先, 基于分布式优化算法 (2), 我们提出第一种事件触发算法, 描述如下:

$$\dot{x}_i(t) = -\sum_{j=1}^{N} L_{ij} x_j(t^j_{k_j(t)}) - \sum_{j=1}^{N} L_{ij} q_j(t^j_{k_j(t)}) - \nabla f_i(x_i(t)),\ \forall x_i(0), \quad t \in [t^i_k, t^i_{k+1}) \tag{4a}$$

$$\dot{q}_i(t) = \sum_{j=1}^{N} L_{ij} x_j(t^j_{k_j(t)}), \quad \forall q_i(0) \tag{4b}$$

其中, $\{t^j_k\}_{k=1}^{\infty}, \forall j \in \mathcal{V}$ 是待定的触发时间序列, 并且 $t^j_{k_j(t)} = \max\{t^j_k : t^j_k \le t\}$.

分布式事件触发算法设计中的关键问题是如何构造触发机制, 以保证提出的算法不存在 Zeno 行为, 并收敛到全局最小值点.

定理 3. 假设无向图是连通的, 并且假设 1 成立. 如果每个智能体 $i \in \mathcal{V}$ 运行分布式事件触发算法 (4), 并通过如下方式确定其触发时间序列:

$$t^i_{k+1} = \max_{t \ge t^i_k} \Big\{ t : \|x_i(t) - x_i(t^i_k)\| \le a_i \mathrm{e}^{-b_i t} \text{ 和 } \|q_i(t) - q_i(t^i_k)\| \le c_i \mathrm{e}^{-d_i t} \Big\}, \quad k = 1, 2, \cdots \tag{5}$$

其中, 所有 $a_i, b_i, c_i, d_i > 0$ 均为设计参数, 则有:

1) 算法 (4) 不存在 Zeno 行为;

2) 每个 $x_i(t)$, $i \in \mathcal{V}$, 渐近收敛到全局最小值点 x^*;

3) 如果假设 2 ~ 3 也成立, 则每个 $x_i(t)$, $i \in \mathcal{V}$, 以不小于 $\frac{\epsilon_2}{\epsilon_4}$ 的速率指数收敛到唯一的全局最小值点 x^*, 其中 ϵ_4 是正常数, 并在证明中给出.

证明. 见附录 C.

接下来, 基于算法 (3), 我们提出了第二种事件触发算法, 描述如下所示:

$$\dot{x}_i(t) = -\sum_{j=1}^{N} L_{ij} x_j(t^j_{k_j(t)}) - q_i(t) - \nabla f_i(x_i(t)),\ \forall x_i(0), \quad t \in [t^i_k, t^i_{k+1}) \tag{6a}$$

$$\dot{q}_i(t) = \sum_{j=1}^{N} L_{ij} x_j(t^j_{k_j(t)}), \quad \sum_{i=1}^{N} q_i(0) = 0 \tag{6b}$$

下面的定理给出了与定理 3 类似的结果:

定理 4. 假设无向图是连通的, 并且假设 1 成立. 如果每个智能体 $i \in \mathcal{V}$ 运行分布式事件触发算法 (6), 并通过如下方式确定其触发时间序列:

$$t^i_{k+1} = \max_{t \ge t^i_k} \Big\{ t : \|x_i(t) - x_i(t^i_k)\| \le a_i \mathrm{e}^{-b_i t} \Big\}, \quad k = 1, 2, \cdots \tag{7}$$

其中, a_i, $b_i > 0$ 均为设计参数, 则有:

1) 算法 (6) 不存在 Zeno 行为;

2) 每个 $x_i(t)$, $i \in \mathcal{V}$, 渐近收敛到全局最小值点 x^*;

3) 如果假设 2 ~ 3 也成立, 则每个 $x_i(t)$, $i \in \mathcal{V}$, 指数收敛到唯一的全局最小值点 x^*.

证明. 该定理的证明与定理 3 的证明相类似, 这里不再赘述.

注 4. 基于算法 (2) 和算法 (3), 提出了对应的事件触发算法 (4) 和算法 (6). 所提出的事件触发通信机制, 受到文献 [37] 中时变触发机制的启发. 与算法 (6) 的触发机制 (7) 相比, 算法 (4) 的触发机制 (5) 更为复杂且需要额外通信开销, 但是算法 (4) 的初始条件 $q_i(0)$ 是可以任意取值的, 更为鲁棒.

注 5. 文献 [29] 中定理 13 要求所有局部成本函数强凸, 而定理 3 和定理 4 只要求全局成本函数有限强凸, 并不要求所有的局部成本函数都如此或强凸. 此外, 我们提出的算法指数收敛到全局最小值点, 而文献 [29] 中提出的算法只能收敛到全局最小值点的一个邻域. 文献 [44−45] 提出的事件触发 ZGS 算法需要特殊的初始化, 而我们提出的算法允许对 $x_i(0)$ 的任意初始化.

5 仿真实验

考虑一个包含 50 个智能体的大规模网络, 其中局部成本函数 f_i 具体描述如下:

$$f_j(x) = 0.5\mathrm{e}^{-0.5x} + 0.4\mathrm{e}^{0.3x}$$

$$f_{5+j}(x) = x^2 \ln(2 + x^2)$$

$$f_{10+j}(x) = 0.5x^2 \ln(1 + x^2) + x^2$$

$$f_{15+j}(x) = x^2 + \mathrm{e}^{0.1x}$$

$$f_{20+j}(x) = \ln(\mathrm{e}^{-0.1x} + \mathrm{e}^{0.3x}) + 0.1x^2$$

$$f_{25+j}(x) = \frac{x^2}{\ln(2 + x^2)}$$

$$f_{30+j}(x) = 0.2\mathrm{e}^{-0.2x} + 0.4\mathrm{e}^{0.4x}$$

$$f_{35+j}(x) = 0.5x^2 \ln(2 + x^2)$$

$$f_{40+j}(x) = \frac{x^2}{\sqrt{x^2 + 1}} + 0.1x^2$$

$$f_{45+j}(x) = x^4 + 2x^2 + 2$$

其中, $j=1,2,\cdots,5$. 函数 $f_i(x)$, $i=6,\cdots,10,\ 36,\cdots,40$, 是非强凸的函数, 而全局成本函数 $\sum_{i=1}^{50} f_i(x)$ 关于点 x^* 有限强凸. 所有局部成本函数均可微, 且具有局部 Lipschitz 梯度. 随机选取一个包含 50 个节点的无向连通图, 并针对该网络拓扑图以及上述定义的成本函数得到以下两部分的仿真结果.

不考虑事件触发时, 为了更好地体现算法 (2), 算法 (3) 与文献 [29−31] 所提算法的区别, 我们对这些算法进行了仿真比较. 通过图 1 可以看出所有算法均为线性收敛. 此外, 算法 (3) 的收敛速度相对较快.

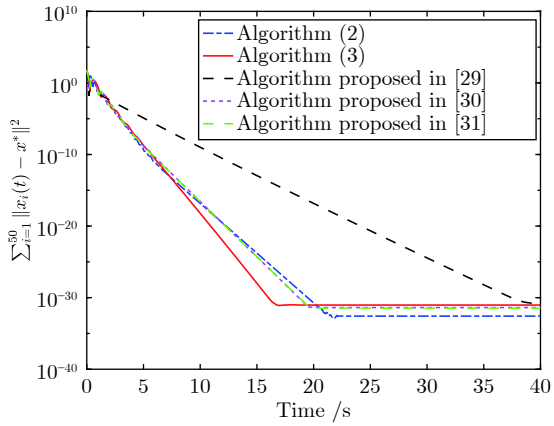


图 1　不同算法中 $\sum_{i=1}^{50}\|x_i(t)-x^*\|^2$ 的演化

Fig. 1　The state evolution of $\sum_{i=1}^{50}\|x_i(t)-x^*\|^2$ in various algorithms

考虑事件触发时, 对于算法 (4), 我们随机选择触发机制 (5) 中的设计参数 a_i, b_i, c_i 和 d_i. 选择采样周期为 0.01 s. 图 2 展示了智能体 6, 16, 26, 36, 46 的状态演化过程, 从中我们可以清楚地看到每个智能体都收敛到全局最小值点 $x^*=-0.01214$. 在 [0, 40 s] 时间段内, 上述 5 个智能体分别被触发了 209, 183, 161, 241, 142 次. 由此可知, 事件触发算法 (4) 在仿真中针对上述 5 个节点避免了大约 95.32% 的通信开销.

