



中国科学院大学
University of Chinese Academy of Sciences

博士学位论文

典型通信信道下的最大似然调制分类算法研究

作者姓名：_____陈文豪_____

指导教师：_____梁旭文 研究员_____

_____上海微小卫星工程中心_____

学位类别：_____工学博士_____

学科专业：_____通信与信息系统_____

培养单位：_____中国科学院上海微系统与信息技术研究所_____

2019 年 6 月

Maximum-Likelihood Modulation Classification in Typical
Communication Channels

A dissertation submitted to
University of Chinese Academy of Sciences
in partial fulfillment of the requirement
for the degree of
Doctor of Philosophy
in Communication and Information System

By

Chen Wenhao

Supervisor: Professor Liang Xuwen

Shanghai Institute of Microsystem and Information Technology,
Chinese Academy of Sciences

June 2019

中国科学院大学

研究生学位论文原创性声明

本人郑重声明：所呈交的学位论文是本人在导师的指导下独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明或致谢。

作者签名：

日 期：

中国科学院大学

学位论文授权使用声明

本人完全了解并同意遵守中国科学院有关保存和使用学位论文的规定，即中国科学院有权保留送交学位论文的副本，允许该论文被查阅，可以按照学术研究公开原则和保护知识产权的原则公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存、汇编本学位论文。

涉密及延迟公开的学位论文在解密或延迟期后适用本声明。

作者签名：

日 期：

导师签名：

日 期：

摘 要

自动调制识别技术是通信对抗和自适应通信的核心技术,长期以来一直是国内外研究的热点。

通信信号传输过程中受到散射、反射、绕射,以及干扰的影响,导致常用的调制识别算法的识别性能大大下降甚至无法进行识别。因此本文重点聚焦于非高斯噪声、衰落信道场景下的自动调制识别算法。基于似然比检验的调制分类算法在分类准确度方面性能最优,但其算法复杂度较高。本文重点研究非高斯噪声、衰落信道下的具有较低算法复杂度的最大似然调制分类算法,主要完成的工作如下:

- 1) 在平坦衰落非高斯噪声信道下,由于采用 Expectation/Conditional Maximization(ECM)算法进行信道参数盲估计时存在收敛速度较慢的问题,提出了基于 Squared Iterative Method(SQUAREM)算法的信道参数盲估计算法。因为基本的 SQUAREM 算法无法保证参数的每一次更新值都满足其限制条件,所以给 SQUAREM 算法添加参数条件检验这一步骤,每当参数更新值不满足其限定条件,便舍弃该更新值,采用 ECM 算法来获得这次迭代的更新值。基于这一核心思想,提出了 Squared Iterative Method coupled with Parameter Checking(SQUAREM-PC)算法,相较于 ECM 算法,其可在不丢失信道估计准确度的前提下,加快信道估计的速度。
- 2) 在平坦衰落非高斯噪声信道下,通过仿真实验比较了基于 SQUAREM-PC 算法进行信道参数盲估计的最大似然调制分类器与基于 ECM 算法进行信道参数盲估计的最大似然调制分类器的分类准确度与分类时间。
- 3) 频率选择性衰落高斯噪声信道下,利用前向迭代算法计算接收信号似然函数值存在当马尔科夫链的状态空间较大时,算法复杂度会非常高的问题。因此本文采用了一种基于前向粒子滤波的接收信号似然函数值近似计算方案,有效地解决了上述问题。
- 4) 在频率选择性衰落高斯噪声信道参数盲估计环节,存在当马尔科夫链的状态空间较大时, Expectation-Maximization coupled with Forward-Backward

Algorithm (EM-FB)算法复杂度高的问题。而目前尝试解决这一问题的 Expectation-Maximization coupled with Joined Two-Filter Smoothing Algorithm (EM-JTFS)算法又存在着信道估计准确度受信道类型影响较大的问题。所以本文提出了一种基于 Rhee-Glynn Smoothing Estimator (RGSE)的信道参数盲估计方案: Expectation-Maximization coupled with Blocked Rhee-Glynn Smoothing Estimator (EM-BRGSE)算法,该算法可以有效地降低信道参数盲估计的复杂度;易于并行化实现的特性可以进一步加快信道参数盲估计的速度;并且信道估计准确度受信道类型影响较小,在最小相位或最大相位信道下均接近 EM-FB 算法。

- 5) 在频率选择性衰落高斯噪声信道下,通过仿真实验研究了利用 EM-BRGSE 算法进行参数盲估计结合接收信号似然函数近似计算的最大似然调制分类器的分类准确度与分类时间,并与利用 EM-JTFS 算法进行参数盲估计结合接收信号似然函数近似计算的最大似然调制分类器作对比。

关键词: 自动调制识别, 非高斯噪声, 多径衰落, 信道盲估计, 期望最大化算法

ABSTRACT

Automatic modulation recognition is the core technique of communication countermeasure and adaptive communication, and has long been a research hotspot at home and abroad.

The communication signal transmission process is affected by scattering, reflection, diffraction and interference, which results in that the recognition performance of the commonly used modulation recognition algorithm is greatly reduced. Therefore, this paper focuses on the automatic modulation recognition algorithm in non-Gaussian noise and fading channel scenarios. The modulation classification algorithm based on likelihood ratio test has the best performance in classification accuracy, but its algorithm complexity is high. In this paper, we focus on the maximum likelihood modulation classification algorithm with low computational complexity in non-Gaussian noise and fading channels, the main work is as follows:

- 1) In flat fading non-Gaussian noise channel, because of the slow convergence speed of ECM algorithm in blind channel parameter estimation, a blind channel parameter estimation algorithm based on SQUAREM algorithm is proposed. Because the basic SQUAREM algorithm can't guarantee that every update value of the parameter meets its limitation, we add the step of parameter checking to SQUAREM algorithm, whenever the update value of the parameter does not meet its limitation, we abandon it and then use ECM algorithm to obtain the update value of this iteration. Based on this idea, SQUAREM-PC algorithm is proposed. Compared with ECM algorithm, SQUAREM-PC algorithm can speed up channel estimation without losing the accuracy.
- 2) In flat fading non-Gaussian noise channel, the classification accuracy and classification time of maximum likelihood modulation classifiers based on

SQUAREM-PC algorithm and ECM algorithm are compared through simulation experiments.

- 3) In frequency selective fading channel with Gaussian noise, the forward iteration algorithm is used to calculate the likelihood value of the received signal. When the state space of the Markov chain is large, the complexity of the algorithm will be very high. In this paper, an approximate calculation scheme of the likelihood function value of the received signal based on forward particle filter is adopted, which effectively solves the above problem.
- 4) In the blind estimation of frequency-selective Gaussian noise channel parameters, the complexity of EM-FB algorithm is also very high when the state space of Markov chain is large. At present, EM-JTFS algorithm, which is trying to solve this problem, has the problem that the accuracy of channel estimation is greatly affected by the type of channel. So in this paper, we propose a RGSE-based blind channel estimation scheme: EM-BRGSE algorithm, which can reduce the complexity of blind channel estimation. It is easy to implement in parallel, which can further speed up the blind estimation of channel parameters. Moreover, the accuracy of channel estimation is less affected by the channel type, which is close to EM-FB algorithm in the minimum or maximum phase channel.
- 5) In frequency selective Gaussian noise channel, the classification accuracy and classification time of maximum likelihood modulation classifier based on EM-BRGSE algorithm for blind parameter estimation and forward particle filter for approximate calculation received signal likelihood value are studied through simulation experiments. And then we compare it with the maximum likelihood modulation classifier based on EM-JTFS algorithm for blind parameter estimation and forward particle filter for approximate calculation of received signal likelihood value.

Key Words: Automatic Modulation Recognition, non-Gaussian Noise, Multipath Fading, Blind Channel Estimation, Expectation-Maximization Algorithm.

目 录

缩略词.....	VII
图目录.....	IX
表目录.....	XI
第一章 绪论.....	1
1.1 研究背景.....	1
1.2 研究现状.....	4
1.2.1 基于似然比检验的调制分类算法.....	4
1.2.2 基于分布检验的调制分类算法.....	7
1.2.3 基于统计模式的调制分类算法.....	8
1.3 面临的问题及本文的主要工作.....	10
1.4 论文主要研究内容和创新点.....	13
第二章 数学基础	17
2.1 EM 算法.....	17
2.2 串行蒙特卡洛方法.....	19
2.3 隐马尔科夫模型和粒子滤波.....	24
2.4 本章小结.....	26
第三章 平坦衰落非高斯噪声信道下的最大似然调制分类器	27
3.1 信号模型.....	27
3.2 最大似然调制分类器.....	28
3.3 信道参数盲估计.....	29
3.4 仿真及分析.....	36
3.5 本章小结.....	43
第四章 频率选择性衰落高斯噪声信道下接收信号似然函数计算	45
4.1 信号模型.....	45

4.2	接收信号似然函数值计算.....	46
4.2.1	枚举法.....	46
4.2.2	前向迭代算法.....	47
4.2.3	前向粒子滤波.....	49
4.3	仿真与分析.....	53
4.4	本章小结.....	57
第五章 频率选择性衰落高斯噪声信道参数盲估计		59
5.1	FB 算法.....	60
5.2	基于粒子平滑近似算法.....	62
5.2.1	FIS 算法.....	63
5.2.2	TFS 算法.....	64
5.2.3	JTFS 算法	68
5.3	基于 RGSE 的近似算法	71
5.3.1	RGSE 基本原理	71
5.3.2	RGSE 实现方法	72
5.3.3	RGSE 方差降低方案	77
5.3.4	BRGSE 算法.....	78
5.4	仿真及分析.....	79
5.5	本章小结.....	90
第六章 总结与展望		91
6.1	论文主要研究成果.....	91
6.2	工作展望.....	92
参考文献.....		93
致 谢.....		101
作者简历及在攻读博士学位期间发表的学术论文与研究成果		103

缩略词

ALRT	Average Likelihood Ratio Test	平均似然比检验
AS	Ancestor Sampling	先代序号采样
AWGN	Additive White Gaussian Noise	加性白高斯噪声
BPF	Bootstrap Particle Filter	自举滤波
BRGSE	Blocked Rhee-Glynn Smoothing Estimator	分块 Rhee-Glynn 平滑估计器
BW	Baum-Welch algorithm	Baum-Welch 算法
CCPF	Coupled Conditional Particle Filter	耦合条件粒子滤波
CPF	Conditional Particle Filter	条件粒子滤波
CRN	Common Random Number	一致随机数方法
ECM	Expectation/Conditional Maximization algorithm	条件期望最大化算法
EM	Expectation-Maximization algorithm	期望最大化算法
EM-BRGSE	Expectation-Maximization coupled with Blocked Rhee-Glynn Smoothing Estimator	结合分块 Rhee-Glynn 平滑估计器的期望最大化算法
EM-FB	Expectation-Maximization coupled with Forward-Backward Algorithm	结合前向-后向迭代的期望最大化算法
EM-JTFS	Expectation-Maximization coupled with Joined Two-Filter Smoothing Algorithm	结合联合双粒子滤波器平滑估计器的期望最大化算法
FB	Forward-Backward algorithm	前向-后向迭代算法
FIS	Fixed-Interval Smoothing	固定间隔平滑算法
GLRT	Generalised Likelihood Ratio Test	广义似然比检验
GMM	Gaussian Mixture Model	混合高斯模型
HLRT	Hybrid Likelihood Ratio Test	混合似然比检验
HMM	Hidden Markov Model	隐马尔科夫模型

ICR	Index-Coupled Resampling	序号耦合重采样
JTFS	Joined Two-Filter Smoothing	联合双粒子滤波器平滑算法
KL	Kullback-Leibler Divergence	KL 散度
KS	Kolmogorov-Smirnov Test	柯尔莫哥洛夫-斯米尔诺夫检验
MSE	Mean Squared Error	均方误差
quasi-ALRT	quasi Average Likelihood Ratio Test	准平均似然比检验
quasi-HLRT	quasi Hybrid Likelihood Ratio Test	准混合似然比检验
RGSE	Rhee-Glynn Smoothing Estimator	Rhee-Glynn 平滑估计器
RB	Rao-Blackwellization Method	Rao-Blackwellization 方法
SQUAREM	Squared Iterative Method	平方迭代方法
SQUAREM-PC	Squared Iterative Method coupled with Parameter Checking	带有参数限制条件检验的平方迭代方法
STEM	Steffensen-type Methods	Steffensen 类方法
TFS	Two-Filter Smoothing	双粒子滤波器平滑算法

图目录

图 1.1 调制分类器在电子对抗中的应用	1
图 1.2 链路自适应系统	3
图 1.3 信道干扰分类	12
图 3.1 SQUAREM-PC 和 ECM 算法的 F_{cnt} 随 SNR 变化图	38
图 3.2 $N = 2$ 时, SQUAREM-PC 和 ECM 算法对信道系数幅度和相位估计的平均偏差	39
图 3.3 $N = 3$ 时, SQUAREM-PC 和 ECM 算法对信道系数幅度和相位估计的平均偏差	39
图 3.4 $N = 2$ 时, 调制分类准确度随 SNR 变化曲线图	40
图 3.5 $N = 3$ 时, 调制分类准确度随 SNR 变化曲线图	41
图 3.6 调制分类准确度随参数 γ 的变化曲线图	42
图 4.1 最小相位信道下, 发射机采用 64QAM 时, 接收信号似然函数近似值分布情况	54
图 4.2 最小相位信道下, 发射机采用 128QAM 时, 接收信号似然函数近似值分布情况	56
图 4.3 最大相位信道下, 发射机采用 128QAM 时, 接收信号似然函数近似值分布情况	57
图 5.1 最小相位信道下, RGSE 对 $E[d(s_{i,T})]$ 的估计结果分布图	82
图 5.2 不同分块数下的 BRGSE 对 $E[d(s_{i,T})]$ 的估计结果分布图	82
图 5.3 最大相位信道下, RGSE 对 $E[d(s_{i,T})]$ 的估计结果分布图	83
图 5.4 最小相位信道下, 三种算法对信道系数估计的 MSE 随着迭代次数的变化情况图	85
图 5.5 最小相位信道下, 三种算法信道系数估计值与真实值之间的误差的平方和的分布情况	85
图 5.6 最大相位信道下, 三种算法对信道系数估计的 MSE 随迭代次数变化情况图	87
图 5.7 最大相位信道下, 三种算法信道系数估计值与真实值之间的误差的平方和的分布情况	87

表目录

表 1. 1 基于似然比检验的调制识别算法研究情况汇总	5
表 3. 1 16QAM 作为发射调制类型, 5000 次实验中, SQUAREM 和 SQUAREM-PC 的 N_{err}	37
表 3. 2 SQUAREM-PC 算法的收敛周期内, 违反限制条件的次数占整个迭代次数的百分比	37
表 3. 3 不同 SNR 条件下的平均调制分类时间	41
表 3. 4 不同 γ 下的平均调制分类时间	43
表 4. 1 最小相位信道下, 发射机采用 64QAM 时, 接收信号似然函数值的近似计算与精确计算时间对比	55
表 4. 2 最小相位信道下, 发射机采用 128QAM 时, 接收信号似然函数值的近似计算与精确计算时间对比	56
表 4. 3 最大相位信道下, 发射机采用 128QAM 时, 接收信号似然函数值的近似计算与精确计算时间对比	57
表 5. 1 不同信道类型下的 $E(\tau)$ 及其所对应的 N	80
表 5. 2 不同信道阶数下的 $E(\tau)$ 及其所对应的 N	80
表 5. 3 不同调制阶数下的 $E(\tau)$ 及其所对应的 N	81
表 5. 4 最小相位信道下, 三种算法对信道盲估计所需的时间	86
表 5. 5 最大相位信道下, 三种算法对信道盲估计所需的时间	88
表 5. 6 最小相位信道下, 基于 EM-BRGSE 和 EM-JTFS 算法的调制分类器平均分类准确度与时间	89
表 5. 7 最大相位信道下, 基于 EM-BRGSE 和 EM-JTFS 算法的调制分类器平均分类准确度与时间	89

第一章 绪论

1.1 研究背景

自动调制识别技术最先应用在军事领域中，电子对抗、威胁评估等都需要应用识别出来的调制方式来确定敌方发射机类型、调制阻塞信号、恢复截获信号。自动一词是相对于传统的人工识别而提出来的，最初的调制识别是由专业人员借助信号处理和观察设备，观察信号波形时域和频域特征来确定的，但这一方法只适用于简单的调制方式，例如只有 AM、FM 信号时。随着无线通信技术，特别是数字通信技术的快速发展，信号调制方式变得越来越复杂，上述方法已无能为力。解决上述问题的唯一技术途径，就是研制能够自动识别通信信号调制方式的设备。

电子对抗（又称电子战）就是敌对双方为削弱、破坏对方电子设备的使用效能、保障己方电子设备发挥效能而采取的各种电子措施和行动。电子对抗分三个部分：电子对抗侦察、电子干扰和电子防御。电子对抗按电子设备的类型可分为雷达对抗、无线电通信对抗、导航对抗、制导对抗、光电对抗和水声对抗等；按配置部位又可分为外层空间对抗、空中对抗、地面（包括海面）对抗和水下对抗。机载电子对抗系统是现代电子对抗的主要手段。随着弹道导弹和卫星的发展，外层空间是一个新的战场，电子对抗在未来的现代化战争中，将对战略攻防起到重要作用[1]。图 1.1 说明了在电子对抗中，调制分类器所起的作用。

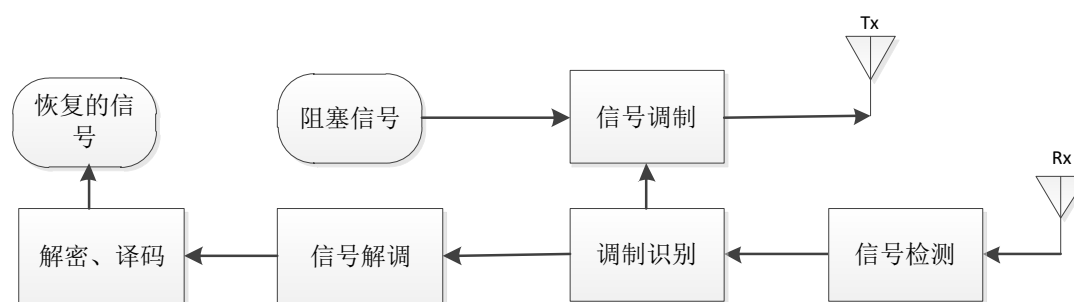


图 1.1 调制分类器在电子对抗中的应用

Figure 1.1 Applications of Modulation Classifier in Electronic Countermeasure

电子对抗侦察的目标是从截获信号中获取有用信息，当接收机成功检测到信

号时，需要应用自动调制识别技术来识别接收信号所采用的调制类型，利用该调制方式进行解调，然后结合解密、译码便可以获得相关情报信息。

在电子干扰方面，阻塞是干扰敌方通信的主要方式。虽然有许多阻塞方式，但是最常用的一种是应用敌方通信采用的调制类型和载波频率发送干扰信号。敌方通信的载波频率可以通过信号检测器获得。通过功率放大器可以将阻塞信号的发射功率调到足够大。为了让干扰程度达到最大化，在满足前面载波频率和功率条件下，需要采用与敌方通信相匹配的调制方式，这里就需要应用自动调制识别技术来获得敌方发射机所采用的调制方式。

在电子防御方面，其目标是保护友方通信，以免受到敌方的通信干扰。如上所述，阻塞器在敌方通信相同的频率段发射大功率干扰信号来覆盖敌方通信。为了达到最大程度的覆盖，其关键点在于采用与敌方通信一致的调制方式。所以有效避免友方通信受到阻塞的方法是及时感知到敌方阻塞信号并做出相应的规避措施，具体来说，友方发射机需要监控阻塞信号的调制方式并且及时改变友方通信所采用的调制方式。

在蜂窝移动通信系统中，一个非常重要的特征是无线信道的时变特性，其中无线信道的时变特性包括传播损耗、平坦衰落、频率选择性衰落以及干扰的变化等因素带来的影响。由于无线信道的变化性，接收端接收到的信号质量也是一个随着无线信道变化的变量，如何有效地利用信道的变化性，如何在有限的带宽上最大限度地提高数据传输速率，从而最大限度地提高频带利用效率，逐渐成为移动通信的研究热点。而链路自适应技术正是由于在提高数据传输速率和频谱利用率方面有很强的优势，从而成为目前和未来移动通信系统的关键技术之一[2]。

链路自适应技术，也叫做自适应调制与编码技术，开辟了一种新的调制方式。它可以根据信道状况自适应选择调制方式，最大程度保证信息传输的高效性和可靠性。由于发射机可以自由选择调制方式，接收机要想正确解调信息，必须知道发射机所用的调制类型。简单的解决方案是将调制信息包含在发射帧中，接收机可以知道发射机所采用的调制方式。但是在信号帧中添加调制信息会降低频谱使用效率，而目前无线频谱资源是极其有限和珍贵的，所以这一种方案虽然简单但还不够有效。因此，自动调制识别成为这一问题很有吸引力的解决方案。

如图 1.2 所示，链路自适应系统中，发射机信号调制单元被自适应调制单元

取代。自适应调制器的功能是从备选调制集中选择调制方式完成信号调制。调制方式的选择是由信道条件所决定的。当信道状况较差时，为了保证通信链路的可靠性，往往会选择调制阶数较低的调制类型，比如 BPSK、QPSK。当信道状况良好时，为了满足高传输速率的要求，往往选择调制阶数高的调制类型，如 16QAM、64QAM 等。在这种通信系统中，发射机和接收机之间唯一的一次合作通信是在系统初始化时，发射机将所采用的备选调制集告知接收机。在之后的通信中，发射机与接收机进行的是非合作通信。在链路自适应系统的接收端，首先进行的是信道状态估计。如果是静态信道，那么信道状态估计只在初始阶段进行。如果是动态信道，那么需要周期地估计信道状态。估计的信道状态信息反馈到发射机，作为调制方式选择的参考。同时，信道状态信息可以辅助自动调制分类。通常情况下，信道状态估计的准确度对于调制分类器性能影响很大。依据调制分类的结果，可以通过可配置信号解调器进行相应的信号解调。如果调制分类的结果正确，可以解调出正确的传输信息。如果调制分类器结果错误，那么整个通信过程将会失效。由此可见自动调制识别技术对于链路自适应系统的重要性。

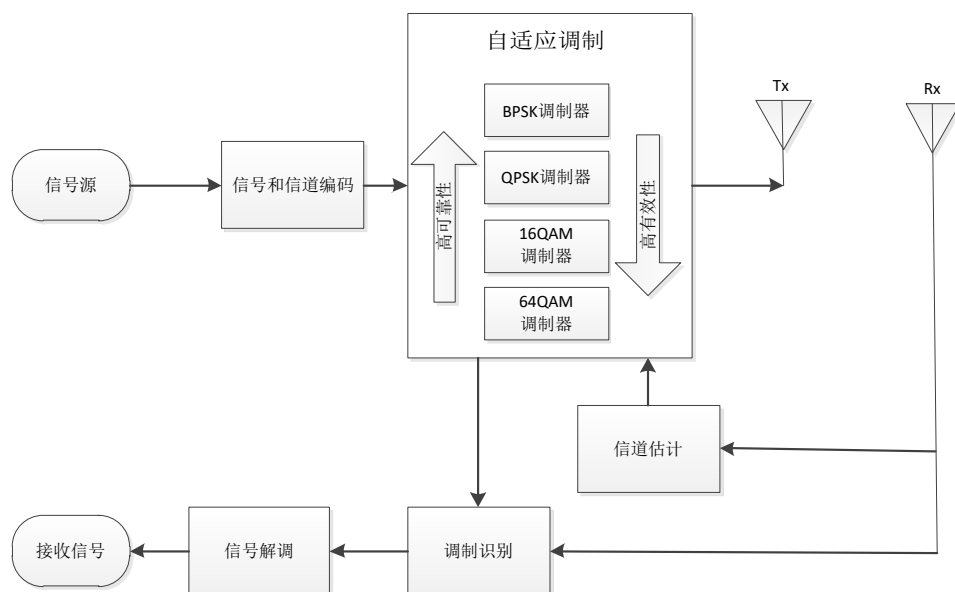


图 1.2 链路自适应系统

Figure 1.2 Link adaptive transmission system

随着微处理器技术的发展，接收机具有更高的计算性能，所以在接收机中进行自动调制识别是可行的。综上所述，自动调制识别技术在军用和民用通信领域有着广泛的应用，是软件无线电和认知无线电等领域的研究基础。因实际通信信

道的复杂性,在现有自动调制识别算法的基础上,研究出适用于实际通信信道的高准确率、低复杂度的自动调制识别算法仍然是一项颇具挑战性的研究课题。

1.2 研究现状

当前对于自动调制识别算法的研究主要分为三大类:基于似然比检验、基于分布检验以及基于特征提取。从分类准确度角度来看,基于似然比检验的调制分类算法性能最优,但从算法复杂度角度来看,基于似然比检验的调制分类算法复杂度高于基于分布检验和特征提取的调制分类算法。下面我们将详细说明这三类调制分类算法的研究现状。

