ARMA 信号的分布式检测

João Domingos 和 João Xavier

arXiv:2304.07225v1 [eess.SP] 14 Apr 2023

摘要——本文研究一种分布式检测设置，其中网络中的智能体希望检测嵌入时间相关噪声中的时变信号。与大多数以往工作仅考虑嵌入时间不相关噪声中的常信号不同，本文中的信号和噪声都可以具有一般的自回归滑动平均（ARMA）时序结构。感兴趣信号是 ARMA 滤波器的脉冲响应，而噪声则是另一个 ARMA 滤波器的输出，该滤波器的输入为白高斯噪声。

针对这一可能引发新现象的扩展问题设置，我们提出了一套全面的解决方案。首先，我们将著名的运行共识检测器（RCD）推广到这种相关设置；然后，利用底层 ARMA 结构设计 RCD 的高效实现；最后，推导 RCD 在这一 ARMA 设置下的理论渐近性能。结果表明，每个智能体的错误概率呈现两种情形之一：要么（a）错误概率以指数速度衰减至零，要么（b）收敛到严格为正的误差底限。情形（a）涵盖了大偏差理论中的经典结果，而情形（b）则是分布式检测中的新现象，由 ARMA 设置引发。我们对这两种情形进行了完整刻画：给出了以底层 ARMA 模型的零点和极点表述的必要且充分条件，以判定各情形何时出现；并为情形（a）的衰减率和情形（b）的正误差底限提供了闭式表达式。我们的分析还表明，ARMA 设置带来了两个新特点：（1）RCD 中使用的阈值会影响错误概率的渐近行为；（2）某些智能体可能具有弱信息性，也就是说，它们的观测不会改善 RCD 的渐近性能，因此可以安全地将其静音，以节省感知资源。数值仿真展示并验证了理论结果。

索引词——分布式检测、二元假设检验、ARMA、时间相关性、RunningConsensus 检测器、乔列斯基分解。

I. 引言

二元检测要解决的问题是：判断两个可能的统计源中，哪一个生成了观测数据。根据数据的观测方式，检测可以在集中式设置或分布式设置中进行。在集中式设置中，单个智能体观测全部数据，并可独自解决决策问题。在分布式设置中，多个智能体共同观测全局数据，但每个智能体只观测其中一部分局部数据。由于每个智能体现在只能观测全局数据的一部分，因此没有任何单个智能体能够独自以最优方式作出决策；智能体通过协作来解决决策问题会更有利。

JoãoDomingos 和 JoãoXavier
○
liveira.domingos@ist.utl.pt,jxavier@isr.tecnico.ulisboa.pt.
任职于葡萄牙里斯本大学 Instituto Superior Técnico 的 Instituto Sistemas e Robótica。

A. 分布式设置

本文考虑分布式设置中的检测，其中全局数据是维度为 N 的离散时间向量值信号。对于这一串全局数据向量，每个智能体 i（其中 $1 \le i \le N$）在每个时间步只在本地观测分量 i。这些智能体希望在以下两个假设之间作出判断：

- 假设 H_1（信号+噪声）：全局数据是嵌入在加性噪声中的给定确定性信号；
- 假设 H_0（仅噪声）：全局数据仅为噪声，

其中，信号和噪声都具有预先指定但除此之外不受限制的自回归移动平均（ARMA）时间结构。

B. 分布式设置中的相关工作

为了更好地将本文置于当前分布式设置下检测领域的研究背景中，下面回顾与本文最接近的工作。

总体而言，分布式设置下的检测主要分为两类：融合中心方法 [1]–[7] 和完全去中心化方法 [8]–[30]:

- 在融合中心方法中，智能体将本地数据的一些检验统计量（通常经过量化）传送给一个中心智能体。中心智能体收到这些统计量后，通过融合规则进行处理，从而得出集体决策 [1]–[7];
- 在去中心化方法中，通过扁平化架构消除了这种层级结构：不存在专门的中心智能体；所有智能体地位平等，并与可用的邻居智能体进行双向通信，以协作得出决策。在每个时间步，每个智能体都会整合新近观测到的（私有）数据，并将更新后的局部决策与邻居的决策相结合，从而更新自己的局部决策。去中心化检测的部分研究关注序贯检测，即只有在能够作出准确度超过可接受阈值的决策时，才停止收集观测 [8]–[10, 13, 14, 26]–[28, 30, 31]。另一部分研究则关注具有复合假设的非序贯检测 [21]–[24]。

本文属于第二类方法——完全去中心化方法。我们考虑一个去中心化的二元检测问题，其中待检测信号是已知但随时间变化的，而观测噪声是时间相关的高斯噪声。我们的方法采用了 [32] 提出并在 [16]中进一步研究的运行一致性检测器（RCD）。尽管关于去中心化检测的文献十分丰富

我们仍未发现其他研究考虑带有时间相关观测噪声的分布式检测。与之最接近的工作是 [19], , 其中 Pang 等人考虑了一个具有时间相关（但随机）检测信号的去中心化二元检测问题；具体而言，他们研究了带有独立同分布白高斯噪声的部分观测马尔可夫过程的检测。

本文假设检测信号是 ARMA 模型的脉冲响应，而观测噪声是另一个由白高斯噪声驱动的 ARMA 滤波器的输出。ARMA 信号在信号处理文献中已得到充分研究，因为它们是有理 $\mathcal{Z}$ 变换 [33]在时间域中的对应形式。由于已有大量文献研究如何使用 ARMA 模型逼近所需的频率响应 [34], , 只要能够准确逼近检测信号的频率响应和噪声的功率谱密度（假设其具有广义平稳性）, ARMA 假设就不会过于严格。另一种方法是在线估计相关参数：在 [35] 中，Kay 研究了如何检测嵌入在参数未知的自回归（AR）噪声中的已知（非 ARMA）信号；研究表明，通过使用最大似然估计迭代更新自回归参数，广义似然比检验（GLRT）在渐近意义下能够达到对数似然比（LLR）检测器的最优性能。这项工作在 [36],中得到了进一步推广，其中检测信号也可以是未知的。

然而需要指出的是，上述研究仅涉及 AR 过程和完全集中式架构。虽然 [35, 36] 提出了在若干参数未知时仍易于实现的集中式检测器，但我们研究的是已知 ARMA 信号在已知加性 ARMA 噪声（不一定局限于 AR 噪声）下的分布式检测器的实现及其渐近性能。

C. 贡献

据我们所知，这是首项研究时间相关观测噪声下分布式检测的工作。具体而言，我们提出了三项明确的贡献：

1) 通过探究 ARMA 滤波器的结构，我们证明，运行共识检测器（RCD）可以高效实现，其计算和内存需求不受观测数量影响；2) 更重要的是，我们全面刻画了 RCD 的渐近性能，并证明虚警概率和漏检概率可能呈现两种渐近情形：(a) 概率以指数速度衰减至零；或 (b) 概率收敛至严格为正的误差下限。情形（b）在分布式场景中是新发现的，它源于一般的 ARMA 设置。我们依据相关 ARMA 模型的零点和极点，明确了触发哪种情形，并推导出指数衰减率和误差下限的闭式表达式；这些表达式表明，RCD 中使用的阈值常数在情形（a）下无关紧要，但会显著影响情形（b）下的极限误差下限。

3) 我们还证明，ARMA 设置也可能产生弱信息型智能体。这类智能体的数据无法改善 RCD 的渐近性能，因此可以舍弃这些数据，而不会损害这一渐近指标。网络压缩方案或许可以利用这一特性：通过剔除大多数（甚至全部）弱信息型智能体，构建一个（可能小得多的）连通网络，并在其中运行 RCD。

D. 论文结构

本文其余部分安排如下。第二节介绍符号约定，并给出 ARMA 模型和 $\mathcal{Z}$变换所需的背景知识。第三节对分布式检测问题进行形式化描述，并介绍感知与通信模型。第四节将通用的运行共识检测器（RCD）扩展至 ARMA 情形。第五节推导加性有色噪声的分布，并运用 ARMA 滤波技术高效实现 RCD。第六节分析 RCD 的渐近性能。第七节探讨在多类有色噪声中检测直流分量的问题，以此结束全文。

II. 预备知识： $\mathcal{Z}$ 变换

为便于后续参考，我们先定义记号并回顾有理 $\mathcal{Z}$ 变换 [33] 的一些关键性质。

A. 记号

索引 $i = 1, \ldots, N$ 表示网络中的智能体，索引 $k = 0, 1, 2, \ldots$ 表示离散时间。

向量和矩阵。诸如 $z(k)$ 的普通向量记号表示网络在时刻 k 的快照：

$$z(k) = \begin{bmatrix} z_1(k) \\ \vdots \\ z_N(k) \end{bmatrix} \in \mathbf{R}^N,$$

其中 $z_i(k) \in \mathrm{R}$ 与智能体 i 相关联。粗体向量记号 $\boldsymbol{z}(k)$ 表示网络从时刻 0 到时刻 k 的历史：

$$\boldsymbol{z}(k) = \begin{bmatrix} \boldsymbol{z}_1(k) \\ \vdots \\ \boldsymbol{z}_N(k) \end{bmatrix} \in \mathbf{R}^{N(k+1)},$$

其中 $\boldsymbol{z}_i(k) \in \mathrm{R}^{k+1}$ 包含智能体 i 的历史，

$$\boldsymbol{z}_i(k) = \begin{bmatrix} z_i(0) \\ z_i(1) \\ \vdots \\ z_i(k) \end{bmatrix} \in \mathbf{R}^{k+1}.$$

请注意，普通的 $z(k)$ 维度固定（与 k 无关），而粗体的 $\boldsymbol{z}_i(k)$ 和 $\boldsymbol{z}(k)$ 的维度会随 k 增长。这一粗体规则也适用于矩阵。

例如， $\boldsymbol{Z}(k)$ 表示网络从 0 到时间 k 的历史，

$$\boldsymbol{Z}(k) = \mathrm{diag}\{\boldsymbol{Z}_1(k), \ldots, \boldsymbol{Z}_N(k)\} \in \mathbf{R}^{N(k+1)\times N(k+1)},$$

其中 $\mathrm{diag}\{\cdot\}$ 表示矩阵的对角堆叠，$\boldsymbol{Z}_i(k) \in \mathbf{R}^{(k+1)\times(k+1)}$ 包含与智能体 i 相关的历史。

若干符号。符号 $\mathbf{1}_n$ 和 I_n 分别表示所有分量均为 1 的 n 维向量和 $n \times n$ 单位矩阵；当上下文能够明确推断时，省略下标 n。向量的欧几里得范数记为 $\|\cdot\|$。对于一个 $m \times n$ 矩阵 X，其各列依次写作 $\begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$，符号 $\mathrm{vec}(X)$ 表示将 X 的各列从左到右堆叠而成的 mn 维列向量：$\mathrm{vec}(X) = \begin{bmatrix} x_1^T & \cdots & x_n^T \end{bmatrix}^T$。

记号 $x \sim \mathcal{N}(\mu, \Sigma)$ 表示随机向量 x 服从均值向量为 μ、协方差矩阵为 Σ 的高斯分布。符号 $\mathbb{E}(x)$ 表示作用于随机向量 x 的期望算子，而 $\mathbb{P}(A)$ 表示事件 A 发生的概率。

序列通常通过大 O 记号 $\mathcal{O}(\cdot)$ 和小 o 记号 $o(\cdot)$ 进行比较。对于序列 $(f(k))_{k\geq 0}$ 和 $(g(k))_{k\geq 0}$，记号 $f(k) = \mathcal{O}(g(k))$ 表示存在一个常数 $M > 0$ 和一个阶数 k_0，使得 $k \geq k_0$ 蕴含 $|f(k)| \leq M|g(k)|$。记号 $f(k) = o(g(k))$ 表示前述常数 M 可以任意小；也就是说，对于任意 $M > 0$，都存在一个阶数 k_0，使得 $k \geq k_0$ 蕴含 $|f(k)| \leq M|g(k)|$（当 $g(k) \neq 0$ 对于 k 足够大时，记号 $f(k) = o(g(k))$ 仅表示 $f(k)/g(k) \to 0$ 当 $k \to \infty$ 时）。