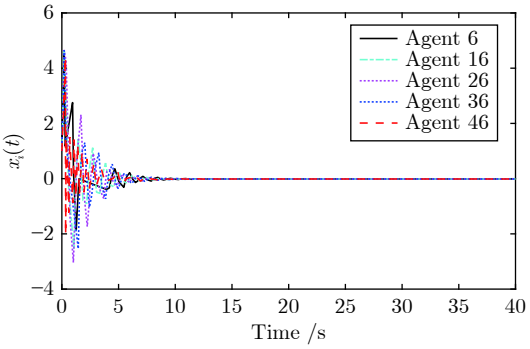


图 2　算法 (4) 中智能体 6, 16, 26, 36, 46 的状态演化

Fig. 2　State evolutions of agents 6, 16, 26, 36, 46 of Algorithm (4)

对于算法 (6), 我们随机选择触发机制 (7) 中的设计参数 a_i 和 b_i. 选择采样周期为 0.01 s. 图 3 展示了智能体 6, 16, 26, 36, 46 的状态演化过程, 从中我们可以清楚地看到所有智能体收敛到全局最小值点 $x^*=-0.01214$. 在 [0, 40 s] 时间段内, 上述 5 个智能体分别被触发了 114, 121, 139, 94, 182 次. 由此可知, 事件触发算法 (6) 在仿真中针对上述 5 个节点避免了大约 96.75% 的通信. 与算法 (4) 所对应的结果相比, 智能体被触发的次数更少, 因此节省了更多的通信和计算量.

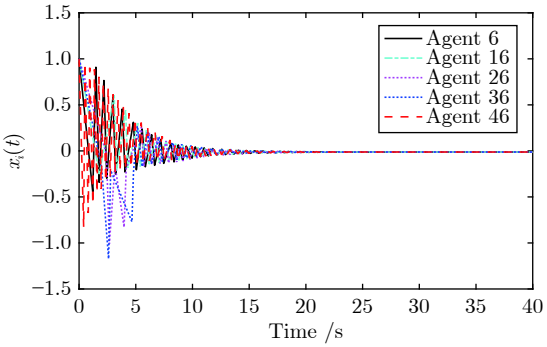


图 3　算法 (6) 中智能体 6, 16, 26, 36, 46 的状态演化

Fig. 3　State evolutions of agents 6, 16, 26, 36, 46 of Algorithm (6)

6　结论

本文考虑了一类分布式优化问题, 针对无向连通图, 基于比例积分策略提出了两类分布式优化算法, 在局部成本函数为可微凸函数的条件下, 证明了所提的分布式优化算法渐近收敛到全局最小值点. 当局部成本函数具有局部 Lipschitz 梯度, 并且全局成本函数对全局最小值点是强凸时, 证明了所提算法指数收敛到唯一的全局最小值点. 此外, 为了避免智能体之间的连续通信和减少通信开销, 提出了两种基于事件触发的分布式优化算法. 证明了所提出的算法不存在 Zeno 行为, 并且在相对应条件下保持了与连接通信下分布式优化算法一样的收敛性. 未来的一个方向是设计分布式优化算法动态事件触发通信机制的条件.

附录 A

下面文献 [46] 的引理, 在后文中指数收敛的证明中起着重要作用.

引理 1. 假设无向图是连通的. 如果令 $K_N=I_N-\frac{1}{N}\mathbf{1}_N\mathbf{1}_N^{\mathrm{T}}$. 则有, Laplacian 矩阵 L 是半正定的, $K_N\mathbf{1}_N=0$,

以及

$$K_N L = L K_N = L, \quad \rho(K_n) = 1$$
$$0 \le \rho_2(L) K_n \le L \tag{A1}$$

如下引理是对文献 [18] 中性质 3.6 的推广, 其中每个局部成本函数梯度局部 Lipschitz 的假设放松了原有局部成本函数梯度全局 Lipschitz 的假设, 在接下来的证明中也很有用.

引理 2. 假设无向图是连通的, 以及假设 1 ~ 3 都成立. 那么对任意 $\varepsilon > 0$ 以及任意紧的凸集 $D \subseteq \mathbf{R}^n$, 以及 $x^* \in D$,

$$(\nabla f(\boldsymbol{x}) - \nabla f(\bar{\boldsymbol{x}}))^{\mathrm{T}}(\boldsymbol{x} - \bar{\boldsymbol{x}}) + \varepsilon \boldsymbol{x}^{\mathrm{T}}(L \otimes I_n)\boldsymbol{x} \ge$$
$$m(\varepsilon)\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2, \quad \forall \boldsymbol{x} \in D^N \tag{A2}$$

其中, $\boldsymbol{x} = [x_1, \cdots, x_N]$, $f(\boldsymbol{x}) = \sum_{i=1}^N f_i(x_i)$, $\bar{\boldsymbol{x}} = \mathbf{1}_N \otimes x^*$, $m(\varepsilon) = \min\left\{m_f - 2M(D)\iota, \frac{\varepsilon\rho_2(L)\iota^2}{1+\iota^2}\right\}$, $M(D) = \max_{i\in\mathcal{V}}\{M_i(D)\}$, 以及 $\iota \in (0, \frac{m_f}{2M(D)})$.

证明. 对任意 $\boldsymbol{x} \in D^N$, 对 $\boldsymbol{x} = \boldsymbol{u} + \boldsymbol{v}$ 进行分解, 且 $\boldsymbol{u} = \mathbf{1}_N \otimes u_0$, $u_0 = \frac{1}{N}\sum_{i=1}^N x_i \in D$ 和 $\boldsymbol{v} = \boldsymbol{x} - \boldsymbol{u}$. 易知, $\boldsymbol{v}^{\mathrm{T}}(\mathbf{1}_N \otimes I_n) = \mathbf{0}_N$. 其余证明与文献 [18] 中性质 3.6 的证明相同, 这里不再赘述. □

附录 B

定理 1 的证明. 1) 在这一部分, 我们使用 Lyapunov 稳定性分析方法. 当假设 **1** 成立时, 每个 $x_i(t)$, $i \in \mathcal{V}$, 都渐近收敛到全局最小值点 x^*, 其可能是不唯一的. 方便起见, 记符号 $\boldsymbol{x} = [x_1, \cdots, x_N]$, $\boldsymbol{q} = [q_1, \cdots, q_N]$, 以及 $\nabla f(\boldsymbol{x}) = [\nabla f_1(x_1), \cdots, f_N(x_N)]$. 则算法 (2) 可以写为如下紧凑形式:

$$\dot{\boldsymbol{x}}(t) = -(L \otimes I_n)\boldsymbol{x}(t) - (L \otimes I_n)\boldsymbol{q}(t) - \nabla f(\boldsymbol{x}(t)) \tag{B1a}$$
$$\dot{\boldsymbol{q}}(t) = (L \otimes I_n)\,\boldsymbol{x}(t) \tag{B1b}$$

考虑如下函数:

$$V_1(\boldsymbol{x}, \boldsymbol{q}) = \frac{1}{2}\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2 + \frac{1}{2}\|\boldsymbol{q} - \bar{\boldsymbol{q}}\|^2 \tag{B2}$$

其中, $\bar{\boldsymbol{x}} = \mathbf{1}_N \otimes x^*$, $\bar{\boldsymbol{q}}$ 满足 $(L \otimes I_n)\bar{\boldsymbol{q}} = -\nabla f(\bar{\boldsymbol{x}})$(根据文献 [28] 中命题 3.2 可知 $\bar{\boldsymbol{q}}$ 存在).