1.2.1 基于似然比检验的调制分类算法

基于似然比检验的调制分类器是迄今为止最为流行的调制分类器,其原因在于当分类器知道正确的信号模型以及准确的信道参数时,基于似然比检验的调制分类器在分类准确度方面是最优的[3][4][7][26][27]。

基于似然比检验的调制分类算法通常分为两个步骤。第一步,基于接收信号计算每个调制类型假设的似然函数值;第二步,比较不同调制类型假设的似然函数值来获得分类结果。常在调制类型假设两两之间采用似然比检验的方法来获得分类结果,这种方法确实能够提高分类的准确度,但是其需要优化门限值的选择,会增加分类器的计算复杂度。更加直观的判决方案便是在所有待识别的调制类型中,选择似然函数值最大的作为分类结果,这种方案不需要选择门限值,更容易实现,这便是最大似然调制分类算法。

在实际的通信过程中,由于无法预先知道相关的信道参数,所以无法直接计算每个调制类型假设的似然函数值。根据处理未知的信道参数的方法的不同,可以将基于似然比检验的调制分类算法分为三大类:平均似然比检验[3]-[16][28][29]、广义似然比检验[17][22][23]和混合似然比检验[17]-[21], [24]-[27][30][77][78]。在文献[5]-[7], [9], [11]-[15]中还提出了基于准平均似然比检验的调制分类算法,在文献[25]中提出了基于准混合似然比检验的调制分类算法。

Polydoros 和 Kim 首先将平均似然比检验的方法应用到调制分类[6],之后,文献[3]-[5], [7]-[16]都采用了这一方法。在处理最大似然调制分类器中的信道参数未知的这一问题时,平均似然比检验方法应用未知参数的概率密度函数对未知

参数求积分来计算每个调制类型假设的似然函数值。由于引入了积分运算，导致平均似然比检验方法的复杂度非常高。同时，需要知道未知参数的概率密度函数，这在实际复杂的通信环境中也是十分困难的。

针对平均似然比检验方法的缺点，Panagiotou, Anastasopoulos 和 Polydoros 提出了广义似然比检验方法[17]。广义似然比检验方法在本质上是最大似然估计器与最大似然分类器的结合。它将信道参数以及发射符号看作是确定的未知量，首先在每一个调制类型假设下，利用最大似然估计的方法得出相关未知量的估计值，然后带入似然函数中求解该假设的似然函数值。但是该方法无法很好地分类嵌套调制类型，如 QAM 和 16QAM，或者 QPSK 与 16PSK 等。

针对广义似然比检验方法无法有效地分类嵌套调制类型的问题，Panagiotou 等又提出了混合似然比检验方法[17]。混合似然比检验方法将未知的发射信号看作是随机变量，将信道参数看作是未知的确定量，所以混合似然比检验方法在处理未知的发射信号时与平均似然比检验方法一致，在处理未知的信道参数时与广义似然比检验方法一致。因此混合似然比检验方法结合了平均似然比检验方法和广义似然比检验方法的优点，能够有效地解决广义似然比检验方法无法准确分类嵌套调制类型的问题。

表格 1.1 列举了几种基于似然比检验的调制识别算法，重点分析了这些算法识别的调制类型集合、使用的信道模型以及信号模型中的未知参数。

表 1.1 基于似然比检验的调制识别算法研究情况汇总

Table 1.1 Summary of research on Modulation Recognition Algorithms based on Likelihood

Ratio Test				
作者	分类器	识别的调制 类型集合	未知参数	信道 类型
Sills[3]	ALRT	BPSK,QPSK,16QAM,V29, 32QAM,64QAM	载波相位	AWGN
Wei 和 Mendel[4]	ALRT	16QAM,V29	--	AWGN

作者	分类器	识别的调制 类型集合	未知参数	信道 类型
Kim 和 Polydoros[5][6]	quasi-ALRT	BPSK,QPSK	载波相位	AWGN
Huan 和 Polydoros[7]	quasi-ALRT	UW,BPSK,QPSK,8PSK,16PSK	载波相位和 定时偏差	AWGN
Sapiano 和 Martin[8]	ALRT	UW,BPSK,QPSK,8PSK	--	AWGN
Long 等[9]	quasi-ALRT	16PSK,16QAM,V29	载波相位	AWGN
Hong 和 Ho[10]	ALRT	BPSK,QPSK	信号电平	AWGN
Beidas 和 Weber[11]	ALRT quasi-ALRT	32FSK,64FSK	相位抖动	AWGN
Beidas 和 Weber[12][14]	ALRT quasi-ALRT	32FSK,64FSK	相位抖动和 定时偏差	AWGN
Panagiotu 等[17]	GLRT HLRT	16PSK,16QAM,V29	载波相位	AWGN
Chugg 等[18]	HLRT	BPSK,QPSK,OQPSK	载波相位、 信号功率和 噪声功率谱 密度	AWGN
Hong 和 Ho[19]	HLRT	BPSK,QPSK	信号电平	AWGN
Hong 和 Ho[20][21]	HLRT	BPSK,QPSK	到达角	AWGN

续表 1.1

作者	分类器	识别的调制 类型集合	未知参数	信道 类型
Dobre 等[24]	HLRT	BPSK,QPSK,8PSK,16PSK, 16QAM,64QAM	信道幅度和 相位	平坦衰落 高斯噪声
Abdi 等[25]	ALRT quasi-HLRT	16QAM,32QAM,64QAM	信道幅度和 相位	平坦衰落 高斯噪声
吴迪等[30]	HLRT	BPSK,QPSK,8PSK,16QAM	信道幅度和 相位	平坦衰落 高斯噪声
Chavali 等[77]	HLRT	BPSK,QPSK,8PSK,16QAM	信道幅度和 相位 噪声 参数	平坦衰落 非高斯噪 声
Puengnim 等[78]	HLRT	BPSK,QPSK,8PSK,16QAM	多径信道系 数	频率选择 性衰落高 斯噪声

1.2.2 基于分布检验的调制分类算法

当接收信号的长度足够长时,利用调制信号的经验累积分布函数值可以构建有效的调制分类器。因为接收信号的经验累积分布函数主要受发射信号采用的调制类型以及信道参数的影响,假设信道参数已经预先估计出来,那么接收信号的经验累积分布函数只和发射符号采用的调制类型有关,因此可以通过分析接收信号的经验累积分布函数来确定发射符号采用的调制类型。

基于分布检验的调制分类器的核心思想是:从待识别调制集合中每一个调制类型对应的理论累积分布函数或经验累积分布函数选出最接近接收信号经验累积分布函数的一个作为分类结果。利用吻合度来衡量两个累积分布函数之间的接近程度。KS 检验是一种吻合度检验的方法[31],利用其来衡量两个累积分布函数之间的接近程度。文献[32]第一次采用 KS 检验来实现调制分类。相比于基于

似然比检验的调制分类器，基于 KS 检验的调制分类器具有相对较低的计算复杂度；相比于基于累积量的调制分类器[34]，基于 KS 检验的调制分类器具有较好的稳健性。文献[35]就计算有效性方面改进了文献[32]中的算法。文献[36]利用调制星座的对称性，提出基于折叠 KS 检验调制分类器，提升了样本点数较少的情形下调制分类的准确度。文献[33]详细地比较了基于分布检验的调制识别算法在采用不同的判决统计量和检验统计量等情况下的识别性能。

1.2.3 基于统计模式的调制分类算法

调制分类中经常使用的信号特征主要分为五类：信号的瞬时域特征、信号的变换域特征、信号的统计特性、调制星座点特征以及信号过零次数。信号瞬时域特征主要包含信号瞬时幅度、相位和频率。变化域特征主要是指信号的傅里叶变化或小波变换后所具有的特征。统计特性主要包含信号的高阶矩、高阶累积量、高阶循环累积量以及循环平稳特性等。下面详细说明各个特征在调制识别中的研究情况。

A. 瞬时域特征

信号的调制主要是通过改变载波频率、载波相位以及载波的幅度来实现的，所以这些参数的瞬时变化特征常常可以用来分类调制类型。在文献中经常采用的信号时域瞬时特征有：信号零中心归一化瞬时幅度绝对值的标准偏差 σ_{aa} 、非弱信号零中心归一化瞬时幅度绝对值的标准偏差 σ_a 、非弱信号瞬时相位零中心非线性成分的标准偏差 σ_{dp} 、非弱信号瞬时相位零中心非线性成分绝对值的标准偏差 σ_{ap} 、信号零中心归一化瞬时频率绝对值的标准偏差 σ_{af} 、非弱信号零中心归一化瞬时频率绝对值的标准偏差 σ_{nf} 、归一化瞬时幅度的峰值 μ_{42}^a 和归一化瞬时频率的峰值 μ_{42}^f 以及频谱对称性。

由于 σ_{aa} 和 σ_a 可以捕捉到调制信号的幅度变化情况，所以文献[37][38][39]采用其作为确定 ASK 调制阶数的特征，文献[40]利用其来区分 ASK 和 PSK 调制方式。与相位相关的瞬时特征 σ_{dp} 和 σ_{ap} 有助于确定 PSK 调制阶数[38][41]，同时，可以利用它们来区分 4PSK/MQAM 和 MFSK/2PSK[40]、MPSK/MQAM 和 MASK。

σ_{af} 和 σ_{nf} 被用来确定 FSK 的调制阶数, 以及区分 FSK 和 PSK[40]。在文献[42]中, μ_{42}^a 被用来区分 2ASK/4ASK 和 AM, μ_{42}^f 被用来区分 2FSK/4FSK 和 FM。另外, 文献[40][41]利用 μ_{42}^a 区分 ASK/QAM 和 FSK/PSK, 文献[41]利用 μ_{42}^f 区分 PSK 和 FSK。频谱对称性在文献[42]中被用来区分 AM/MASK 和 VSB、FM/MFSK/DSB/MPSK 和 LSB/USB、LSB 和 USB。

虽然基于瞬时特征的调制分类器易于实现, 但是这些瞬时特征极易受到噪声的影响, 导致分类器的抗噪声性能较差。

B. 变换域特征

在傅里叶变换域中, 零中心归一化瞬时幅度功率谱密度的最大值 γ_{\max} 经常被用作分类调制方式的特征。由于 γ_{\max} 反映了信号幅度的变化情况, 所以在模拟调制类型中或者数字调制类型中, 经常使用其来区分幅度类型的调制方式和非幅度类型的调制方式。在识别模拟调制类型时, 文献[37]利用其来区分 AM 和 CW/FM。在识别数字调制类型时, 文献[40][41][42]利用其来区分 FSK/PSK 和 MQAM/MASK。另外, 接收信号解析形式的 k 次方离散傅里叶变换后的幅度的最大值 Γ_k 因其受载波频率偏移和时偏的影响较小的原因在文献[43]中被用来识别 PSK 类调制方式。

小波变换因其能够降低噪声对特征提取影响的优势被广泛应用于自动调制识别领域[37][38], [44]-[51]。因为从小波变化域中提取的特征既包含了信号的时域信息, 也包含了信号的频域信息, 所以基于小波变换域中的特征的调制分类器的稳健性优于基于傅里叶变换域中特征的调制分类器。

C. 统计特性

高阶统计量主要包括高阶矩和高阶累积量, 被广泛应用于分类各种 ASK、PSK 和 QAM 类调制方式[47][48][53][54]。信号的高阶统计量的特征主要有三个优势: 反映了信号的高阶统计特性、消除了噪声对特征提取的影响以及这些特征受相位旋转的影响较小。文献[49]利用高阶累积量构成多维空间来判断信号的调制方式。文献[52]利用高阶累积量来消除多径效应的影响, 实现多径信道下的调制识别。

调制信号具有循环平稳特性，并且这一特征不受载波频率、相位偏移以及定时偏差的影响，所以在文献[55][56]中采用这一特征进行盲调制方式识别。但是要想提升分类的准确度，还需依赖稳健的循环频率估计。

D. 调制星座形状特征

PSK 和 QAM 信号星座图的几何特征表征了每一种调制方案，这些几何特征主要包含星座图中星座点数以及星座点的位置，一般用到原点的距离以及相位角来定义星座点的位置。在文献[57]中，利用星座旋转技术将星座点的相位-幅度分布转化为一维分布。文献[58]通过将截获信号的星座图与参考星座图比较来进行调制分类，研究了噪声、载波频偏以及载波相偏对分类准确度的影响。

E. 信号过零次数

对截获信号的过零次数进行计数已经被用于调制分类研究中[59][60]。很明显地，PSK 信号的过零率在观察信号中是恒定的，而 FSK 信号的过零率在观察信号中是变化的。这一特性被用来确定 FSK 信号的调制阶数以及区分 PSK 和 FSK[61]。

F. 特征学习

特征学习类算法直接利用原始接收数据而非精心挑选的信号特征进行自动调制识别。文献[62]提出采用卷积神经网络进行自动调制识别，时域同相分量和正交分量样本直接作为卷积神经网络的输入信号，在不同的信噪比下，合适的匹配滤波器可以被自动学习出来。在文献[63]中，研究了不同尺寸、深度的卷积神经网络的调制分类准确度；同时，作者还将卷积神经网络和长短期记忆网络相结合来提高分类准确度。文献[64]构建了 5 层卷积神经网络来识别高频信号，频率偏移以及噪声对该分类器性能的影响较大。为了提高识别准确率，文献[65]提出一个新的模型来获得频率偏移以及相位噪声的估计值。为了进一步提高文献[62]中提出的基于卷积神经网络的调制分类器准确度，作者设计了信号失真矫正模块，实验结果显示卷积神经网络结合信号失真矫正模块可以提高调制分类准确度。

1.3 面临的问题及本文的主要工作

在实际的通信场景中，由于无线电波的传输介质或者传输路径的改变会引起接收信号功率的变化，这一现象称作衰落。在发射机和接收机相对固定的场景中，大气条件如降雨、闪电等会引起信号衰落；在发射机和接收机相对移动的场景中，

信号传输路径上随时间变化的障碍物对传输信号产生复杂的传输效应从而引起信号的衰落。

在无线通信系统中，考虑到传输信道的各种影响，信号衰落可以分为：

1) 大尺度衰落

大尺度衰落主要包括路径损耗和阴影效应。路径损耗是由发射机辐射的功率损耗以及信号传输信道的影响引起的；阴影效应是由发射机和接收机之间会吸收部分信号功率的障碍物所引起的。信号传输距离较大时（100-1000 米）会产生较明显的路径损耗，而信号传输距离在与障碍物尺度成正比的距离时（室外 10-100 米，室内 10-100 米以下）会产生较明显的阴影效应。

2) 小尺度衰落

小尺度衰落是指在短距离和短时间内接收信号强度的快速波动，其主要是由于经过多径传输到达接收机的信号相互叠加所形成的。根据多径时延扩展与符号周期的大小关系可以将小尺度衰落分为平坦衰落和频率选择性衰落。

A. 平坦衰落

当发射符号周期远远大于多径时延扩展时，信号在传输信道中经历平坦衰落。在平坦衰落信道中，无线信道在大于发射信号带宽的带宽上具有恒定增益和线性相位响应。在这种类型的衰落中，接收信号的所有频率分量同时以相同的比例波动。

B. 频率选择性衰落

当发射符号周期小于多径时延扩展时，信号在传输信道中经历频率选择性衰落。在这种类型的衰落中，接收信号的不同频率分量以不同的比例波动。

无线电波在信道中传输时，除了受到信道衰落的影响，还会受到各种噪声的干扰。在实际通信环境中，无线电与各种干扰其工作的电磁辐射源共存。这些将附加噪声引入通信系统的源可能包括大气噪声[66]、无线网络中的同信道干扰[67]以及各种人造设备（如汽车点火装置[81]、电梯[68]、微波炉[69]、计算机时钟[70]以及 LCD 监视器[71]等）的附带辐射。在各种环境下进行的各种测量一致认为，这种加性噪声可以由非高斯过程很好地建模的，并且在大多数情况下表现出脉冲特性[68]-[75]。

因此，无线电波在无线信道中传播受到信道衰落和噪声的干扰，我们将其统

称为信道干扰。无线电波在信道中受到的信道干扰情况如图 1.3 所示。

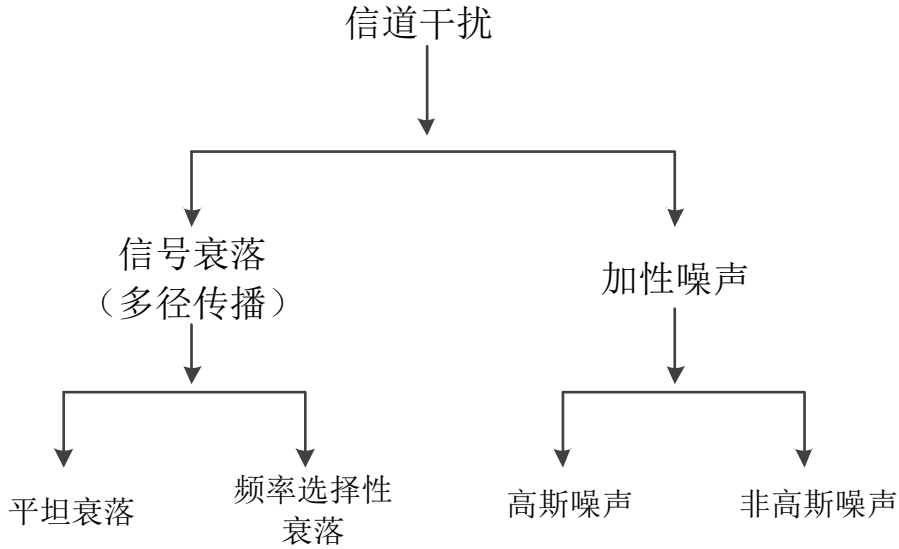


图 1.3 信道干扰分类

Figure 1.3 Classification of channel interference

从自动调制识别算法的研究现状来看，大多数算法都是在假设无线信道为 AWGN 信道的条件下提出的。然而，适用于 AWGN 信道的调制识别算法在非高斯噪声信道下性能下降严重，因此需要研究适用于非高斯噪声信道下的自动调制识别算法。

因为基于似然比检验的调制识别算法在分类准确度方面性能最优，所以本文首先研究非高斯噪声信道下的最大似然调制识别算法。目前这方面的研究主要有文献[76][77]。文献[76]假设调制分类器已知了信道状态和噪声分布，而在实际通信信道下，调制分类器很难预先知道这些信息。文献[77]研究了平坦衰落非高斯噪声信道下的盲调制识别算法，作者用混合高斯模型建模非高斯噪声，然后利用 ECM 算法对信道参数进行盲估计，最后将估计出来的参数带入接收信号的似然函数中来计算每一个调制假设的似然函数值并作决策。由于 ECM 算法较慢的收敛速度，导致了该调制分类算法的复杂度很高，因此本文提出了一种加速 ECM 算法收敛速度的方案，加快了平坦衰落非高斯噪声信道下信道参数盲估计的速度，进而降低了调制分类算法的复杂度。

另外，基于 AWGN 假设下提出的调制分类器的性能在频率选择性衰落信道下将会极大地降低，因此本文还研究了频率选择性衰落高斯噪声信道下的最大似

然调制识别算法。

文献[78]提出采用隐马尔科夫模型建模频率选择性衰落高斯噪声信道下的接收信号,将每个调制假设下信道参数的估计等效为隐马尔科夫模型参数估计问题,利用 EM-FB 算法进行参数估计,随后利用最大似然调制分类算法进行调制分类。由于当调制阶数 M 很大或者信道阶数 L 很大时,EM-FB 算法的复杂度将会非常高。文献[91]提出基于粒子平滑的信道参数盲估计算法来解决当 M 很大或者 L 很大时,EM-FB 算法复杂度太高的问题。但是[91]提出的算法在最大相位信道下,存在参数估计的准确度较差的问题,所以本文基于 RGSE 算法,提出一种在最小、最大相位信道下,参数估计的准确度均接近 EM-FB 算法的低复杂度信道盲估计算法,并且所提出的算法易于并行实现,可以进一步加快信道盲估计的速度。另外,用来计算接收信号似然函数值的前向迭代算法同样存在当 M 很大或者 L 很大时,算法复杂度会非常高的问题,本文采用前向粒子滤波算法来近似计算接收信号似然函数值,有效地解决了这一问题。

1.4 论文主要研究内容和创新点

本文以降低基于似然函数的调制分类算法复杂度为出发点,围绕平坦衰落非高斯噪声信道和频率选择性衰落高斯噪声信道环境下的最大似然调制分类算法展开研究。全文内容安排如下:

第一章: 绪论

主要介绍论文的研究背景、研究意义,调制分类算法的研究现状、存在的问题,以及本文的主要研究内容和创新点。

第二章: 数学基础

主要介绍了本研究中涉及到的相关算法: EM 算法、串行蒙特卡洛方法以及隐马尔科夫模型和粒子滤波算法。

第三章: 平坦衰落非高斯噪声信道下的最大似然调制分类算法研究

首先介绍了平坦衰落非高斯噪声信道下,接收信号的生成模型;其次,介绍了每种调制类型假设的似然函数值计算方案;最后,介绍了该信道环境下信道参数盲估计算法,分析了 ECM 算法的缺点,提出了 SQUAREM-PC 算法,并通过仿真实验验证所提出算法性能的优越性。

第四章: 频率选择性衰落高斯噪声信道下接收信号似然函数计算

首先介绍了频率选择性衰落高斯噪声信道下接收信号的生成模型；其次，基于信号模型，介绍了接收信号似然函数值的计算方案：枚举法、前向迭代算法以及基于前向粒子滤波的近似计算，从算法复杂度方面比较了这三种算法。

第五章：频率选择性衰落高斯噪声信道参数盲估计

基于第四章的信号模型，介绍了频率选择性衰落高斯噪声信道下参数盲估计算法：EM-FB 算法、EM-JTFS 算法，基于这两种算法仍存在的问题，提出了 EM-BRGSE 算法，并介绍了降低 RGSE 算法估计方差的方案。最后通过仿真实验比较 EM-FB、EM-JTFS 和 EM-BRGSE 三种算法在信道估计准确度和信道估计速度方面的性能。同时，结合第四章中的似然函数近似计算方案，比较基于 EM-BRGSE 算法和 EM-JTFS 算法的最大似然调制分类器的分类准确度和分类速度。

第六章：总结与展望

总结了本文研究所取得的成果，并分析研究中仍存在的问题，为下一步的研究指明方向。

本文的主要创新点总结如下：

- 1) 在平坦衰落非高斯噪声信道下，由于采用 ECM 算法进行信道参数盲估计时存在收敛速度较慢的问题，提出了基于 SQUAREM 算法的信道参数盲估计算法。因为基本的 SQUAREM 算法无法保证参数的每一次更新值都满足其限制条件，所以给 SQUAREM 算法添加参数条件检验这一步骤，每当参数更新值不满足其限定条件，便舍弃该更新值，采用 ECM 算法来获得这次迭代的更新值。基于这一核心思想，提出了 SQUAREM-PC 算法，相较于 ECM 算法，其可在不丢失信道估计准确度的前提下，加快信道估计的速度。
- 2) 频率选择性衰落高斯噪声信道下，利用前向迭代算法计算接收信号似然函数值存在当马尔科夫链的状态空间较大时，算法复杂度会非常高的问题。因此本文采用了一种基于前向粒子滤波的接收信号似然函数值近似计算方案，有效地解决了上述问题。
- 3) 在频率选择性衰落高斯噪声信道参数盲估计环节，当马尔科夫链的状态空间较大时，EM-FB 算法复杂度同样非常高，而目前尝试解决这一问题的 EM-JTFS 算法又存在着信道估计准确度受信道类型影响较大的问题，所以本

文提出了一种基于 RGSE 的多径信道盲估计方案：EM-BRGSE 算法，该算法可以有效地降低信道参数盲估计的复杂度；易于并行化实现的特性可以进一步加快信道参数盲估计的速度；并且信道估计准确度受信道类型影响较小，在最小相位或最大相位信道下均接近 EM-FB 算法。

第二章 数学基础

由于采用 HLRT 的框架进行不同信道条件下的调制分类研究,所以每种信道条件下的调制分类大致分为信道建模、接收信号似然函数计算以及信道参数盲估计这三个步骤。在信道参数盲估计这一环节,采用最大似然估计的框架对信号模型中未知参数进行估计。由于接收信号似然函数值对未知参数直接求导的复杂性,这里选择采用 EM 算法进行未知参数最大似然解的求解。在频率选择性衰落信道条件下,利用隐马尔科夫模型对接收信号进行建模。由于当隐马尔科夫模型的状态空间很大时,EM 算法求期望这一步算法复杂度会很高,所以这里利用串行蒙特卡洛方法进行近似计算,来降低 EM 算法求期望步骤的复杂度。本章重点介绍 EM 算法、串行蒙特卡洛方法、隐马尔科夫模型以及粒子滤波算法(串行蒙特卡洛方法的一个特例),为后面的研究打下坚实的基础。