B. $\mathcal{Z}$ 变换

序列 $(x(k))_{k\geq 0}$ 的 $\mathcal{Z}$ 变换是一个复值函数

$$X(z^{-1}) = \sum_{k\geq 0} x(k)\, z^{-k}, \quad z \in \mathcal{ROC} \subseteq \mathbf{C},$$

其中 $\mathcal{ROC} := \{z \in \mathbf{C}\colon \sum_{k\geq 0} |x(k)\ z^{-k}| < \infty\}$ 表示其收敛域（ROC）。本文仅考虑有理的 $\mathcal{Z}$ 变换，

$$X(z^{-1}) = A\, \frac{\prod_{j=1}^{q} (1 - z_j\, z^{-1})}{\prod_{j=1}^{p} (1 - p_j\, z^{-1})}, \quad |z| > \rho, \tag{1}$$

其中 A 为常数，分子多项式的根 $\{z_j\}_{j=1}^q$ 是 $X(z^{-1})$ 的零点，而分母多项式的根 $\{p_j\}_{j=1}^p$ 是 $X(z^{-1})$ 的极点；在这种情况下，$\mathcal{ROC} := \{z\colon |z| > \rho\}$，其中 $\rho := \max\{|p_j|\colon j\}$。

部分分式展开。众所周知，可以从其 $\mathcal{Z}$ 变换中恢复序列 $(x(k))_{k\geq 0}$。对于有理 $\mathcal{Z}$ 变换，可以采用部分分式展开这一标准技术进行恢复：假设式（1）中的 $X(z^{-1})$ 是严格真有理的（阶数 p、q 满足 $p > q$），并且仅有单极点（$p_m \neq p_n$ 对于 $m \neq n$），首先将 $X(z^{-1})$ 表示为

$$X(z^{-1}) = \sum_{j=1}^{p} \frac{r_j}{1 - p_j\, z^{-1}},$$

留数 r_j 为

$$r_j = (1 - p_j\, z^{-1})\, X(z^{-1})\Big|_{z=p_j};$$

现在，将每个 p 简单分式求逆，得到

$$x(k) = \sum_{j=1}^{p} r_j\, p_j^k, \tag{2}$$

这表明序列 $(x(k))_{k\geq 0}$ 是 p 个离散时间复指数信号 p_j^k 的和。

ARMA 滤波器。考虑一个系数为 $\{b(j)\}_{j=1}^p, \{a(j)\}_{j=1}^q$ 的自回归移动平均（ARMA）(p, q) 滤波器，并设 $(x(k))_{k\geq 0}$ 为输入信号，对应的输出信号为 $(y(k))_{k\geq 0}$。输入信号和输出信号在时域中由以下线性递推关系联系：

$$y(k) = \sum_{j=1}^{p} b(j)\, y(k-j) + x(k) + \sum_{j=1}^{q} a(j)\, x(k-j), \tag{3}$$

（约定 $x(k) = 0$ 对 $k < 0$ 成立），并在 $\mathcal{Z}$ 域中表示为

$$Y(z^{-1}) = \frac{1 - \sum_{j=1}^{q} b(j)\, z^{-j}}{1 + \sum_{j=1}^{p} a(j)\, z^{-j}}\, X(z^{-1}). \tag{4}$$

之所以称为自回归移动平均，是因为在（3）中，我们有 p 个关于 $y(k)$ 的自回归项，以及 q 个关于 $x(k)$ 的移动平均项。方程（3）和（4）是 ARMA(p, q) 滤波器的等价表示，其中时域表示（3）更便于实现，而 $\mathcal{Z}$ 表示（4）更便于分析。

III. 问题表述

假设 N 个智能体收集局部观测，这些观测的统计特性因两个可能的世界状态而有所不同，分别记为 H_1 或 H_0。我们要解决的问题是设计一种协作算法，使智能体通过彼此间的局部交互逐步推进，并共同判定当前成立的是哪个假设：H_1 还是 H_0。

A. 感知与通信模型

具体而言，在假设 H_1 下，每个智能体 i 观测到已知确定性信号 $\theta_i(k)$ 的一个含噪版本；而在假设 H_0 下，每个智能体 i 仅观测到噪声：

$$\begin{aligned} H_1&:\ y_i(k) = \theta_i(k) + n_i(k),\ i = 1, \ldots, N \\ H_0&:\ y_i(k) = n_i(k),\ i = 1, \ldots, N, \end{aligned} \tag{5}$$

对于 $k \geq 0$ 和 $1 \leq i \leq N$。

我们假设噪声 $(n_i(k))$ 和信号 $(\theta_i(k))$ 都遵循 ARMA 结构，即将它们建模为 ARMA 滤波器的输出。对于噪声模型，相应滤波器的输入为 $\sigma\,\epsilon_i(k)$，其中 $\sigma > 0$ 和 $(\epsilon_i(k))$ 是白高斯噪声（WGN）（因此，$\sigma\,\epsilon_i(0)\ldots,\sigma\,\epsilon_i(k) \overset{i.i.d}{\sim} \mathcal{N}(0,\sigma^2)$）；对于信号模型，相应滤波器的输入为 $A\,\delta(k)$，其中 $A \neq 0$ 和 δ 表示单位冲激（$\delta(0) = 1$ 且 $\delta(k) = 0$ 对于 $k > 0$）。根据 (3)，可得

$$n_i(k) = \sum_{j=1}^{p_i} a_i(j)\, n_i(k-j) + \sigma\,\epsilon_i(k) + \sigma \sum_{j=1}^{q_i} b_i(j)\,\epsilon_i(k-j) \tag{6}$$

$$\theta_i(k) = \sum_{j=1}^{\bar{p}_i} \bar{a}_i(j)\,\theta_i(k-j) + A\,\delta(k) + A \sum_{j=1}^{\bar{q}_i} \bar{b}_i(j)\,\delta(k-j), \tag{7}$$

在已知 A、σ 和滤波器系数 $a_i(j)\,\bar{a}_i(j)\;b_i(j)\;\bar{b}_i(j)$ 的情况下。

通信模型。除感知外，智能体还可以通过在网络上交换信息与邻居通信。该网络建模为无向图 $G = (V, E)$，其中 $V = \{1,\ldots,N\}$ 是智能体集合，E 是通信信道集合。边 $(i,j) \in E$ 表示智能体 i 和 j 可以相互双向交换信息。我们用 $\mathcal{N}_i$ 表示智能体 i 的邻居：$\mathcal{N}_i := \{j \in V\colon (i,j) \in E\}$。

每个智能体 i 必须通过按时间 k 顺序处理其观测 $y_i(k)$，在假设 H_0 和 H_1 之间作出决策，同时利用通信网络从邻近节点 $j \in \mathcal{N}_i$ 接收一些局部决策变量 $l_j(k)$（每个智能体 i 如何计算其局部变量 $l_i(k)$，详见第四节）。我们假设局部测量 $y_i(k)$ 与广播同步进行。因此，在每个时间步 k，智能体 i 同时观测 $y_i(k)$、向邻居广播其局部决策变量 $l_i(k)$，并接收邻居变量 $l_j(k)$，其中 $j \in \mathcal{N}_i$。

B. 假设

我们采用以下三个假设。

假设 1。（噪声在空间上相互独立）对于每个时刻 k，加性噪声 $n_i(k)$ 在各智能体之间相互独立，这意味着向量 $n(k) = \{n_i(k)\}_{i=1}^N$ 具有相互独立的分量。不过要注意，每个分量 $(n_i(k))_{k\geq 0}$ 在时间上可以相关；

假设 2。（网络是连通的）通信网络 G 是一个连通图，也就是说，任意两个智能体之间都存在一条路径（可能包含多跳）。

这是一个常见（且必要）的假设，否则就无法将孤立智能体的观测纳入考虑。为便于后续引用，我们为 G 关联一个预先选定的对称权重矩阵 $W \in \mathbb{R}^{N\times N}$，该矩阵符合网络的稀疏性（$W_{i,j} = 0$ 当且仅当 $j \notin \mathcal{N}_i$ 时），且具有如下形式的特征值分解

$$W = \underbrace{\begin{pmatrix} \frac{1}{\sqrt{N}}\mathbf{1}_N & S \end{pmatrix}}_{Q_W} \underbrace{\begin{pmatrix} 1 & \\ & D \end{pmatrix}}_{\Lambda_W} \underbrace{\begin{pmatrix} \frac{1}{\sqrt{N}}\mathbf{1}_N^T \\ S^T \end{pmatrix}}_{Q_W^T}, \tag{8}$$

其中 $D \in \mathbb{R}^{(N-1)\times(N-1)}$ 是一个对角矩阵，其对角元素为 d_n，满足 $|d_n| < 1$，对于 $1 \leq n \leq N-1$，并且$S \in \mathbb{R}^{N\times(N-1)}$ 具有正交归一列（$S^TS = I$），每一列都与 $\mathbf{1}_N$正交，即 $S^T\mathbf{1}_N = 0$。

请注意，W因此既是行随机（$W\mathbf{1}_N = \mathbf{1}_N$）矩阵，也是列随机（$\mathbf{1}_N^T W = \mathbf{1}_N^T$）矩阵。众所周知，此类矩阵确实存在，甚至可以通过分布式方式获得，例如当 W 填入Metropolis权重时。

假设 3。（ARMA 模型）对于每个智能体 i，有理$\mathcal{Z}$-变换

$$\frac{A}{\sigma}\frac{1 - \sum_{j=1}^{q_i} a_i(j)\, z^{-j}}{1 + \sum_{j=1}^{p_i} b_i(j)\, z^{-j}}\;\frac{1 + \sum_{j=1}^{\bar{q}_i} \bar{b}_i(j)\, z^{-j}}{1 - \sum_{j=1}^{\bar{p}_i} \bar{a}_i(j)\, z^{-j}} \tag{9}$$

被假设为不存在零极点抵消（这不失一般性）。令 $Q_i = q_i + \overline{q}_i$ 和 $P_i = p_i + \overline{p}_i$ 分别表示其零点数和极点数，其中 P_i 个极点记为 $\{p_{i,1}, p_{i,2}, \ldots, p_{i,P_i}\}$。我们假设 $P_i > Q_i$（$\mathcal{Z}$-变换是严格真有理的），所有极点均为单极点（$p_{i,\,j} \neq p_{i,\,l}$对于 $j \neq l$），并且

$$1 \geq \rho_i = |p_{i,1}| > |p_{i,2}| \geq \cdots \geq |p_{i,P_i}|. \tag{10}$$

因此，每个智能体的 $\mathcal{Z}$-变换都是稳定的，其极点的最大绝对值记为 ρ_i，并且该最大值由唯一极点（记为$p_{i,\,1}$）取得。特别地，（9）的收敛域为 $\mathcal{ROC}_i = \{z \in \mathbb{C}\colon\ |z| > \rho_i\}$。

IV. 将运行共识检测器（RCD）扩展到 ARMA 设置

我们提出，智能体通过采用运行共识检测器（RCD）的扩展版本 [16, 32, 37]，协作解决全局决策问题。该扩展保留了标准 RCD 的通用模板，具体如下：在每个时刻 k，每个智能体 i 通过以仿射方式组合邻居 $j \in \mathcal{N}_i$ 传递的最新决策变量 $l_j(k-1)$ 与自身局部观测的历史，形成更新后的局部决策变量 $l_i(k) \in \mathbb{R}$

$$\boldsymbol{y}_i(k) = \begin{bmatrix} y_i(0) \\ y_i(1) \\ \vdots \\ y_i(k) \end{bmatrix} \in \mathbf{R}^{k+1}.$$