$V_1(\boldsymbol{x}, \boldsymbol{q})$ 沿轨迹 (B1) 的导数满足:

$$\begin{aligned}\dot{V}_1(\boldsymbol{x}, \boldsymbol{q}) =& (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}[-(L \otimes I_n)\boldsymbol{x} - (L \otimes I_n)\boldsymbol{q} - \\ & \nabla f(\boldsymbol{x})] + (\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(L \otimes I_n)\boldsymbol{x} = \\ & (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}[-(L \otimes I_n)\boldsymbol{x} - (L \otimes I_n)(\boldsymbol{q} - \bar{\boldsymbol{q}}) + \\ & \nabla f(\bar{\boldsymbol{x}}) - \nabla f(\boldsymbol{x})] + (\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(L \otimes I_n)\boldsymbol{x} = \\ & -\boldsymbol{x}^{\mathrm{T}}(L \otimes I_n)\boldsymbol{x} - \\ & (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}[\nabla f(\boldsymbol{x}) - \nabla f(\bar{\boldsymbol{x}})] \le 0\end{aligned} \tag{B3}$$

其中, 用到了 $(L \otimes I_n)\bar{\boldsymbol{q}} = -\nabla \boldsymbol{f}(\bar{\boldsymbol{x}})$ 和 $\bar{\boldsymbol{x}}^{\mathrm{T}}(L \otimes I_n) = 0$. 由式 (B3) 可得, $V_1(\boldsymbol{x}, \boldsymbol{q})$ 径向无界, 从而由 LaSalle′s 不变集原理[53] 可知, $\boldsymbol{x}(t)$ 渐近收敛于 $\{\mathbf{1}_N \otimes x^1 \in \mathbf{R}^{nN} : \sum_{i=1}^n (x^1 - x^*) \times (\nabla f_i(x^1) - \nabla f_i(x^*)) = 0\}$ 因为 $\sum_{i=1}^n \nabla f_i(x^*) = 0$, 因此可知 $\sum_{i=1}^n (x^1 - x^*)(\nabla f_i(x^1) - \nabla f_i(x^*)) = 0$ 等价于 $\sum_{i=1}^n \nabla f_i(x^1) = 0$, 即 x^1 是全局最小值点. 因此, $\boldsymbol{x}(t)$ 渐近收敛于 $\mathbf{1}_N \otimes x^*$, 即每个 $x_i(t)$, $i \in \mathcal{V}$, 渐近收敛于 x^*.

2) 在这一部分, 我们证明当假设 2 和 3 成立时, 算法的指数收敛性.

我们首先证明对于任意的初始状态 $\boldsymbol{x}(0)$ 和 $\boldsymbol{q}(0)$, 存在凸紧集 $\mathcal{C} \subseteq \mathbf{R}^n$, 使得 $x^* \in \mathcal{C}$ 和 $x_i(t) \in \mathcal{C}$, $\forall t \ge 0$, $\forall i \in \mathcal{V}$. 集合 $\mathcal{C}$ 的具体形式依赖于 $\boldsymbol{x}(0)$, $\boldsymbol{q}(0)$ 和 x^*, 将在后文中给出.

由式 (B2) 和 (B3), 对任意 $t \ge 0$ 和 $i \in \mathcal{V}$, 可得:

$$\|x_i(t) - x^*\|^2 \le \|\boldsymbol{x}(t) - \bar{\boldsymbol{x}}\|^2 \le 2V_1(\boldsymbol{x}(t), \boldsymbol{q}(t)) \le 2V_1(\boldsymbol{x}(0), \boldsymbol{q}(0))$$

因此, $\mathcal{C} = \{x \in \mathbf{R}^n : \|x - x^*\|^2 \le 2V_1(\boldsymbol{x}(0), \boldsymbol{q}(0))\}$ 是我们要寻找的凸紧集.

接下来, 考虑如下函数:

$$V_2(\boldsymbol{x}, \boldsymbol{q}) = (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}(K_N \otimes I_n)(\boldsymbol{q} - \bar{\boldsymbol{q}}) \tag{B4}$$

$$V_3(\boldsymbol{x}, \boldsymbol{q}) = \frac{\epsilon_1 + 2}{2}(\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(K_N \otimes I_n)(\boldsymbol{q} - \bar{\boldsymbol{q}}) + \frac{\epsilon_1 + 1}{2}\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2 \tag{B5}$$

其中, $\epsilon_1 = \frac{2(M(\mathcal{C}))^2}{m_1\rho_2(L)}$, $m_1 = \min\left\{m_f - 2M(\mathcal{C})\iota, \frac{\rho_2(L)\iota^2}{2(1+\iota^2)}\right\}$, $M(\mathcal{C}) = \max_{i\in\mathcal{V}}\{M_i(\mathcal{C})\}$, 以及 $\iota \in (0, \frac{m_f}{2M(\mathcal{C})})$.

$V_2(\boldsymbol{x}, \boldsymbol{q})$ 沿轨迹 (B1) 的导数满足:

$$\begin{aligned}\dot{V}_2(\boldsymbol{x}, \boldsymbol{q}) =& (\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(K_N \otimes I_n)[-(L \otimes I_n)\boldsymbol{x} - \\ & (L \otimes I_n)\boldsymbol{q} - \nabla f(\boldsymbol{x})] + \\ & (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}(L \otimes I_n)\boldsymbol{x} = \\ & (\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(K_N \otimes I_n)[-(L \otimes I_n)\boldsymbol{x} - \\ & (L \otimes I_n)(\boldsymbol{q} - \bar{\boldsymbol{q}}) - (\nabla f(\boldsymbol{x}) - \nabla f(\bar{\boldsymbol{x}}))] + \\ & (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}(L \otimes I_n)\boldsymbol{x} \le \\ & -\frac{3}{4}(\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(L \otimes I_n)(\boldsymbol{q} - \bar{\boldsymbol{q}}) - \\ & (\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(L \otimes I_n)\boldsymbol{x} + \\ & (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}(L \otimes I_n)\boldsymbol{x} + \\ & \frac{\rho_2(L)}{4}(\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(K_N \otimes I_n)(\boldsymbol{q} - \bar{\boldsymbol{q}}) + \\ & \frac{1}{\rho_2(L)}\|\nabla f(\boldsymbol{x}) - \nabla f(\bar{\boldsymbol{x}})\|^2 \le \\ & -\frac{1}{2}(\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(L \otimes I_n)(\boldsymbol{q} - \bar{\boldsymbol{q}}) - \\ & (\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(L \otimes I_n)\boldsymbol{x} + \\ & (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}(L \otimes I_n)\boldsymbol{x} + \\ & \frac{1}{\rho_2(L)}\|\nabla f(\boldsymbol{x}) - \nabla f(\bar{\boldsymbol{x}})\|^2 \le \\ & -\frac{1}{2}(\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(L \otimes I_n)(\boldsymbol{q} - \bar{\boldsymbol{q}}) - \\ & (\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(L \otimes I_n)\boldsymbol{x} + \\ & (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}(L \otimes I_n)\boldsymbol{x} + \\ & \frac{(M(\mathcal{C}))^2}{\rho_2(L)}\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2\end{aligned} \tag{B6}$$

其中, 第二个等式用到了性质 $(L\otimes I_n)\bar{\boldsymbol{q}}=-\nabla \boldsymbol{f}(\bar{\boldsymbol{x}})$, 以及引理 1 中 (A1) 给出的性质 $K_NL=L$ 和 Cauchy-Schwarz 不等式; 第二个不等式利用引理 1 中 (A1) 给出的性质 $\rho_2(L)K_N\leq L$; 最后一个不等式用到了由假设 3 得到的 $f(\boldsymbol{x})$ 具有局部 Lipschitz 梯度的性质.

类似地, $V_3(\boldsymbol{x},\boldsymbol{q})$ 沿轨迹 (B1) 的导数满足:

$$\begin{aligned}\dot{V}_3(\boldsymbol{x},\boldsymbol{q})=&(\epsilon_1+1)(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}[-(L\otimes I_n)\boldsymbol{x}-\\&(L\otimes I_n)\boldsymbol{q}-\nabla f(\boldsymbol{x})]\\&(\epsilon_1+2)(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(L\otimes I_n)\boldsymbol{x}=\\&(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(L\otimes I_n)\boldsymbol{x}-(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}(L\otimes I_n)\boldsymbol{x}-\\&\epsilon_1(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}(L\otimes I_n)(\boldsymbol{x}-\bar{\boldsymbol{x}})-\\&(\epsilon_1+1)(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}[(\nabla f(\boldsymbol{x})-\nabla f(\bar{\boldsymbol{x}}))]\end{aligned}$$

进而可得,

$$\begin{aligned}\dot{V}_3(\boldsymbol{x},\boldsymbol{q})\leq&(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(L\otimes I_n)\boldsymbol{x}-(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}(L\otimes I_n)\boldsymbol{x}-\\&\frac{\epsilon_1}{2}\boldsymbol{x}^{\mathrm{T}}(L\otimes I_n)\boldsymbol{x}-\\&(\epsilon_1+1)(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}[(\nabla f(\boldsymbol{x})-\nabla f(\bar{\boldsymbol{x}}))]+\\&\frac{1}{4}(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(L\otimes I_n)(\boldsymbol{q}-\bar{\boldsymbol{q}})\end{aligned}\tag{B7}$$

由于假设 1 ~ 3 成立, 由引理 2 中式 (A2) 可得,

$$\begin{aligned}&(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}\big(\nabla f(\boldsymbol{x})-\nabla f(\bar{\boldsymbol{x}})\big)\geq\\&m_1\|\boldsymbol{x}-\bar{\boldsymbol{x}}\|^2-\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}(L\otimes I_n)\boldsymbol{x}\end{aligned}\tag{B8}$$