2.1 EM 算法

EM 算法是一类通过迭代进行最大似然估计的优化算法,通常作为牛顿迭代法的替代用于对包含隐变量或缺失数据的概率模型进行参数估计。下面我们介绍 EM 算法具体实现步骤以及收敛特性。

假设对于非完整随机数据(观察数据) \mathcal{X} , 存在未观察到的随机数据 \mathcal{Y} 使得 $\mathcal{Z} = (\mathcal{X}, \mathcal{Y})$ 可以被看作是完整随机数据。这里的非完整和完整的概念根据实际的应用场景来确定。

为了方便 EM 算法的数学表达,我们用 $f(\mathbf{z}; \boldsymbol{\theta}) = f(\mathbf{x}, \mathbf{y}; \boldsymbol{\theta})$ 表示随机变量 \mathcal{Z} 和 \mathcal{Y} 的联合概率密度函数, $g(\mathbf{x}; \boldsymbol{\theta})$ 表示随机变量 \mathcal{X} 的边缘概率密度函数,同时, $k(\mathbf{y}; \boldsymbol{\theta} | \mathbf{x})$ 表示在给定 $\mathcal{X} = \mathbf{x}$ 条件下, \mathcal{Y} 的条件概率密度函数。EM 算法的目标是从参数空间中选出使非完整数据的对数似然函数值(也就是 $\ln g(\mathbf{x}; \boldsymbol{\theta})$)最大化的参数值 $\hat{\boldsymbol{\theta}}$ 。其主要通过两个步骤来实现:求期望和最大化。求期望是指在给定观

察数据 \mathcal{X} 和当前参数估计值 θ^c 的条件下, 求完整数据的对数似然函数值的期望, 数学表达式如公式 (2.1) 所示。

$$Q(\theta; \theta^c) = E[\ln f(\mathbf{x}, \mathbf{y}; \theta) | \mathbf{x}, \theta^c] \quad \dots (2.1)$$

这里 $E[\cdot]$ 表示求期望。最大化是指寻找 θ^+ 最大化公式 (2.1), 其数学表达式如公式 (2.2) 所示。

$$\theta^+ = \arg \max_{\theta} Q(\theta; \theta^c) \quad \dots (2.2)$$

EM 算法的迭代过程可以归纳为:

- 1) 初始化参数值 θ^c ;
- 2) 求期望: 计算 $Q(\theta; \theta^c)$;
- 3) 最大化: 计算 $\theta^+ = \arg \max_{\theta} Q(\theta; \theta^c)$;
- 4) 将 θ^+ 作为新的当前参数估计值: $\theta^c = \theta^+$;
- 5) 重复步骤 2) 3) 4) 直到满足一定的收敛条件。

EM 算法的基本步骤如前面所述, 下面证明该算法的收敛性。正如前文所述, EM 算法的目的是最大化非完整数据对数似然函数, 这里非完整数据对数似然函数我们用 $L(\theta) \triangleq \ln g(\mathbf{x}; \theta)$ 来表示。根据贝叶斯定理, 可以得到公式 (2.3)。

$$f(\mathbf{x}, \mathbf{y}; \theta) = k(\mathbf{y}; \theta | \mathbf{x}) g(\mathbf{x}; \theta) \quad \dots (2.3)$$

对公式 (2.3) 两边求对数可得公式 (2.4)。

$$\ln g(\mathbf{x}; \theta) = \ln f(\mathbf{x}, \mathbf{y}; \theta) - \ln k(\mathbf{y}; \theta | \mathbf{x}) \quad \dots (2.4)$$

利用期望运算的线性性质, 在给定观察数据 \mathbf{x} 和参数空间中的任意参数值 θ' 条件下, 在公式 (2.4) 的等号两边分别对未观察数据 \mathbf{y} 求期望可以得到公式 (2.5)。

$$E[\ln g(\mathbf{x}; \theta) | \mathbf{x}, \theta'] = E[\ln f(\mathbf{x}, \mathbf{y}; \theta) | \mathbf{x}, \theta'] - E[\ln k(\mathbf{y}; \theta | \mathbf{x}) | \mathbf{x}, \theta'] \quad \dots (2.5)$$

由于公式 (2.5) 等号的左边项可以看作是和 \mathbf{y} 无关的常数项, 所以我们可以得到公式 (2.6)。

$$\begin{aligned}\ln g(\mathbf{x}; \boldsymbol{\theta}) &= E[\ln f(\mathbf{x}, \mathbf{y}; \boldsymbol{\theta}) | \mathbf{x}, \boldsymbol{\theta}'] - E[\ln k(\mathbf{y}; \boldsymbol{\theta} | \mathbf{x}) | \mathbf{x}, \boldsymbol{\theta}'] \\ \Rightarrow L(\boldsymbol{\theta}) &= Q(\boldsymbol{\theta}; \boldsymbol{\theta}') - H(\boldsymbol{\theta}; \boldsymbol{\theta}')\end{aligned}\quad \dots (2.6)$$

其中 $H(\boldsymbol{\theta}; \boldsymbol{\theta}') \triangleq E[\ln k(\mathbf{y}; \boldsymbol{\theta} | \mathbf{x}) | \mathbf{x}, \boldsymbol{\theta}']$, $Q(\boldsymbol{\theta}; \boldsymbol{\theta}')$ 的定义如公式 (2.1) 所示。至此, 我们建立了 $L(\boldsymbol{\theta})$ 和 $Q(\boldsymbol{\theta}; \boldsymbol{\theta}')$ 之间的关系。

基于公式 (2.6) 发现如果满足下面的两个条件, EM 算法将具有收敛属性:

(1) $Q(\boldsymbol{\theta}^+; \boldsymbol{\theta}^c) \geq Q(\boldsymbol{\theta}^c; \boldsymbol{\theta}^c)$, (2) $H(\boldsymbol{\theta}^+; \boldsymbol{\theta}^c) \leq H(\boldsymbol{\theta}^c; \boldsymbol{\theta}^c)$, 这里 $\boldsymbol{\theta}^c$ 表示当前参数估计值, $\boldsymbol{\theta}^+$ 的定义如公式 (2.2) 所示。

从 $\boldsymbol{\theta}^+$ 的定义来看, 条件 (1) 是满足的。为了证明条件 (2) 成立, 我们用 $H(\boldsymbol{\theta}^+; \boldsymbol{\theta}^c) - H(\boldsymbol{\theta}^c; \boldsymbol{\theta}^c)$ 可得公式 (2.7)。

$$\begin{aligned}H(\boldsymbol{\theta}^+; \boldsymbol{\theta}^c) - H(\boldsymbol{\theta}^c; \boldsymbol{\theta}^c) &= E[\ln k(\mathbf{y}; \boldsymbol{\theta}^+ | \mathbf{x}) | \mathbf{x}, \boldsymbol{\theta}^c] - E[\ln k(\mathbf{y}; \boldsymbol{\theta}^c | \mathbf{x}) | \mathbf{x}, \boldsymbol{\theta}^c] \\ &= \int_{\mathbf{y}} \left[\ln \frac{k(\mathbf{y}; \boldsymbol{\theta}^+ | \mathbf{x})}{k(\mathbf{y}; \boldsymbol{\theta}^c | \mathbf{x})} \right] k(\mathbf{y}; \boldsymbol{\theta}^c | \mathbf{x}) d\mathbf{y}\end{aligned}\quad \dots (2.7)$$

根据琴生不等式[79], 可以得到如公式 (2.8) 所示的不等式。

$$\int_{\mathbf{y}} \left[\ln \frac{k(\mathbf{y}; \boldsymbol{\theta}^+ | \mathbf{x})}{k(\mathbf{y}; \boldsymbol{\theta}^c | \mathbf{x})} \right] k(\mathbf{y}; \boldsymbol{\theta}^c | \mathbf{x}) d\mathbf{y} \leq \ln \int_{\mathbf{y}} \frac{k(\mathbf{y}; \boldsymbol{\theta}^+ | \mathbf{x})}{k(\mathbf{y}; \boldsymbol{\theta}^c | \mathbf{x})} k(\mathbf{y}; \boldsymbol{\theta}^c | \mathbf{x}) d\mathbf{y} \quad \dots (2.8)$$

因为不等式 (2.8) 中不等号的右边项等于 0, 所以可以得到条件 (2) 成立。根据公式 (2.6) 和条件 (1) 和 (2) 可以得出 $L(\boldsymbol{\theta}^+) \geq L(\boldsymbol{\theta}^c)$, 也就是说, EM 算法的每一次迭代都能够保证非完整数据的似然函数值不变小, 所以 EM 算法最终会收敛到非完整数据对数似然函数的局部最大值。

2.2 串行蒙特卡洛方法

串行蒙特卡洛方法是一种广义的蒙特卡洛采样方法, 其依次对目标概率分布序列 $\pi_1(x_1), \pi_2(x_{1:2}), \dots, \pi_n(x_{1:n}), \dots$ 进行采样, 其中 $\pi_n(x_{1:n})$ 是定义在空间 \mathbf{X}^n 上的概率密度函数, 可表示为公式 (2.9) 所示的形式。

$$\pi_n(x_{1:n}) = \frac{\gamma_n(x_{1:n})}{Z_n} \quad \dots (2.9)$$

这里仅仅要求函数 $\gamma_n : X^n \rightarrow \mathbb{R}^+$ 在定义域内的每一个值是已知的。归一化常数 Z_n 的表达式如公式 (2.10) 所示,

$$Z_n = \int \gamma_n(x_{1:n}) dx_{1:n} \quad \dots (2.10)$$

其值可以未知。串行蒙特卡洛方法在时刻 1 得出 $\pi_1(x_1)$ 的近似分布以及 Z_1 的估计值, 在时刻 2 得出 $\pi_2(x_{1:2})$ 的近似分布以及 Z_2 的近似值, 依此类推。

为了介绍串行蒙特卡洛方法, 首先说明对于特定的 n 时, 如何用蒙特卡洛方法来近似概率分布 $\pi_n(x_{1:n})$ 。假设从分布 $\pi_n(x_{1:n})$ 中独立采样出 N 个样本点:

$X_{1:n}^i \sim \pi_n(x_{1:n}), i=1, \dots, N$, 那么利用蒙特卡洛方法可以得到分布 $\pi_n(x_{1:n})$ 的近似值如公式 (2.11) 所示。

$$\hat{\pi}_n(x_{1:n}) = \frac{1}{N} \sum_{i=1}^N \delta_{X_{1:n}^i}(x_{1:n}) \quad \dots (2.11)$$

其中 $\delta_{x_0}(x)$ 表示狄拉克函数。基于这一近似, 可以得到任何边缘分布 $\pi_n(x_k)$ 的近似值如公式 (2.12) 所示。

$$\hat{\pi}_n(x_k) = \frac{1}{N} \sum_{i=1}^N \delta_{X_k^i}(x_k) \quad \dots (2.12)$$

同时, 任何函数 $\varphi_n : X^n \rightarrow \mathbb{R}$ 的期望 $I_n(\varphi_n) \triangleq \int \varphi_n(x_{1:n}) \pi_n(x_{1:n}) dx_{1:n}$ 可以由公式 (2.13) 近似计算。

$$\begin{aligned} I_n^{\text{MC}}(\varphi_n) &= \int \varphi_n(x_{1:n}) \hat{\pi}_n(x_{1:n}) dx_{1:n} \\ &= \frac{1}{N} \sum_{i=1}^N \varphi_n(X_{1:n}^i) \end{aligned} \quad \dots (2.13)$$

不难得出 $I_n^{\text{MC}}(\varphi_n)$ 是 $I_n(\varphi_n)$ 的无偏估计, 并且可以得出 $I_n^{\text{MC}}(\varphi_n)$ 的方差如公式 (2.14) 所示。

$$\text{var}[I_n^{\text{MC}}(\varphi_n)] = \frac{1}{N} \left(\int \varphi_n^2(x_{1:n}) \pi_n(x_{1:n}) dx_{1:n} - I_n^2(\varphi_n) \right) \quad \dots (2.14)$$

蒙特卡洛方法的主要优势在于不管空间 X^n 的维度大小, 近似误差的方差以

$O(1/N)$ 的速率减小。然而, 关于蒙特卡洛方法, 有两个问题需要解决:

1) 如果 $\pi_n(x_{1:n})$ 是复杂的高维概率分布, 将无法直接从该分布中进行采样。

2) 即使知道如何精确地从 $\pi_n(x_{1:n})$ 中进行采样, 采样算法的复杂度也至少与变量数 n 成正比。所以当依次从 $\pi_1(x_1), \pi_2(x_{1:2}), \dots, \pi_n(x_{1:n}), \dots$ 分布中进行采样时, 算法的复杂度至少会随着 n 的增加而呈线性增加。

首先, 利用重要性采样算法来解决问题 1)。重要性采样算法依赖引入满足条件: $\pi_n(x_{1:n}) > 0 \Rightarrow q_n(x_{1:n}) > 0$ 的重要性分布 $q_n(x_{1:n})$ 来解决不易从分布 $\pi_n(x_{1:n})$ 中采样的问题。基于重要性分布, 可以将公式 (2.9) 和 (2.10) 改写成公式 (2.15) 和 (2.16) 的形式。

$$\pi_n(x_{1:n}) = \frac{w_n(x_{1:n})q_n(x_{1:n})}{Z_n} \quad \dots (2.15)$$

$$Z_n = \int w_n(x_{1:n})q_n(x_{1:n})dx_{1:n} \quad \dots (2.16)$$

其中 $w_n(x_{1:n}) = \gamma_n(x_{1:n}) / q_n(x_{1:n})$ 表示未归一化权重函数。

这里可以选择易于采样的重要性分布 $q_n(x_{1:n})$, 从 $q_n(x_{1:n})$ 中独立采样 N 个样本点: $X_{1:n}^i \sim q_n(x_{1:n}), i=1, \dots, N$, 然后将 $q_n(x_{1:n})$ 的蒙特卡洛近似分布带入到公式 (2.15) 和 (2.16) 中可以得到 $\pi_n(x_{1:n})$ 和 Z_n 的近似值如公式 (2.17) 和 (2.18) 所示。

$$\hat{\pi}_n(x_{1:n}) = \sum_{i=1}^N W_n^i \delta_{X_{1:n}^i}(x_{1:n}) \quad \dots (2.17)$$

$$\hat{Z}_n = \frac{1}{N} \sum_{i=1}^N w_n(X_{1:n}^i) \quad \dots (2.18)$$

其中 $W_n^i = w_n(X_{1:n}^i) / \sum_{j=1}^N w_n(X_{1:n}^j)$ 。

重要性采样算法可以得到归一化常数 Z_n 的无偏估计, 并且估计器的相对方差如公式 (2.19) 所示。

$$\frac{\text{var}[\hat{Z}_n]}{Z_n^2} = \frac{1}{N} \left(\int \frac{\pi_n^2(x_{1:n})}{q_n(x_{1:n})} dx_{1:n} - 1 \right) \quad \dots (2.19)$$

同时, 利用重要性采样算法可以得到 $I_n(\varphi_n)$ 的估计值如公式 (2.20) 所示。

$$I_n^{\text{IS}}(\varphi_n) = \sum_{i=1}^N W_n^i \varphi_n(X_{1:n}^i) \quad \dots (2.20)$$

不同于 $I_n^{\text{MC}}(\varphi_n)$ ，对于有限的采样点， $I_n^{\text{IS}}(\varphi_n)$ 是 $I_n(\varphi_n)$ 的有偏估计。但是其具有一致性，并且渐进偏差如公式 (2.21) 所示。

$$\lim_{N \rightarrow \infty} N(I_n^{\text{IS}}(\varphi_n) - I_n(\varphi_n)) = - \int \frac{\pi_n^2(x_{1:n})}{q_n(x_{1:n})} (\varphi_n(x_{1:n}) - I_n(\varphi_n)) dx_{1:n} \quad \dots (2.21)$$

进一步， $I_n^{\text{IS}}(\varphi_n)$ 满足以渐进方差为公式 (2.22) 的中心极限定理。

$$\frac{1}{N} \int \frac{\pi_n^2(x_{1:n})}{q_n(x_{1:n})} (\varphi_n(x_{1:n}) - I_n(\varphi_n))^2 dx_{1:n} \quad \dots (2.22)$$

其次，寻找合适的算法使得在每个时刻采样的复杂度保持恒定来解决问题 2)。这里只要限制重要性分布满足公式 (2.23) 所示的结构便可。

$$\begin{aligned} q_n(x_{1:n}) &= q_{n-1}(x_{1:n-1}) q_n(x_n | x_{1:n-1}) \\ &= q_1(x_1) \prod_{k=2}^n q_k(x_k | x_{1:k-1}) \end{aligned} \quad \dots (2.23)$$

这就意味着，要想在时刻 n 采样 $X_{1:n}^i \sim q_n(x_{1:n})$ ，只需要在时刻 1 采样 $X_1^i \sim q_1(x_1)$ ，然后在时刻 $k = 2, \dots, n$ 采样 $X_k^i \sim q_k(x_k | X_{1:k-1}^i)$ 即可。未归一化权重函数可以由公式 (2.24) 的分解迭代计算。

$$\begin{aligned} w_n(x_{1:n}) &= \frac{\gamma_n(x_{1:n})}{q_n(x_{1:n})} \\ &= \frac{\gamma_{n-1}(x_{1:n-1})}{q_{n-1}(x_{1:n-1})} \frac{\gamma_n(x_{1:n})}{\gamma_{n-1}(x_{1:n-1}) q_n(x_n | x_{1:n-1})} \end{aligned} \quad \dots (2.24)$$

因此，未归一化权重函数可以写成公式 (2.25) 所示的形式。

$$\begin{aligned} w_n(x_{1:n}) &= w_{n-1}(x_{1:n-1}) \alpha_n(x_{1:n}) \\ &= w_1(x_1) \prod_{k=2}^n \alpha_k(x_{1:k}) \end{aligned} \quad \dots (2.25)$$

其中权重增量函数 $\alpha_n(x_{1:n})$ 的表达式如公式 (2.26) 所示。

$$\alpha_n(x_{1:n}) = \frac{\gamma_n(x_{1:n})}{\gamma_{n-1}(x_{1:n-1}) q_n(x_n | x_{1:n-1})} \quad \dots (2.26)$$

我们称重要性分布满足公式 (2.23)，依次进行重要性采样的算法为串行重

要性采样算法，其算法流程如算法 2.1 所示。

算法 2.1 串行重要性采样算法

1: 在时刻 $n=1$ 时

采样 $X_1^i \sim q_1(x_1), i=1, \dots, N$;

计算权重 $w_1(X_1^i)$ 以及 $W_1^i \propto w_1(X_1^i)$ 。

2: 在时刻 $n \geq 2$ 时

采样 $X_n^i \sim q_n(x_n | X_{1:n-1}^i), i=1, \dots, N$;

计算权重 $w_n(X_{1:n}^i) = w_{n-1}(X_{1:n-1}^i) \alpha_n(X_{1:n}^i)$ 以及 $W_n^i \propto w_n(X_{1:n}^i)$ 。

在任何时刻 n ，根据算法 2.1 采样到的样本点，结合公式 (2.17) 和 (2.18) 可以得到 $\pi_n(x_{1:n})$ 和 Z_n 的近似值。

在一些场景中，如果从 $q_n(x_n | x_{1:n-1})$ 进行采样以及计算 $\alpha_n(x_{1:n})$ 所需要的时间均与 n 无关时，串行重要性采样算法便解决了上述的问题 2)。但是串行重要性采样算法存在着一个致命的缺陷：对于标准的重要性采样算法，估计结果的方差随着 n 的增大呈指数增大，而串行重要性采样算法仅仅是要求重要性分布满足公式 (2.23) 的特殊的重要性采样算法，所以串行重要性采样算法依旧存在着估计值方差随着 n 的增大呈指数增大的问题。

重采样技术作为串行蒙特卡洛方法的重要组成部分可以有效地解决上述问题。虽然重采样是一个直觉意义上的概念，但是它却有着巨大的理论和实用价值。首先来看一下重要性采样算法对目标分布 $\pi_n(x_{1:n})$ 的近似分布 $\hat{\pi}_n(x_{1:n})$ ，这一近似是基于对重要性分布 $q_n(x_{1:n})$ 中的样本点进行加权而得来，而并没有获得近似分布于 $\pi_n(x_{1:n})$ 的样本点。要想获得近似分布于 $\pi_n(x_{1:n})$ 的样本点，可以对重要性采样算法获得的近似分布 $\hat{\pi}_n(x_{1:n})$ 采样，也就是依据概率 W_n^i 采样 $X_{1:n}^i$ 。这一操作便称作重采样，因为采样的对象 $\hat{\pi}_n(x_{1:n})$ 本身也是通过采样而得。从 $\hat{\pi}_n(x_{1:n})$ 中采样出 N 个样本点，在得到的新的样本中， $X_{1:n}^i$ 出现了 N_n^i 次，其中 $\{N_n^i\}_{i=1}^N$ 服从以 $(N, \{W_n^i\}_{i=1}^N)$ 为参数的多项分布，并且新的样本中每个样本点的权重均为 $1/N$ 。这

样，可以利用重采样获得的新的样本点来近似 $\hat{\pi}_n(x_{1:n})$ ，如公式 (2.27) 所示。

$$\bar{\pi}_n(x_{1:n}) = \sum_{i=1}^N \frac{N_n^i}{N} \delta_{X_{1:n}^i}(x_{1:n}) \quad \dots (2.27)$$

由于 $E[N_n^i | \{W_n^i\}_{i=1}^N] = N W_n^i$ ，所以 $\bar{\pi}_n(x_{1:n})$ 是 $\hat{\pi}_n(x_{1:n})$ 的无偏估计。

重采样技术的一个非常重要的优势在于：它可以以较大概率去除掉权重较小的样本点。在迭代近似的框架中，我们不希望利用权重较小的样本点进行迭代更新，而希望将计算资源集中在权重较大的样本点上。

串行蒙特卡洛方法就是串行重要性采样和重采样技术的结合。在时刻 1，利用重要性采样算法获得 $\pi_1(x_1)$ 的近似值 $\hat{\pi}_1(x_1)$ ，然后利用重采样技术得到 N 个等概率的重采样样本，记为 $\{1/N, \bar{X}_1^i\}$ 。之后，在时刻 2，根据串行重要性采样算法，采样 $X_2^i \sim q_2(x_2 | \bar{X}_1^i)$ ，所以样本 (\bar{X}_1^i, X_2^i) 近似分布于 $\pi_1(x_1)q_2(x_2 | x_1)$ 。根据公式 (2.24) 可以得到重要性权重 $w_2(x_{1:2})$ 等于权重增量 $\alpha_2(x_{1:2})$ 。之后再根据归一化权重值进行重采样，以此类推便是串行蒙特卡洛方法。该方法的步骤如算法 2.2 所示。

算法 2.2 串行蒙特卡洛方法

1: 在时刻 $n=1$ 时

采样 $X_1^i \sim q_1(x_1), i=1, \dots, N$;

计算权重 $w_1(X_1^i)$ 以及 $W_1^i \propto w_1(X_1^i)$;

重采样样本及权重集合 $\{W_1^i, X_1^i\}_{i=1}^N$ 得到 N 个等概率样本集合 $\{1/N, \bar{X}_1^i\}_{i=1}^N$ 。

2: 在时刻 $n \geq 2$ 时

采样 $X_n^i \sim q_n(x_n | \bar{X}_{1:n-1}^i)$ 以及设置 $X_{1:n}^i = (\bar{X}_{1:n-1}^i, X_n^i)$ ，其中 $i=1, \dots, N$;

计算权重 $\alpha_n(X_{1:n}^i)$ 以及 $W_n^i \propto \alpha_n(X_{1:n}^i)$;

重采样样本及权重集合 $\{W_n^i, X_{1:n}^i\}_{i=1}^N$ 得到 N 个等概率样本集合 $\{1/N, \bar{X}_{1:n}^i\}_{i=1}^N$ 。

2.3 隐马尔科夫模型和粒子滤波

首先考虑一个满足 $X_1 \sim \mu(x_1)$ 以及 $X_n | (X_{n-1} = x_{n-1}) \sim f(x_n | x_{n-1})$ 的离散时间

马尔科夫过程 $\{X_n\}_{n \geq 1}$ ，其中 $\mu(x_1)$ 表示概率密度函数， $f(x|x')$ 表示从状态 x' 转移到状态 x 的概率密度函数。一般情况下，主要研究在给定观察信号过程 $\{Y_n\}_{n \geq 1}$ 的条件下对 $\{X_n\}_{n \geq 1}$ 的估计问题。这里假设在给定 $\{X_n\}_{n \geq 1}$ 条件下， $\{Y_n\}_{n \geq 1}$ 是统计独立的，并且边缘概率密度函数 $Y_n | (X_n = x_n) \sim g(y_n | x_n)$ 。这样的信号模型便是隐马尔科夫模型。