具体而言，将局部决策变量 $l_i(k)$、$1 \leq i \leq N$ 堆叠成全局决策向量 $l(k) \in \mathbb{R}^N$；我们建议该向量遵循以下动态方程

$$l(k) = W\,l(k-1) + \eta(k), \tag{11}$$

对于 $k \geq 0$（初始条件为 $l(-1) = 0$），其中 $\eta(k) \in \mathrm{R}^N$ 是一个创新向量，其每个分量 i 都以仿射方式依赖于智能体 i的局部观测历史，即

$$\eta_i(k) = \mathcal{A}_{i,k}\big(\boldsymbol{y}_i(k)\big). \tag{12}$$

这里，$\mathcal{A}_{i,k}$: $\mathrm{R}^{k+1} \mapsto \mathrm{R}$ 是一个仿射映射，将在下一节 V 中详细说明。动态方程 (11) 有一个简单的解释：右侧第一项在邻居之间融合局部决策变量，第二项则将每个智能体的历史观测与其最新观测进行融合。

每个智能体 i 形成更新后的局部决策变量 $l_i(k)$ 后，即可立即利用它对全局环境作出局部决策 $D_i(k) \in \{0,1\}$：

$$D_i(k) = \begin{cases} 1, & \text{if } l_i(k) \geq \gamma \\ 0, & \text{if } l_i(k) < \gamma, \end{cases} \tag{13}$$

其中，阈值 $\gamma \in \mathrm{R}$ 是一个设计参数；直观而言，降低 γ 会使决策偏向 H_1，反之亦然。我们将在阐明检测器 (13) 的长期性能后讨论 γ 的选择（参见定理 3 后的紧接讨论）。目前，γ 只是一个固定参数。

V. 高效实现扩展 RCD

本节首先提出一个用于 (12) 的仿射映射 $\mathcal{A}_{i,k}$；这一思路源于只有 $N = 1$ 个智能体的网络这一特殊情形。由于 $\mathcal{A}_{i,k}$ 作用于一个不断扩展的历史向量 $\boldsymbol{y}_i(k) \in \mathrm{R}^{k+1}$，朴素的实现方式需要每个智能体都配备同样不断扩展的内存，很快便会变得不可行。本节将通过给出一种可行的实现方式来绕过这一问题，该方式所需内存有限且不随时间 k变化。

A. 通过单智能体网络引出 $\mathcal{A}_{i,k}$

考虑一个包含 $N = 1$ 个智能体的网络。在时间 k，唯一的智能体 $i = 1$ 拥有观测历史 $\boldsymbol{y}_i(k)$，并且必须在以下选项中作出决定：

$$\begin{aligned} H_1 &: \ \boldsymbol{y}_i(k) = \boldsymbol{\theta}_i(k) + \boldsymbol{n}_i(k) \\ H_0 &: \ \boldsymbol{y}_i(k) = \boldsymbol{n}_i(k), \end{aligned}$$

其中

$$\boldsymbol{\theta}_i(k) = \begin{bmatrix} \theta_i(0) \\ \theta_i(1) \\ \vdots \\ \theta_i(k) \end{bmatrix}, \qquad \boldsymbol{n}_i(k) = \begin{bmatrix} n_i(0) \\ n_i(1) \\ \vdots \\ n_i(k) \end{bmatrix}. \tag{14}$$

注意，噪声向量 $\boldsymbol{n}_i(k)$服从高斯分布，即$\boldsymbol{n}_i(k) \sim \mathcal{N}(0, \boldsymbol{\Sigma}_i(k))$，其中 $\boldsymbol{\Sigma}_i(k) \in \mathrm{R}^{(k+1)\times(k+1)}$ 表示其协方差矩阵。统计学中的标准结果 [38], [39]表明，在贝叶斯风险准则下，最优检测器是对数似然函数的阈值判决，即

$$D_i(k) = \begin{cases} 1, & \text{if } l_i(k) \geq \gamma \\ 0, & \text{if } l_i(k) < \gamma, \end{cases}$$

其中

$$l_i(k) = \boldsymbol{\theta}_i(k)^T \boldsymbol{\Sigma}_i(k)^{-1} \boldsymbol{y}_i(k) - \boldsymbol{\theta}_i(k)^T \boldsymbol{\Sigma}_i(k)^{-1} \boldsymbol{\theta}_i(k)/2, \tag{15}$$

暂且假设矩阵 $\boldsymbol{\Sigma}_i(k)$可逆（定理 1 将证明这一假设确实成立）。现在有一个关键观察：这些决策变量的序列 $(l_i(k))_{k\geq 0}$ 可以递归地按如下方式得到：令

$$\boldsymbol{\Sigma}_i(k) = \boldsymbol{L}_i(k)\, \boldsymbol{L}_i(k)^T \tag{16}$$

为乔列斯基分解 [40]，，其中 $\boldsymbol{L}_i(k) \in \mathrm{R}^{(k+1)\times(k+1)}$ 是一个对角元素为正的下三角矩阵；将该分解代入式 (15)，得到

$$l_i(k) = \widehat{\boldsymbol{\theta}}_i(k)^T \widehat{\boldsymbol{y}}_i(k) - \widehat{\boldsymbol{\theta}}_i(k)^T \widehat{\boldsymbol{\theta}}_i(k)/2, \tag{17}$$

其中

$$\widehat{\boldsymbol{\theta}}_i(k) = \boldsymbol{L}_i(k)^{-1} \boldsymbol{\theta}_i(k) \tag{18}$$

and

$$\widehat{\boldsymbol{y}}_i(k) = \boldsymbol{L}_i(k)^{-1} \boldsymbol{y}_i(k). \tag{19}$$

由于 $\boldsymbol{\Sigma}_i(k-1)$ 嵌套于 $\boldsymbol{\Sigma}_i(k)$ 中，形式如下：

$$\boldsymbol{\Sigma}_i(k) = \begin{bmatrix} \boldsymbol{\Sigma}_i(k-1) & \star \\ \star & \star \end{bmatrix},$$

Cholesky 因子 $\boldsymbol{L}_i(k-1)$ 也嵌套于 $\boldsymbol{L}_i(k)$ 中，形式如下：

$$\boldsymbol{L}_i(k) = \begin{bmatrix} \boldsymbol{L}_i(k-1) & 0 \\ \star & \star \end{bmatrix}.$$

因此，

$$\widehat{\boldsymbol{\theta}}_i(k) = \begin{bmatrix} \boldsymbol{L}_i(k-1)^{-1} \boldsymbol{\theta}_i(k-1) \\ \widehat{\theta}_i(k) \end{bmatrix} \tag{20}$$

and

$$\widehat{\boldsymbol{y}}_i(k) = \begin{bmatrix} \boldsymbol{L}_i(k-1)^{-1} \boldsymbol{y}_i(k-1) \\ \widehat{y}_i(k) \end{bmatrix}; \tag{21}$$

现在将 (20) 和 (21) 代入 (17)，得到递归形式

$$l_i(k) = l_i(k-1) + \widehat{\theta}_i(k)\, \widehat{y}_i(k) - \widehat{\theta}_i(k)^2/2. \tag{22}$$

比较更新式 (22) 和 (11)，可知应定义 $\mathcal{A}_{i,\,k}$ 为

$$\mathcal{A}_{i,k}\big(\boldsymbol{y}_i(k)\big) = \widehat{\theta}_i(k)\ \widehat{y}_i(k) - \frac{1}{2}\widehat{\theta}_i(k)^2, \tag{23}$$

目前先考虑一般的 $N \geq 1$和智能体 $i = 1 \ldots$、N。此定义确保对于 $N = 1$，RDC 能够精确地恢复对数似然方案，因为 (11) 中的权重矩阵 W ^在 $N = 1$时会化简为 $W = 1$。在 (23) 中，标量 $\theta_i(k)$ 和~$\widehat{y}_i(k)$ 表示 (18) 和 (19) 中定义的向量 $\boldsymbol{\theta}_i(k)$ 和$\widehat{\boldsymbol{y}}_i(k)$ 的最后一个分量（每个向量的大小为 $k+1$）。

总之，定义

$$\widehat{y}(k) = \begin{bmatrix} \widehat{y}_1(k) \\ \vdots \\ \widehat{y}_N(k) \end{bmatrix}, \quad \widehat{\theta}(k) = \begin{bmatrix} \widehat{\theta}_1(k) \\ \vdots \\ \widehat{\theta}_N(k) \end{bmatrix}, \tag{24}$$

^ ^以及 $N \times N$ 对角矩阵 $\Theta(k) = \mathrm{diag}(\theta(k))$，我们有更新（11）由以下部分组成

$$l(k) = W l(k-1) + \widehat{\Theta}(k)\, \widehat{y}(k) - \frac{1}{2}\widehat{\Theta}(k)^2 \mathbf{1}. \tag{25}$$

本节接下来将介绍一种递归计算^compute $\theta_i(k)$ 和 $\widehat{y}_i(k)$ 的高效方法。第一步是递归求出（16）的 Cholesky因子 $\boldsymbol{L}_i(k)$，并证明这些矩阵可逆，因为它们具有正对角线。由（16）可知， $\boldsymbol{\Sigma}_i(k)$ 也可逆。

B. 递归求得 $\boldsymbol{\Sigma}_i(k)$的乔列斯基分解

利用 ARMA 等信号结构来设计递归算法，一直是统计信号处理中的一个关键思想，例如基于创新过程 [41]的滤波。在 [41] 第 4.4 节中，Kailath 等人给出了指数相关过程的 L-D-L 分解和乔列斯基分解的高效实现。正如下一个定理所示，这里也可以采用类似方法递归求得式 (16) 中的 Cholesky 因子 $\boldsymbol{L}_i(k)$。

定理1。（Cholesky 因子的递推关系）式 (16) 中的 Cholesky 因子 $\boldsymbol{L}_i(k)$为

$$\boldsymbol{L}_i(k) = \sigma \left(I_{k+1} - \boldsymbol{A}_i(k)\right)^{-1} \boldsymbol{B}_i(k),$$

其中，严格下三角矩阵 $\boldsymbol{A}_i(k)$和下三角矩阵 $\boldsymbol{B}_i(k)$满足以下递推关系

$$\begin{aligned}\boldsymbol{A}_i(k+1) &= \begin{pmatrix} \boldsymbol{A}_i(k) & 0 \\ \boldsymbol{\alpha}_i(k+1)^T & 0 \end{pmatrix}, \\ \boldsymbol{B}_i(k+1) &= \begin{pmatrix} \boldsymbol{B}_i(k) & 0 \\ \boldsymbol{\beta}_i(k+1)^T & 1 \end{pmatrix}. \end{aligned} \tag{26}$$

式 (26) 中的向量 $\boldsymbol{\alpha}_i(k)$ 和 $\boldsymbol{\beta}_i(k)$以 $\boldsymbol{\alpha}_i(0) = 0$ 和$\boldsymbol{\beta}_i(0) = 1$ 为初值，并由下式给出

$$\boldsymbol{\alpha}_i(k+1)^T = \begin{cases} \begin{pmatrix} a_i(k+1) \ \dots \ a_i(1) \end{pmatrix} & , k+1 \le p_i \\ \begin{pmatrix} 0_{k+1-p_i}^T \ a_i(p_i) \ \dots \ a_i(1) \end{pmatrix} & , k+1 > p_i, \end{cases}$$

$$\boldsymbol{\beta}_i(k+1)^T = \begin{cases} \begin{pmatrix} b_i(k+1) \ \dots \ b_i(1) \end{pmatrix} & , k+1 \le q_i \\ \begin{pmatrix} 0_{k+1-q_i}^T \ b_i(q_i) \ \dots \ b_i(1) \end{pmatrix} & , k+1 > q_i. \end{cases} \tag{27}$$

证明。由（6）可知，（14）中的向量 $\boldsymbol{n}_i(k)$满足

$$(I_{k+1} - \boldsymbol{A}_i(k))\boldsymbol{n}_i(k) = \sigma\, \boldsymbol{B}_i(k)\, \boldsymbol{\epsilon}_i(k),$$