因此, 由 $(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}[(\nabla f(\boldsymbol{x})-\nabla f(\bar{\boldsymbol{x}}))]\geq 0$, $\epsilon_1=\frac{2(M(\mathcal{C}))^2}{m_1\rho_2(L)}$, 由式 (B6) ~ (B8) 可得

$$\begin{aligned}\dot{V}_2(\boldsymbol{x},\boldsymbol{q})+\dot{V}_3(\boldsymbol{x},\boldsymbol{q})\leq&-\frac{1}{4}(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(L\otimes I_n)(\boldsymbol{q}-\bar{\boldsymbol{q}})-\\&\epsilon_1m_1\|\boldsymbol{x}-\boldsymbol{x}^*\|^2\end{aligned}\tag{B9}$$

考虑如下候选 Lyapunov 函数

$$W_0(\boldsymbol{x},\boldsymbol{q})=V_2(\boldsymbol{x},\boldsymbol{q})+V_3(\boldsymbol{x},\boldsymbol{q})\tag{B10}$$

由 (B9) 可知, $W_0(\boldsymbol{x},\boldsymbol{q})$ 沿 (B1) 的轨迹的导数, 满足如下不等式,

$$\begin{aligned}\dot{W}_0(\boldsymbol{x},\boldsymbol{q})\leq&-\frac{1}{4}(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(L\otimes I_n)(\boldsymbol{q}-\bar{\boldsymbol{q}})-\\&\epsilon_1m_1\|\boldsymbol{x}-\boldsymbol{x}^*\|^2\leq\\&-\frac{\rho_2(L)}{4}(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(K_N\otimes I_n)(\boldsymbol{q}-\bar{\boldsymbol{q}})\\&-\epsilon_1m_1\|\boldsymbol{x}-\boldsymbol{x}^*\|^2\leq\\&-\epsilon_2[(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(K_N\otimes I_n)(\boldsymbol{q}-\bar{\boldsymbol{q}})+\\&\|\boldsymbol{x}-\boldsymbol{x}^*\|^2]\leq 0\end{aligned}\tag{B11}$$

其中, $\epsilon_2=\min\{\frac{\rho_2(L)}{4},\epsilon_1m_1,1\}>0$.

令 $\epsilon_3=\frac{\epsilon_1+3}{2}$, 可得,

$$\begin{aligned}W_0(\boldsymbol{x},\boldsymbol{q})\leq&\epsilon_3[(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(K_N\otimes I_n)(\boldsymbol{q}-\bar{\boldsymbol{q}})+\\&\|\boldsymbol{x}-\boldsymbol{x}^*\|^2]\end{aligned}\tag{B12}$$

从而, 由式 (B11) 和式 (B12) 可得 $\dot{W}_0(\boldsymbol{x},\boldsymbol{q})\leq-\frac{\epsilon_2}{\epsilon_3}W_0(\boldsymbol{x},\boldsymbol{q})$. 因此, $W_0(\boldsymbol{x}(t),\boldsymbol{q}(t))\leq W_0(\boldsymbol{x}(0),\boldsymbol{q}(0))\mathrm{e}^{-\frac{\epsilon_2}{\epsilon_3}t}$. 由 $\|\boldsymbol{x}-\bar{\boldsymbol{x}}\|^2\leq\frac{2}{\epsilon_1}W_0(\boldsymbol{x},\boldsymbol{q})$. 从而可知, $x_i(t)$ 以不小于 $\frac{\epsilon_2}{2\epsilon_3}>0$ 的速度指数收敛到全局唯一最小值点 x^*.

附录 C

定理 3 的证明. 我们首先给出一些必要的符号说明. 假定 $t_1^j=0$, $\forall j\in\mathcal{V}$. 简便起见, 我们分别定义 $\hat{x}_j(t)=x_j(t^j_{k_j(t)})$, $\hat{q}_j(t)=q_j(t^j_{k_j(t)})$, $e_j^x(t)=\hat{x}_j(t)-x_j(t)$, $e_j^q(t)=\hat{q}_j(t)-q_j(t)$, $\boldsymbol{x}(t)=[x_1(t),\cdots,x_N(t)]^{\mathrm{T}}$, 以及 $\boldsymbol{q}(t)=[q_1(t),\cdots,q_N(t)]^{\mathrm{T}}$.

1) 在这一部分中, 我们通过反证法证明算法 (4) 不存在 Zeno 行为. 假设算法 (4) 存在 Zeno 行为, 则存在一个智能体 $i\in\mathcal{V}$, 使得 $\lim_{k\to\infty}t_k^i=T_0$, 其中 $T_0>0$ 是一个常数. 注意到 $x_i(t)$ 和 $q_i(t)$ 都是连续的. 因此存在常数 $P_1>0$ 和 $P_2>0$, 使得 $\|x_i(t)\|\leq P_1$ 和 $\|q_i(t)\|\leq P_2$ 对所有 $i\in\mathcal{V}$ 和所有 $t\in[0,T_0]$ 都成立.

根据假设 1 可知 $f(\boldsymbol{x})$ 连续可微, 另外 $\|x_i(t)\|\leq P_1$, $\forall i\in\mathcal{V}$, $\forall t\in[0,T_0]$. 因此, 存在一个常数 $P_3>0$ 使得 $\|\nabla f(\boldsymbol{x})\|\leq P_3$, $\forall t\in[0,T_0]$.

令 $C_1=2L_{ii}P_1+2L_{ii}P_2+P_3$ 以及 $C_2=2L_{ii}P_1$. 那么, 由式 (4) 可得,

$$\begin{aligned}\|\dot{x}_i(t)\|\leq&\|\sum_{j=1}^N L_{ij}x_j(t^j_{k_j(t)})\|+\|\sum_{j=1}^N L_{ij}q_j(t^j_{k_j(t)})\|+\\&\|\nabla f_i(x_i(t))\|\leq C_1\end{aligned}\tag{C1}$$

$$\|\dot{q}_i(t)\|\leq\|\sum_{j=1}^N L_{ij}x_j(t^j_{k_j(t)})\|\leq C_2\tag{C2}$$

记 $a_0=\min\{a_1,\cdots,a_N\}$, $b_0=\max\{b_1,\cdots,b_N\}$, $c_0=\min\{c_1,\cdots,c_N\}$, $d_0=\max\{d_1,\cdots,d_N\}$, 以及 $\epsilon_0=\min\left\{\frac{a_0\mathrm{e}^{-b_0T_0}}{2C_1},\frac{c_0\mathrm{e}^{-d_0T_0}}{2C_2}\right\}>0$. 根据极限的性质可知, 存在正整数 $N(\epsilon_0)$, 使得,

$$t_k^i\in[T_0-\epsilon,T_0],\quad\forall k\geq N(\epsilon_0)\tag{C3}$$

对于给定的触发机制 (5), 可得 $\forall t\geq 0$,

$$\begin{aligned}\|\mathrm{e}_i^x(t)\|=\|x_i(t)-x_i(t_k^i)\|\leq a_i\mathrm{e}^{-b_it}\\\|\mathrm{e}_i^q(t)\|=\|q_i(t)-q_i(t_k^i)\|\leq c_i\mathrm{e}^{-d_it}\end{aligned}\tag{C4}$$

因为对于任何触发时间 t_k^i, 都有 $\hat{x}_i(t_k^i)=x_i(t_k^i)$. 因此由式 (C1) ~ (C2) 可得如下使得式 (C4) 成立的充分条件,

$$C_1(t-t_k^i)\leq a_0\mathrm{e}^{-b_0t}和C_2(t-t_k^i)\leq c_0\mathrm{e}^{-d_0t}\tag{C5}$$

假设我们已经确定了智能体 i 的第 $N(\epsilon)$ 次触发时间, 记为 $t^i_{N(\epsilon_0)}$. 记 $t^i_{N(\epsilon_0)+1}$ 和 $\tilde{t}^i_{N(\epsilon_0)+1}$ 分别为由式 (5) 和 (C5) 所确定的下一次触发时间. 则有, $t^i_{N(\epsilon_0)+1}\geq\tilde{t}^i_{N(\epsilon_0)+1}$ 以及

$$\begin{aligned}\tilde{t}^i_{N(\epsilon_0)+1}=&t^i_{N(\epsilon_0)}+\min\left\{\frac{a_0\mathrm{e}^{-b_0t^i_{N(\epsilon_0)}}}{C_1},\frac{c_0\mathrm{e}^{-d_0t^i_{N(\epsilon_0)}}}{C_2}\right\}\geq\\&t^i_{N(\epsilon_0)}+\min\left\{\frac{a_0\mathrm{e}^{-b_0T_0}}{C_1},\frac{c_0\mathrm{e}^{-d_0T_0}}{C_2}\right\}=\\&t^i_{N(\epsilon_0)}+2\epsilon_0\end{aligned}$$

其中, 不等号的成立用到了 $t^i_{N(\epsilon_0)} \le T_0$. 从而可得,

$$t^i_{N(\epsilon_0)+1} - t^i_{N(\epsilon_0)} \ge \tilde{t}^i_{N(\epsilon_0)+1} - t^i_{N(\epsilon_0)} \ge 2\epsilon_0$$

这与式 (C3) 相矛盾. 因此, 事件触发算法 (4) 不存在 Zeno 行为.