上述描述定义了一个贝叶斯概率模型，其中 $\{X_n\}_{n \geq 1}$ 的先验概率密度以及 $\{Y_n\}_{n \geq 1}$ 的似然函数分别如公式 (2.28) 和 (2.29) 所示。

$$p(x_{1:n}) = \mu(x_1) \prod_{k=2}^n f(x_k | x_{k-1}) \quad \dots (2.28)$$

$$p(y_{1:n} | x_{1:n}) = \prod_{k=1}^n g(y_k | x_k) \quad \dots (2.29)$$

在该贝叶斯模型下，在给定 $\{Y_n\}_{n \geq 1}$ 条件下，推理 $\{X_n\}_{n \geq 1}$ 主要依靠公式 (2.30) 所示的后验概率。

$$p(x_{1:n} | y_{1:n}) = \frac{p(x_{1:n}, y_{1:n})}{p(y_{1:n})} \quad \dots (2.30)$$

其中 $p(x_{1:n}, y_{1:n}) = p(y_{1:n} | x_{1:n}) p(x_{1:n})$ ， $p(y_{1:n}) = \int p(x_{1:n}, y_{1:n}) dx_{1:n}$ 。

对于有限状态空间隐马尔科夫模型，上述的积分运算可以用有限项求和来替代，因此所有的概率值都可以被精确计算。对于线性高斯模型，很容易发现后验概率 $p(x_{1:n} | y_{1:n})$ 同样也服从高斯分布，可以用卡尔曼技术[80]求解其均值和方差。但是对于大多数非线性非高斯模型，无法精确地计算相关概率密度值，需要利用数值方法来近似计算。粒子方法是一类灵活有效的基于仿真的方法，可以得到后验概率分布 $p(x_{1:n} | y_{1:n})$ 和归一化常数 $p(y_{1:n})$ 的近似值，它们隶属于串行蒙特卡洛方法。

当我们想要依次近似概率分布 $\{p(x_{1:n} | y_{1:n})\}_{n \geq 1}$ 以及归一化常数 $\{p(y_{1:n})\}_{n \geq 1}$ 时，可以将公式 (2.9) 中的 $\gamma_n(x_{1:n})$ 设置为 $p(x_{1:n}, y_{1:n})$ ， Z_n 设置为 $p(y_{1:n})$ ，然后利用

串行蒙特卡洛方法进行近似计算,这种串行蒙特卡洛方法的特例便是粒子滤波算法。

2.4 本章小结

本章主要介绍了 EM 算法的实现步骤,证明了 EM 算法的收敛特性,介绍了串行蒙特卡洛方法、隐马尔科夫模型的定义以及粒子滤波算法,为下面的章节打好基础。

第三章 平坦衰落非高斯噪声信道下的最大似然调制分类器

大量实验表明实际的通信信道中存在着自然噪声和人为噪声,这两种噪声的结合呈现出明显的非高斯特性。传统的基于高斯噪声假设下得到的调制分类器的性能在非高斯噪声下性能会大幅度下降,因此本章重点研究非高斯噪声信道下的调制分类算法。由于基于似然比检验的调制分类算法在分类准确度方面性能最优,所以本章主要研究非高斯噪声信道下的最大似然调制分类算法。按照最大似然调制分类算法的一般框架,本章从非高斯噪声信道建模、接收信号似然函数计算以及信道参数盲估计这三个方面着手研究。

3.1 信号模型

用 \mathbf{M} 表示发射机采用的线性调制星座点集合, $m=|\mathbf{M}|$ 表示调制星座点数。从 \mathbf{M} 中均匀采样出发射符号,这里用 $\{a_i\}_{i \geq 1}$ 来表示发射符号。经过预处理后,在匹配滤波器输出端以符号速率采样可以获得接收信号的等效复基带形式,如公式 (3.1) 所示。

$$y_i = \beta a_i + \varepsilon_i \quad \dots (3.1)$$

其中, β 是一个复数,其包括了平坦衰落信道以及发射信号的功率和载波相位对于接收信号的影响,这里我们统称 β 为信道系数。假设 $\{\varepsilon_i\}_{i \geq 1}$ 是独立同分布的复随机变量,表示信道中的加性非高斯噪声。由于混合高斯分布具有较好的近似 Middleton's Class A 分布[81][82]和其他的对称分布[83]的特性,所以选择混合高斯分布来表示 ε_i 的概率分布。这里我们假设 ε_i 的实部和虚部独立同分布。将 ε_i 看作是由其实部和虚部构成的二维向量,其概率密度分布表达式如公式 (3.2) 所示。

$$p(\mathbf{x}) = \sum_{n=1}^N \frac{\lambda_n}{2\pi |\boldsymbol{\Sigma}_n|^{1/2}} \exp\left(-\frac{1}{2} \mathbf{x}^T \boldsymbol{\Sigma}_n^{-1} \mathbf{x}\right) \quad \dots (3.2)$$

其中 λ_n 表示 ε_i 服从混合高斯分布中第 n 个成份所表示的高斯分布的概率大小,满

足限制条件： $0 \leq \lambda_n \leq 1$ 和 $\sum_{n=1}^N \lambda_n = 1$ 。 $\Sigma_n = (\sigma_n^2 / 2) \mathbf{I}_{2 \times 2}$ 表示混合高斯分布中第 n 个成份所表示的高斯分布的协方差矩阵，其中 σ_n^2 表示该成分的平均能量。

3.2 最大似然调制分类器

在平坦衰落非高斯噪声信道模型中，由于发射符号集合 $\{a_t\}_{t=1}^T$ 中的每一个元素独立同分布，噪声集合 $\{\varepsilon_t\}_{t=1}^T$ 中的每一个元素也是独立同分布，并且 a_t 和 ε_t 相互独立，所以在给定调制类型假设 H_i 下，接收信号的条件概率可由公式 (3.3) 计算得到。

$$p(y_{1:T}; \boldsymbol{\theta} | H_i) = \prod_{t=1}^T \sum_{a_t} \frac{1}{m_i} p(y_t; \boldsymbol{\theta} | a_t) \quad \dots (3.3)$$

这里 $\boldsymbol{\theta}$ 表示信道参数构成的向量， m_i 表示假设 H_i 所对应的调制类型 M_i 的调制阶数。根据 3.1 节中的信号模型， $p(y_t; \boldsymbol{\theta} | a_t)$ 的计算方式如公式 (3.4) 所示。

$$p(y_t; \boldsymbol{\theta} | a_t) = \sum_{n=1}^N \frac{\lambda_n}{\pi \sigma_n^2} \exp\left(-\frac{|y_t - \beta a_t|^2}{\sigma_n^2}\right) \quad \dots (3.4)$$

其中 $\lambda_n, \sigma_n^2, \beta$ 的定义与 3.1 节信号模型中的定义一致。

由于在实际通信过程中，无法预先知道信道中的相关参数，所以无法直接利用公式 (3.4) 计算接收信号的似然函数值。这里采用混合假设检验机制[85]来解决这一问题：在每一个调制假设 H_i 下，首先利用最大似然估计准则估计出信道参数 $\hat{\boldsymbol{\theta}}_i = \{\hat{\beta}_i, \hat{\lambda}_1^i, \dots, \hat{\lambda}_N^i, (\hat{\sigma}_1^i)^2, \dots, (\hat{\sigma}_N^i)^2\}$ ，然后将估计得来的参数带入公式 (3.3) 和 (3.4) 中计算每个假设的似然函数值，从而得到最大似然调制分类器如公式 (3.5) 所示。

$$\begin{aligned} \hat{H} &= \arg \max_{H_i} \ln p(y_{1:T}; \hat{\boldsymbol{\theta}}_i | H_i) \\ &= \arg \max_{H_i} \sum_{t=1}^T \ln \left(\sum_{a_t} \frac{1}{m_i} \sum_{n=1}^N \frac{\hat{\lambda}_n^i}{\pi (\hat{\sigma}_n^i)^2} \exp\left(-\frac{|y_t - \hat{\beta}_i a_t|^2}{(\hat{\sigma}_n^i)^2}\right) \right) \end{aligned} \quad \dots (3.5)$$

因此，为了得到最大似然调制分类器，首先需要在每个调制类型假设下，对信道参数进行估计，由于没有导频序列可以利用，所以需要进行信道参数的盲估

计。在下面一节，将详细介绍平坦衰落非高斯噪声信道下的信道参数的盲估计算法。

3.3 信道参数盲估计

最大似然估计的一般解法为：利用目标似然函数对待估计参数求导，并令导数为零，求解方程便可得到相应的最大似然估计值。由公式 (3.3) (3.4) 可得到接收信号的似然函数，因为利用其对待估计参数求导十分复杂，所以这里不适合采用一般地最大似然估计方案。EM 算法利用数值迭代的方式估计最大似然解，研究发现利用该方法可使得这一场景下的最大似然估计变得十分简单。

EM 算法通过添加隐藏数据，与现有的观察数据构成完整数据。在某一些场合下，观察数据的似然函数对参数求导十分复杂，而完整数据的似然函数对参数求导却十分简单，这时，可以利用 EM 算法来进行参数的最大似然估计。

下面我们将具体介绍在调制类型假设 H_i 下，采用期望最大化算法进行信道参数盲估计的具体过程。为了表达简洁，省略参数估计值的下标 i 。

针对平坦衰落非高斯噪声信道下的信道参数盲估计这一具体问题，利用 EM 算法进行最大似然估计的第一步便是定义隐藏数据。首先定义变量 $\{z_t\}_{t=1}^T$ ， z_t 的值表示接收信号 y_t 的噪声由混合高斯模型中的第 z_t 个成份所表示的高斯噪声建模，所以 $z_t \in \{1, \dots, N\}$ 。基于上述变量的定义，将隐藏数据定义为 $\mathbf{h} = \{a_t, z_t\}_{t=1}^T$ ，观察数据定义为 $\mathbf{o} = \{y_t\}_{t=1}^T$ ，所以完整数据 $\mathbf{c} = \{\mathbf{h}, \mathbf{o}\} = \{a_t, z_t, y_t\}_{t=1}^T$ 。定义好隐藏数据和观察数据后，便可根据 EM 算法的标准步骤进行相关量的表示。这里我们用 $\boldsymbol{\theta} = \{\beta, \lambda_1, \dots, \lambda_N, \sigma_1^2, \dots, \sigma_N^2\}$ 表示信号模型中待估计的相关参数集合， $\boldsymbol{\theta}^p = \{\beta^p, \lambda_1^p, \dots, \lambda_N^p, (\sigma_1^p)^2, \dots, (\sigma_N^p)^2\}$ 表示当前参数估计值。利用 EM 算法求解 $\boldsymbol{\theta}$ 的最大似然估计值的具体步骤如下文所述。

1) 求期望。首先要求在给定观察数据 \mathbf{o} 和当前参数估计值 $\boldsymbol{\theta}^p$ 的条件下，完整数据 \mathbf{c} 的对数似然函数的期望，该期望的具体数学表达如公式 (3.6) 所示。

$$\begin{aligned} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^p | H_i) &= E[\ln p(\mathbf{c}; \boldsymbol{\theta} | H_i) | \mathbf{o}, \boldsymbol{\theta}^p, H_i] \\ &= \sum_{\mathbf{h}} \ln(p(\mathbf{c}; \boldsymbol{\theta} | H_i)) P(\mathbf{h}; \boldsymbol{\theta}^p | \mathbf{o}, H_i) \end{aligned} \quad \dots (3.6)$$

因为 $\{y_t\}_{t=1}^T$, $\{a_t\}_{t=1}^T$ 和 $\{z_t\}_{t=1}^T$ 均为独立同分布变量集合, 所以公式 (3.6) 中的 $\ln p(\mathbf{c}; \boldsymbol{\theta} | \mathbf{H}_i)$ 可以写为公式 (3.7) 的形式。

$$\begin{aligned} \ln p(\mathbf{c}; \boldsymbol{\theta} | \mathbf{H}_i) &= \ln \prod_{t=1}^T p(y_t, a_t, z_t; \boldsymbol{\theta} | \mathbf{H}_i) \\ &= \sum_{t=1}^T \ln(p(y_t; \boldsymbol{\theta} | a_t, z_t, \mathbf{H}_i) P(a_t, z_t; \boldsymbol{\theta} | \mathbf{H}_i)) \quad \dots (3.7) \\ &= \sum_{t=1}^T \ln\left(\frac{\lambda_{z_t}}{m_i} p(y_t; \boldsymbol{\theta} | a_t, z_t, \mathbf{H}_i)\right) \end{aligned}$$

这里利用了随机变量 a_t 和 z_t 相互独立得到 $P(a_t, z_t; \boldsymbol{\theta} | \mathbf{H}_i) = P(a_t | \mathbf{H}_i) P(z_t; \boldsymbol{\theta})$, 因为发射符号 a_t 是从调制星座点集合 \mathbf{M}_i 中等概率采样得来, 所以 $P(a_t | \mathbf{H}_i) = 1/m_i$ 。

接收信号 y_t 的噪声由混合高斯模型的第 z_t 个成份建模的概率 $P(z_t; \boldsymbol{\theta}) = \lambda_{z_t}$ 。

利用贝叶斯公式, 公式 (3.6) 中的隐藏数据后验概率 $P(\mathbf{h}; \boldsymbol{\theta}^p | \mathbf{o}, \mathbf{H}_i)$ 可以表示为公式 (3.8) 所示的形式。

$$\begin{aligned} P(\mathbf{h}; \boldsymbol{\theta}^p | \mathbf{o}, \mathbf{H}_i) &= \prod_{t=1}^T P(a_t, z_t; \boldsymbol{\theta}^p | y_t, \mathbf{H}_i) \\ &= \prod_{t=1}^T \frac{P(y_t, a_t, z_t; \boldsymbol{\theta}^p | \mathbf{H}_i)}{p(y_t; \boldsymbol{\theta}^p | \mathbf{H}_i)} \quad \dots (3.8) \\ &= \prod_{t=1}^T \frac{\lambda_{z_t}^p}{m_i} \frac{P(y_t; \boldsymbol{\theta}^p | a_t, z_t, \mathbf{H}_i)}{p(y_t; \boldsymbol{\theta}^p | \mathbf{H}_i)} \end{aligned}$$

将公式 (3.7) 和公式 (3.8) 带入公式 (3.6) 中并交换求和顺序可得 $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^p | \mathbf{H}_i)$ 函数的形式如公式 (3.9) 所示。

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}^p | \mathbf{H}_i) = \sum_{t=1}^T \sum_{\mathbf{h}} \ln\left(\frac{\lambda_{z_t}}{m_i} p(y_t; \boldsymbol{\theta} | a_t, z_t, \mathbf{H}_i)\right) \prod_{\tau=1}^T \frac{\lambda_{z_\tau}^p}{m_i} \frac{P(y_\tau; \boldsymbol{\theta}^p | a_\tau, z_\tau, \mathbf{H}_i)}{p(y_\tau; \boldsymbol{\theta}^p | \mathbf{H}_i)} \quad \dots (3.9)$$

如果将公式 (3.9) 的第一个求和运算的第 l 项记为 S_l , 则 $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^p | \mathbf{H}_i) = \sum_{l=1}^T S_l$ 。

通过整理计算, S_l 可以简化为公式 (3.10) 所示的形式。

$$\begin{aligned}
 S_l &= \sum_{\{a_t\}_{t=1}^T} \sum_{\{z_t\}_{t=1}^T} \ln\left(\frac{\lambda_{z_l}}{m_i} p(y_l; \boldsymbol{\theta} | a_l, z_l, \mathbf{H}_i)\right) \prod_{\tau=1}^T P(a_\tau, z_\tau; \boldsymbol{\theta}^p | y_\tau, \mathbf{H}_i) \\
 &= \sum_{\{a_t\}_{t=1}^T} \sum_{\{z_t\}_{t=1}^T} \ln\left(\frac{\lambda_{z_l}}{m_i} p(y_l; \boldsymbol{\theta} | a_l, z_l, \mathbf{H}_i)\right) (P(a_l, z_l; \boldsymbol{\theta}^p | y_l, \mathbf{H}_i)) \\
 &\quad \times \prod_{\tau=1, \tau \neq l}^T P(a_\tau, z_\tau; \boldsymbol{\theta}^p | y_\tau, \mathbf{H}_i) \quad \dots (3.10) \\
 &= \sum_{a_l} \sum_{z_l} \ln\left(\frac{\lambda_{z_l}}{m_i} p(y_l; \boldsymbol{\theta} | a_l, z_l, \mathbf{H}_i)\right) P(a_l, z_l; \boldsymbol{\theta}^p | y_l, \mathbf{H}_i) \\
 &\quad \times \sum_{\{a_t\}_{t=1}^T - a_l} \sum_{\{z_t\}_{t=1}^T - z_l} \prod_{\tau=1, \tau \neq l}^T P(a_\tau, z_\tau; \boldsymbol{\theta}^p | y_\tau, \mathbf{H}_i)
 \end{aligned}$$

其中 $\{a_t\}_{t=1}^T - a_l$ 表示从集合 $\{a_t\}_{t=1}^T$ 中去除掉元素 a_l ，同样地， $\{z_t\}_{t=1}^T - z_l$ 表示从集合 $\{z_t\}_{t=1}^T$ 中去除掉元素 z_l 。在公式 (3.10) 中，根据全概率公式可以得到

$\sum_{\{a_t\}_{t=1}^T - a_l} \sum_{\{z_t\}_{t=1}^T - z_l} \prod_{\tau=1, \tau \neq l}^T P(a_\tau, z_\tau; \boldsymbol{\theta}^p | y_\tau, \mathbf{H}_i) = 1$ ，所以 $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^p | \mathbf{H}_i)$ 函数的形式可以简化为公式 (3.11) 所示的形式。

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}^p | \mathbf{H}_i) = \sum_{t=1}^T \sum_{a_t} \sum_{z_t} \ln\left(\frac{\lambda_{z_t}}{m_i} p(y_t; \boldsymbol{\theta} | a_t, z_t, \mathbf{H}_i)\right) \frac{\lambda_{z_t}^p}{m_i} \frac{P(y_t; \boldsymbol{\theta}^p | a_t, z_t, \mathbf{H}_i)}{p(y_t; \boldsymbol{\theta}^p | \mathbf{H}_i)} \quad \dots (3.11)$$

根据 3.1 节中的信号模型可知公式 (3.11) 中的相关 pdf 的形式如公式 (3.12) (3.13) (3.14) 所示。

$$p(y_t; \boldsymbol{\theta} | a_t, z_t, \mathbf{H}_i) = \frac{1}{\pi \sigma_{z_t}^2} \exp\left(-\frac{|y_t - \beta a_t|^2}{\sigma_{z_t}^2}\right) \quad \dots (3.12)$$

$$p(y_t; \boldsymbol{\theta}^p | a_t, z_t, \mathbf{H}_i) = \frac{1}{\pi (\sigma_{z_t}^p)^2} \exp\left(-\frac{|y_t - \beta^p a_t|^2}{(\sigma_{z_t}^p)^2}\right) \quad \dots (3.13)$$

$$p(y_t; \boldsymbol{\theta}^p | \mathbf{H}_i) = \frac{1}{m_i} \sum_{a_t} \sum_{n=1}^N \frac{\lambda_n^p}{\pi (\sigma_n^p)^2} \exp\left(-\frac{|y_t - \beta^p a_t|^2}{(\sigma_n^p)^2}\right) \quad \dots (3.14)$$

2) 最大化。 $p+1$ 时刻的参数估计值 $\boldsymbol{\theta}^{p+1}$ 可以通过在参数空间中寻找使 $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^p | \mathbf{H}_i)$ 最大化的 $\boldsymbol{\theta}$ 值得来，如公式 (3.15) 所示。

$$\boldsymbol{\theta}^{p+1} = \arg \max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^p | \mathbf{H}_i) \quad \dots (3.15)$$

由于信道系数 β 与混合高斯模型中参数的耦合关系，寻找能够同时最大化

$Q(\mathbf{\theta}, \mathbf{\theta}^p | \mathbf{H}_i)$ 的 β 与 $\{\lambda_n, \sigma_n^2\}_{n=1}^N$ 非常困难，因此，采用 EM 算法的一种变形，ECM 算法来求解公式 (3.15)。ECM 算法的求期望这一步与 EM 算法一样，仅仅将 EM 算法中复杂的最大化步骤用计算简单的条件最大化代替。

采用 ECM 算法求解公式 (3.15) 时，将待估计参数 $\mathbf{\theta}$ 分为两个集合 $\mathbf{\theta}_1 = \{\beta\}$ 和 $\mathbf{\theta}_2 = \{\lambda_n, \sigma_n^2\}_{n=1}^N$ 。为了表达简洁，用 φ_{mj}^p 表示在求期望步骤中求得的给定 y_i 和 $\mathbf{\theta}^p$ 条件下， a_i 和 z_i 的后验概率值，也就是 $\varphi_{mj}^p = P(a_i = s_j, z_i = n; \mathbf{\theta}^p | y_i, \mathbf{H}_i)$ 。其中 s_j 表示调制类型假设 H_i 所对应的调制星座点集合 \mathbf{M}_i 中第 j 个元素。

为了寻找使得 $Q(\mathbf{\theta}, \mathbf{\theta}^p | \mathbf{H}_i)$ 最大化的参数值 $\mathbf{\theta}$ ，首先保持 $\mathbf{\theta}_2$ 为当前参数估计值 $\mathbf{\theta}_2^p$ 不变，用 $Q(\mathbf{\theta}, \mathbf{\theta}^p | \mathbf{H}_i)$ 对 β 求导，并令导数为零可得 β^{p+1} 的更新如公式 (3.16) 所示。

$$\beta^{p+1} = \frac{\sum_{t=1}^T \sum_{n=1}^N \sum_{j=1}^{m_i} \frac{y_t (s_j)^*}{(\sigma_n^p)^2} \varphi_{mj}^p}{\sum_{t=1}^T \sum_{n=1}^N \sum_{j=1}^{m_i} \frac{|s_j|^2}{(\sigma_n^p)^2} \varphi_{mj}^p} \quad \dots (3.16)$$

然后保持 $\mathbf{\theta}_1$ 为 $\mathbf{\theta}_1^{p+1} = \beta^{p+1}$ 不变，用 $Q(\mathbf{\theta}, \mathbf{\theta}^p | \mathbf{H}_i)$ 分别对 λ_n 和 σ_n^2 求导，并令导数为零，求解方程便可获得 $\mathbf{\theta}_2$ 的更新值 $\mathbf{\theta}_2^{p+1}$ 。注意这里在求解 λ_n^{p+1} 时，需要考虑约束条件 $\sum_{n=1}^N \lambda_n^{p+1} = 1$ ，利用拉格朗日乘子法可以获得 λ_n^{p+1} 的更新公式 (3.17)。

$$\lambda_n^{p+1} = \frac{1}{T} \sum_{t=1}^T \sum_{j=1}^{m_i} \varphi_{mj}^p \quad \dots (3.17)$$

直接利用 $Q(\mathbf{\theta}, \mathbf{\theta}^p | \mathbf{H}_i)$ 对 σ_n^2 求导，并令导数为零可得到 $(\sigma_n^{p+1})^2$ 的计算公式 (3.18)。

$$(\sigma_n^{p+1})^2 = \frac{\sum_{t=1}^T \sum_{j=1}^{m_i} |y_t - \beta^{p+1} s_j|^2 \varphi_{mj}^p}{\sum_{t=1}^T \sum_{j=1}^{m_i} \varphi_{mj}^p} \quad \dots (3.18)$$

综上，利用 ECM 算法更新 $\mathbf{\theta}^p$ 的过程如算法 3.1 所示。虽然 ECM 算法可以稳定地收敛到对数似然函数局部最大值点，但是其收敛速度较慢。由于 ECM 算法的算法复杂度与迭代次数成正比，所以收敛速度越慢，导致 ECM 算法复杂度

越高。为了降低信道参数盲估计的算法复杂度，我们采用 SQUAREM[86]来加快 ECM 算法的收敛速度。

算法 3.1 采用 ECM 算法更新 θ^p

- 1: 输入 θ^p .
- 2: 利用公式 (3.13) (3.14) 计算 $\varphi_{mj}^p, t=1, \dots, T, n=1, \dots, N, j=1, \dots, m_i$.
- 3: 利用公式 (3.16) 计算 β^{p+1} .
- 4: 利用公式 (3.17) (3.18) 分别更新 $\lambda_n^{p+1}, (\sigma_n^{p+1})^2, n=1, \dots, N$.
- 5: 输出 θ^{p+1} .