其中，矩阵 $\boldsymbol{A}_i(k)$和 $\boldsymbol{B}_i(k)$如上所定义，$\boldsymbol{\epsilon}_i(k)$表示一个零均值白高斯向量： $\boldsymbol{\epsilon}_i(k) \sim \mathcal{N}(0, I_{k+1})$。由于 $I_{k+1} - \boldsymbol{A}_i(k)$是一个主对角线元素均为 1 的下三角矩阵，因此它的逆矩阵存在（且也是主对角线元素均为 1 的下三角矩阵）。由此可得

$$\boldsymbol{n}_i(k) = \sigma \left(I_{k+1} - \boldsymbol{A}_i(k)\right)^{-1} \boldsymbol{B}_i(k)\, \boldsymbol{\epsilon}_i(k).$$

因此， $\boldsymbol{n}_i(k)$ 的协方差为 $\boldsymbol{\Sigma}_i(k) = \boldsymbol{L}_i(k)\, \boldsymbol{L}_i(k)^T$，其中 $\boldsymbol{L}_i(k) = \sigma \left(I_{k+1} - \boldsymbol{A}_i(k)\right)^{-1} \boldsymbol{B}_i(k)$，如所述。最后请注意， $\boldsymbol{L}_i(k)$ 是两个对角线元素为正的下三角矩阵的乘积，因此自身也是这样的矩阵；所以，所提及的 $\boldsymbol{L}_i(k)$ 确实是 $\boldsymbol{\Sigma}_i(k)$ 的 Cholesky 因子。 □

C. 高效实现映射 $\mathcal{A}_{i,k}$

在实现式 (23) 中的映射 $\mathcal{A}_{i,\ k}$ 时，主要的计算步骤是 (18) 和 (19)，二者都由以下通用运算构成

$$\boldsymbol{z}_i(k) = \begin{bmatrix} z_i(0) \\ z_i(1) \\ \vdots \\ z_i(k) \end{bmatrix} \quad \mapsto \quad \widehat{\boldsymbol{z}}_i(k) = \begin{bmatrix} \widehat{z}_i(0) \\ \widehat{z}_i(1) \\ \vdots \\ \widehat{z}_i(k) \end{bmatrix},$$

其中 $\widehat{\boldsymbol{z}}_i(k) = \boldsymbol{L}_i(k)^{-1} \boldsymbol{z}_i(k)$。如果我们在每个时刻 k都显式存储并求逆不断增长的矩阵 $\boldsymbol{L}_i(k) \in \mathrm{R}^{(k+1)\times(k+1)}$，很快就会遇到内存和计算方面的限制。下一个定理表明，可以通过 ARMA 滤波器从 $\boldsymbol{z}_i(k)$高效地求得 $\widehat{\boldsymbol{z}}_i(k)$。

定理 2。（ARMA 滤波器解释）向量 $\widehat{\boldsymbol{z}}_i(k)$ 的各分量，定义为 $\boldsymbol{L}_i(k)^{-1} \boldsymbol{z}_i(k)$，是 ARMA($p_i, q_i$) 滤波器的前 $k+1$ 个输出样本，该滤波器的输入为 $\boldsymbol{z}_i(k)$，即

$$\widehat{z}_i(k) = -\sum_{j=1}^{q_i} b_i(j)\, \widehat{z}_i(k-j) + \frac{1}{\sigma} z_i(k) - \sum_{j=1}^{p_i} \frac{a_i(j)}{\sigma}\, z_i(k-j), \tag{28}$$

在零初始条件下 $\widehat{z}_i(k) = z_i(k) = 0$，其中 $k < 0$。

证明。由

$$\underbrace{\sigma \left(I_{k+1} - \boldsymbol{A}_i(k)\right)^{-1} \boldsymbol{B}_i(k)}_{\boldsymbol{L}_i(k)} \widehat{\boldsymbol{z}}_i(k) = \boldsymbol{z}_i(k)$$

可得ws

$$\boldsymbol{B}_i(k)\, \widehat{\boldsymbol{z}}_i(k) = \frac{1}{\sigma} \left(I_{k+1} - \boldsymbol{A}_i(k)\right) \boldsymbol{z}_i(k). \tag{29}$$

将 (26) 和 (27) 代入 (29)，即可得到 (28)。 □

(28) 中的滤波器实际上就是噪声滤波器 (6) 的逆滤波器；因此，将其应用于 $\boldsymbol{z}_i(k) = \boldsymbol{n}_i(k)$ 时，输出为零均值白高斯信号 $\widehat{\boldsymbol{z}}_i(k) = \boldsymbol{\epsilon}_i(k)$。所以，滤波器 (28) 本质上对 $\boldsymbol{n}_i(k)$进行了白化。

总之，智能体 i 在每个时间^步 k通过获取所需的 $\theta_i(k)$ 和 $\widehat{y}_i(k)$ 来实现映射 $\mathcal{A}_{i,k}$，它们分别是其两个滤波器（28）本地副本的输出：一个副本接收感兴趣信号 $\theta_i(k)$；另一个副本接收局部观测流 $y_i(k)$。

^为计算滤波信号 $\widehat{y}_i(k)$ $\theta_i(k)$，智能体 i 存储^滤波信号 $\widehat{\boldsymbol{y}}_i(k)$、 $\boldsymbol{\theta}_i(k)$ 的过去 q_i 个样本，以及未滤波信号 $\boldsymbol{y}_i(k)$、 $\boldsymbol{\theta}_i(k)$ 的过去 p_i 个样本，也就是说，

$$\{\widehat{y}_i(k), \dots, \widehat{y}_i(k-q_i)\},\ \{\widehat{\theta}_i(k), \dots, \widehat{\theta}_i(k-q_i)\}$$
$$\{y_i(k), \dots, y_i(k-p_i)\},\ \{\theta_i(k), \dots, \theta_i(k-p_i)\},$$

在所有信号均满足零初始条件的情况下，即 $\widehat{y}_i(k)$ ^ $= y_i(k) = \theta_i(k) = \theta_i(k) = 0$ 对任意 $k < 0$ 均成立。因此，智能体 i 的内存需求随其在通信网络 G 中邻居集合的大小呈 $\mathcal{O}(p_i + q_i + |\mathcal{N}(i)|)$ 增长|$\mathcal{N}(i)$| （智能体 i 还需要存储矩阵 W 的第 i 行中的非零项）。利用 (28) 计算 $\theta_i(k)$ 和 $\widehat{y}_i(k)$，很容易得出，实现 (13) 也需要 $\mathcal{O}(p_i + q_i + |\mathcal{N}(i)|)$ 次运算（求和或乘积）。

VI. RCD 的理论分析

本节分析局部 RCD 检测器 (13) 的性能。具体而言，对于智能体 i 的局部检测器，我们研究两种错误概率的长期行为（$k \to \infty$），即虚警和漏检的概率，

$$\begin{aligned}\mathbb{P}_{F,i}(k) &:= \mathbb{P}_{H_0}(D_i(k) = 1),\\ \mathbb{P}_{M,i}(k) &:= \mathbb{P}_{H_1}(D_i(k) = 0).\end{aligned} \tag{30}$$

这些概率的渐近行为将用时间尺度序列 $(f(k))_{k\geq 0}$ 表示，该序列取决于状态 (a) 或情形（b）中哪一种被激活：在状态 (a) 中，$f(k)$ 是线性增长的序列，且 $f_a(k) = k+1$；而在情形（b）中，它是常数序列，且 $f_b(k) = 1$。下文中，我们分别用 $\mathcal{F} = \{(f_a(k))_k \text{ 和 } (f_b(k))_k\}$ 表示前面提到的两种时间缩放。

哪种 情形被激活取决于增长率 e of

$$\left(\left\|\widehat{\boldsymbol{\theta}}(k)\right\|^2\right), \tag{31}$$

其中

$$\widehat{\boldsymbol{\theta}}(k) = \begin{bmatrix}\widehat{\boldsymbol{\theta}}_1(k)\\ \vdots \\ \widehat{\boldsymbol{\theta}}_N(k)\end{bmatrix} \in \mathbf{R}^{N(k+1)}, \tag{32}$$

$\widehat{\ }$与 $\boldsymbol{\theta}_i(k)$ 在式 (18) 中定义。注意

$$\left\|\widehat{\boldsymbol{\theta}}(k)\right\|^2 = \sum_{m=0}^{k}\left\|\widehat{\theta}(m)\right\|^2,$$

其中 $\widehat{\theta}(k)$ 如式 (24) 所示；因此，$\left\|\widehat{\boldsymbol{\theta}}(k)\right\|^2$ 可解释为从 $m=0$ 到 $m=k$ 的白化网络信号 $(\boldsymbol{\theta}(m))$ 的累计能量$\widehat{\ }$。

我们将错误概率的分析分为两个步骤，概述如下:

1) 暂且假设序列 (31) 按如下方式增长

$$\left\|\widehat{\boldsymbol{\theta}}(k)\right\|^2 = \alpha f(k) + o\left(f(k)\right), \tag{33}$$

其中 α 是一个正数，序列 $(f(k))$ 是 $\mathcal{F}$ 中的一种时间尺度。定理 3 表明

$$-\frac{\log \mathbb{P}_{F,i}(k)}{f(k)} \to \beta_F \tag{34}$$

$$-\frac{\log \mathbb{P}_{M,i}(k)}{f(k)} \to \beta_M, \tag{35}$$

当 $k \to \infty$ 时，对于某些显式常数 $\beta_F, \beta_M > 0$。因此，在状态 (a) 中，概率$\mathbb{P}_{F,i}(k)$ 和 $\mathbb{P}_{M,i}(k)$ 会像标准大偏差结果中那样指数级快速衰减至零 [42], ，其近似表现为 $\mathbb{P}_{F,i}(k) \simeq \exp(-\beta_F(k+1))$ 和 $\mathbb{P}_{M,i}(k) \simeq \exp(-\beta_M(k+1))$。概率的这种指数级消失可归因于 $\boldsymbol{\theta}(k)$ 的能量线性增长；这是一种普遍效应，在集中式设置中早已得到注意 [43], [44], [45]。与之形成鲜明对比的是，在新颖的情形（b）中，错误概率会趋于一个正常的误差底限，因此当 k 较大时，$\mathbb{P}_{F,i}(k) \simeq \exp(-\beta_F)$ 且 $\mathbb{P}_{M,i}(k) \simeq \exp(-\beta_M)$;

2）在定理 4 中，我们通过明确给出正常数 α 和时间缩放（$f(k)$）$\in \mathcal{F}$，证明假设（33）确实成立；这些量均由底层 ARMA 滤波器决定。

现在我们执行分析的两个步骤。

A. 步骤 1：$\mathbb{P}_{F,\ i}(k)$ 和 $\mathbb{P}_{M,\ i}(k)$ 的渐近性能

关键结果是接下来要介绍的定理 3。

定理 3。（局部检测器 (13) 的渐近性能）若第 III-B 节中的假设 1 和假设 2 成立。

若增长率条件 (33) 也成立，则智能体 i 处局部检测器 (13) 的虚警概率 $\mathbb{P}_{F,i}(k)$ 和漏检概率$\mathbb{P}_{M,i}(k)$ 满足 (34) 和 (35)，其中

$$\beta_F = \begin{cases}\dfrac{\alpha}{8} & ,\ if\ f = f_a\\ -\log\ \mathcal{Q}\left(\dfrac{\sqrt{\alpha}}{2} + \dfrac{\gamma N}{\sqrt{\alpha}}\right) & ,\ if\ f = f_b,\end{cases} \tag{36}$$

$$\beta_M = \begin{cases}\dfrac{\alpha}{8} & ,\ if\ f = f_a\\ -\log\ \mathcal{Q}\left(\dfrac{\sqrt{\alpha}}{2} - \dfrac{\gamma N}{\sqrt{\alpha}}\right) & ,\ if\ f = f_b,\end{cases} \tag{37}$$