2) 在这一部分, 与定理 **1** 中 1) 部分的证明相类似. 我们使用 Lyapunov 稳定性分析证明, 当假设 **1** 成立时, 每个 $x_i(t)$, $i \in \mathcal{V}$ 都渐近收敛到全局最小值点 x^* . 方便起见, 记符号 $\hat{\boldsymbol{x}} = [\hat{x}_1, \cdots, \hat{x}_N]$, $\hat{\boldsymbol{q}} = [\hat{q}_1, \cdots, \hat{q}_N]$, $\boldsymbol{e}^x = [e^x_1, \cdots, e^x_N]$ 以及 $\boldsymbol{e}^q = [e^q_1, \cdots, e^q_N]$. 则算法 (4) 可以写成如下紧凑形式:

$$\dot{\boldsymbol{x}}(t) = -(L \otimes I_n)\hat{\boldsymbol{x}}(t) - (L \otimes I_n)\hat{\boldsymbol{q}}(t) - \nabla f(\boldsymbol{x}(t)) \tag{C6a}$$

$$\dot{\boldsymbol{q}}(t) = (L \otimes I_n)\,\hat{\boldsymbol{x}}(t) \tag{C6b}$$

考虑 Lyapunov 函数 (B2), 可得 $V_1(\boldsymbol{x}, \boldsymbol{q})$ 沿轨迹 (C6) 的导数满足:

$$\begin{aligned}
\dot{V}_1(\boldsymbol{x}, \boldsymbol{q}) =& (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}[-(L \otimes I_n)(\boldsymbol{x} + \boldsymbol{e}^x) - (L \otimes I_n)(\boldsymbol{q} + \boldsymbol{e}^q) - \\
& \nabla f(\boldsymbol{x})] + (\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(L \otimes I_n)(\boldsymbol{x} + \boldsymbol{e}^x) = \\
& (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}[-(L \otimes I_n)(\boldsymbol{x} + \boldsymbol{e}^x) \\
& - (L \otimes I_n)(\boldsymbol{q} - \bar{\boldsymbol{q}} + \boldsymbol{e}^q) + \\
& \nabla f(\bar{\boldsymbol{x}}) - \nabla f(\boldsymbol{x})] + (\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(L \otimes I_n)(\boldsymbol{x} + \boldsymbol{e}^x) = \\
& - \boldsymbol{x}^{\mathrm{T}}(L \otimes I_n)\boldsymbol{x} - \boldsymbol{x}^{\mathrm{T}}(L \otimes I_n)\boldsymbol{e}^x - \\
& \boldsymbol{x}^{\mathrm{T}}(L \otimes I_n)\boldsymbol{e}^q - (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}[\nabla f(\boldsymbol{x}) - \nabla f(\bar{\boldsymbol{x}})] + \\
& (\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(L \otimes I_n)\boldsymbol{e}^x \le \\
& - \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}(L \otimes I_n)\boldsymbol{x} + \rho(L)\|\boldsymbol{e}^x\|^2 + \rho(L)\|\boldsymbol{e}^q\|^2 + \\
& \rho(L)\|\boldsymbol{q} - \bar{\boldsymbol{q}}\|\|\boldsymbol{e}^x\| - \\
& (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}\big(\nabla f(\boldsymbol{x}) - \nabla f(\bar{\boldsymbol{x}})\big) \le
\end{aligned} \tag{C7}$$

$$\begin{aligned}
& - \frac{1}{2}\boldsymbol{x}^{\top}(L \otimes I_n)\boldsymbol{x} + \rho(L)(\|\boldsymbol{e}^x\|^2 + \|\boldsymbol{e}^q\|^2) + \\
& \frac{b}{2}\|\boldsymbol{q} - \bar{\boldsymbol{q}}\|^2 + \frac{(\rho(L))^2}{2b}\|\boldsymbol{e}^x\|^2 - \\
& (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\top}\big(\nabla f(\boldsymbol{x}) - \nabla f(\bar{\boldsymbol{x}})\big)
\end{aligned} \tag{C8}$$

其中, 第一个不等式由 $\boldsymbol{e}^x = \hat{\boldsymbol{x}} - \boldsymbol{x}$ 和 $\boldsymbol{e}^q = \hat{\boldsymbol{q}} - \boldsymbol{q}$ 得到; 根据 $(L \otimes I_n)\bar{\boldsymbol{q}} = -\nabla \boldsymbol{f}(\bar{\boldsymbol{x}})$ 可得第二个不等式; 最后一个不等式由 $\bar{\boldsymbol{x}}^{\mathrm{T}}(L \otimes I_n) = 0$, Young′s 不等式, Cauchy-Schwarz 不等式诱导的不等式, 以及 $b = \min\{b_1, \cdots, b_N, d_1, \cdots, d_N\} > 0$ 得到.

同时, 定义 $a = \max\{a_1, \cdots, a_N, c_1, \cdots, c_N\} > 0$. 那么由式 (5) 可得,

$$\|\boldsymbol{e}^x\|^2 \le Na^2\mathrm{e}^{-2bt} \text{ 和 } \|\boldsymbol{e}^q\|^2 \le Nc^2\mathrm{e}^{-2dt} \tag{C9}$$

由式 (B2) 可知,

$$\|\boldsymbol{q} - \bar{\boldsymbol{q}}\|^2 \le 2V_1(\boldsymbol{x}, \boldsymbol{q}) \tag{C10}$$

根据式 (C8), (C9) 和 (C10) 可得如下不等式

$$\dot{V}_1(\boldsymbol{x}, \boldsymbol{q}) \le bV_1(\boldsymbol{x}, \boldsymbol{q}) + 2\rho(L)Na^2\mathrm{e}^{-2bt} + \frac{1}{2b}(\rho(L))^2Na^2\mathrm{e}^{-2bt}$$

因此

$$V_1(\boldsymbol{x}(t), \boldsymbol{q}(t)) \le \gamma_1\mathrm{e}^{bt} \tag{C11}$$

其中

$$\gamma_1 = V_1(\boldsymbol{x}(0), \boldsymbol{q}(0)) + \frac{2}{3b}\rho(L)Na^2 + \frac{1}{6b}(\rho(L))^2Na^2$$

由式 (C10) 可得,

$$\|\boldsymbol{q} - \bar{\boldsymbol{q}}\| \le \sqrt{2\gamma_1}\mathrm{e}^{\frac{b}{2}t} \tag{C12}$$

根据式 (C7), (C9) 和 (C12) 可得如下不等式

$$\begin{aligned}
\dot{V}_1(\boldsymbol{x}, \boldsymbol{q}) \le & - \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}(L \otimes I_n)\boldsymbol{x} + \gamma_2 N\mathrm{e}^{-\frac{b}{2}t} - \\
& (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}\big(\nabla f(\boldsymbol{x}) - \nabla f(\bar{\boldsymbol{x}})\big)
\end{aligned} \tag{C13}$$

其中, $\gamma_2 = a(2a + \sqrt{\frac{2\gamma_1}{N}})\rho(L)$.