因为 ECM 算法隐式地定义了一个映射 F ，满足 $\theta^{p+1} = F(\theta^p), p=0, 1, \dots$ ，所以 ECM 算法可以看作是求解函数 $g(\theta) = F(\theta) - \theta$ 的零点。因此，为了加速 ECM 算法的收敛速度，需要寻找更加快速的求解方程 $g(\theta) = 0$ 的解的方法。首先，定义向量序列 $\{\mathbf{u}_i\}, i \geq 0$ ，这个序列是由 ECM 算法从 $\mathbf{u}_0 = \theta^p$ 开始迭代产生的，也就是 $\mathbf{u}_{i+1} = F(\mathbf{u}_i), i \geq 0$ 。公式 (3.19) 展示了利用牛顿法求解 $g(\theta) = 0$ 的过程。

$$\theta^{p+1} = \mathbf{u}_0 - (M^p)^{-1} g(\mathbf{u}_0) \quad \dots (3.19)$$

其中 $M^p = J(\theta^p) - I$ ， $J(\theta^p)$ 是函数 $F(\theta)$ 的雅克比矩阵在 $\theta = \theta^p$ 时的值。牛顿法实际上是通过求解 $g(\theta)$ 在 $\theta = \mathbf{u}_0$ 附近的线性近似函数 $\tilde{g}(\theta) = g(\mathbf{u}_0) + M^p(\theta - \mathbf{u}_0)$ 的零点得来。对于求解标量方程的零点问题，Steffensen 法在不用计算函数的导数的条件下，可达到与牛顿法一样的收敛速度。对于求解向量方程的零点问题，Steffensen-type 法近似计算雅克比矩阵的方法为：分别在 $\theta = \mathbf{u}_0$ 和 $\theta = \mathbf{u}_1$ 处，写出 $g(\theta)$ 的线性近似函数 $g_0(\theta)$ 和 $g_1(\theta)$ 如公式 (3.20) (3.21) 所示。

$$g_0(\theta) = g(\mathbf{u}_0) + \frac{1}{\alpha^p}(\theta - \mathbf{u}_0) \quad \dots (3.20)$$

$$g_1(\theta) = g(\mathbf{u}_1) + \frac{1}{\alpha^p}(\theta - \mathbf{u}_1) \quad \dots (3.21)$$

这里用 $(1/\alpha^p)I$ 近似 $J(\boldsymbol{\theta}^p) - I$ 和 $J(F(\boldsymbol{\theta}^p)) - I$ 。函数 $g_0(\boldsymbol{\theta})$ 和 $g_1(\boldsymbol{\theta})$ 的零点如公式 (3.22) (3.23) 所示。

$$t_0^{p+1} = \mathbf{u}_0 - \alpha^p g(\mathbf{u}_0) \quad \dots (3.22)$$

$$t_1^{p+1} = \mathbf{u}_1 - \alpha^p g(\mathbf{u}_1) \quad \dots (3.23)$$

因为 t_0^{p+1} 和 t_1^{p+1} 是函数 $g(\boldsymbol{\theta})$ 零点的两个不同的近似值，所以希望 t_0^{p+1} 和 t_1^{p+1} 能尽可能地接近。这里选择 $-\|t_1^{p+1} - t_0^{p+1}\|^2 / \alpha^p$ 作为衡量两个零点接近程度的标准，限制 α^p 为负数，最小化 $-\|t_1^{p+1} - t_0^{p+1}\|^2 / \alpha^p$ 可以得到步长如公式 (3.24) 所示。

$$\alpha^p = -\frac{\|\mathbf{r}^p\|}{\|\mathbf{v}^p\|} \quad \dots (3.24)$$

其中 $\mathbf{r}^p = \mathbf{u}_1 - \mathbf{u}_0 = F(\boldsymbol{\theta}^p) - \boldsymbol{\theta}^p$ ， $\mathbf{v}^p = g(\mathbf{u}_1) - g(\mathbf{u}_0) = F(F(\boldsymbol{\theta}^p)) - 2F(\boldsymbol{\theta}^p) + \boldsymbol{\theta}^p$ 。将公式 (3.24) 的步长带入到公式 (3.22) 可以得到寻找 ECM 算法映射函数对应的方程 $F(\boldsymbol{\theta}) = \boldsymbol{\theta}$ 的解的新的迭代机制，这里称其为 STEM，如公式 (3.25) 所示。

$$\boldsymbol{\theta}^{p+1} = \boldsymbol{\theta}^p - \alpha^p (F(\boldsymbol{\theta}^p) - \boldsymbol{\theta}^p) \quad \dots (3.25)$$

定义 $F(\boldsymbol{\theta}) = \boldsymbol{\theta}$ 的解为 $\boldsymbol{\theta}^*$ ，以泰勒级数的方式在 $\boldsymbol{\theta}^*$ 处展开函数 $F(\boldsymbol{\theta})$ ，

$F^{(2)}(\boldsymbol{\theta}) (= F(F(\boldsymbol{\theta})))$ 等，并且计算展开式在 $\boldsymbol{\theta} = \boldsymbol{\theta}^p$ 处的值可得公式 (3.26)。

$$F^{(j)}(\boldsymbol{\theta}^p) = \boldsymbol{\theta}^* + \mathbf{J}^j \mathbf{e}^p + o(\mathbf{e}^p), \quad j=1, 2, \dots \quad \dots (3.26)$$

其中 $\mathbf{e}^p = \boldsymbol{\theta}^p - \boldsymbol{\theta}^*$ ， \mathbf{J} 表示 $F(\boldsymbol{\theta})$ 的雅克比矩阵在 $\boldsymbol{\theta} = \boldsymbol{\theta}^*$ 处的值。由公式 (3.26) 可以得到两个十分有用的等式 (3.27) 和 (3.28)。

$$\mathbf{r}^p = (\mathbf{J} - I)\mathbf{e}^p + o(\mathbf{e}^p) \quad \dots (3.27)$$

$$\mathbf{v}^p = (\mathbf{J} - I)^2 \mathbf{e}^p + o(\mathbf{e}^p) \quad \dots (3.28)$$

同时，还可以利用公式 (3.25) 和公式 (3.26) 得到误差传递公式，如公式 (3.29) 所示。

$$\mathbf{e}^{p+1} = [I - \alpha^p (\mathbf{J} - I)]\mathbf{e}^p + o(\mathbf{e}^p) \quad \dots (3.29)$$

在 STEM 迭代机制的基础上，结合“squaring”的思想[87]，可以得到收敛速度快于 STEM 的算法，也就是 SQUAREM 算法。SQUAREM 算法要求误差传递公式为公式（3.30）所示的形式。

$$\mathbf{e}^{p+1} = [I - \alpha^p (\mathbf{J} - I)]^2 \mathbf{e}^p + o(\mathbf{e}^p) \quad \dots (3.30)$$

结合公式（3.27）和（3.28），可以得到满足公式（3.30）所要求的误差传递公式的迭代机制如公式（3.31）所示。

$$\boldsymbol{\theta}^{p+1} = \boldsymbol{\theta}^p - 2\alpha^p \mathbf{r}^p + (\alpha^p)^2 \mathbf{v}^p \quad \dots (3.31)$$

利用 SQUAREM 算法加速算法 3.1 的收敛速度时，由于其未考虑相关参数的限制条件，例如 $0 \leq \lambda_n \leq 1, \sigma_n^2 > 0$ ，所以利用公式（3.31）更新参数时会导致相关参数违反其限制条件，从而导致整个迭代算法无法收敛。比如当某一时刻的 $(\sigma_n^p)^2 < 0$ 时，公式（3.13）中某些 $\exp(-|y_i - \alpha^p s_j|^2 / (\sigma_n^p)^2)$ 可能会太大以至于超过了计算机的精度，从而导致相应的 φ_{mj}^p 无意义。因此，每当利用公式（3.31）获得参数更新值 $\boldsymbol{\theta}^{p+1}$ 时，需要检验其是否满足参数的限定条件：

$0 \leq \lambda_n^{p+1} \leq 1, (\sigma_n^{p+1})^2 > 0$ 。如果 $\boldsymbol{\theta}^{p+1}$ 违反了任何一条限定条件，则放弃 $\boldsymbol{\theta}^{p+1}$ ，采用 ECM 算法来得到更新值 $\boldsymbol{\theta}^{p+1} = F(F(\boldsymbol{\theta}^p))$ 。这就意味着，每当 $\boldsymbol{\theta}^{p+1}$ 没有违反限定条件，我们接受采用步长 α^p 来加快收敛速度，一旦 $\boldsymbol{\theta}^{p+1}$ 违反了限定条件，则放弃步长 α^p ，利用 ECM 算法来更新参数。这里称添加了参数限定条件检验步骤的 SQUAREM 算法为 SQUAREM-PC。因此，SQUAREM-PC 算法的收敛速度会快于 ECM 算法，同时，可以保证收敛到观察数据对数似然函数的局部最大值点。这里定义 SQUAREM-PC 的一次完整的迭代过程为：在获得 $\boldsymbol{\theta}^{p+1}$ 的基础上，再做一次迭代，将利用 SQUAREM-PC 算法获得的下一时刻参数估计值 $\boldsymbol{\theta}^{p+1}$ 定义为 $\boldsymbol{\theta}^{p+1} = F(\boldsymbol{\theta}^{p+1})$ 。具体的 SQUAREM-PC 算法流程如算法 3.2 所示。SQUAREM-PC 算法的复杂度为 $O(m_i N T f)$ ，其中 f 表示 SQUAREM-PC 算法达到收敛时计算映射函数 F 的次数。

算法 3.2 SQUAREM-PC 算法

- 1: 输入 θ^0 和 ρ .
- 2: 赋值 $p = 0$.
- 3: 依次计算 $\mathbf{u}_1 = F(\theta^p); \mathbf{u}_2 = F(\mathbf{u}_1); \mathbf{r}^p = \mathbf{u}_1 - \theta^p; \mathbf{v}^p = \mathbf{u}_2 - 2\mathbf{u}_1 + \theta^p$.
- 4: 计算步长 $\alpha^p = -\|\mathbf{r}^p\| / \|\mathbf{v}^p\|$.
- 5: 计算 $\theta' = \theta^p - 2\alpha^p \mathbf{r}^p + (\alpha^p)^2 \mathbf{v}^p$.
- 6: 判断 θ' 是否违反参数限制条件, 如果违反某条参数限制条件, 将 \mathbf{u}_2 赋值给 θ' : $\theta' = \mathbf{u}_2$.
- 7: 得到 $p+1$ 时刻参数更新值: $\theta^{p+1} = F(\theta')$.
- 8: 判断是否满足收敛条件: $\|\theta^{p+1} - \theta^p\| < \rho$. 如果没有满足收敛条件, 则将 p 值加 1, 重复步骤 3~8 直到满足收敛条件为止。

3.4 仿真及分析

在实验部分, 真实的信道参数的设置为: 假设信道系数 β 的幅度服从瑞利分布, 并且满足 $E[|\beta|^2] = 2$; β 的相位服从于 $(0, 2\pi]$ 之间的均匀分布。高斯混合模型的成份数 N 分别取 2 和 3 时进行研究。当 $N = 2$ 时, 设置 $\lambda_1 = 0.9, \sigma_2^2 / \sigma_1^2 = 100$; 当 $N = 3$ 时, 设置 $\lambda_1 = 0.85, \lambda_2 = 0.1, \sigma_2^2 / \sigma_1^2 = 50, \sigma_3^2 / \sigma_1^2 = 100$ 。观察信号序列的长度 $T = 500$ 。在调制星座点的能量归一化的基础上, 信噪比 SNR 定义为

$$\text{SNR} = 10 \log_{10}(E[|\beta|^2] / \sigma_1^2)。$$

由于 ECM 算法和 SQUAREM-PC 算法的收敛结果都受到初始值设置的影响, 所以这里在作对比实验时, 保持两者采用相同的初始化方式[83]:

- 1) 设置 $\lambda_1^0 = 0.8, \lambda_2^0 = \dots = \lambda_N^0 = 0.2 / (N-1)$ 。
- 2) 取观察信号的模的中位数 $\bar{\sigma} = \text{median}\{|y_t| : t = 1, \dots, T\}$, 设置 $(\sigma_1^0)^2 = (\bar{\sigma} / 0.7)^2$, 其余高斯成份的方差满足 $(\sigma_n^0)^2 = \gamma(\sigma_{n-1}^0)^2, n = 2, \dots, N-1$, 并且第 N 个高斯成份的方差设置为 $(\sigma_N^0)^2 = \max\{|y_t|^2 : t = 1, \dots, T\}$, 其中 $\gamma = [(\sigma_N^0)^2 / (\sigma_1^0)^2]^{1/(N-1)}$ 。

ECM 算法和 SQUAREM-PC 算法的收敛标记 ρ 都设置为 10^{-6} ，ECM 算法的最大迭代次数设置为 6000，因为在 SQUAREM-PC 算法的一次迭代中需要计算三次映射函数 F ，于是将 SQUAREM-PC 算法的最大迭代次数设置为 2000。

首先，我们研究 SQUAREM 算法与 SQUAREM-PC 算法。通过统计两个算法的参数估计值中出现无意义估计结果的次数（用 N_{cnt} 表示）来说明参数限制条件检验的必要性。表格 3.1 展示了 5000 次实验中，两个算法参数估计值中出现无意义估计结果的次数 N_{cnt} 。从表 3.1 中可以明显看出采用了参数限制条件检验步骤的 SQUAREM-PC 算法的优越性。紧接着，我们研究了在 SQUAREM-PC 算法的一个收敛周期内，违反参数限制条件的次数占整个迭代次数的百分比，如表 3.2 所示。

表 3.1 16QAM 作为发射调制类型，5000 次实验中，SQUAREM 和 SQUAREM-PC 的 N_{cnt}

Table 3.1 N_{cnt} of SQUAREM and SQUAREM-PC algorithm when the modulation type is

16QAM and the number of experiments is 5000

<div> N_{cnt} 算法 </div>	SNR(dB)				
	-3	1	5	9	13
SQUAREM	271	428	780	1270	1661
SQUAREM-PC	0	0	0	0	0

表 3.2 SQUAREM-PC 算法的收敛周期内，违反限制条件的次数占整个迭代次数的百分比

Table 3.2 The average percentage of violating times in one whole iterations of

SQUAREM-PC

SNR(dB)	-3	1	5	9	13
百分比	0.26%	0.66%	1.70%	5.11%	9.88%

其次，我们比较 SQUAREM-PC 算法与 ECM 算法在估计信道参数时的性能。这里主要考虑参数估计准确度以及算法收敛速度两个性能指标。用映射函数 F 的平均计算次数（用 F_{cnt} 来表示）作为比较两个算法收敛快慢的依据。用信道系数幅度与相位的估计值与真实值之间的平均偏差作为比较两个算法对于参数估计

准确度的依据。这里估计值 \hat{X} 与真实参数 X 的平均偏差定义为 $E[|\hat{X} - X|]$ 。在该实验中，选择 16QAM 作为发射调制类型。图 3.1 显示了在 $N = 2, 3$ 时，SQUAREM-PC 和 ECM 算法的 F_{cnt} 随 SNR 变化的曲线图。从图 3.1 中可以看出，SQUAREM-PC 算法的收敛速度快于 ECM 算法。图 3.2 和图 3.3 分别显示了在 N 取 2 和 3 时，SQUAREM-PC 算法和 ECM 算法对信道系数 β 的幅度和相位估计的平均偏差。从图 3.2 和图 3.3 中可以发现利用 SQUAREM-PC 算法对信道系数估计的准确度接近 ECM 算法的准确度，甚至在 SNR 较大时，SQUAREM-PC 算法的估计准确度还优于 ECM 算法的准确度，这可能是因为在高 SNR 时，SQUAREM-PC 算法收敛于接收信号对数似然函数的全局最大值的概率高于 ECM 算法，具体的理论解释这一现象仍需研究。综合来看，SQUAREM-PC 算法在没丢失估计准确度的情形下，加快了 ECM 算法的收敛速度。

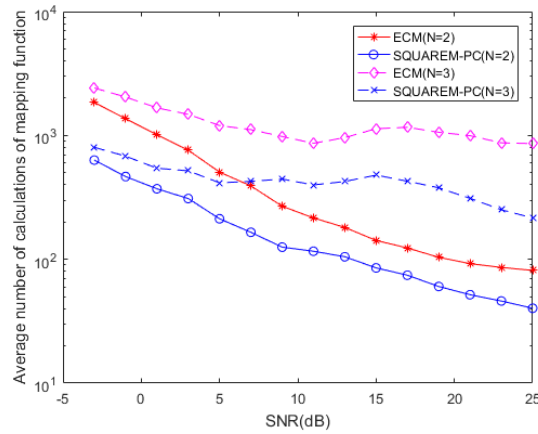


图 3.1 SQUAREM-PC 和 ECM 算法的 F_{cnt} 随 SNR 变化图

Figure 3.1 F_{cnt} for SQUAREM-PC and ECM algorithm versus SNR when $N = 2, 3$

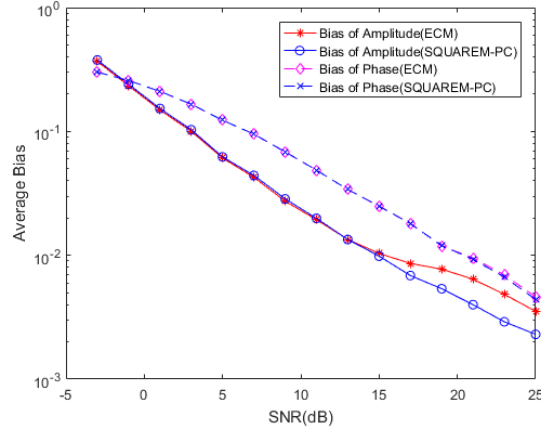


图 3.2 $N = 2$ 时, SQUAREM-PC 和 ECM 算法对信道系数幅度和相位估计的平均偏差

Figure 3.2 Average deviation of amplitude and phase of channel coefficient by
SQUAREM-PC and ECM when $N = 2$

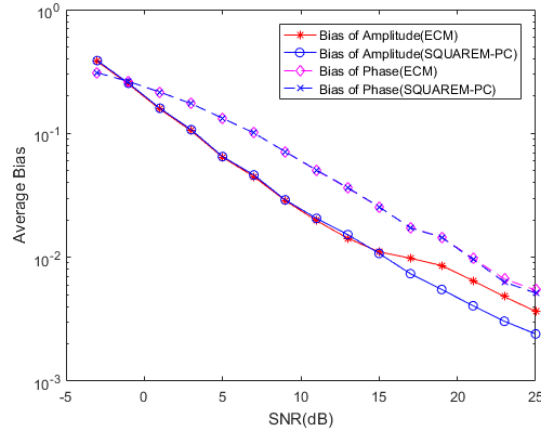


图 3.3 $N = 3$ 时, SQUAREM-PC 和 ECM 算法对信道系数幅度和相位估计的平均偏差

Figure 3.3 Average deviation of amplitude and phase of channel coefficient by
SQUAREM-PC and ECM when $N = 3$

最后, 比较基于 SQUAREM-PC 算法进行信道盲估计的最大似然调制分类器与基于 ECM 算法的最大似然调制分类器性能, 主要比较对象为调制分类器的分类准确度以及分类时间。用平均正确分类概率 P_{cc} 作为衡量分类准确度的指标, 其定义如公式 (3.32) 所示。

$$P_{cc} = \frac{1}{C} \sum_{i=1}^C P(\hat{H}_i = H_i | H_i) \quad \dots (3.32)$$

这里 H_i 表示发射机采用的调制方式为假设 H_i 所对应的调制类型。用平均调制分类时间 t_{mc} 作为衡量调制分类器快慢的指标，其定义如公式 (3.33) 所示。

$$t_{mc} = \frac{1}{C} \sum_{i=1}^C t_{mc}^i \quad \dots (3.33)$$

其中 t_{mc}^i 表示发射机采用假设 H_i 所对应的调制类型时，进行调制分类所需要的时间。这里我们选择的待分类的调制集合为 BPSK、QPSK、8-PSK 和 16-QAM，发射机以相等概率从调制集合中选择调制方式。图 3.4 和图 3.5 分别是 $N = 2$ 和 $N = 3$ 时，已知信道参数情形下的最大似然调制分类器分类准确度、基于 ECM 算法进行盲信道估计的最大似然调制分类器分类准确度和基于 SQUAREM-PC 算法进行盲信道估计的最大似然调制分类器分类准确度随 SNR 变化的曲线图。从图 3.4 和图 3.5 中可以看出，基于 SQUAREM-PC 算法进行盲信道估计的最大似然调制分类器分类准确度与基于 ECM 算法的准确度基本一致，在 SNR 较高时，前者的分类准确度甚至优于后者，这与图 3.2 和图 3.3 中，在 SNR 较高时，SQUAREM-PC 算法对参数估计的准确度高于 ECM 算法是一致的。表 3.3 展示了基于 ECM 算法的调制分类器与基于 SQUAREM-PC 算法的调制分类器在不同的 SNR 下的平均调制分类时间 t_{mc} 。从表 3.3 中可以看出基于 SQUAREM-PC 算法的调制分类器的时间明显小于基于 ECM 算法的调制分类器。综合上述性能对比，可以发现基于 SQUAREM-PC 算法进行信道盲估计的最大似然调制分类器能够在不丢失分类准确度的条件下，加快基于 ECM 算法的调制分类器的分类速度。

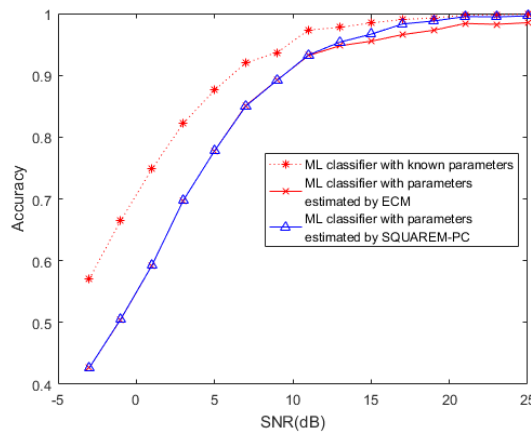


图 3.4 $N = 2$ 时，调制分类准确度随 SNR 变化曲线图

Figure 3.4 The accuracy of classifier versus SNR when $N = 2$

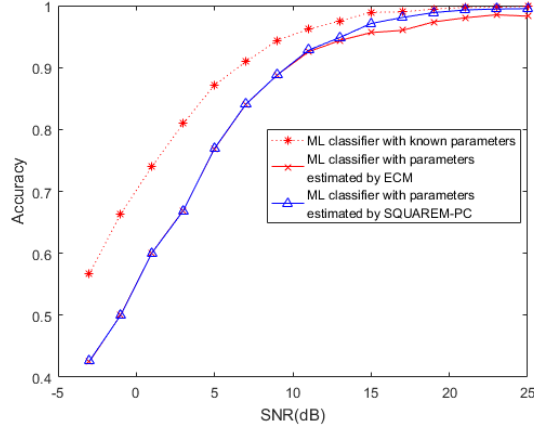

 图 3.5 $N=3$ 时，调制分类准确度随 SNR 变化曲线图

 Figure 3.5 The accuracy of classifier versus SNR when $N=3$

表 3.3 不同 SNR 条件下的平均调制分类时间

Table 3.3 Average time of Classifier versus SNR

时间(s)	SNR(dB)						
		-3	1	5	9	13	17
算法							
ECM($N = 2$)		10.72	6.87	3.93	2.00	1.15	0.67
SQUAREM-PC($N = 2$)		3.83	2.33	1.35	0.70	0.46	0.30
ECM($N = 3$)		15.92	11.52	8.35	6.15	5.02	5.42
SQUAREM-PC($N = 3$)		5.65	3.73	2.79	2.17	1.80	1.78

上述实验都是基于假设非高斯噪声服从高斯混合分布下进行的，而实际场景中，噪声分布可能不是高斯混合分布。一般常用的非高斯噪声的概率分布还有 Middleton's Class A 模型和对称 α 分布，这里我们比较当实际的噪声分布服从对称 α 分布时，基于 SQUAREM-PC 算法与基于 ECM 算法的分类器的性能。首先假设非高斯噪声服从位置参数 $\delta=0$ 的对称稳定 α 分布，由于采用等效复基带模型分析接收信号，于是假设噪声样本的实部和虚部服从二元各向同性的对称稳定 α 分布[84]。由于 α 分布没有闭式概率密度函数，而混合高斯分布具有很好的近似其他分布的能力，所以利用混合高斯分布的概率密度函数来近似 α 分布的概率密度函数来进行最大似然调制分类。在具体的实验中，利用二元各向同性的对称稳定 α 分布来生成噪声样本，同时发射信号经历同上述实验一致的瑞利衰落信道。在这样的信道条件下，研究基于上述两种算法的最大似然调制分类器的性能。在