其中，$\mathcal{Q}$ 函数用于计算标准高斯分布的尾概率，$\mathcal{Q}(t) = \int_t^{+\infty} \exp(-\frac{x^2}{2})dx/\sqrt{2\pi}$。

在转向定理证明之前，值得注意的是，正如 (36) 和 (37) 所示，智能体 i 的错误概率 $\mathbb{P}_{F,i}(k)$ 和 $\mathbb{P}_{M,i}(k)$ 的渐近性能并不取决于具体的智能体 i。因此，无论时间相关性如何，所有采用 RDC 的智能体都将具有相同的渐近性能。

还需注意，假设 3 为检测信号和测量噪声都指定了 ARMA 结构，但此处并未用到该假设。这意味着定理 3 实际上比所需的更强，因为它涵盖的时间相关信号与噪声远不止 ARMA 类型；这一结果或许本身也值得关注。

证明。我们只证明漏检概率 (35) 的渐近行为（证明结果 (34) 的方法类似）。因此，从现在起，我们假设测量值 $y(k)$ 是根据假设 H_1 生成的，见 (5)。

由于 (30) 中的漏检概率取决于局部决策 $D_i(k)$ 的分布，而后者又取决于 $l_i(k)$ 的分布（回顾 (13)），我们首先仔细考察按 (25) 演化的向量 $l(k)$ 的分布。

由于 $\hat{y}(k)$ 是一个高斯随机向量，由 (25) 可知，$l(k)$ 同样是一个高斯随机向量:

$$l(k) \sim \mathcal{N}(\mu(k), \Omega(k)),$$

符号 $\mu(k)$ 和 $\Omega(k)$ 分别表示 $l(k)$ 的均值和协方差。事实上，由 (25) 可推导出均值和协方差的递推关系，使用以下 H_1 假设:

$$\mu(k+1) = W\mu(k) + \frac{1}{2}\widehat{\Theta}(k)^2\mathbf{1} \tag{38}$$

$$\Omega(k+1) = W\Omega(k)W^T + \widehat{\Theta}(k)^2, \tag{39}$$

其中，为了得到 (38)，我们使用了 $\mathbb{E}(\widehat{y}(k)) = \Theta(k)$111（因为假设了 H_1）；而对于 (39)，我们利用了这样一个事实：$(\widehat{y}(k))$ 是一个（白化的）不相关随机向量序列，其协方差矩阵为单位矩阵（参见 (16) 和 (19)）。

下一步，我们将利用递推关系 (38) 和 (39)，确定 $\mu(k)$ 和 $\Omega(k)$ 的渐近行为。为此，下面的引理至关重要。

引理 1. 假设 n 维向量序列 $(v(k))_{k\geq 0}$ 满足递推关系

$$v(k+1) = \mathcal{W}v(k) + u(k), \tag{40}$$

对于 $k \geq 0$（其中 $v(-1) = 0$），$\mathcal{W}$ 是一个对称矩阵，其特征值分解为

$$\mathcal{W} = \begin{bmatrix} \frac{1}{\sqrt{n}}\mathbf{1}_n & \mathcal{S} \end{bmatrix} \begin{bmatrix} 1 & \\ & \mathcal{D} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{n}}\mathbf{1}_n^T \\ \mathcal{S}^T \end{bmatrix}, \tag{41}$$

其中 $\mathcal{D} \in \mathbb{R}^{(n-1)\times(n-1)}$ 是一个对角矩阵，其对角元素满足 $|d_j| < 1$，对于 $1 \leq j \leq n-1$。

若 $(u(k))$ 是一个非负向量序列 $(u_j(k) \geq 0$，对于 $1 \leq j \leq n)$，且其增长满足

$$\sum_{m=0}^{k} \mathbf{1}_n^T u(m) = \theta f(k) + o(f(k)), \tag{42}$$

其中 $\theta > 0$，且 $(f(k))$ 是 $\mathcal{F}$ 中的一个时间尺度，则

$$\frac{v(k)}{f(k)} \to \frac{\theta}{n}\mathbf{1}_n, \tag{43}$$

如 $k \to \infty$。

引理 1 在附录中证明。现在我们将其使用两次：

- 首先，注意到将 $v(k) = N(k)$、$\widehat{\Theta}(k)$、$u(k) = \mathcal{W} = W$ 和 $(\Omega(k))^2\mathbf{1}/2$ 分别对应起来后，(38) 符合模板 (40)；事实上，特征值分解 () 验证了 (41)，而 (33) 在 $\theta = \alpha/2$ 的条件下验证了 (42)。因此，

$$\frac{\mu(k)}{f(k)} \to \frac{\alpha}{2N}\mathbf{1}_N. \tag{44}$$

- 其次，注意到 (39) 也符合模板 (40)，只需改用 (39) 的向量化形式：

$$\operatorname{vec}(\Omega(k+1)) = (W \otimes W)\operatorname{vec}(\Omega(k)) + \operatorname{vec}\left(\widehat{\Theta}^2(k)\right), \tag{45}$$

其中 $\otimes$ 表示克罗内克积，并且利用了通用性质 $\operatorname{vec}(AXB) = (B^T \otimes A)\operatorname{vec}(X)$，从而由 (39) 得到 (45)。递推关系 (45) 显然符合模板 (40)。由于 (8) 意味着 $\mathcal{W} = W \otimes W$ 具有如下特征值分解，因此条件 (41) 得到验证

$$\mathcal{W} = \underbrace{(Q_W \otimes Q_W)}_{Q_{\mathcal{W}}} \underbrace{(\Lambda_W \otimes \Lambda_W)}_{\Lambda_{\mathcal{W}}} \underbrace{(Q_W \otimes Q_W)^T}_{Q_{\mathcal{W}}^T},$$

满足式 (41) 中特征值和特征向量要求的分解。此外，根据式 (33)，条件 (42) 由 $\theta = \alpha$ 得到验证。因此，

$$\frac{\operatorname{vec}(\Omega(k))}{f(k)} \to \frac{\alpha}{N^2}\mathbf{1}_{N^2}. \tag{46}$$

现在我们可以推导式 (35)，即.....的渐近行为

$$-\frac{\log \mathbb{P}_{M,i}(k)}{f(k)}.$$

鉴于 $\mathbb{P}_{M,\,i}(k) = \mathbb{P}_{H_1}(l_i(k) < \gamma)$ 且 $l_i(k) \sim \mathcal{N}(\mu_i(k), \Omega_{ii}(k))$，其中 $\mu_i(k)$ 是向量 $\mu(k)$ 的第 i 个分量，而 $\Omega_{ii}(k)$ 是矩阵 $\Omega(k)$ 的第 i 个对角元素，我们有

$$-\frac{\log \mathbb{P}_{M,i}(k)}{f(k)} = -\frac{\log \mathcal{Q}\left(\frac{-\gamma+\mu_i(k)}{\sqrt{\Omega_{ii}(k)}}\right)}{f(k)}. \tag{47}$$

我们研究 $k \to \infty$ 时式 (47) 在以下两种可能情形下的极限：

- 在状态 (a) 下，我们有 $f(k) = k+1$，这意味着其变为 (47)

$$-\frac{\log \mathbb{P}_{M,i}(k)}{f(k)} = -\frac{\log \mathcal{Q}\left(\frac{-\gamma+\mu_i(k)}{\sqrt{\Omega_{ii}(k)}}\right)}{k+1}. \tag{48}$$

记为

$$\frac{-\gamma+\mu_i(k)}{\sqrt{\Omega_{ii}(k)}} = \underbrace{\frac{-\gamma/(k+1)+\mu_i(k)/(k+1)}{\sqrt{\Omega_{ii}(k)/(k+1)}}}_{u(k)} \underbrace{\sqrt{k+1}}_{\sqrt{v(k)}},$$

我们有 $u(k) \to \sqrt{\alpha}/2$（利用式 (44) 和 (46)），且 $v(k) \to \infty$。因此，我们可以应用以下辅助引理2，其证明见附录。

引理 2. 假设 $(u(k))$ 和 $(v(k))$ 是正项序列。若 $u(k) \to \theta > 0$ 且 $v(k) \to \infty$，则

$$-\frac{\log \mathcal{Q}\left(u(k)\sqrt{v(k)}\right)}{v(k)} \to \frac{\theta^2}{2}. \tag{49}$$

在 (65) 中应用引理 2，可得 (35)，其中 $\beta_M = \alpha/8$；

- 在情形（b）下，我们有 $f(k) = 1$。因此，根据 (44) 和 (46)，再结合函数 $\mathcal{Q}$ 的连续性，可得极限 (35)，其中 $\beta_M = -\log \mathcal{Q}(\sqrt{\alpha}/2 - \gamma N/\sqrt{\alpha})$。

□

关于阈值 γ。如 (36) 和 (37) 所示，局部检测器 (13) 的阈值 γ 不会影响状态 (a) 下错误概率的渐近行为。这与分布式检测中的已有结果一致 [16]，其中 γ 通常设为零点。相比之下，在情形（b）下，阈值 γ 会产生显著影响。它成为一个设计参数，可用于权衡渐近虚警概率 $\mathbb{P}_F(\gamma) := \lim_k \mathbb{P}_{F,\,i}(k)$ 与渐近漏检概率 $\mathbb{P}_M(\gamma) := \lim_k \mathbb{P}_{M,\,i}(k)$。事实上，较大的 $\gamma \geq 0$ 会使检测器倾向于选择假设 H_0，从而降低虚警概率并提高漏检概率：

$$\lim_{\gamma\to+\infty} \mathbb{P}_{F,i}(\gamma) = 0, \quad \lim_{\gamma\to+\infty} \mathbb{P}_{M,i}(\gamma) = 1.$$

当 $\gamma < 0$ 时，同样可以采用对称论证。

定理 3 的一个特例。在 $N = 1$智能体检测白高斯噪声 $n(k) \sim \mathcal{N}(0$中的常信号 $\theta(k) = A$ 这一特例中，由于可以验证式 (33) 在 $\alpha = A^2/\sigma^2$和 $f(k) = k + 1$ 的条件下成立，因此会启用状态 (a)。因此，定理 3 断言错误概率以共同指数 $\beta_F = \beta_M = \beta$呈指数衰减，其中 $\beta = A^2/(8\sigma^2)$。实际上，这样的 β 是分布 H_1 和 H_0 [16]下测量值 $y(k)$之间的切尔诺夫距离。

B. 步骤 2：ARMA 模型的能量 $||\widehat{\boldsymbol{\theta}}(k)||^2$ 增长

本节推导式 (32) 中定义的 $\boldsymbol{\theta}(k)$ 的能量增长率。具体而言，我们通过从底层 ARMA 模型中求得 α 和时间尺度 ($f(k)$)，证明式 (33) 成立。

我们首先注意到，式 (32) 意味着能量在各智能体之间相互独立，即

$$\left\|\widehat{\boldsymbol{\theta}}(k)\right\|^2 = \sum_{i=1}^{N} \left\|\widehat{\boldsymbol{\theta}}_i(k)\right\|^2, \tag{50}$$

这表明应考察每个智能体处白化信号的能量增长率。因此，以下我们聚焦于一个一般的智能体 i，并研究

$$\left\|\widehat{\boldsymbol{\theta}}_i(k)\right\|^2 = \sum_{m=0}^{k} \widehat{\theta}_i(m)^2.$$

回顾一下，$\theta_i(k)$ 是输入为 $\theta_i(k)$ 的 ARMA 滤波器 (28) 的输出，而 $\theta_i(k)$ 是 ARMA 滤波器 (7) 的脉冲响应。因此，$\theta_i(k)$ 的 $\mathcal{Z}$ 变换为

$$\widehat{\Theta}_i(z^{-1}) = \frac{A}{\sigma} \frac{1 - \sum_{j=1}^{q_i} a_i(j)\, z^{-j}}{1 + \sum_{j=1}^{p_i} b_i(j)\, z^{-j}} \frac{1 + \sum_{j=1}^{\bar{q}_i} \bar{b}_i(j)\, z^{-j}}{1 - \sum_{j=1}^{\bar{p}_i} \bar{a}_i(j)\, z^{-j}}. \tag{51}$$