关于 $t \ge 0$ 和所有的 $i \in \mathcal{V}$ 定义 $z_i(t) = \mathrm{e}^{-\frac{b}{4}t}$, 记 $\boldsymbol{z}(t) = [z_1(t), \cdots, z_N(t)]$. 考虑如下候选 Lyapunov 函数:

$$W_1(\boldsymbol{x}, \boldsymbol{q}, \boldsymbol{z}) = V_1(\boldsymbol{x}, \boldsymbol{q}) + \frac{4\gamma_2}{b}\|\boldsymbol{z}\|^2 \tag{C14}$$

那么, 由式 (C13) 可知, $W_1(\boldsymbol{x}, \boldsymbol{q}, \boldsymbol{z})$ 沿轨迹 (C6) 的导数满足如下不等式,

$$\begin{aligned}
\dot{W}_1(\boldsymbol{x}, \boldsymbol{q}, \boldsymbol{z}) \le & - \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}(L \otimes I_n)\boldsymbol{x} - \gamma_2\|\boldsymbol{z}\|^2 - \\
& (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}\big(\nabla f(\boldsymbol{x}) - \nabla f(\bar{\boldsymbol{x}})\big) \le 0
\end{aligned} \tag{C15}$$

因此, $W_1(\boldsymbol{x}, \boldsymbol{q}, \boldsymbol{z})$ 径向无界. 由 LaSalle′s 不变集原理[53] 可知, $\boldsymbol{x}(t)$ 渐近收敛于 $\{\boldsymbol{1}_N \otimes x^1 \in \mathbf{R}^{nN} : \sum_{i=1}^n (x^1 - x^*)(\nabla f_i(x^1) - \nabla f_i(x^*)) = 0\}$. 因为 $\sum_{i=1}^n \nabla f_i(x^*) = 0$, 从而可知 $\sum_{i=1}^n (x^1 - x^*)(\nabla f_i(x^1) - \nabla f_i(x^*)) = 0$ 等价于 $\sum_{i=1}^n \nabla f_i(x^1) = 0$, 即 x^1 是全局最小值点. 因此, $\boldsymbol{x}(t)$ 渐近收敛于 $\boldsymbol{1}_N \otimes x^*$, 即每个 $x_i(t)$, $i \in \mathcal{V}$, 渐近收敛于 x^* .

3) 在这一部分, 证明当假设 2 ~ 3 成立时, 算法的指数收敛性. 与定理 **1** 中 2) 的证明相类似.

由式 (B2), (C14) 和 (C15), 可得对于所有的 $t \ge 0$ 和 $i \in \mathcal{V}$, 有如下不等式成立:

$$\|x_i(t) - x^*\|^2 \le 2W_1(\boldsymbol{x}(0), \boldsymbol{q}(0), \boldsymbol{z}(0))$$

因此, 所要寻找的凸紧集为 $\mathcal{C} = \{x \in \mathbf{R}^n : \|x - x^*\|^2 \le 2W_1(\boldsymbol{x}(0), \boldsymbol{q}(0), \boldsymbol{z}(0))\}$.

接下来, 由式 (B4) 可得 V_2 沿轨迹 (C6) 的导数满足

$$\begin{aligned}
\dot{V}_2(\boldsymbol{x}, \boldsymbol{q}) =& (\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(K_N \otimes I_n)[-(L \otimes I_n)\hat{\boldsymbol{x}} - (L \otimes I_n)\hat{\boldsymbol{q}} - \\
& \nabla f(\boldsymbol{x})] + (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}(L \otimes I_n)\hat{\boldsymbol{x}} = \\
& (\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(K_N \otimes I_n)[-(L \otimes I_n)\hat{\boldsymbol{x}} - \\
& (L \otimes I_n)(\boldsymbol{q} - \bar{\boldsymbol{q}} + \boldsymbol{e}^q) - (\nabla f(\boldsymbol{x}) - \nabla f(\bar{\boldsymbol{x}}))] + \\
& (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}(L \otimes I_n)\hat{\boldsymbol{x}} \le \\
& - \frac{3}{4}(\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(L \otimes I_n)(\boldsymbol{q} - \bar{\boldsymbol{q}}) + \rho(L)\|\boldsymbol{e}^q\|^2 - \\
& (\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(L \otimes I_n)\hat{\boldsymbol{x}}(t) + (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}(L \otimes I_n)\hat{\boldsymbol{x}} + \\
& \frac{\rho_2(L)}{4}(\boldsymbol{q} - \bar{\boldsymbol{q}})^{\mathrm{T}}(K_N \otimes I_n)(\boldsymbol{q} - \bar{\boldsymbol{q}}) +
\end{aligned}$$

$$\frac{1}{\rho_2(L)}\|\nabla f(\boldsymbol{x})-\nabla f(\bar{\boldsymbol{x}})\|^2\le$$
$$-\frac{1}{2}(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(L\otimes I_n)(\boldsymbol{q}-\bar{\boldsymbol{q}})+\rho(L)\|\boldsymbol{e}^q\|^2-$$
$$(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(L\otimes I_n)\hat{\boldsymbol{x}}+(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}(L\otimes I_n)\hat{\boldsymbol{x}}+$$
$$\frac{1}{\rho_2(L)}\|\nabla f(\boldsymbol{x})-\nabla f(\bar{\boldsymbol{x}})\|^2\le$$
$$-\frac{1}{2}(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(L\otimes I_n)(\boldsymbol{q}-\bar{\boldsymbol{q}})+\rho(L)\|\boldsymbol{e}^q\|^2-$$
$$(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(L\otimes I_n)\hat{\boldsymbol{x}}+(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}(L\otimes I_n)\hat{\boldsymbol{x}}+$$
$$\frac{(M(\mathcal{C}))^2}{\rho_2(L)}\|\boldsymbol{x}-\bar{\boldsymbol{x}}\|^2 \tag{C16}$$

其中, 第二个等式用到了 $\boldsymbol{e}^q$的定义, 性质 $(L\otimes I_n)\bar{\boldsymbol{q}}=-\nabla\boldsymbol{f}(\bar{\boldsymbol{x}})$, 以及引理 1 中 (A1) 给出的性质 $K_NL=L$ 和 Cauchy-Schwarz 不等式; 第二个不等式利用引理 1 中 (A1) 给出的性质 $\rho_2(L)K_N\le L$; 最后一个不等式用到了由假设 3 得到的 $f(\boldsymbol{x})$ 具有局部 Lipschitz 梯度的事实.

类似地, 由 (B15) 可得 V_3 沿轨迹 (C6) 的导数满足

$$\begin{aligned}\dot{V}_3(\boldsymbol{x},\boldsymbol{q})=&(\epsilon_1+1)(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}[-(L\otimes I_n)\hat{\boldsymbol{x}}-\\&(L\otimes I_n)\hat{\boldsymbol{q}}-\nabla f(\boldsymbol{x})]\\&(\epsilon_1+2)(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(L\otimes I_n)\hat{\boldsymbol{x}}=\\&(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(L\otimes I_n)\hat{\boldsymbol{x}}-(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}(L\otimes I_n)\hat{\boldsymbol{x}}-\\&\epsilon_1(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}(L\otimes I_n)(\boldsymbol{x}-\bar{\boldsymbol{x}}+\boldsymbol{e}^x)-\\&(\epsilon_1+1)(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}(L\otimes I_n)\boldsymbol{e}^q+\\&(\epsilon_1+1)(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(L\otimes I_n)\boldsymbol{e}^x-\\&(\epsilon_1+1)(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}[(\nabla f(\boldsymbol{x})-\nabla f(\bar{\boldsymbol{x}}))]\end{aligned}$$

进而可得,

$$\begin{aligned}\dot{V}_3(\boldsymbol{x},\boldsymbol{q})\le&(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(L\otimes I_n)\hat{\boldsymbol{x}}-(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}(L\otimes I_n)\hat{\boldsymbol{x}}-\\&\frac{\epsilon_1}{2}\boldsymbol{x}^{\mathrm{T}}(L\otimes I_n)\boldsymbol{x}+\epsilon_1\rho(L)\|\boldsymbol{e}^x\|^2+\\&\frac{(\epsilon_1+1)^2\rho(L)}{\epsilon_1}\|\boldsymbol{e}^q\|^2+(\epsilon_1+1)^2\rho(L)\|\boldsymbol{e}^x\|^2+\\&\frac{1}{4}(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(L\otimes I_n)(\boldsymbol{q}-\bar{\boldsymbol{q}})-\\&(\epsilon_1+1)(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}[(\nabla f(\boldsymbol{x})-\nabla f(\bar{\boldsymbol{x}}))]\end{aligned}\tag{C17}$$

因此, 由 $(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}}[(\nabla f(\boldsymbol{x})-\nabla f(\bar{\boldsymbol{x}}))]\ge 0$, $\epsilon_1=\frac{2(M(\mathcal{C}))^2}{m_1\rho_2(L)}$, 式 (B8) 以及式 (C9) 和 (C16) ~ (C17) 可得

$$\begin{aligned}\dot{V}_2(\boldsymbol{x},\boldsymbol{q})+\dot{V}_3(\boldsymbol{x},\boldsymbol{q})\le&-\\&\frac{1}{4}(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(L\otimes I_n)(\boldsymbol{q}-\bar{\boldsymbol{q}})-\\&\epsilon_1m_1\|\boldsymbol{x}-\boldsymbol{x}^*\|^2+\gamma_3N\mathrm{e}^{-2bt}\end{aligned}\tag{C18}$$

其中 $\gamma_3=a^2\rho(L)[1+\epsilon_1+(1+\epsilon_1)^2(1+\frac{1}{\epsilon_1})]$.