进行调制分类时，采用上述的衰落信道下的信号模型进行分析，假设噪声服从成份数 $N = 3$ 的高斯混合分布，图 3.6 展示了对称稳定 α 分布的特征指数

$\alpha = 1.2, 1.5, 1.8$ 时，基于两种算法的最大似然调制分类器分类准确度随对称稳定 α 分布的离差 γ 的变化曲线图。从图 3.6 中可以看到基于 SQUAREM-PC 算法的最大似然调制分类器分类准确度与基于 ECM 算法的调制分类准确度基本保持一致。表 3.4 展示了基于两种算法的调制分类器的平均调制分类时间。从表 3.4 中可以看出基于 SQUAREM-PC 算法的最大似然调制分类器比基于 ECM 算法的分类器的分类速度快。综上可以看出，当噪声分布为对称 α 分布时，基于 SQUAREM-PC 算法的调制分类器依旧可以在不损失分类精度的条件下，加快基于 ECM 算法的调制分类器的分类速度。

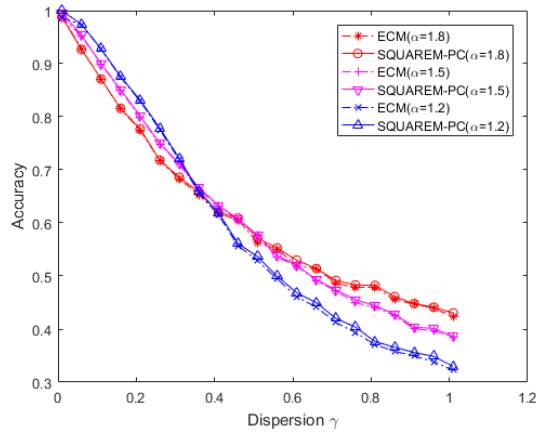


图 3.6 调制分类准确度随参数 γ 的变化曲线图

Figure 3.6 The accuracy of classifier versus γ

表 3.4 不同 γ 下的平均调制分类时间Table 3.4 Average time of classifier versus γ

时间(s) \ γ 算法	0.01	0.21	0.41	0.61	0.81	1.01
ECM($\alpha = 1.8$)	4.49	9.68	12.68	14.99	16.33	17.77
SQUAREM-PC($\alpha = 1.8$)	1.65	3.49	5.01	5.96	6.91	7.73
ECM($\alpha = 1.5$)	1.35	6.91	11.85	15.07	17.59	19.73
SQUAREM-PC($\alpha = 1.5$)	0.60	3.33	5.98	8.45	10.26	11.77
ECM($\alpha = 1.2$)	0.74	6.48	13.95	19.22	23.44	25.22
SQUAREM-PC($\alpha = 1.2$)	0.52	3.52	8.50	12.58	16.04	17.93

3.5 本章小结

本章首先采用 GMM 来建模非高斯噪声；结合信号模型，利用贝叶斯定理推导出接收信号似然函数值计算方法；在分析了 ECM 算法在信道参数盲估计时所存在的缺点的基础上，提出了基于 SQUAREM-PC 算法的信道盲估计算法；最后，通过仿真实验验证了 SQUAREM-PC 算法相较于 ECM 算法，能够在不丢失信道估计准确度的前提下，加快信道参数盲估计的速度；同时，也验证了基于 SQUAREM-PC 算法的最大似然调制分类器相较于基于 ECM 算法的最大似然调制分类器，能够在不丢失分类准确度的前提下，加快调制分类的速度。

第四章 频率选择性衰落高斯噪声信道下接收信号似然函数计算

在研究频率选择性衰落信道下自动调制识别算法时，我们首先建立了该信道下，接收信号的生成模型。由于采用最大似然调制分类算法框架，所以需要计算每个调制类型假设下，接收信号的似然函数值。本章主要介绍这两方面的内容。

4.1 信号模型

在频率选择性衰落信道下的信号模型描述中，同样用 \mathbf{M} 表示发射机采用的线性调制星座点集合， $m = |\mathbf{M}|$ 表示调制星座点数。从 \mathbf{M} 中均匀采样出发射符号，这里用 $\{a_t\}_{t \geq 1}$ 来表示发射符号。经过预处理后，在匹配滤波器输出端以符号速率采样可以获得接收信号的等效复基带形式，如公式（4.1）所示。

$$y_t = \sum_{l=0}^{L-1} a_{t-l} h_l + \varepsilon_t \quad \dots (4.1)$$

其中 L 表示多径信道的阶数， $\mathbf{h} = [h_0, \dots, h_{L-1}]^T$ 表示多径信道系数。通信信道中存在大量独立的方差有限的噪声源，根据中心极限定理[89]，这些独立的噪声源叠加后近似高斯分布，因此，这里假设加性噪声服从高斯分布。所以这里 $\{\varepsilon_t\}_{t \geq 1}$ 是独立复高斯噪声。 ε_t 的实部和虚部相互独立，且服从相同的高斯分布，所以可以将 ε_t 看作是由其实部和虚部构成二维向量，服从二元高斯分布，其概率分布如公式（4.2）所示。

$$p(\mathbf{x}) = \frac{1}{2\pi |\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2} \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x}\right) \quad \dots (4.2)$$

其中 $\boldsymbol{\Sigma} = (\sigma^2 / 2) \mathbf{I}_{2 \times 2}$ ，这里 σ^2 表示复高斯噪声的能量， $\mathbf{I}_{2 \times 2}$ 表示二维单位矩阵。

将该信号模型改写为矩阵形式后，便可以得到如公式（4.3）所示的隐马尔科夫模型。

$$\begin{aligned} \mathbf{s}_t &= \mathbf{Q} \mathbf{s}_{t-1} + \mathbf{w}_t \\ y_t &= \mathbf{h}^T \mathbf{s}_t + \varepsilon_t \end{aligned} \quad \dots (4.3)$$

其中 $\mathbf{s}_t = [a_t, \dots, a_{t-L+1}]^T$, $\mathbf{w}_t = [a_t, 0, \dots, 0]^T$ 。 \mathbf{Q} 是移位矩阵, 其元素 $Q_{ij} = \delta_{i,j+1}$, 这里 δ_{ij} 是克罗内克符号, 当 $i = j$ 时, $\delta_{ij} = 1$; 当 $i \neq j$ 时, $\delta_{ij} = 0$ 。由此, 该隐马尔科夫模型具有以下特性:

- 1) 马尔科夫链 $\{\mathbf{s}_t\}_t$ 的状态空间为 $S = \mathbf{M}^L$, 该离散空间中元素个数为 m^L 。
- 2) 从状态空间 S 中均匀采样可得马尔科夫链的初始状态。
- 3) 状态空间 S 中的状态 $\mathbf{s} = [a_1, \dots, a_L]^T$ 到状态 $\mathbf{s}' = [a'_1, \dots, a'_L]^T$ 的转移概率如公式 (4.4) 所示。

$$P(\mathbf{s}' | \mathbf{s}) = \begin{cases} \frac{1}{m} & \text{当 } [a'_2, \dots, a'_L]^T = [a_1, \dots, a_{L-1}]^T \text{ 时} \\ 0 & \text{其他} \end{cases} \quad \dots (4.4)$$

- 4) 在给定状态 \mathbf{s} 的条件下, 观察信号 y 的条件概率密度如公式 (4.5) 所示。其中, $\boldsymbol{\theta} = [\mathbf{h}^T, \sigma^2]^T$ 表示信号模型中的信道参数向量。

$$p(y; \boldsymbol{\theta} | \mathbf{s}) = \frac{1}{\pi \sigma^2} \exp\left(-\frac{|y - \mathbf{h}^T \mathbf{s}|^2}{\sigma^2}\right) \quad \dots (4.5)$$

4.2 接收信号似然函数值计算

4.2.1 枚举法

根据贝叶斯全概率公式, 结合 4.1 节的信号模型, 在调制类型假设 H_i 下, 接收信号条件概率可由公式 (4.6) 计算得到。

$$p(y_{1:T}; \boldsymbol{\theta} | H_i) = \sum_{\mathbf{s}_{1:T}} p(y_{1:T}; \boldsymbol{\theta} | \mathbf{s}_{1:T}) P(\mathbf{s}_{1:T} | H_i) \quad \dots (4.6)$$

这里假设在给定隐藏状态 $\mathbf{s}_{1:T}$ 条件下, 观察序列 $y_{1:T}$ 相互独立, 所以

$$p(y_{1:T}; \boldsymbol{\theta} | \mathbf{s}_{1:T}) = \prod_{t=1}^T p(y_t; \boldsymbol{\theta} | \mathbf{s}_t)。$$
 同时, 利用状态转移概率, 可以得到

$$P(\mathbf{s}_{1:T} | H_i) = P(\mathbf{s}_1) \prod_{t=2}^T P(\mathbf{s}_t | \mathbf{s}_{t-1})。$$

公式 (4.6) 通过枚举所有可能的隐藏状态序列来计算接收信号的条件概率。对于多径信号模型, 一共有 m_i^{L+T-1} 种可能的隐藏状态序列组合, 其中 m_i 表示假设

H_i 所对应的调制类型 M_i 的调制阶数。所以这种计算接收信号条件概率的方案算法复杂度非常高。

4.2.2 前向迭代算法

利用 FB 算法[88]中前向迭代部分可以得到复杂度较低的接收信号条件概率计算方法。定义 $\alpha_i(n) \triangleq p(y_{1:t}, \mathbf{s}_t = S_i(n); \boldsymbol{\theta} | H_i)$ ，其中 $S_i(n)$ 表示在假设 H_i 下，隐马尔科夫模型状态空间 S_i 中的第 n 个状态。利用 $\alpha_{t+1}(k)$ 和 $\alpha_t(n)$ 之间的迭代关系可以计算接收信号条件概率，迭代关系的具体推导过程如公式 (4.7) 所示。

$$\begin{aligned}
 \alpha_{t+1}(k) &= p(y_{1:t+1}, \mathbf{s}_{t+1} = S_i(k); \boldsymbol{\theta} | H_i) \\
 &= p(y_{t+1}; \boldsymbol{\theta} | \mathbf{s}_{t+1} = S_i(k)) p(y_{1:t}, \mathbf{s}_{t+1} = S_i(k); \boldsymbol{\theta} | H_i) \\
 &= p(y_{t+1}; \boldsymbol{\theta} | \mathbf{s}_{t+1} = S_i(k)) \sum_n p(y_{1:t}, \mathbf{s}_t = S_i(n), \mathbf{s}_{t+1} = S_i(k); \boldsymbol{\theta} | H_i) \quad \dots (4.7) \\
 &= p(y_{t+1}; \boldsymbol{\theta} | \mathbf{s}_{t+1} = S_i(k)) \sum_n P(\mathbf{s}_{t+1} = S_i(k) | \mathbf{s}_t = S_i(n)) \alpha_t(n)
 \end{aligned}$$

利用上述迭代关系计算接收信号条件概率的完整过程可整理为：

1) 初始化：

$$\alpha_1(n) = p(y_1; \boldsymbol{\theta} | \mathbf{s}_1 = S_i(n)) P(\mathbf{s}_1 = S_i(n)), \quad n = 1, 2, \dots, m_i^L$$

2) 迭代：

$$\begin{aligned}
 \alpha_{t+1}(k) &= p(y_{t+1}; \boldsymbol{\theta} | \mathbf{s}_{t+1} = S_i(k)) \sum_n P(\mathbf{s}_{t+1} = S_i(k) | \mathbf{s}_t = S_i(n)) \alpha_t(n), \\
 &\quad k = 1, 2, \dots, m_i^L \\
 &\quad t = 1, 2, \dots, T-1
 \end{aligned}$$

3) 终止：

$$p(y_{1:T}; \boldsymbol{\theta} | H_i) = \sum_{n=1}^{m_i^L} \alpha_T(n)$$

因为该隐马尔科夫模型的状态转移概率具有稀疏特性（状态转移概率矩阵中大多数元素为 0），所以在求解 $\alpha_{t+1}(k)$ 时，只需要 $\{\alpha_t(n)\}_{n=1}^{m_i^L}$ 集合中的 m_i 个元素参与计算。我们称利用前向迭代的方式计算接收信号条件概率的方法为前向迭代算法，该算法复杂度为 $O(Tm_i^{L+1})$ 。相较于之前的枚举法，前向迭代算法有效地降低了计算复杂度。

由于 $\alpha_t(n) = \sum_{\mathbf{s}_{1:t-1}} P(\mathbf{s}_1) \prod_{\tau=1}^{t-1} P(\mathbf{s}_{\tau+1} | \mathbf{s}_\tau) \prod_{\tau=1}^t p(y_\tau; \boldsymbol{\theta} | \mathbf{s}_\tau)$, 其中 $\mathbf{s}_t = S_t(n)$, 我们发现 $\alpha_t(n)$ 是由很多项的和构成的, 其中每一个被加项都远远小于 0, 随着 t 的增加, 每一个被加项都会以指数的速度趋近于 0, 所以当 t 很大时, $\alpha_t(n)$ 的计算会超出计算机的计算精度, 导致计算出错。这里采用缩放的方法来有效解决这一问题。

定义 $c_t \triangleq 1 / \sum_{n=1}^{m_t^L} \tilde{\alpha}_t(n)$, 其中 $\tilde{\alpha}_t(n)$ 表示缩放过程中产生的中间量。采用缩放的前向迭代算法的完整过程为:

1) 初始化:

计算 $\alpha_1(n)$, 并将 $\alpha_1(n)$ 赋值给中间量 $\tilde{\alpha}_1(n)$: $\tilde{\alpha}_1(n) = \alpha_1(n)$, 其中 $n = 1, 2, \dots, m_1^L$ 。

根据定义计算 c_1 , 用 c_1 对 $\tilde{\alpha}_1(n)$ 进行缩放得 $\hat{\alpha}_1(n) = c_1 \tilde{\alpha}_1(n)$ 。

2) 迭代:

根据迭代公式 (4.7), 利用缩放后的 $\hat{\alpha}_{t-1}(k)$ 迭代产生中间量 $\tilde{\alpha}_t(n)$, 然后计算 c_t , 用 c_t 对 $\tilde{\alpha}_t(n)$ 进行缩放得 $\hat{\alpha}_t(n) = c_t \tilde{\alpha}_t(n)$ 。

3) 终止:

根据上述计算过程可以看出缩放后的 $\hat{\alpha}_t(n)$ 和未缩放的 $\alpha_t(n)$ 的关系为:

$\hat{\alpha}_t(n) = \prod_{\tau=1}^t c_\tau \alpha_\tau(n)$, 同时由缩放的运算可知 $\sum_{n=1}^{m_t^L} \hat{\alpha}_t(n) = 1$, 所以最终接收信号的似然函数为: $p(y_{1:T}; \boldsymbol{\theta} | H_i) = 1 / \prod_{\tau=1}^T c_\tau$ 。

因为前向迭代算法的计算复杂度为 $O(Tm_i^{L+1})$, 所以当 m_i 或者 L 较大时, 该算法的复杂度依旧非常高。这就导致了在调制分类时, 当待识别的调制集合中存在调制阶数较高的调制方式或者通信信道阶数较高时, 采用前向迭代算法精确计算每个调制假设的似然函数值复杂度较高。为了有效地解决这一问题, 我们采用粒子滤波算法来近似计算接收信号似然函数值。下面详细介绍这一算法。

4.2.3 前向粒子滤波

为了表达方便, 用 $\pi_{t|t}(\cdot; \boldsymbol{\theta})$ 表示 $P(\mathbf{s}_t; \boldsymbol{\theta} | y_{1:t})$ 。利用粒子滤波方法可以得到 \mathbf{s}_t 的边缘滤波分布 $\pi_{t|t}(\cdot; \boldsymbol{\theta}^c)$, $1 \leq t \leq T$ 的近似值。下面详细描述粒子滤波算法[91]的过程。

假设在时刻 t , 通过粒子滤波方法得到 $\pi_{t|t}(\cdot; \boldsymbol{\theta})$ 的近似分布, 表示为 $\hat{\pi}_{t|t}(\cdot; \boldsymbol{\theta})$, 其主要由 N 个粒子 (状态空间 S 中的元素) 和粒子所对应的权重构成。 $\hat{\pi}_{t|t}(\cdot; \boldsymbol{\theta})$ 的数学表达式如公式 (4.8) 所示。

$$\hat{\pi}_{t|t}(\mathbf{s}; \boldsymbol{\theta}) = \sum_{i=0}^{N-1} w_t^i \delta_{\xi_t^i}(\mathbf{s}) \quad \dots (4.8)$$

其中粒子 $\xi_t^i \in S$, 其权重为 w_t^i , N 个粒子的权重满足 $\sum_{i=0}^{N-1} w_t^i = 1$ 。当 $\mathbf{s} = \xi_t^i$ 时, $\delta_{\xi_t^i}(\mathbf{s}) = 1$; 当 $\mathbf{s} \neq \xi_t^i$ 时, $\delta_{\xi_t^i}(\mathbf{s}) = 0$ 。由于 $t+1$ 时刻的滤波分布 $\pi_{t+1|t+1}(\cdot; \boldsymbol{\theta})$ 可分解成公式 (4.9) 所示的形式,

$$\begin{aligned} \pi_{t+1|t+1}(\mathbf{s}; \boldsymbol{\theta}) &\propto \sum_{\mathbf{x}} p(y_{t+1}, \mathbf{s}_{t+1} = \mathbf{s}, \mathbf{s}_t = \mathbf{x}; \boldsymbol{\theta}) \\ &\propto \sum_{\mathbf{x}} p(y_{t+1}; \boldsymbol{\theta} | \mathbf{s}_{t+1} = \mathbf{s}) P(\mathbf{s}_{t+1} = \mathbf{s} | \mathbf{s}_t = \mathbf{x}) \pi_{t|t}(\mathbf{x}; \boldsymbol{\theta}) \end{aligned} \quad \dots (4.9)$$

所以将 t 时刻的滤波分布近似值 $\hat{\pi}_{t|t}(\cdot; \boldsymbol{\theta})$ 带入到公式 (4.9) 中可以得到 $t+1$ 时刻的滤波分布近似值 $\hat{\pi}_{t+1|t+1}(\cdot; \boldsymbol{\theta})$, 如公式 (4.10) 所示。

$$\begin{aligned} \hat{\pi}_{t+1|t+1}(\mathbf{s}; \boldsymbol{\theta}) &\propto \sum_{\mathbf{x}} p(y_{t+1}; \boldsymbol{\theta} | \mathbf{s}_{t+1} = \mathbf{s}) P(\mathbf{s}_{t+1} = \mathbf{s} | \mathbf{s}_t = \mathbf{x}) \hat{\pi}_{t|t}(\mathbf{x}; \boldsymbol{\theta}) \\ &= \sum_{i=0}^{\tilde{N}-1} \sum_{j=0}^{N-1} w_t^j P(\mathbf{s}_{t+1} = \tilde{\xi}_{t+1}^i | \mathbf{s}_t = \xi_t^j) p(y_{t+1}; \boldsymbol{\theta}^c | \mathbf{s}_{t+1} = \tilde{\xi}_{t+1}^i) \delta_{\tilde{\xi}_{t+1}^i}(\mathbf{s}) \quad \dots (4.10) \\ &= \sum_{i=0}^{\tilde{N}-1} \tilde{w}_{t+1}^i \delta_{\tilde{\xi}_{t+1}^i}(\mathbf{s}) \end{aligned}$$

其中 $\{\tilde{\xi}_{t+1}^i\}_{i < \tilde{N}}$ 表示由 t 时刻的粒子集合 $\{\xi_t^i\}_{i < N}$ 产生的 $t+1$ 时刻所有可能的粒子集合, 根据状态转移概率可知 $\tilde{N} \leq Nm$ 。由于从 t 时刻的粒子集合 $\{\xi_t^i\}_{i < N}$ 到 S 中除掉粒子集合 $\{\tilde{\xi}_{t+1}^i\}_{i < \tilde{N}}$ 外的其他元素的转移概率均为零, 所以 S 中除掉粒子集合 $\{\tilde{\xi}_{t+1}^i\}_{i < \tilde{N}}$ 外的其他元素所对应的权重均为零, 因此利用 $\{\tilde{\xi}_{t+1}^i\}_{i < \tilde{N}}$ 可以充分近似 $t+1$

时刻的滤波分布。 $t+1$ 时刻的粒子 $\tilde{\xi}_{t+1}^i$ 的未归一化权重 \tilde{w}_{t+1}^i 可由公式(4.11)更新计算。

$$\begin{aligned}\tilde{w}_{t+1}^i &= \sum_{j=0}^{N-1} w_t^j P(\mathbf{s}_{t+1} = \tilde{\xi}_{t+1}^i | \mathbf{s}_t = \xi_t^j) p(y_{t+1}; \boldsymbol{\theta}^c | \mathbf{s}_{t+1} = \tilde{\xi}_{t+1}^i) \\ &= \sum_{j \in \mathcal{P}(\tilde{\xi}_{t+1}^i)} \frac{1}{m} w_t^j p(y_{t+1}; \boldsymbol{\theta}^c | \mathbf{s}_{t+1} = \tilde{\xi}_{t+1}^i)\end{aligned}\quad \dots (4.11)$$

这里我们称 t 时刻的状态 \mathbf{s}_t 是 $t+1$ 时刻的状态 \mathbf{s}_{t+1} 的先代, 反之, 称 $t+1$ 时刻的状态 \mathbf{s}_{t+1} 为 t 时刻的状态 \mathbf{s}_t 的后代。 $\mathcal{P}(\tilde{\xi}_{t+1}^i)$ 表示粒子 $\tilde{\xi}_{t+1}^i$ 所有可能的先代粒子的序号集合。例如当 $\tilde{\xi}_{t+1}^i = [a_{t+1}, \dots, a_{t-L+2}]^T$ 时, $\mathcal{P}(\tilde{\xi}_{t+1}^i) = \{j | \xi_t^j = [a_t, \dots, a_{t-L+2}, x]^T, x \in \mathbf{M}\}$ 。所以当 $\tilde{\xi}_{t+1}^i$ 和 $\tilde{\xi}_{t+1}^j$ 只有第一个元素不同时, $\mathcal{P}(\tilde{\xi}_{t+1}^i) = \mathcal{P}(\tilde{\xi}_{t+1}^j)$ 。

为了描述由 t 时刻的粒子和权重迭代产生 $t+1$ 时刻的粒子和相应的权重的简便方法, 在调制星座点集合 \mathbf{M} 中人为定义出大小顺序, 记为 $\mathbf{M} = \{a^0, \dots, a^{m-1}\}$, 用上标的大小表示元素之间的大小关系。同时, 也可以利用上标的数字来检索 \mathbf{M} 中的元素。粒子 ξ_t^i 表示为 $\xi_t^i = [a_{t,0}^i, \dots, a_{t,L-1}^i]^T \in \mathbf{M}^L$, 其中 $a_{t,0}^i$ 表示 t 时刻的发射符号, $a_{t,L-1}^i$ 表示 $t-L+1$ 时刻的发射符号。基于此, 计算 $t+1$ 时刻的粒子和权重的方法为:

a) .去掉 t 时刻每一个粒子的最后一个元素, 例如去掉 ξ_t^i 的最后一个元素得到

$p_t^i = [a_{t,0}^i, \dots, a_{t,L-2}^i]^T \in \mathbf{M}^{L-1}$ 。如果 $p_t^i = p_t^j$, 那么 ξ_t^i 和 ξ_t^j 将会有相同的子代集合。

b) .对 $p_t^i, i=0, \dots, N-1$ 排序得到 $p_t^{(i)}, i=0, \dots, N-1$ 。排序的原则是: 首先按 p_t^i 的第一个元素进行升序排列, 对于第一个元素相同的, 再按第二个元素进行升序排列, 依此类推。

c) .去掉 $\{p_t^{(i)}\}_{i < N}$ 中重复出现的元素, 留下 d_t 个互不相同的元素, 用 $\{u_t^i\}_{i < d_t}$ 表示。

用 j_t^i 表示能使 $u_t^i = p_t^{(j_t^i)}$ 的最小的序号 (定义 $j_t^{d_t} = N$)。

d) .计算 $v_t^i = (1/m) \sum_{j=j_t^i}^{j_t^{i+1}-1} w_t^j, i=0, \dots, d_t-1$ 。

e) .生成 $\tilde{N} = d_t m$ 个子代粒子和每个粒子对应的未归一化权重:

$$\begin{aligned} \tilde{\xi}_{t+1}^0 &= [a^0, u_t^0]^T, & \tilde{w}_{t+1}^0 &= v_t^0 p(y_{t+1}; \boldsymbol{\theta}^c | \mathbf{s}_{t+1} = \tilde{\xi}_{t+1}^0), \\ &\vdots & & \\ \tilde{\xi}_{t+1}^{m-1} &= [a^{m-1}, u_t^{m-1}]^T, & \tilde{w}_{t+1}^{m-1} &= v_t^{m-1} p(y_{t+1}; \boldsymbol{\theta}^c | \mathbf{s}_{t+1} = \tilde{\xi}_{t+1}^{m-1}), \\ \tilde{\xi}_{t+1}^m &= [a^0, u_t^1]^T, & \tilde{w}_{t+1}^m &= v_t^1 p(y_{t+1}; \boldsymbol{\theta}^c | \mathbf{s}_{t+1} = \tilde{\xi}_{t+1}^m), \\ &\vdots & & \\ &\vdots & & \\ \tilde{\xi}_{t+1}^{\tilde{N}-1} &= [a^{m-1}, u_t^{d_t-1}]^T, & \tilde{w}_{t+1}^{\tilde{N}-1} &= v_t^{d_t-1} p(y_{t+1}; \boldsymbol{\theta}^c | \mathbf{s}_{t+1} = \tilde{\xi}_{t+1}^{\tilde{N}-1}). \end{aligned}$$