这正是式 (9) 中的表达式，适用第 III-B 节中的假设 3。因此，根据部分分式展开结果 (2)，我们有

$$\widehat{\theta}_i(k) = \sum_{j=1}^{P_i} r_{i,j}\, p_{i,j}^k, \tag{52}$$

其中 $r_{i,\,j}$ 表示与极点 i 相关联的留数，即

$$r_{i,j} = (1 - p_{i,j}\, z^{-1})\, \widehat{\Theta}_i(z^{-1})\Big|_{z=p_{i,j}}.$$

现在我们来计算 $\boldsymbol{\theta}_i(k)$ 的能量，这主要取决于 ρ_i，即极点绝对值的最大值（回顾公式 (10)）。

定理 4。（信号 $\boldsymbol{\theta}_i(k)$的渐近能量）在假设 3 下，我们有

$$\left\|\widehat{\boldsymbol{\theta}}_i(k)\right\|^2 = \begin{cases} |r_{i,1}|^2(k+1) + o(k+1) & ,\ \textit{if } \rho_i = 1 \\ \sum_{j,l=1}^{P_i} \dfrac{r_{i,j} r_{i,l}^\star}{1 - p_{i,j} p_{i,l}^\star} + o(1) & ,\ \textit{if } \rho_i < 1, \end{cases} \tag{53}$$

其中 $z^\star$ 表示复数 z 的复共轭。

证明。首先考虑 $\rho_i = 1$ 的情形，我们有

$$1 = |p_{i,1}| > r \geq |p_{i,2}| \geq \cdots \geq |p_{i,P_i}|, \tag{54}$$

对于某个 $1 > r \geq 0$。引入记号

$$\Delta(k) = \sum_{j=2}^{P_i} r_{i,j} p_{i,j}^k$$

我们将 (52) 重写为 $\widehat{\theta}_i(k) = r_{i,\,1} p_{i,\,1}^k + \Delta(k)$。由此得到

$$\begin{aligned} \left\|\widehat{\boldsymbol{\theta}}_i(k)\right\|^2 = \sum_{m=0}^{k} |\widehat{\theta}_i(m)|^2 &= \sum_{m=0}^{k} \left| r_{i,1} p_{i,1}^m + \Delta(m) \right|^2 \\ &= |r_{i,1}|^2 (k+1) + u(k), \end{aligned} \tag{55}$$

其中

$$u(k) = \sum_{m=0}^{k} r_{i,1} p_{i,1}^m \Delta(m)^\star + r_{i,1}^\star (p_{i,1}^\star)^m \Delta(m) + |\Delta(m)|^2.$$

现在，注意到 $u(k)$具有如下形式

$$u(k) = \sum_{l=1}^{L} \left(\sum_{m=0}^{k} v_l w_l^m \right),$$

（对于某些 L 和常数 v_l），其中每个 w_l 都是两个极点的乘积（共轭情形除外），且这两个极点不全等于 $p_{i,\,1}$；具体而言，$w_l \in \{p_{i,j_1}\, p_{i,j_2},\, p_{i,j_1}^*\, p_{i,j_2},\, p_{i,j_1}\, p_{i,j_2}^*,\, p_{i,j_1}^*\, p_{i,j_1}^*\}$对于某个 $(j_1, j_2) \neq (1, 1)$。因此，由 (54) 可得 $|w_l| \leq r < 1$，这意味着

$$|u(k)| \leq \sum_{l=1}^{L} \left(\sum_{m=0}^{k} |v_l| r^m \right) \leq \sum_{l=1}^{L} |v_l| \frac{1}{1-r}.$$

因此，序列 ($u(k)$) 有界，进而满足 $u(k) = o(k+1)$。这一观察结合 (55)，得到了 (53) 中的第一种情况。

现在我们转向 $\rho_i < 1$ 的情况。我们有

$$\left\|\widehat{\boldsymbol{\theta}}_i(k)\right\|^2 = \sum_{m=0}^{k} |\widehat{\theta}_i(m)|^2 = \sum_{j,l=1}^{P_i} \left(\sum_{m=0}^{k} r_{i,j} r_{i,l}^\star (p_{i,j} p_{i,l}^\star)^m \right).$$

由于 $|p_{i,\,j} p_{i,\,l}^\star| \leq \rho_i^2 < 1$，因此可得

$$\sum_{m=0}^{k} r_{i,j} r_{i,l}^\star (p_{i,j} p_{i,l}^\star)^m \to \frac{r_{i,j} r_{i,l}^\star}{1 - p_{i,j} p_{i,l}^\star}.$$

因此，

$$\sum_{m=0}^{k} |\widehat{\theta}_i(m)|^2 \to \sum_{j,l=1}^{P_i} \frac{r_{i,j} r_{i,l}^\star}{1 - p_{i,j} p_{i,l}^\star},$$

这对应于式 (53) 中的第二种情况。 □

利用（50）和定理 4，我们得出（33）成立。事实上，令

$$\rho = \max\{\rho_i \colon 1 \leq i \leq N\} \tag{56}$$

表示各智能体中极点（51）的绝对值最大值。出现两种情形：

- 如果 $\rho = 1$ （因此， $\rho_i = 1$ 至少对一个智能体而言 i），则状态 (a) 被激活。在这种情况下，

$$\alpha = \sum_{i \in \mathcal{I}} |r_{i,\ 1}|^2, \tag{57}$$

其中 $\mathcal{I} = \{i \in \{1, \dots, N\} : \ \rho_i = 1\}$，且 $f\ (k)\ = k+1$；

- 如果 $\rho < 1$ （因此， $\rho_i < 1$ 对所有 i），则情形（b）被激活，其

$$\alpha = \sum_{i \in \mathcal{I}} \sum_{j,l=1}^{P_i} \frac{r_{i,j} r^{\star}_{i,l}}{1 - p_{i,j} p^{\star}_{i,l}} \tag{58}$$

where $\mathcal{I} = \{1, 2, \dots, N\}$, and $f(k) = 1$.

C. 弱信息型智能体

如式 (57) 所示，在状态 (a) 中，只有子集$\mathcal{I}$ 中的智能体会对所有智能体所具有的错误概率渐近指数衰减率作出贡献。因此，子集 $\mathcal{I}$可视为强信息智能体子集，而不属于该集合的智能体（如果存在）则可视为弱信息型智能体——它们的测量值无法帮助提高指数衰减率。这一源自 ARMA 设置的现象意味着，我们可以免除弱信息型智能体获取测量值的任务，从而节省可能价格昂贵的传感器，同时保持原有的渐近检测性能。

请注意，在情形（b）中，所有智能体都是强信息型智能体，因为如式 (58) 所示，它们都会对概率误差底限作出贡献。

VII. 数值实验：嵌入有色噪声中的直流分量

为验证理论结果，我们考虑一个由 $N = 40$ 个智能体通过通信网络连接的检测问题（该网络取样自 Erdős-Rényi 模型），如图 1 所示。式 (11) 中的矩阵 W 选为分布式

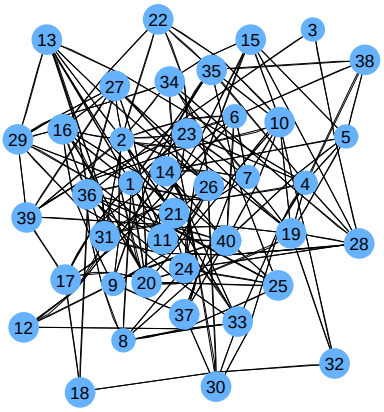

图 1：第 VII 节所用的通信网络是一个包含 $N = 40$ 个节点、连接概率为 $2\log(N)/N$ 的 Erdős-Rényi 图样本；该网络样本有 154 条边。

平均矩阵 [47]

$$W = I_N - \frac{1}{\lambda_2 + \lambda_N} \mathcal{L},$$

其中， $\mathcal{L}$ 是图 1 中图的拉普拉斯矩阵，$0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_N$ 表示拉普拉斯矩阵 $\mathcal{L}$ 的特征值（按升序排列）。所有智能体均采用检测器 (13)，其中 $\gamma = 0$。

我们考虑这样一种检测设置：智能体希望检测被有色噪声掩盖的一个公共直流电平信号，即

$$H_1 : \ y_i(k) = A + n_i(k)$$
$$H_0 : \ y_i(k) = n_i(k),$$

其中 $A > 0$ 为直流分量的幅度。尽管我们的框架也足以涵盖 ARMA 信号，但我们考虑直流分量，以使推导出的速率简单易懂。噪声序列 $(n_i(k))$ 遵循 ARMA 模型，其

$$n_i(k) = n_i(k-1) + \sigma\, \epsilon_i(k) + \sigma\, b_i(1)\, \epsilon_i(k-1),$$

这意味着 (51) 变为

$$\widehat{\Theta}_i(z^{-1}) = \frac{A/\sigma}{1 + b_i(1)\, z^{-1}}. \tag{59}$$

ARMA 滤波器的系数 $b_i(1)$ 因智能体而异，并经过选择以触发状态 (a) 和情形（b），具体内容将在以下各节中详述。

A. 状态 (a)：错误概率呈指数级快速衰减

我们固定了 $A = 1$ 和 $\sigma = 10$，首先从区间 $(-1,1)$ 中均匀随机生成系数 $b_i(1)$：这使得对于所有 i，式 (59) 中极点的绝对值都严格小于一。随后，为了激活状态 (a)，我们将智能体 3 的系数修改为 $b_3(1) = 1$，这使得$\rho_3 = 1$，因此也使得 $\rho = 1$ （回顾定义 (56)）。

因此，强信息智能体的子集是单元素集合 $\mathcal{I} = \{3\}$。这意味着，从渐近意义上说，式 (50) 中网络信号 $\boldsymbol{\theta}(k)$ 的能量来自局部信号 $\boldsymbol{\theta}_3(k)$ 的能量，而后者按照定理 4 的规律表现为

$$\left\| \widehat{\boldsymbol{\theta}}_3(k) \right\|^2 = \frac{A^2}{\sigma^2}(k+1) + o(k+1).$$

（式 (53) 中出现的留数$r_{3,\ 1}$ 在当前情形 (59) 下等于 A/σ ）。因此，

$$\left\| \widehat{\boldsymbol{\theta}}(k) \right\|^2 = \alpha\,(k+1) + o(k+1),$$

其中 $\alpha = A^2/\sigma^2$。

图 2 展示了两个智能体的漏检经验概率：信息性很强的智能体 3，以及作为弱信息性智能体示例的智能体 18。

与定理 3 的预测一致，图 2 显示概率 $\mathbb{P}_{M,\ i}(k)$ 以指数速度快速衰减，并且满足

$$-\frac{\log \mathbb{P}_{M,i}(k)}{k} \to \beta_M = \frac{\alpha}{8} = \frac{A^2}{8\sigma^2} \approx 0.0013. \tag{60}$$

（在时刻 $k = 4000$，速率$\{-\log \mathbb{P}_{M,i}(k)/k$的标准差为 $1 \leq i \leq N\}$，约为 10^{-4}；均值速率 $-\log\mathbb{P}_{M,i}(k)/k$ 与 (60) 中极限β_M之间的偏差约为 5×10^{-4}。）(60) 中的衰减率很直观，因为它取决于信噪比 A/σ。