关于 $t\ge 0$ 和所有 $i\in\mathcal{V}$, 定义 $\zeta_i(t)=\mathrm{e}^{-bt}$ 和 $\boldsymbol{\zeta}(t)=[\zeta_1(t),\cdots,\zeta_N(t)]$. 考虑如下候选 Lyapunov 函数

$$W_2(\boldsymbol{x},\boldsymbol{q},\boldsymbol{\zeta})=V_2(\boldsymbol{x},\boldsymbol{q})+V_3(\boldsymbol{x},\boldsymbol{q})+\frac{\gamma_3}{b}\|\boldsymbol{\zeta}\|^2 \tag{C19}$$

由式 (C18) 和 (C19) 可知, $W_2(\boldsymbol{x},\boldsymbol{q},\boldsymbol{\zeta})$ 沿轨迹 (C6) 的导数满足如下不等式:

$$\begin{aligned}\dot{W}_2(\boldsymbol{x},\boldsymbol{q},\boldsymbol{\zeta})\le&-\frac{1}{4}(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(L\otimes I_n)(\boldsymbol{q}-\bar{\boldsymbol{q}})-\\&\epsilon_1m_1\|\boldsymbol{x}-\boldsymbol{x}^*\|^2-\gamma_3\|\boldsymbol{\zeta}\|^2\le\\&-\frac{\rho_2(L)}{4}(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(K_N\otimes I_n)(\boldsymbol{q}-\bar{\boldsymbol{q}})-\\&\epsilon_1m_1\|\boldsymbol{x}-\boldsymbol{x}^*\|^2-\gamma_3\|\boldsymbol{\zeta}\|^2\le\\&-\epsilon_2[(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(K_N\otimes I_n)(\boldsymbol{q}-\bar{\boldsymbol{q}})+\\&\|\boldsymbol{x}-\boldsymbol{x}^*\|^2+\gamma_3\|\boldsymbol{\zeta}\|^2]\end{aligned}\tag{C20}$$

令 $\epsilon_4=\max\{\frac{\epsilon_1+3}{2},\frac{1}{b}\}$, 则有:

$$\begin{aligned}W_2(\boldsymbol{x},\boldsymbol{q},\boldsymbol{\zeta})\le&\epsilon_4\big[(\boldsymbol{q}-\bar{\boldsymbol{q}})^{\mathrm{T}}(K_N\otimes I_n)(\boldsymbol{q}-\bar{\boldsymbol{q}})+\\&\|\boldsymbol{x}-\boldsymbol{x}^*\|^2+\gamma_3\|\boldsymbol{\zeta}\|^2\big]\end{aligned}\tag{C21}$$

那么, 由式 (C20) 和 (C21) 可得, $\dot{W}_2(\boldsymbol{x},\boldsymbol{q},\boldsymbol{\zeta})\le-\frac{\epsilon_2}{\epsilon_4}W_2(\boldsymbol{x},\boldsymbol{q},\boldsymbol{\zeta})$. 因此, $W_2(\boldsymbol{x}(t),\boldsymbol{q}(t),\boldsymbol{\zeta}(t))\le W_2(\boldsymbol{x}(0),\boldsymbol{q}(0),\boldsymbol{\zeta}(0))\mathrm{e}^{-\frac{\epsilon_2}{\epsilon_4}t}$. 由 $\|\boldsymbol{x}-\bar{\boldsymbol{x}}\|^2\le\frac{2}{\epsilon_1}W_2(\boldsymbol{x},\boldsymbol{q},\boldsymbol{\zeta})$ 可知, $x_i(t)$ 以不小于 $\frac{\epsilon_2}{2\epsilon_4}>0$ 的速度指数收敛到全局唯一最小值点 x^*.

References

1 Tsitsiklis J N. Problems in decentralized decision making and computation [Ph. D. dissertation], MIT, Cambridge, MA, 1984

2 Tsitsiklis J N, Bertsekas D P, Athans M. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Transactions on Automatic Control*, 1986, **31**(9): 803−812

3 Hong Yi-Guang, Zhang Yan-Qiong. Distributed optimization: algorithm design and convergence analysis. *Control Theory & Applications*, 2014, **31**: 850−857
(洪奕光, 张艳琼. 分布式优化: 算法设计和收敛性分析. 控制理论与应用, 2014, **31**: 850−857)

4 Yi Peng, Hong Yi-Guang. Distributed cooperative optimization and its applications. *Scientia Sinica Mathematica*, 2016, **46**(10): 1547−1564
(衣鹏, 洪奕光. 分布式合作优化及其应用. 中国科学: 数学, 2016, **46**(10): 1547−1564)

5 Xie Pei, You Ke-You, Hong Yi-Guang, Xie Li-Hua. A survey of distributed convex optimization algorithms over networks. *Control Theory & Application*, 2018, **35**(7): 918−927
(谢佩, 游科友, 洪奕光, 谢立华. 网络化分布式凸优化算法研究进展. 控制理论与应用, 2018, **35**(7): 918−927)

6 Nedić A, Olshevsky A, Rabbat M G. Network topology and communication-computation tradeoffs in decentralized optimization. *Proceedings of the IEEE*, 2018, **106**(5): 953−976

7 Wang Long, Lu Kai-Hong, Guan Yong-Qiang. Distributed optimization via multi-agent systems. *Control Theory & Applications*, 2019, **36**(11): 1820−1883
(王龙, 卢开红, 关永强. 分布式优化的多智能体方法. 控制理论与应用, 2019, **36**(11): 1820−1883)

8 Yang T, Yi X L, Wu J F, Yuan Y, Wu D, Meng Z Y, et al. A survey of distributed optimization. *Annual Reviews in Control*, 2019, **47**: 278−305

9 Khan U A, Bajwa W U, Nedić A, Rabbat M G, Sayed A H. Optimization for Data-Driven Learning and Control. *Proceedings of the IEEE*, 2020, **108**(11): 1863−1868

10 Yang Tao, Chai Tian-You. Research status and prospects of dis-

tributed collaborative optimization. *Scientia Sinica Technologica*, 2020, **50**(11): 1414−1425
(杨涛, 柴天佑. 分布式协同优化的研究现状与展望. 中国科学: 技术科学, 2020, **50**(11): 1414−1425)

11 Johansson B, Keviczky T, Johansson M, Johansson K H. Subgradient methods and consensus algorithms for solving convex optimization problems. In: Proceedings of *the IEEE Conference on Decision and Control*, Cancun, Mexico: IEEE, 2008. 4185−4190

12 Nedić A, Ozdaglar A. Distributed subgradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control*, 2009, **54**(1): 48−61

13 Zhu M, Martínez S. On distributed convex optimization under inequality and equality constraints. *IEEE Transactions on Automatic Control*, 2012, **57**(1): 151−164

14 Nedić A, Olshevsky A. Distributed optimization over time-varying directed graphs. *IEEE Transactions on Automatic Control*, 2015, **60**(3): 601−615

15 Yang T, Lu J, Wu D, Wu J, Shi G, Meng Z, Johansson K H. A distributed algorithm for economic dispatch over time-varying directed networks with delays. *IEEE Transactions on Industrial Electronics*, 2017, **64**(6): 5095−5106

16 Matei I, Baras J S. Performance evaluation of the consensus-based distributed subgradient method under random communication topologies. *IEEE Journal of Selected Topics in Signal Processing*, 2011, **5**(4): 754−771

17 Yuan K, Ling Q, Yin W. On the convergence of decentralized gradient descent. *SIAM Journal on Optimization*, 2015, **26**(3): 1835−1854

18 Shi W, Ling Q, Wu G, Yin W. EXTRA: An exact first-order algorithm for decentralized consensus optimization. *SIAM Journal on Optimization*, 2015, **25**(2): 944−966

19 Yao L, Yuan Y, Sundaram S, Yang T. Distributed finite-time optimization. In: Proceedings of the 14th International Conference on Control and Automation. Anchorage, AK, USA: IEEE, 2018. 147−154

20 Qu G, Li N. Harnessing smoothness to accelerate distributed optimization. *IEEE Transactions on Control of Network Systems*, 2018, **5**(3): 1245−1260

21 Xu J, Zhu S, Soh Y C, Xie L. Augmented distributed gradient methods for multi-agent optimization under uncoordinated constant stepsizes. In: Proceedings of the 54th IEEE Conference on Decision and Control. Osaka, Japan: IEEE, 2015. 2055−2060