在上述的迭代过程中, b) 步骤的排序的复杂度是 $O(N \log_2 N)$, a)、c) 和 d) 步骤的算法复杂度都是 $O(N)$, 最后一步的算法复杂度是 $O(Nm)$ 。由于一般情况下 $\log_2 N \ll m$, 所以整个迭代过程的算法复杂度由 $O(Nm)$ 主导。

通过上述迭代, 由于考虑了 t 时刻的 N 个粒子的所有子代粒子, $t+1$ 时刻的粒子数为 \tilde{N} , 随着时间的推移, 近似滤波分布的粒子会越来越多, 导致算法复杂度上升。所以在每一次迭代后, 比如从 t 时刻到 $t+1$ 时刻, 结合粒子的权重, 需要按一定准则从 \tilde{N} 个粒子中采样出 N 个粒子来近似滤波分布。下面我们详细说明如何采样, 能够保证信息损失量最小。

在每一次迭代更新过程中, 都会产生 $\tilde{N} \leq Nm$ 个新的粒子, 所以需要从这 \tilde{N} 个粒子中选出 N 个粒子来近似前向滤波分布, 这样可以保证算法复杂度不会随着前向迭代的进行而上升。这一选择的问题可以描述为: 给定离散概率分布 $\mathbf{w} = \{w_i\}_{i \leq \tilde{N}-1}$, 寻找离散概率分布 $\mathbf{W} = \{W_i\}_{i \leq \tilde{N}-1}$ 在某种程度上近似 \mathbf{w} , 这里要求 \mathbf{W} 中非零值的平均个数小于等于 N 。

一个自然的想法就是采用确定选择的方法。首先将 \mathbf{w} 按从大到小排列, 得到排序后的向量 $\{w_{(i)}\}_{i \leq (\tilde{N}-1)}$, 这时候按照公式 (4.12) 对 \mathbf{W} 向量中的元素进行赋值。

$$W_i = \begin{cases} w_{(i)} / \sum_{j=0}^{N-1} w_{(j)}, & \text{当 } i \leq N-1 \\ 0, & \text{其他} \end{cases} \quad \dots (4.12)$$

但是这种方法计算出来的新的权重是有偏的。下面介绍无偏的随机选择方法, 这里的 W_i 满足两个条件:

- 1) $E[W_i] = w_i$;
- 2) $\sum_{i=0}^{M-1} P(W_i > 0) = N$ 。

第一个条件保证了 W_i 的无偏性，第二个条件保证了 \mathbf{W} 中的非零值的平均个数等于 N 。满足这两个条件的 \mathbf{W} 有很多，这里同时需要 \mathbf{W} 能够最小化 $E[d(\mathbf{W}, \mathbf{w})]$ ，其中 $d(\cdot, \cdot)$ 表示距离（或者伪距离）函数，随着 $d(\cdot, \cdot)$ 的不同，将会得到不同的 \mathbf{W} 向量。

当 $d(\mathbf{W}, \mathbf{w}) = \sum_{i=0}^{M-1} W_i \ln(W_i / w_i)$ 表示 KL 散度时，根据[91]中定理 1，可以得到随机选择算法 1：首先求出方程 $\sum_{i=0}^{M-1} \min\{w_i / \lambda, 1\} = N$ 的解 λ 。对于所有满足 $w_i \geq \lambda$ 的 i ， $W_i = w_i$ ；对于所有满足 $w_i < \lambda$ 的 i ， W_i 的取值如公式（4.13）所示。

$$W_i = \begin{cases} \lambda, & \text{以 } w_i / \lambda \text{ 的概率} \\ 0, & \text{以 } 1 - w_i / \lambda \text{ 的概率} \end{cases} \quad \dots (4.13)$$

最后将所得的 \mathbf{W} 向量归一化便可得到 \mathbf{w} 的近似值。总的来说，这个方法的思想是保留权重较大的粒子，并且不改变它们所对应的权重值，对于权重相对较小的粒子，按照正比于它们原先权重的概率选择保留或舍弃，保留下来的粒子，给它们赋值新的权重 λ 。

当 $d(\mathbf{W}, \mathbf{w}) = \sum_{i=0}^{M-1} (W_i - w_i)^2 / w_i$ 表示卡方伪距离时，根据[91]中定理 1，可以得到随机选择算法 2：首先求出方程 $\sum_{i=0}^{M-1} \min\{\sqrt{w_i / \lambda}, 1\} = N$ 的解 λ 。对于所有满足 $w_i \geq \lambda$ 的 i ， $W_i = w_i$ ；对于所有满足 $w_i < \lambda$ 的 i ， W_i 的取值如公式（4.14）所示。

$$W_i = \begin{cases} \sqrt{w_i \lambda}, & \text{以 } \sqrt{w_i / \lambda} \text{ 的概率} \\ 0, & \text{以 } 1 - \sqrt{w_i / \lambda} \text{ 的概率} \end{cases} \quad \dots (4.14)$$

最后将所得的 \mathbf{W} 向量归一化便可得到 \mathbf{w} 的近似值。这个方法的思想是保留权重较大的粒子，并且不改变它们所对应的权重值，对于权重相对较小的粒子，按照正比于它们原先权重的平方根的概率选择保留或舍弃，进一步加大了原先权重较小的粒子保留下来的概率。

上面介绍的确定选择方法，对于权重值较小的粒子，没有概率被选中，将会被舍弃；对于随机选择方法 1，权重值较小的粒子有正比于其权重值的概率被保留下来；对于随机选择方法 2，权重值较小的粒子有正比于其权重值平方根的概率被保留下来，所以对于这三种选择方法，权重值较小的粒子被保留下来的概率依次上升。

利用前向粒子滤波近似计算接收信号似然函数值的过程如算法 4.1 所示。

算法 4.1 前向粒子滤波

1: 初始化部分:

从状态空间 S 中均匀采样出 ξ_1^i ，并计算 $\tilde{w}_1^i = p(y_1; \theta^c | \xi_1^i)$ ，其中 $i = 0, \dots, N-1$;

计算 $l_1 = N^{-1} \sum_{i=0}^{N-1} \tilde{w}_1^i$ ，计算归一化权重 $w_1^i = (Nl_1)^{-1} \tilde{w}_1^i$ 。

2: 迭代部分:

for $t = 1:T-1$

根据 t 时刻的粒子集合 $\{\xi_t^i\}_{i=0}^{N-1}$ 产生粒子集合 $\{\xi_{t+1}^i\}_{i=0}^{\tilde{N}_{t+1}-1}$;

利用公式 4.11 计算 \tilde{w}_{t+1}^i ;

计算 $l_{t+1} = \sum_{i=0}^{\tilde{N}_{t+1}-1} \tilde{w}_{t+1}^i$ ，以及 $\bar{w}_{t+1}^i = (l_{t+1})^{-1} \tilde{w}_{t+1}^i$;

利用确定选择方法、随机选择方案 1 或 2 从 $\{\xi_{t+1}^i, \bar{w}_{t+1}^i\}_{i=0}^{\tilde{N}_{t+1}-1}$ 中选出 $\{\xi_{t+1}^i, w_{t+1}^i\}_{i=0}^{N-1}$ 。

end

利用算法 4.1 中的 $\{l_t\}_{t=1}^T$ 可以得到接收信号的似然函数近似值

$$\hat{p}(y_{1:T}; \theta | H_i) = \prod_{t=1}^T l_t。$$

4.3 仿真与分析

在最小相位信道下，取信道系数 $\mathbf{h} = [0.63, 0.05, -0.3]^T$ ， $\text{SNR} = 15\text{dB}$ ， $T = 200$ 发射机采用 64QAM 调制方式。首先，随机生成一个观察信号序列，利用前向迭代算法精确计算其似然函数值；其次，在 $N = 1000, 3000, 5000, 10000$ 时，分别多次采用前向粒子滤波算法近似计算接收信号的似然函数值，其结果分布情况如图 4.1 所示，图中左边四个箱线图代表采用随机选择方法 1 的前向粒子滤波算法近似计算结果，右边四个箱线图则表示采用确定选择方法的前向粒子滤波算法，虚

线表示接收信号似然函数精确值。从图 4.1 中可以看出随着 N 的增大, 似然函数估计值的方差随之减小。同时, 从图 4.1 中还可以发现在最小相位信道下, 当 N 相同时, 采用确定选择方法的前向粒子滤波算法和采用随机选择方法 1 的前向粒子滤波算法对接收信号似然函数值估计的方差相近。表 4.1 展示了接收信号似然函数值的近似计算与精确计算所用时间, 仿真实验采用 Matlab 软件, 运行在 Intel(R) Xeon(R) Platinum 8163 主频为 2.50GHz 的 CPU 上。从表中可以明显看出, 近似计算接收信号似然函数值的速度快于精确计算。由于确定选择方法无需求解 λ , 所以基于确定选择方法的粒子滤波算法速度快于基于随机选择方法 1 的粒子滤波算法。

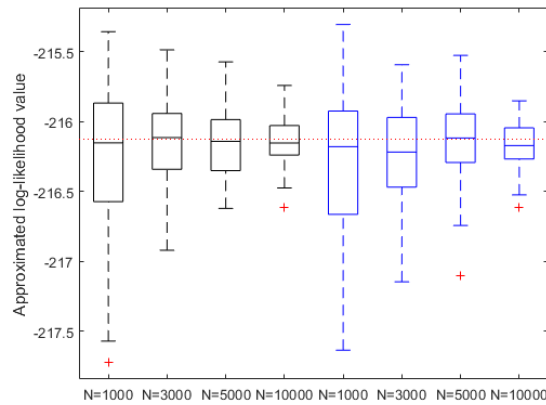


图 4.1 最小相位信道下, 发射机采用 64QAM 时, 接收信号似然函数近似值分布情况

Figure 4.1 Distributions of estimates of likelihood value of received signal when the channel type is Minimum phase and the type of modulation is 64QAM

表 4.1 最小相位信道下，发射机采用 64QAM 时，接收信号似然函数值的近似计算与精确计算时间对比

Table 4.1 Comparison of the time of approximate calculation and exact calculation of likelihood value of received signal when the channel type is Minimum phase and the type of modulation is 64QAM

算法	前向粒子滤波 (随机选择方法 1)		前向迭代算法	前向粒子滤波 (确定选择方法)	
	N=5000	N=10000		N=5000	N=10000
时间（秒）	16.67	35.68	286.99	2.54	5.17

在保持最小相位信道下相关信道参数不变的条件下，我们研究发射机选择 128QAM 作为调制方式时，接收信号似然函数近似计算与精确计算时的性能对比。同样，首先利用前向迭代算法计算出接收信号的准确似然函数值；然后，利用前向粒子滤波算法分别在 $N = 1000, 3000, 5000, 10000$ 时近似接收信号的似然函数值。图 4.2 显示了近似值的分布情况，图中左边四个箱线图代表采用随机选择方法 1 的前向粒子滤波算法近似计算结果，右边四个箱线图则表示采用确定选择方法的前向粒子滤波算法。同样，随着随着 N 的增大，似然函数估计值的方差随之减小。表 4.2 展示了接收信号似然函数值近似计算与精确计算所需要的时间情况。从表 4.1 和 4.2 中可以看出，当调制阶数上升时，前向迭代算法所需时间会大幅度提升。采用前向滤波算法近似计算接收信号似然函数值所需时间远远小于采用前向迭代算法进行精确计算所需的时间。

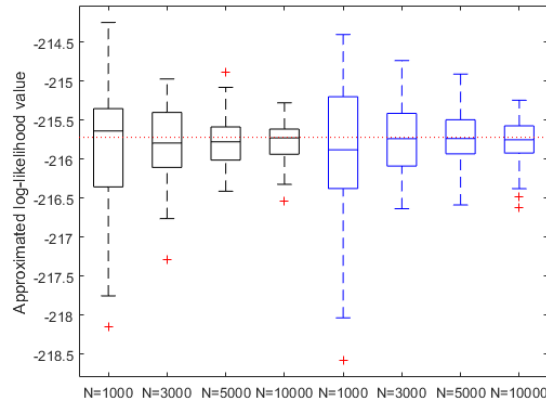


图 4.2 最小相位信道下，发射机采用 128QAM 时，接收信号似然函数近似值分布情况

Figure 4.2 Distributions of estimates of likelihood value of received signal when the channel type is Minimum phase and the type of modulation is 128QAM

表 4.2 最小相位信道下，发射机采用 128QAM 时，接收信号似然函数值的近似计算与精确计算时间对比

Table 4.2 Comparison of the time of approximate calculation and exact calculation of likelihood value of received signal when the channel type is Minimum phase and the type of modulation is 128QAM

算法	前向粒子滤波 (随机选择方法 1)		前向迭代算法	前向粒子滤波 (确定选择方法)	
	N=5000	N=10000		N=5000	N=10000
时间（秒）	102.23	381.61	6958.30	8.10	13.78

在最大相位信道下，取信道系数 $\mathbf{h} = [-0.3, 0.05, 0.8]^T$ ， $\text{SNR} = 15\text{dB}$ ， $T = 200$ 发射机采用 128QAM 调制方式。首先，随机生成一个观察信号序列，利用前向迭代算法精确计算其似然函数值；其次，在 $N = 90000, 120000, 150000$ 时，分别多次采用前向滤波算法（确定选择方法）近似计算接收信号的似然函数值，其结果如图 4.3 所示。从图 4.3 中可以看出随着 N 的增大，似然函数估计值的方差随之减小。结合图 4.2 与 4.3，可以看出在最大相位信道下，需要更大的 N 值才能使

得似然函数近似值方差较小。表 4.3 展示了接收信号似然函数值近似计算与精确计算所需要的时间情况。从表 4.3 中可以看出，采用前向滤波算法（确定选择方法）近似计算接收信号似然函数值所需时间远远小于采用前向迭代算法进行精确计算所需的时间。

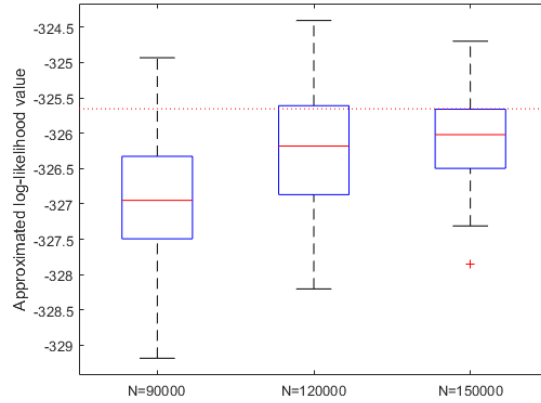


图 4.3 最大相位信道下，发射机采用 128QAM 时，接收信号似然函数近似值分布情况

Figure 4.3 Distributions of estimates of likelihood value of received signal when the channel type is Maximum phase and the type of modulation is 128QAM

表 4.3 最大相位信道下，发射机采用 128QAM 时，接收信号似然函数值的近似计算与精确计算时间对比

Table 4.3 Comparison of the time of approximate calculation and exact calculation of likelihood value of received signal when the channel type is Maximum phase and the type of modulation is 128QAM

算法	前向粒子滤波 (确定选择方法)			前向迭代算法
	N=90000	N=120000	N=150000	
时间（秒）	240.22	259.95	284.49	6979.42

4.4 本章小结

本章首先建立了频率选择性衰落信道下接收信号的生成模型；根据信号模型，

总结了两种精确计算接收信号似然函数值的方案：枚举法和前向迭代算法；在详细分析这两种算法的缺点的基础上，介绍了基于粒子滤波的接收信号似然函数近似计算方案；最后，通过仿真实验分析了粒子滤波算法对接收信号似然函数近似性能，以及比较了近似计算方案与精确计算方案在计算接收信号似然函数值所需要的时间。

第五章 频率选择性衰落高斯噪声信道参数盲估计

第四章中计算接收信号似然函数值的过程需要知道信号模型中的参数值，而在实际通信过程中无法预先知道信号模型中的相关参数值，所以在采用最大似然调制分类方案进行调制识别时，需要在每个假设下进行信道参数的盲估计，然后将估计出的参数值代入上述接收信号似然函数计算的算法中得到每个调制假设的似然函数值。这一章将详细介绍频率选择性衰落高斯噪声信道条件下的信道参数盲估计的相关算法。

在介绍频率选择性衰落高斯噪声信道下的信道参数盲估计算法时，所有的分析都是在假设 H_i 下进行的，为了表达简洁，将在相关表达式中省略 H_i 以及相应的下标 i 。同 3.3 节一样的原因，这里采用 EM 算法来求解参数的最大似然解。

按照 EM 算法的一般思路，首先进行 EM 算法中的求期望这一步，也就是获得 Q 函数的表达式。这里将隐马尔科夫模型的隐藏状态序列 $\mathbf{s}_{1:T}$ 看作隐藏变量，根据 Q 函数的定义以及 4.1 节的信号模型，可以得到 Q 函数的形式如公式 (5.1) 所示。

$$\begin{aligned} Q(\boldsymbol{\theta}; \boldsymbol{\theta}^c) &= E[\ln p(y_{1:T}, \mathbf{s}_{1:T}; \boldsymbol{\theta}) | y_{1:T}, \boldsymbol{\theta}^c] \\ &= \sum_{\mathbf{s}_{1:T}} [\ln p(y_{1:T}; \boldsymbol{\theta} | \mathbf{s}_{1:T}) + \ln P(\mathbf{s}_{1:T})] P(\mathbf{s}_{1:T}; \boldsymbol{\theta}^c | y_{1:T}) \end{aligned} \quad \dots (5.1)$$

其中 $\ln p(y_{1:T}; \boldsymbol{\theta} | \mathbf{s}_{1:T})$ 的表达式如公式 (5.2) 所示。

$$\ln p(y_{1:T}; \boldsymbol{\theta} | \mathbf{s}_{1:T}) = \sum_{t=1}^T \left[-\ln(\pi\sigma^2) - \frac{|y_t - \mathbf{h}^T \mathbf{s}_t|^2}{\sigma^2} \right] \quad \dots (5.2)$$

$\ln P(\mathbf{s}_{1:T})$ 是与待估计参数无关的常数项，其表达式如公式 (5.3) 所示。

$$\ln P(\mathbf{s}_{1:T}) = -(L+T-1) \ln m \quad \dots (5.3)$$

公式 (5.1) - (5.3) 中的字母的定义同 4.1 节的信号模型中的字母， $\boldsymbol{\theta}^c$ 表示当前信道参数的估计值。 $P(\mathbf{s}_{1:T}; \boldsymbol{\theta}^c | y_{1:T})$ 表示在给定接收数据 $y_{1:T}$ 和当前参数估计值 $\boldsymbol{\theta}^c$

时, $\mathbf{s}_{1:T}$ 的后验概率, 关于该后验概率的计算, 是频率选择性衰落高斯噪声信道下参数盲估计的核心, 将在下文中重点说明。

EM 算法的第二步是在参数空间中寻找能够最大化 Q 函数的参数值。利用 $Q(\boldsymbol{\theta}; \boldsymbol{\theta}^c)$ 分别对 \mathbf{h}, σ^2 求偏导, 并令导数为零求解方程可得 \mathbf{h}, σ^2 的更新公式, 如公式 (5.4) (5.5) 所示。

$$\left(\sum_{\mathbf{s}_{1:T}} g(\mathbf{s}_{1:T}) P(\mathbf{s}_{1:T}; \boldsymbol{\theta}^c | y_{1:T}) \right) (\mathbf{h}^+)^* = \sum_{\mathbf{s}_{1:T}} z(\mathbf{s}_{1:T}) P(\mathbf{s}_{1:T}; \boldsymbol{\theta}^c | y_{1:T}) \quad \dots (5.4)$$

$$(\sigma^2)^+ = \frac{1}{T} \sum_{\mathbf{s}_{1:T}} d(\mathbf{s}_{1:T}) P(\mathbf{s}_{1:T}; \boldsymbol{\theta}^c | y_{1:T}) \quad \dots (5.5)$$

其中 $g(\mathbf{s}_{1:T}) \triangleq \sum_{t=1}^T \mathbf{s}_t \mathbf{s}_t^H$, $z(\mathbf{s}_{1:T}) \triangleq \sum_{t=1}^T \mathbf{s}_t y_t^*$, $d(\mathbf{s}_{1:T}) \triangleq \sum_{t=1}^T |y_t - (\mathbf{h}^+)^T \mathbf{s}_t|^2$ 。参数加上标“+”表示更新后的参数值, 上标“*”表示共轭运算, 上标“H”表示共轭转置运算。

由此可以看出计算隐藏状态序列的后验概率 $P(\mathbf{s}_{1:T}; \boldsymbol{\theta}^c | y_{1:T})$ 对于信道参数估计的重要性, 下面我们将介绍三种算法来计算该值: FB 算法[90]、基于粒子平滑近似算法[91]和基于 RGSE 近似算法[92][93]。

5.1 FB 算法

以分析公式 (5.4) 等号左边项 $\sum_{\mathbf{s}_{1:T}} \left(\sum_{t=1}^T \mathbf{s}_t \mathbf{s}_t^H \right) P(\mathbf{s}_{1:T}; \boldsymbol{\theta}^c | y_{1:T})$ 为例来说明 FB 算法的具体步骤。通过调整求和顺序和利用全概率公式, 可以得到公式 (5.6)。

$$\begin{aligned} & \sum_{\mathbf{s}_{1:T}} \left(\sum_{t=1}^T \mathbf{s}_t \mathbf{s}_t^H \right) P(\mathbf{s}_{1:T}; \boldsymbol{\theta}^c | y_{1:T}) \\ &= \sum_{t=1}^T \sum_{\mathbf{s}_{1:T}} \mathbf{s}_t \mathbf{s}_t^H P(\mathbf{s}_{1:T}; \boldsymbol{\theta}^c | y_{1:T}) \\ &= \sum_{t=1}^T \sum_{\mathbf{s}_t} \mathbf{s}_t \mathbf{s}_t^H \left[\sum_{\mathbf{s}_1} \dots \sum_{\mathbf{s}_{t-1}} \sum_{\mathbf{s}_{t+1}} \dots \sum_{\mathbf{s}_T} P(\mathbf{s}_{1:T}; \boldsymbol{\theta}^c | y_{1:T}) \right] \\ &= \sum_{t=1}^T \sum_{\mathbf{s}_t} \mathbf{s}_t \mathbf{s}_t^H P(\mathbf{s}_t; \boldsymbol{\theta}^c | y_{1:T}) \end{aligned} \quad \dots (5.6)$$

其中 \mathbf{s}_t 的取值为隐马尔科夫模型的状态空间 S 中的所有状态，根据 4.1 节信号模型可知 S 中的状态数为 m^L ，我们用 $S(n)$ 表示 S 中的第 n 个状态，公式 (5.6) 可以整理为公式 (5.7)。

$$\sum_{\mathbf{s}_{1:T}} \left(\sum_{t=1}^T \mathbf{s}_t \mathbf{s}_t^H \right) P(\mathbf{s}_{1:T}; \boldsymbol{\theta}^c | y_{1:T}) = \sum_{t=1}^T \sum_{n=1}^{m^L} S(n) S(n)^H P(\mathbf{s}_t = S(n); \boldsymbol{\theta}^c | y_{1:T}) \quad \dots (5.7)$$

利用 FB 算法计算公式 (5.7) 中 $P(\mathbf{s}_t = S(n); \boldsymbol{\theta}^c | y_{1:T})$ 时，首先定义前向迭代变量 $\alpha_t(n) \triangleq p(y_{1:t}, \mathbf{s}_t = S(n); \boldsymbol{\theta}^c)$ 和后向迭代变量 $\beta_t(n) \triangleq p(y_{t+1:T}; \boldsymbol{\theta}^c | \mathbf{s}_t = S(n))$ ，利用贝叶斯定理可以得到 $P(\mathbf{s}_t = S(n); \boldsymbol{\theta}^c | y_{1:T})$ 的计算方法，如公式 (5.8) 所示。

$$P(\mathbf{s}_t = S(n); \boldsymbol{\theta}^c | y_{1:T}) = \frac{\alpha_t(n) \beta_t(n)}{\sum_{n=1}^{m^L} \alpha_t(n) \beta_t(n)} \quad \dots (5.8)$$