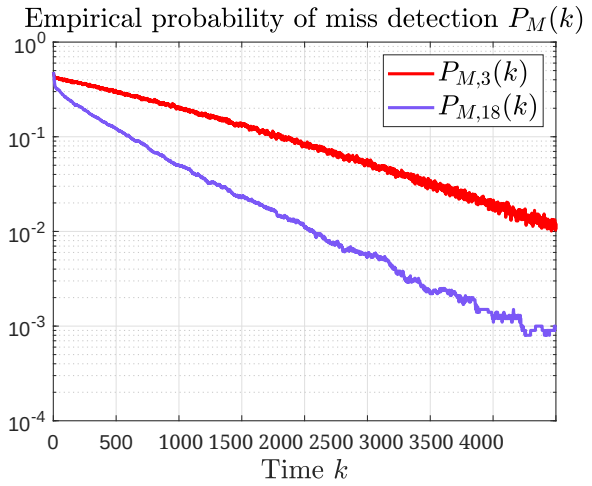

图 2：智能体 3（强信息型）和智能体 18（弱信息型）的漏检经验概率。所有曲线均为 10^5 次蒙特卡洛试验的平均值。其余 38 个弱信息型智能体的曲线$\mathbb{P}_{M,\,i}(k)$未予显示，因为它们与$\mathbb{P}_{M,\,18}(k)$几乎完全相同。

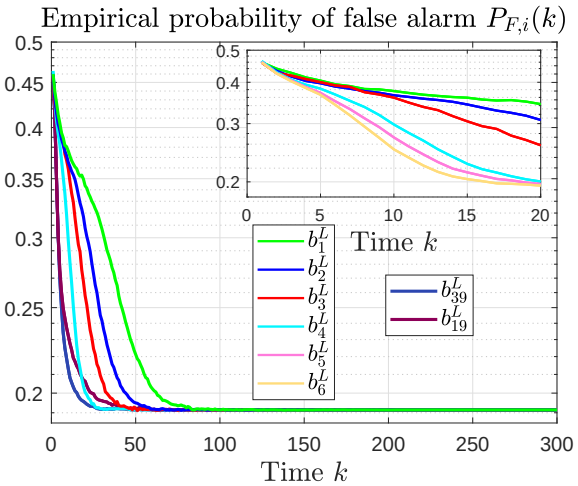

图 3：部分智能体的虚警经验概率。所有曲线均为 10^5 次蒙特卡洛试验的平均值。系数 $b_1^L \geq b_2^L \cdots \geq b_N^L$ 表示{b_i(1) 的绝对值： $1 \leq i \leq N$}，并按非递减顺序排列（因此，$b_1^L = \max_{1\leq i\leq N} |b_i(1)|$ ，且 $b_N^L = \min_{1\leq i\leq N} |b_i(1)|$）。收敛到极限 (61) 最慢的六个智能体，是绝对值 $|b_i(1)|$ 最大的智能体。

B. 情形（b）：错误概率稳定在正误差下限

我们固定了 $A = 1$ 和 $\sigma = 5$，并保持所有参数 b_i(1) 不变（也就是说，与前一节 VII-A 中相同），唯一的例外是，我们在区间 (−1,1) 内均匀随机生成了 b_3(1)。这意味着 (59) 中的极点现在绝对值严格小于 1，也就是说，(56) 中的 ρ 满足 $\rho < 1$；事实上，对于我们这组 b_i(1)，有 $\rho = 0.94$。因此，情形（b）被激活，且所有智能体都具有强信息性。

因此，定理 33 预言错误概率不再衰减至零，而是稳定在一个正误差下限。例如，定理 33 断言

$$\mathbb{P}_{F,i}(k) \to \mathcal{Q}\left(\frac{\sqrt{\alpha}}{2}\right) \tag{61}$$

对于所有智能体 i；这里，α 由式 (58) 给出，根据式 (59)，可写为

$$\alpha = \sum_{i=1}^{N} \frac{A^2}{\sigma^2} \frac{1}{1 - b_1(i)^2}.$$

对于我们这组 b_i(1)，可得增长率 $\alpha \simeq 3.0489$ 和极限虚警概率 $\mathcal{Q}\left(\frac{\sqrt{\alpha}}{2}\right) \approx 0.1913$。

这些预测可由数值模拟予以验证，如图 3 所示。图中给出了部分智能体的虚警经验概率。（在时刻 $k = 300$，{$\mathbb{P}_{F,i}$(k) 的标准差： $1 \leq i \leq N$}约为 10^{-16}，而平均速率$\mathbb{P}_{F,\,i}(k)$ 与极限 (61) 之间的偏差约为 10^{-4}。）

VIII. 结论

本文研究了一种分布式检测设置，其中信号和噪声均可遵循通用 ARMA 模型。这种灵活性使我们能够通过逼近 $\mathcal{Z}$变换，在噪声信号中塑造几乎任意所需的时间相关性。

为处理这种 ARMA 设置，我们提出了对运行共识检测器（RCD）的扩展，并证明该扩展可在每个智能体上高效实现，其计算需求与时间无关，且与通信网络的稀疏性成正比。

正如理论分析所示，ARMA 设置可能使错误概率的渐近行为出现一种新情形。具体而言，在简单的标准（非 ARMA ）设置中，局部检测器的错误概率通常会以指数速度衰减至零；然而，ARMA 设置可能引发另一种情形，此时错误概率会收敛至一个正的误差底限。我们通过推导闭式表达式，全面刻画这两种情形下的指数衰减率和正误差底限，并揭示它们与底层 ARMA 模型之间的关系。我们的分析还揭示了标准检测文献中未曾出现的两个特征。首先，RCD 中使用的阈值常数会极大影响另一种情形下的误差底限。其次，ARMA 设置可能导致智能体分为强信息性智能体和弱信息性智能体；如有需要，后者的测量值可以安全地舍弃（或根本不进行测量），因为这些测量值不会影响 RCD 的渐近性能。数值仿真验证了我们的理论发现。

参考文献

[1] J. Lu、S. Zhou、T. Feng 和 Z. Tong，《一种基于卡方近似的分布式检测加权方法》，载于 2022 IEEE雷达会议（RadarConf22）。IEEE，2022 年，第 1–6 页。[2] Z. Sun、S. Zou、R. Zhang 和 Q. Li，《匿名异构传感器网络中的最快变化检测》，《IEEE信号处理汇刊》，第 70 卷，第 1041–1055 页，2022 年。[3] Y. Inan、M. Kayaalp、A. H. Sayed 和 E. Telatar，《无记忆量化下分布式假设检验的基本极限》，载于 ICC 2022-IEEE国际通信会议。IEEE，2022 年，第 4824–4829 页。

[4] M. A. Wahdan 和 M. A. Altınkaya ，“基于最大平均熵的局部观测量化及其在分布式检测中的应用”，《数字信号处理》，第123卷，文章编号103427，2022年。[5] X. Sun，“基于相关观测的去中心化窄带与宽带频谱感知”，博士论文，密西西比大学，2021年。[6] D. Ciuonzo、P. S. Rossi 和 P. K. Varshney，“乘性衰落无线传感器网络中的分布式检测：基于广义得分检验”，IEEE物联网期刊，第8卷，第11期，第9059–9071页，2021年。[7] A. Mohammadi、D. Ciuonzo、A. Khazaee 和 P. S. Rossi，“分布式稀疏信号检测的广义局部最强检验”，IEEE网络信号与信息处理汇刊，第8卷，第528–542页，2022年。[8] Z. Li、Y. Mo 和 F. Hao，“基于拜占庭传感器的分布式序贯假设检验”，IEEE信号处理汇刊，第69卷，第3044–3058页，2021年。[9] S. Li 和 X. Wang，“完全分布式序贯假设检验：算法与渐近分析”，IEEE信息论汇刊，第64卷，第4期，第2742–2758页，2018年。[10] A. K. Sahu 和 S. Kar，“高斯二元假设检验的分布式序贯检测：异构网络”，载于2014年第48届阿西洛马会议：信号、系统与计算机。IEEE，2014年，第723–727页。[11] D. Li、S. Kar 和 S. Cui，“基于双层大偏差分析的传感器网络分布式最快检测”，IEEE物联网期刊，第5卷，第2期，第930–942页，2018年。[12] D. Reinhard、M. Fauß 和 A. M. Zoubir，“分布式联合检测与估计：一种序贯方法”，载于2020年信息科学与系统年会（CISS）。IEEE，2020年，第1–6页。[13] A. K. Sahu 和 S. Kar，“复合假设检验的递归分布式检测：加性高斯噪声中的非线性观测模型”，IEEE信息论汇刊，第63卷，第8期，第4797–4828页，2017年。[14] D. Bajovic、D. Jakovetic、A. K. Sahu 和 S. Kar，“非独立同分布随机矩阵乘积的大偏差及其在分布式检测中的应用”，载于2018年IEEE国际信息论研讨会（ISIT）。IEEE，2018年，第1061–1065页。[15] D. Jakoveti´c、J. M. Moura 和 J. Xavier，“共识创新检测：通信噪声下的相变”，载于2012年第50届Allerton通信、控制与计算年会。IEEE，2012年，第1559–1563页。[16] D. Bajovic、J. Xavier、B. Sinopoli、J. M. Moura等，“基于高斯运行一致性的分布式检测：大偏差渐近分析”，IEEE信号处理汇刊，第59卷，第9期，第4381–4396页，2011年。[17] A. E. Feitosa、V. H. Nascimento 和 C. G. Lopes，“分布式网络的低复杂度最大后验概率检测器”，载于ICASSP 2020——2020年IEEE国际声学、语音与信号处理会议。IEEE，2020年，第5920–5924页。[18] ——，“使用最大似然检测器的分布式网络自适应检测”，IEEE信号处理快报，第25卷，第7期，第974–978页，2018年。[19] F. Pang、K. Doˇgançay 和 Q. Zhang，“基于扩散卡尔曼滤波的高斯–马尔可夫信号分布式检测”，《信号处理》，第153卷，第368–378页，2018年。[20] M. R. Leonard 和 A. M. Zoubir，“用于检测非高斯噪声中随机信号的鲁棒分布式序贯假设检验”，载于2017年第25届欧洲信号处理会议（EUSIPCO）。IEEE，2017年，第166–170页。[21] M. Kayaalp、Y. Inan、E. Telatar 和 A. H. Sayed，“论分布式推断中信念的算术与几何融合”，arXiv预印本 arXiv:2204.13741，2022年。[22] J. A. Maya 和 L. R. Vega，“传感器网络中一般参数数据分布下的完全分布式复合检验”，IEEE网络信号与信息处理汇刊，第7卷，第509–521页，2021年。[23] A. Nedi´c、A. Olshevsky 和 C. A. Uribe，“分布式（非贝叶斯）学习教程：问题、算法与结果”，载于2016年IEEE第55届决策与控制会议（CDC）。IEEE，2016年，第6795–6801页。[24] A. Lalitha 和 T. Javidi，“基于网络主动假设检验的学习”，载于2017年IEEE信息论研讨会（ITW）。IEEE，2017年，第374–378页。[25] M. R. Leonard、M. Stiefel、M. Fauß 和 A. M. Zoubir，“分布式传感器网络中的多重假设鲁棒序贯检验”，载于2018年IEEE声学、语音与信号处理国际会议（ICASSP）。IEEE，2018年，第4394–4398页。[26] A. K. Sahu和S. Kar，“分布式广义似然比检验：基本极限与权衡”，载于2016年IEEE国际声学、语音与信号处理会议（ICASSP）。IEEE，2016年，第4573–4577页。[27] K. Liu和Y. Mei，“分布式序贯检测中cisprt算法性能特性的改进”，《信号处理》，第172卷，第107573页，2020年。[28] A. K. Sahu和S. Kar，“高斯均值偏移假设检验的分布式序贯检测”，IEEE信号处理汇刊，第64卷，第1期，第89–103页，2015年。[29] M. R. Leonard和A. M. Zoubir，“分布式传感器网络中的鲁棒序贯检测”，IEEE信号处理汇刊，第66卷，第21期，第5648–5662页，2018年。[30] S. Li和X. Wang，“采用量化消息交换的分布式序贯假设检验”，IEEE信息论汇刊，第66卷，第1期，第350–367页，2019年。[31] W. Hou、M. R. Leonard和A. M. Zoubir，“通过鲁棒估计实现鲁棒分布式序贯检测”，载于2017年第25届欧洲信号处理会议（EUSIPCO）。IEEE，2017年，第2051–2055页。[32] P. Braca、S. Marano、V. Matta和P. Willett，“运行共识在二元假设检验中的渐近最优性”，IEEE信号处理汇刊，第58卷，第2期，第814–825页，2009年。[33] A. V. Oppenheim、R. W. Schafer和J. R. Buck，《离散时间信号处理》（第2版）。美国：Prentice-Hall, Inc.，1999年。[34] T. W. Parks和C. S. Burrus，《数字滤波器设计》。Wiley-Interscience，1987年。[35] S. Kay，“通过自回归建模，在未知有色噪声中的渐近最优检测”，IEEE声学、语音与信号处理汇刊，第31卷，第4期，第927–940页，1983年。[36] B. Tang、H. He和S. Kay，“未知有色高斯噪声中的自适应信号检测与参数估计”，2016年。[37] P. Braca、S. Marano和V. Matta，“在无线传感器网络中监测环境的同时确保共识”，IEEE信号处理汇刊，第56卷，第7期，第3375–3380页，2008年。[38] M. Schervish，《统计学理论》，Springer统计学丛书。Springer New York，1996年。[在线]。网址：https://books.google.pt/books?id=F9A9af4It10C [39] S. M. Kay，《统计信号处理基础：估计理论》。美国：Prentice-Hall, Inc.，1993年。[40] R. A. Horn和C. R. Johnson，《矩阵分析》，第2版。美国纽约州纽约市：剑桥大学出版社，2012年。[41] T. Kailath、A. H. Sayed和B. Hassibi，《线性估计》。Prentice Hall，2000年，编号：BOOK。[42] R. S. Ellis，《熵、大偏差与统计力学》。Taylor & Francis，2006年，第1431卷，第821期。[43] H. V. Poor，《信号检测与估计导论》（第2版）。德国柏林、海德堡：Springer-Verlag，1994年。[44] B. C. Levy，《信号检测与参数估计原理》，第1版。Springer Publishing Company, Incorporated，2010年。[45] H. L. Van Trees和K. L. Bell，《检测、估计与调制理论》，第2版。新泽西州霍博肯：Wiley，2013年。[在线]。网址：https://cds.cern.ch/record/1018795 [46] A. Dembo和O. Zeitouni，《大偏差技术与应用》，2010年1月，第95卷。[47] L. Xiao和S. Boyd，“用于分布式平均的快速线性迭代”，《系统与控制快报》，第53卷，第1期，第65–78页，2004年。