22 Yang S, Tan S, Xu J X. Consensus based approach for economic dispatch problem in a smart grid. *IEEE Transactions on Power Systems*, 2013, **28**(4): 4416−4426

23 Du W, Yao L, Wu D, Li X, Liu G, Yang T. Accelerated distributed energy management for microgrids. In: Proceedings of the 2018 IEEE Power & Energy Society General Meeting. Portland, OR, USA: IEEE, 2018. 1−5

24 Pu S, Shi W, Xu J, Nedić A. A push-pull gradient method for distributed optimization in networks. In: Proceedings of the 57th IEEE Conference on Decision and Control. Miami, FL, USA: IEEE, 2018. 3385−3390

25 Xin R, Khan U A. A linear algorithm for optimization over directed graphs with geometric convergence. *IEEE Control Systems Letters*, 2018, **2**(3): 325−330

26 Zhu M, Martínez S. Discrete-time dynamic average consensus. *Automatica*, 2010, **46**(2): 322−329

27 Wang J, Elia N. Control approach to distributed optimization. In: Proceedings of the 48th Annual Allerton Conference on Communication, Control, and Computing. Allerton, Illinois, USA: IEEE, 2010. 557−561

28 Gharesifard B, Cortés J. Distributed continuous-time convex optimization on weight-balanced digraphs. *IEEE Transactions on Automatic Control*, 2014, **59**(3): 781−786

29 Kia S S, Cortés J, Martínez S. Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication. *Automatica*, 2015, **55**: 254−264

30 Lu J, Tang C Y. Zero-gradient-sum algorithms for distributed convex optimization: The continuous-time case. *IEEE Transactions on Automatic Control*, 2012, **57**(9): 2348−2354

31 Varagnolo D, Zanella F, Cenedese A, Pillonetto G, Schenato L. Newton-Raphson consensus for distributed convex optimization. *IEEE Transactions on Automatic Control*, 2016, **61**(4): 994−1009

32 Wei E, Ozdaglar A, Jadbabaie A. A distributed Newton method for network utility maximization-I: Algorithm. *IEEE Transactions on Automatic Control*, 2013, **58**(9): 2162−2175

33 Aström K J, Bernhardsson B M. Comparison of Riemann and Lebesgue sampling for first order stochastic systems. In: Proceedings of the 41st IEEE Conference on Decision and Control, Las Vegas, NV, USA: IEEE, 2002. 2011−2016

34 Tabuada P. Event-triggered real-time scheduling of stabilizing control tasks. *IEEE Transactions on Automatic Control*, 2007, **52**(9): 1680−1685

35 Girard A. Dynamic triggering mechanisms for event-triggered control. *IEEE Transactions on Automatic Control*, 2015, **60**(7): 1992−1997

36 Dimarogonas D V, Frazzoli E, Johansson K H. Distributed event-triggered control for multi-agent systems. *IEEE Transactions on Automatic Control*, 2012, **57**(5): 1291−1297

37 Seyboth G S, Dimarogonas D V, Johansson K H. Event-based broadcasting for multi-agent average consensus. *Automatica*, 2013, **49**(1): 245−252

38 Meng X, Xie L, Soh Y C, Nowzari C, Pappas G J. Periodic event-triggered average consensus over directed graphs. In: Proceedings of the 54th IEEE Transactions on Decision and Control. Osaka, Japan: IEEE, 2015. 4151−4156

39 Meng X, Xie L, Soh Y C. Asynchronous periodic event-triggered consensus for multi-agent systems. *Automatica*, 2017, **84**: 214−220

40 Yi X. Resource-constrained multi-agent control systems: Dynamic event-triggering, input saturation, and connectivity preservation. [Master thesis], Royal Institute of Technology, Sweden, 2017

41 Nowzari C, Cortés J, Pappas G. Event-triggered control for multi-agent average consensus. *Cooperative Control of Multi-Agent Systems. John Wiley & Sons, Ltd*, 2018, 177−208

42 Yi X, Yang T, Wu J, Johansson K H. Distributed event-triggered control for global consensus of multi-agent systems with input saturation. *Automatica*, 2019, **100**: 1−9

43 Liu S, Xie L, Quevedo D E. Event-triggered quantized communication-based distributed convex optimization. *IEEE Transactions on Control of Network Systems*, 2018, **5**(1): 167−178

44 Chen W, Ren W. Event-triggered zero-gradient-sum distributed consensus optimization over directed networks. *Automatica*, 2016, **65**: 90−97

45 Du W, Yi X, Jemin G, Johansson K H, Yang T. Distributed optimization with dynamic event-triggered mechanisms. In: Proceedings of the 57th IEEE Conference on Decision and Control, Miami, FL, USA: IEEE, 2018. 969−974

46 Yi X, Yao L, Yang T, George J, Johansson K H. Distributed optimization for second-order multi-agent systems with dynamic event-triggered communication. In: Proceedings of the 57th IEEE Conference on Decision and Control, Miami, FL, USA: IEEE, 2018. 3397−3402

47 Wang D, Gupta V, Wang W. An event-triggered protocol for distributed optimal coordination of double-integrator multi-agent systems. *Neurocomputing*, 2018, **319**(30): 34−41

48 Liu C, Li H, Shi Y, Xu D. Event-triggered broadcasting for distributed smooth optimization. In: Proceedings of the 58th IEEE Conference on Decision and Control, Nice, France: IEEE, 2019. 716−721

49 Liu C, Li H, Shi Y, Xu D. Distributed event-triggered gradient method for constrained convex minimization. *IEEE Transactions on Automatic Control*, 2020, **65**(2): 778−785

50 Li M, Su L, Liu T. Distributed optimization with event-triggered communication via input feedforward passivity. *IEEE Control Systems Letters*, 2020, **5**(1): 283−288

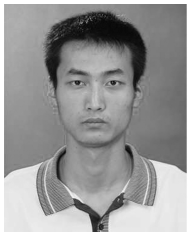
51 Johansson K H, Egerstedt M, Lygeros J, Sastry S. On the regularization of Zeno hybrid automata. *Systems & Control Letters*, 1999, **38**(3): 141−150

52 Godsi C, Royle G F, *Algebraic Graph Theory*, ser. Graduate Texts in Mathematics. New York: Springer-Verlag, 2001, 207

53 Khalil H K, *Nonlinear Systems*, 3rd ed. Prentice-Hall, New Jersey, 2002

杨 涛 东北大学流程工业综合自动化国家重点实验室教授. 主要研究方向为工业人工智能, 信息物理系统, 分布式协同控制和优化.
E-mail: yangtao@mail.neu.edu.cn
(**YANG Tao** Professor at the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, China. His research interest covers industrial artificial intelligence, cyber physical system, distributed collaborative control and optimization.)

徐 磊 东北大学流程工业综合自动化国家重点实验室博士研究生. 主要研究方向为分布式控制及优化, 网络化系统, 马尔科夫跳变系统.
E-mail: 2010345@stu.neu.edu.cn
(**XU Lei** Ph. D. candidate at the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, China. His research interest covers distributed control and optimization, network system, and Markovian jump systems.)



易新蕾 瑞典皇家理工学院电气工程与计算机科学学院博士后. 主要研究方向为在线优化, 分布式优化, 事件驱动控制.
E-mail: xinleiy@kth.se
(**YI Xin-Lei** Postdoctor at the School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Sweden. His research interest covers online optimization, distributed optimization, and event-triggered control.)

张圣军 北德州大学电气工程专业博士研究生. 主要研究方向为分布式优化, 统计学习, 稀疏主成分分析.
E-mail: ShengjunZhang@my.unt.edu
(**ZHANG Sheng-Jun** Ph. D. candidate in the Department of Electrical Engineering, University of North Texas, USA. His research interest covers distributed optimization, statistical learning, and Sparse PCA.)

陈蕊娟 华中科技大学人工智能与自动化学院博士研究生. 主要研究方向为基于动力系统的优化算法的设计和理论分析.
E-mail: ruijuancheni@hust.edu.cn
(**CHEN Rui-Juan** Ph. D. candidate at the School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, China. Her research interest covers the design and theoretical analysis of optimization algorithm based on dynamic system.)

李渝哲 东北大学流程工业综合自动化国家重点实验室教授. 主要研究方向为网络化系统, 信息物理系统, 人工智能与信息安全. 本文通信作者.
E-mail: yuzheli@mail.neu.edu.cn
(**LI Yu-Zhe** Professor at the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, China. His research interest covers network system, cyber physical system, artificial intelligence, and information security. Corresponding author of this paper.)