附录

我们从以下基本结果开始，该结果将用于引理 1 的证明：如果 $(a(k))$ 是一个各项非负的序列（即，对于所有 k，$a(k) \geq 0$），并且满足

$$a(k) \leq r\, a(k-1) + b(k), \quad k \geq 0, \tag{62}$$

（其中 $a(-1) = 0$）若 $|r| < 1$ 且 $b(k) \to 0$，则 $a(k) \to 0$。这个初等结果可证明如下：由于 $(b(k))$ 是一个收敛序列，因此它有界，设 $|b(k)| \leq B$ 对所有 $k \geq 0$ 均成立；这意味着序列 $(a(k))$ 本身有界：$a(k) \leq A$ 对所有 $k \geq 0$ 均成立（例如，可以通过归纳法证明 $A = B/(1 - |r|)$ 即可）；因此，我们有

$L = $上极限$_k\, a(k)$ 有限（回想一下，上极限$_k\, a(k) = \inf_k A(k)$, 其中 $A(k) = \sup\{a(k)\ a(k+1)\ a(k+2)\ \ldots\}$）；现在，由 (62) 可得 $a(k) \le |r|a(k-1) + |b(k)|$，由上极限的次可加性（上极限$_k(\alpha(k) + \beta(k)) \le$ 上极限$_k\, \alpha(k) +$ 上极限$_k\, \beta(k)$）及其平移不变性（上极限$_k\, \alpha(k) =$ 上极限$_k\, \alpha(k-1)$），再加上上极限$_k\, b(k) = 0$，可得 $L \le |r|L$；最后，将 $L \le |r|L$ 与 $|r| < 1$ 和 $L \ge 0$ 相交，得到 $L = 0$，这等价于 $a(k) \to 0$。

现在转而证明引理 1。特征值分解 (41) 为 R^n 提供了一个自然的标准正交基，即正交矩阵 $\left[\frac{1}{\sqrt{n}}\mathbf{1}_n \quad \mathcal{S}\right]$ 的各列，因此任意向量 $x \in \mathrm{R}^n$ 都可以唯一地写成

$$x = \frac{1}{\sqrt{n}}\mathbf{1}_n\overline{x} + \mathcal{S}\widetilde{x},$$

其中 $\overline{x} = \mathbf{1}_n^T x/\sqrt{n}$ 和 $\widetilde{x} = \mathcal{S}^T x$。在这个坐标系中，证明 (43) 等价于证明

$$\frac{\overline{v}(k)}{f(k)} \to \frac{\theta}{\sqrt{n}} \tag{63}$$

$$\frac{\widetilde{v}(k)}{f(k)} \to 0. \tag{64}$$

(63) 的证明。将递推关系 (40) 两边左乘 $\mathbf{1}_n^T\sqrt{n}$，并利用 $\mathbf{1}_n^T\mathcal{W} = \mathbf{1}_n^T$，得到递推关系

$$\overline{v}(k) = \overline{v}(k-1) + \mathbf{1}_n^T u(k)/\sqrt{n}. \tag{65}$$

回想一下，$\overline{v}(-1) = v(-1) = 0$，由 (65) 可得

$$\frac{\overline{v}(k)}{f(k)} = \frac{1}{\sqrt{n}}\frac{1}{f(k)}\sum_{m=0}^{k}\mathbf{1}_n^T u(m). \tag{66}$$

现在在 (66) 中使用增长率假设 (42)，即可得到 (63)。

(64) 的证明。将递推关系 (40) 两边左乘 $\mathcal{S}^T$，并利用 $\mathcal{S}^T\mathcal{W} = \mathcal{D}\mathcal{S}^T$，得到递推关系

$$\widetilde{v}(k) = \mathcal{D}\,\widetilde{v}(k-1) + \mathcal{S}^T u(k),$$

或者，用分量表示，

$$\widetilde{v}_j(k) = d_j\,\widetilde{v}_j(k-1) + s_j^T u(k), \quad 1 \le j \le n-1, \tag{67}$$

其中，$\widetilde{v}_j(k)$表示向量 $\widetilde{v}(k)$的第 j个分量，d_j是对角矩阵 $\mathcal{D}$ 的第 j个对角元素，而 s_j 是 $\mathcal{S}$的第 j列。

将（67）除以 $f(k)$，得到

$$\frac{\widetilde{v}_j(k)}{f(k)} = d_j\,\frac{f(k-1)}{f(k)}\frac{\widetilde{v}_j(k-1)}{f(k-1)} + \frac{s_j^T u(k)}{f(k)}, \tag{68}$$

这意味着

$$\begin{aligned}\left|\frac{\widetilde{v}_j(k)}{f(k)}\right| &\le |d_j|\left|\frac{\widetilde{v}_j(k-1)}{f(k-1)}\right| + \left|\frac{s_j^T u(k)}{f(k)}\right| \\ &\le |d_j|\left|\frac{\widetilde{v}_j(k-1)}{f(k-1)}\right| + \|s_j\|_\infty\,\frac{\mathbf{1}_n^T u(k)}{f(k)},\end{aligned} \tag{69}$$

其中我们使用了 $f(k) > 0$和 $f(k-1) \le f(k)$（这在以下任一情形下都成立：$f(k) = 1$对所有 k都成立，或 $f(k) = k+1$对所有 k都成立）

为得到第一个不等式；为得到第二个不等式，我们对一般向量 $a = (a_1, \ldots, a_n)$ 和 $b = (b_1, \ldots, b_n)$ 使用了 $a^T b \le \|a\|_\infty \|b\|_1$（其中 $\|a\|_\infty = \max\{|a_1|, \ldots, |a_n|\}$ 且 $\|b\|_1 = |b_1| + \cdots + |b_n|$）（注意，$\|u(k)\|_1 = \mathbf{1}_n^T u(k)$，因为假设每个 $u(k)$ 都是非负向量）。现在请注意

$$\frac{\mathbf{1}_n^T u(k)}{f(k)} \to 0. \tag{70}$$

事实上，若对所有 k都有 $f(k) = 1$，则由（42）可知，级数 $\sum_{k\ge 0}\mathbf{1}_n^T u(k)$收敛（至 θ）；这意味着 $\mathbf{1}_n^T u(k) \to 0$，从而得到（70）。反之，若对所有 k都有 $f(k) = k+1$，则

$$\frac{\mathbf{1}_n^T u(k)}{k+1} = \underbrace{\frac{1}{k+1}\sum_{m=0}^{k}\mathbf{1}_n^T u(m)}_{\alpha(k)} - \frac{k}{k+1}\underbrace{\left(\frac{1}{k}\sum_{m=0}^{k-1}\mathbf{1}_n^T u(m)\right)}_{\beta(k)}. \tag{71}$$

由于由（42）可知 $\alpha(k) \to \theta$ 且 $\beta(k) \to \theta$，在（71）中取极限 $k \to \infty$，即可得到（70）。

最后，利用 (70) 和假设 $|d_j| < 1$，我们看到 (69) 符合模板 (62)，由此得出$\widetilde{v}_j(k)/f(k) \to 0$ 对 $1 \le j \le n-1$成立；也就是说，我们得到 (64)。

A. *Proof of Lemma 2*

我们的证明使用了已知界

$$\frac{t}{1+t^2}\exp\left(-\frac{t^2}{2}\right) \le \sqrt{2\pi}\,\mathcal{Q}(t) \le \frac{1}{t}\exp\left(-\frac{t^2}{2}\right),$$

对 $t > 0$成立，或者等价地，

$$\frac{t^2}{2} + \alpha(t) \le -\log\mathcal{Q}(t) \le \frac{t^2}{2} + \beta(t), \tag{72}$$

其中 $\alpha(t) = \log\left(\sqrt{2\pi}t\right)$ 且 $\beta(t) = \log\left(\sqrt{2\pi}(t + 1/t)\right)$ 对 $t > 0$成立。

(72) 中的不等式给出

$$\frac{u(k)^2}{2} + \frac{\alpha\left(u(k)\sqrt{v(k)}\right)}{v(k)} \le -\frac{\log\mathcal{Q}\left(u(k)\sqrt{v(k)}\right)}{v(k)}, \tag{73}$$

其左侧验证了

$$\frac{u(k)^2}{2} + \frac{\alpha\left(u(k)\sqrt{v(k)}\right)}{v(k)} \to \frac{\theta^2}{2} \tag{74}$$

因为 $u(k) \to \theta$ 和 $v(k) \to \infty$。

（72）中的第二个不等式给出

$$-\frac{\log\mathcal{Q}\left(u(k)\sqrt{v(k)}\right)}{v(k)} \le \frac{u(k)^2}{2} + \frac{\beta\left(u(k)\sqrt{v(k)}\right)}{v(k)}. \tag{75}$$

现在，令 $\beta(t) = \log\left(\sqrt{2\pi}t\left(1 + 1/t^2\right)\right)$，并使用一般不等式 $\log(1+r) \le r$ for $r > 0$，可得

$$\log\left(\sqrt{2\pi}t\right) \le \beta(t) \le \log\left(\sqrt{2\pi}t\right) + 1/t^2,$$

由此很容易推出 $\beta\left(u(k)\sqrt{v(k)}\right)/v(k) \to 0$. 因此，（75）的右侧满足

$$\frac{u(k)^2}{2} + \frac{\beta\left(u(k)\sqrt{v(k)}\right)}{v(k)} \to \frac{\theta^2}{2}. \tag{76}$$

结果（73）、（74）、（75）和（76）意味着（49）。