前向迭代变量 $\alpha_t(n)$ 的迭代公式如公式 (4.7) 所示，后向迭代变量 $\beta_t(n)$ 的迭代公式如公式 (5.9) 所示。

$$\begin{aligned} \beta_t(n) &= p(y_{t+1:T}; \boldsymbol{\theta}^c | \mathbf{s}_t = S(n)) \\ &= \sum_{\mathbf{s}_{t+1}} p(y_{t+1:T}, \mathbf{s}_{t+1}; \boldsymbol{\theta}^c | \mathbf{s}_t = S(n)) \\ &= \sum_k p(y_{t+1}; \boldsymbol{\theta}^c | \mathbf{s}_{t+1} = S(k)) P(\mathbf{s}_{t+1} = S(k) | \mathbf{s}_t = S(n)) \beta_{t+1}(k) \end{aligned} \quad \dots (5.9)$$

参照 4.2.2 节中的缩放前向迭代算法计算前向迭代变量 $\alpha_t(n)$ ，得到缩放后相关的量 $c_t, \hat{\alpha}_t(n)$ ， $t=1, \dots, T, n=1, \dots, m^L$ ，这些符号的含义与 4.2.2 节相同。同样由于 $\beta_t(n)$ 的计算也会超出计算机的计算精度，所以这里也对 $\beta_t(n)$ 进行缩放计算。我们采用 c_t 对 $\beta_t(n)$ 进行缩放，其计算过程为：

1) 初始化：

初始化 $\beta_T(n) = 1$ ，并将 $\beta_T(n)$ 赋值给 $\tilde{\beta}_T(n)$ ： $\tilde{\beta}_T(n) = \beta_T(n)$ ，其中 $n=1, \dots, m^L$ 。

用 c_T 对 $\tilde{\beta}_T(n)$ 缩放得 $\hat{\beta}_T(n) = c_T \tilde{\beta}_T(n)$ 。

2) 迭代：

根据公式 (5.9)，利用 $\hat{\beta}_{t+1}(n)$ 迭代计算得到 $\tilde{\beta}_t(n)$ ，用 c_t 对 $\tilde{\beta}_t(n)$ 缩放得

$\hat{\beta}_t(n) = c_t \tilde{\beta}_t(n)$ 。

由 $\hat{\beta}_t(n)$ 的迭代计算过程可以得出其与未经缩放的 $\beta_t(n)$ 数值关系为:

$\hat{\beta}_t(n) = \prod_{\tau=t}^T c_\tau \beta_t(n)$ 。将缩放后的前向迭代变量 $\hat{\alpha}_t(n)$ 和后向迭代变量 $\hat{\beta}_t(n)$ 代入公式 (5.8) 等号的右边项, 可得公式 (5.10)。

$$\begin{aligned} \frac{\hat{\alpha}_t(n) \hat{\beta}_t(n)}{\sum_{n=1}^{m^L} \hat{\alpha}_t(n) \hat{\beta}_t(n)} &= \frac{\prod_{\tau=1}^t c_\tau \alpha_t(n) \prod_{\tau=t}^T c_\tau \beta_t(n)}{\sum_{n=1}^{m^L} \prod_{\tau=1}^t c_\tau \alpha_t(n) \prod_{\tau=t}^T c_\tau \beta_t(n)} \\ &= \frac{\alpha_t(n) \beta_t(n)}{\sum_{n=1}^{m^L} \alpha_t(n) \beta_t(n)} \quad \dots (5.10) \\ &= P(\mathbf{s}_t = S(n); \boldsymbol{\theta}^c | y_{1:T}) \end{aligned}$$

公式 (5.10) 的等量关系说明可以直接利用缩放后的前向后向迭代变量计算隐藏状态的后验概率 $P(\mathbf{s}_t = S(n); \boldsymbol{\theta}^c | y_{1:T})$ 。

综上所述, 采用 FB 算法计算 $P(\mathbf{s}_t = S(n); \boldsymbol{\theta}^c | y_{1:T})$ 的具体步骤为:

- 1) 计算缩放后的前向迭代变量 $\hat{\alpha}_t(n)$, 以及保存缩放系数 c_t , $t=1, \dots, T$ 。
- 2) 利用缩放系数 c_t 计算缩放后的后向迭代变量 $\hat{\beta}_t(n)$, $t=T, T-1, \dots, 1$ 。
- 3) 根据公式 (5.10) 利用 $\hat{\alpha}_t(n)$ 和 $\hat{\beta}_t(n)$ 计算隐藏状态 \mathbf{s}_t 的后验概率

$$P(\mathbf{s}_t = S(n); \boldsymbol{\theta}^c | y_{1:T})。$$

FB 算法的算法复杂度主要集中在 1) 和 2) 两个步骤中, 由 4.2.2 节中的分析可知

1) 和 2) 的算法复杂度均为 $O(Tm^{L+1})$, 因此利用 FB 算法计算隐藏状态的后验概率的算法复杂度为 $O(2Tm^{L+1})$, 所以当调制阶数 m 或者信道阶数 L 较大时, FB 算法的复杂度将会非常高, 于是[91]提出采用近似计算隐藏状态 \mathbf{s}_t 的后验概率的方法来解决这一问题。

5.2 基于粒子平滑近似算法

由 5.1 节的分析可知, 当状态空间 S 较大时, 利用 FB 算法计算

$P(\mathbf{s}_t = S(n); \boldsymbol{\theta}^c | y_{1:T})$ 算法复杂度会非常高。为了有效地解决这一问题, [91]提出基于串行蒙特卡洛方法近似计算 $P(\mathbf{s}_t = S(n); \boldsymbol{\theta}^c | y_{1:T})$ 的方案。根据后验概率近似方

法不同，主要分为 FIS、TFS 和 JTFS 三种方法，下面将详细介绍这三种方法的原理和实现方法。

为了简化公式表达，这里引入了一些符号：当 $1 \leq k \leq j \leq T$ 时， $\mathbf{s}_{k:j} = (\mathbf{s}_k, \dots, \mathbf{s}_j)$ 表示隐马尔可夫模型中隐藏状态序列， $y_{k:j} = (y_k, \dots, y_j)$ 表示接收信号序列。由于介绍粒子滤波方法时会频繁使用到概率表达式，引入缩写

$\pi_{k|j}(\mathbf{s}; \boldsymbol{\theta}^c) \triangleq P(\mathbf{s}_k = \mathbf{s}; \boldsymbol{\theta}^c | y_{1:j})$ 以及缩写 $\pi_{k|j:n}(\mathbf{s}; \boldsymbol{\theta}^c) \triangleq P(\mathbf{s}_k = \mathbf{s}; \boldsymbol{\theta}^c | y_{1:n})$ 来简化表示。

5.2.1 FIS 算法

FIS 算法首先通过粒子滤波方法得到 \mathbf{s}_t 的边缘滤波分布 $\pi_{t|t}(\cdot; \boldsymbol{\theta}^c)$, $1 \leq t \leq T$ 的近似值，然后利用反向迭代矫正滤波分布中每个粒子对应的权重得到 \mathbf{s}_t 的平滑分布 $\pi_{t|T}(\cdot; \boldsymbol{\theta}^c)$ 的近似值。

这里前向粒子滤波部分的实现方法同 4.2.3 节的前向粒子滤波，这里不再赘述。

因为边缘平滑分布 $\pi_{t|T}(\cdot; \boldsymbol{\theta}^c)$ 可以分解成如公式 (5.11) 所示的形式，

$$\begin{aligned}
 \pi_{t|T}(\mathbf{s}; \boldsymbol{\theta}^c) &= \sum_{\mathbf{x}} P(\mathbf{s}_t = \mathbf{s}, \mathbf{s}_{t+1} = \mathbf{x}; \boldsymbol{\theta}^c | y_{1:T}) \\
 &= \sum_{\mathbf{x}} \pi_{t+1|T}(\mathbf{x}; \boldsymbol{\theta}^c) P(\mathbf{s}_t = \mathbf{s}; \boldsymbol{\theta}^c | \mathbf{s}_{t+1} = \mathbf{x}, y_{1:t}) \\
 &= \sum_{\mathbf{x}} \pi_{t+1|T}(\mathbf{x}; \boldsymbol{\theta}^c) \frac{P(\mathbf{s}_t = \mathbf{s}, \mathbf{s}_{t+1} = \mathbf{x}; \boldsymbol{\theta}^c | y_{1:t})}{P(\mathbf{s}_{t+1} = \mathbf{x}; \boldsymbol{\theta}^c | y_{1:t})} \quad \dots (5.11) \\
 &= \pi_{t|t}(\mathbf{s}; \boldsymbol{\theta}^c) \sum_{\mathbf{x}} \pi_{t+1|T}(\mathbf{x}; \boldsymbol{\theta}^c) \frac{P(\mathbf{s}_{t+1} = \mathbf{x} | \mathbf{s}_t = \mathbf{s})}{\sum_{\mathbf{z}} P(\mathbf{s}_{t+1} = \mathbf{x} | \mathbf{s}_t = \mathbf{z}) \pi_{t|t}(\mathbf{z}; \boldsymbol{\theta}^c)}
 \end{aligned}$$

所以可以利用 t 时刻的滤波分布 $\pi_{t|t}(\mathbf{s}; \boldsymbol{\theta}^c)$ 和 $t+1$ 时刻的平滑分布 $\pi_{t+1|T}(\mathbf{s}; \boldsymbol{\theta}^c)$ 计算出 t 时刻的平滑分布 $\pi_{t|T}(\mathbf{s}; \boldsymbol{\theta}^c)$ 。通过前向粒子滤波，已经得到了近似滤波分布 $\hat{\pi}_{t|t}(\mathbf{s}; \boldsymbol{\theta}^c)$, $t=1, \dots, T$ ，将其和近似平滑分布 $\hat{\pi}_{t+1|T}(\mathbf{s}; \boldsymbol{\theta}^c)$ 带入公式 (5.11) 可以得到近似平滑分布迭代公式。我们将 t 时刻的近似平滑分布表示为公式 (5.12) 所示的形式。

$$\hat{\pi}_{t|T}(\mathbf{s}; \boldsymbol{\theta}^c) = \sum_{i=0}^{N-1} w_{t|T}^i \delta_{\mathbf{s}_t^i}(\mathbf{s}) \quad \dots (5.12)$$

其中 $\{\xi_t^i\}_{i < N}$ 为近似 t 时刻的滤波分布时所用的粒子，根据公式 (5.11)， ξ_t^i 所对应的平滑权重 $w_{t|T}^i$ 可由迭代公式 (5.13) 计算得到。

$$\begin{aligned} w_{t|T}^i &= w_t^i \sum_{j=0}^{N-1} \frac{w_{t+1|T}^j P(\mathbf{s}_{t+1} = \xi_{t+1}^j | \mathbf{s}_t = \xi_t^i)}{\sum_{l=0}^{N-1} w_t^l P(\mathbf{s}_{t+1} = \xi_{t+1}^l | \mathbf{s}_t = \xi_t^i)} \\ &= w_t^i \sum_{j \in \mathcal{Q}(\xi_t^i)} \frac{w_{t+1|T}^j}{\sum_{l \in \mathcal{P}(\xi_{t+1}^j)} w_t^l} \end{aligned} \quad \dots (5.13)$$

其中 $\mathcal{Q}(\xi_t^i)$ 表示在 $t+1$ 时刻的所有粒子中， ξ_t^i 的子代粒子的序号集合， $\mathcal{P}(\xi_{t+1}^j)$ 表示 t 时刻的所有粒子中， ξ_{t+1}^j 的先代粒子的序号集合。在更新平滑分布粒子权重时，同样可以通过对 t 时刻粒子去掉最后一个元素后排序， $t+1$ 时刻粒子去掉第一个元素后排序的方法找到每个粒子对应的子代集合或先代集合，便于计算。排序的算法复杂度为 $O(N \log_2 N)$ ，明显优于一般的状态模型中的反向迭代算法复杂度 $O(N^2)$ 。

综上，FIS 算法近似隐藏状态后验概率主要包括前向滤波和反向矫正权重两个步骤。前向滤波部分， $t=1$ 时的初始化时刻的算法复杂度为 $O(N)$ ；在之后的迭代过程中，每一步的算法复杂度为 $O(Nm)$ ；在反向迭代更新粒子权重的部分，每一步的算法复杂度为 $O(N \log_2 N)$ ，因为一般情况下 $\log_2 N \ll m$ ，所以反向更新系数的算法复杂度低于前向迭代近似滤波分布的算法复杂度。所以 FIS 算法在近似计算 $P(\mathbf{s}_t = ; \boldsymbol{\theta}^c | y_{1:T}), t=1, \dots, T$ 的算法复杂度为 $O(TNm)$ 。

5.2.2 TFS 算法

FIS 算法有一个缺点：近似平滑分布的粒子完全由前向滤波分布的粒子决定。所以当准确的平滑分布中某些权重较大的粒子所对应的滤波分布权重较低时，在前向滤波迭代中，很可能采样不到这些粒子，从而导致平滑分布近似误差较大。

在我们所研究的离散马尔科夫模型中，不仅 $(\mathbf{s}_t, y_t), t=1, \dots, T$ 形成隐马尔科夫链，同时反向时间序列 $(\mathbf{s}_t, y_t), t=T, \dots, 1$ 也是一个隐马尔科夫链，其原因在于给定 \mathbf{s}_{t+1} 时， $P(\mathbf{s}_t | \mathbf{s}_{t+1})$ 是关于 \mathbf{s}_t 的概率分布，满足 $\sum_{\mathbf{s}_t} P(\mathbf{s}_t | \mathbf{s}_{t+1}) = 1$ 。所以可以针对反

向时间序列采用粒子滤波，近似其滤波分布 $\pi_{t|t:T}(\mathbf{s}; \boldsymbol{\theta}^c)$, $t = T, \dots, 1$ ，也称其为反向滤波分布，可借此分析未来的观察序列对当前隐藏状态的影响。基于公式(5.14)的分解可以迭代获得反向滤波分布。

$$\begin{aligned}
 & \pi_{t|t:T}(\mathbf{s}; \boldsymbol{\theta}^c) \\
 &= \sum_{\mathbf{x}} P(\mathbf{s}_t = \mathbf{s}, \mathbf{s}_{t+1} = \mathbf{x}; \boldsymbol{\theta}^c | y_{t:T}) \\
 &\propto \sum_{\mathbf{x}} P(\mathbf{s}_t = \mathbf{s}, \mathbf{s}_{t+1} = \mathbf{x}, y_{t:T}; \boldsymbol{\theta}^c) \quad \dots (5.14) \\
 &\propto \sum_{\mathbf{x}} p(y_{t+1:T}; \boldsymbol{\theta}^c | \mathbf{s}_{t+1} = \mathbf{x}) P(\mathbf{s}_{t+1} = \mathbf{x} | \mathbf{s}_t = \mathbf{s}) p(y_t; \boldsymbol{\theta}^c | \mathbf{s}_t = \mathbf{s}) \\
 &\propto p(y_t; \boldsymbol{\theta}^c | \mathbf{s}_t = \mathbf{s}) \sum_{\mathbf{x}} \pi_{t+1|t+1:T}(\mathbf{x}; \boldsymbol{\theta}^c) P(\mathbf{s}_{t+1} = \mathbf{x} | \mathbf{s}_t = \mathbf{s})
 \end{aligned}$$

假设利用前向粒子滤波迭代得到粒子以及粒子对应的权重集合，表示为 $\{\xi_{t|t}^i, w_{t|t}^i\}_{i < N_F}$ ，其中 N_F 表示前向滤波采用的粒子数。所以前向滤波分布 $\pi_{t|t}(\cdot; \boldsymbol{\theta}^c)$ 可近似为 $\hat{\pi}_{t|t}(\cdot; \boldsymbol{\theta}^c) = \sum_{i=0}^{N_F-1} w_{t|t}^i \delta_{\xi_{t|t}^i}^i$, $t = 1, \dots, T$ 。同样，假设利用反向粒子滤波迭代得到粒子及其权重集合 $\{\xi_{t|t:T}^i, w_{t|t:T}^i\}_{i < N_B}$ ，其中 N_B 表示反向滤波采用的粒子数。于是反向滤波分布 $\pi_{t|t:T}(\cdot; \boldsymbol{\theta}^c)$ 可近似为 $\hat{\pi}_{t|t:T}(\cdot; \boldsymbol{\theta}^c) = \sum_{i=0}^{N_B-1} w_{t|t:T}^i \delta_{\xi_{t|t:T}^i}^i$, $t = T, \dots, 1$ 。这里前向粒子滤波和反向粒子滤波是完全独立进行的。

TFS 基于前向粒子滤波和反向粒子滤波，根据平滑分布分解方式的不同，可以使近似平滑分布的粒子完全与前向滤波分布中的粒子一致或者完全与反向滤波分布中的粒子一致。

使得近似平滑分布的粒子完全与前向滤波分布中的粒子一致的分解方式如公式(5.15)所示。

$$\begin{aligned}
 \pi_{t|T}(\mathbf{s}; \boldsymbol{\theta}^c) &\propto p(y_{1:T}, \mathbf{s}_t = \mathbf{s}; \boldsymbol{\theta}^c) \\
 &\propto p(y_{t+1:T}; \boldsymbol{\theta}^c | \mathbf{s}_t = \mathbf{s}) P(\mathbf{s}_t = \mathbf{s}; \boldsymbol{\theta}^c | y_{1:t}) \\
 &\propto p(\mathbf{s}_t = \mathbf{s}; \boldsymbol{\theta}^c | y_{t+1:T}) P(\mathbf{s}_t = \mathbf{s}; \boldsymbol{\theta}^c | y_{1:t}) \quad \dots (5.15) \\
 &= \pi_{t|t}(\mathbf{s}; \boldsymbol{\theta}^c) \sum_{\mathbf{x}} P(\mathbf{s}_t = \mathbf{s} | \mathbf{s}_{t+1} = \mathbf{x}) \pi_{t+1|t+1:T}(\mathbf{x}; \boldsymbol{\theta}^c)
 \end{aligned}$$

将 $\pi_{t|t}(\cdot; \boldsymbol{\theta}^c)$ 的近似值 $\hat{\pi}_{t|t}(\cdot; \boldsymbol{\theta}^c)$ 和 $\pi_{t+1|t+1:T}(\cdot; \boldsymbol{\theta}^c)$ 的近似值 $\hat{\pi}_{t+1|t+1:T}(\cdot; \boldsymbol{\theta}^c)$ 带入到公式

(5.15)中便可得到平滑分布 $\pi_{t|T}(\cdot; \boldsymbol{\theta}^c)$ 的近似值 $\hat{\pi}_{t|T}(\cdot; \boldsymbol{\theta}^c)$ ，其可表示为公式(5.16)。

$$\begin{aligned}
 \hat{\pi}_{t|T}(\mathbf{s}; \boldsymbol{\theta}) &\propto \sum_{i=0}^{N_F-1} w_{t|t}^i \delta_{\xi_{t|t}^i}(\mathbf{s}) \sum_{j=0}^{N_B-1} P(\mathbf{s}_t = \mathbf{s} | \mathbf{s}_{t+1} = \xi_{t+1|t+1:T}^j) w_{t+1|t+1:T}^j \\
 &= \sum_{i=0}^{N_F-1} w_{t|T}^i \delta_{\xi_{t|t}^i}(\mathbf{s})
 \end{aligned} \quad \dots (5.16)$$

其中 $P(\mathbf{s}_t = \mathbf{s} | \mathbf{s}_{t+1} = \mathbf{x})$ 的计算方式同 $P(\mathbf{s}_{t+1} = \mathbf{x} | \mathbf{s}_t = \mathbf{s})$ ，近似平滑分布的粒子的权重（未归一化） $w_{t|T}^i = w_{t|t}^i \sum_{j=0}^{N_B-1} P(\mathbf{s}_{t+1} = \xi_{t+1|t+1:T}^j | \mathbf{s}_t = \xi_{t|t}^i) w_{t+1|t+1:T}^j$ 。这里近似平滑分布所用的粒子与前向滤波中的粒子一致，权重的更新方式结合了前向滤波分布与反向滤波分布，这一点不同于 FIS 算法。

类似地，依据公式（5.17）对平滑分布 $\pi_{t|T}(\mathbf{s}; \boldsymbol{\theta}^c)$ 进行分解可以使近似平滑分布的粒子与反向滤波分布中的粒子一致。

$$\begin{aligned}
 \pi_{t|T}(\mathbf{s}; \boldsymbol{\theta}^c) &\propto p(y_{1:t-1}, y_{t:T}, \mathbf{s}_t = \mathbf{s}; \boldsymbol{\theta}^c) \\
 &\propto p(y_{t:T}; \boldsymbol{\theta}^c | \mathbf{s}_t = \mathbf{s}) P(\mathbf{s}_t = \mathbf{s}; \boldsymbol{\theta}^c | y_{1:t-1}) \\
 &\propto p(\mathbf{s}_t = \mathbf{s}; \boldsymbol{\theta}^c | y_{t:T}) P(\mathbf{s}_t = \mathbf{s}; \boldsymbol{\theta}^c | y_{1:t-1}) \\
 &= \pi_{t|t:T}(\mathbf{s}; \boldsymbol{\theta}^c) \sum_{\mathbf{x}} P(\mathbf{s}_t = \mathbf{s} | \mathbf{s}_{t-1} = \mathbf{x}) \pi_{t-1|t-1}(\mathbf{x}; \boldsymbol{\theta}^c)
 \end{aligned} \quad \dots (5.17)$$

同样将 $\pi_{t|t:T}(\cdot; \boldsymbol{\theta}^c)$ 的近似值 $\hat{\pi}_{t|t:T}(\cdot; \boldsymbol{\theta}^c)$ 和 $\pi_{t-1|t-1}(\cdot; \boldsymbol{\theta}^c)$ 的近似值 $\hat{\pi}_{t-1|t-1}(\cdot; \boldsymbol{\theta}^c)$ 带入到公式（5.17）可以得到 $\pi_{t|T}(\cdot; \boldsymbol{\theta}^c)$ 的近似值 $\hat{\pi}_{t|T}(\cdot; \boldsymbol{\theta}^c)$ ，如公式（5.18）所示。

$$\begin{aligned}
 \hat{\pi}_{t|T}(\mathbf{s}; \boldsymbol{\theta}^c) &\propto \sum_{i=0}^{N_B-1} w_{t|t:T}^i \delta_{\xi_{t|t:T}^i}(\mathbf{s}) \sum_{j=0}^{N_F-1} P(\mathbf{s}_t = \xi_{t|t:T}^i | \mathbf{s}_{t-1} = \xi_{t-1|t-1}^j) w_{t-1|t-1}^j \\
 &= \sum_{i=0}^{N_B-1} \bar{w}_{t|T}^i \delta_{\xi_{t|t:T}^i}(\mathbf{s})
 \end{aligned} \quad \dots (5.18)$$

其中权重（未归一化） $\bar{w}_{t|T}^i = w_{t|t:T}^i \sum_{j=0}^{N_F-1} P(\mathbf{s}_t = \xi_{t|t:T}^i | \mathbf{s}_{t-1} = \xi_{t-1|t-1}^j) w_{t-1|t-1}^j$ 。

TFS 算法的这两种平滑分布分解的方法，都是基于独立的前向滤波和反向滤波，权重的计算结合了前向滤波分布和反向滤波分布。但是这两种分解方法使得近似平滑分布的粒子要么基于前向滤波粒子，要么基于反向滤波粒子，依旧会出现平滑分布的估计误差较大的情形。

TFS 算法的计算量主要存在于前向滤波和反向滤波，前向滤波和反向滤波的计算复杂度同 FIS 中所分析的，分别为 $O(TN_F m)$ 和 $O(TN_B m)$ 。

由于该信号模型中转移矩阵的稀疏特性，利用上述的 TFS 算法估计得到的近似平滑分布在某些时刻 t 会出现平滑权重都为 0 的现象，例如对任意 j, i 的组合， $P(\mathbf{s}_{t+1} = \xi_{t+1|t+1:T}^j | \mathbf{s}_t = \xi_{t|t}^i) = 0$ ，此时便会出现平滑权重都为 0 的情形。定义 $\xi_{t|t}^i \triangleq [a_{t,0}^i, \dots, a_{t,L-1}^i]^T$ ， $\xi_{t+1|t+1:T}^j \triangleq [b_{t+1,0}^j, \dots, b_{t+1,L-1}^j]^T$ ，只有当 $\xi_{t|t}^i$ 的前 $L-1$ 个元素与 $\xi_{t+1|t+1:T}^j$ 的后 $L-1$ 个元素一致时，也就是 $[a_{t,0}^i, \dots, a_{t,L-2}^i] = [b_{t+1,1}^j, \dots, b_{t+1,L-1}^j]$ 转移概率才为非零，这就是某些时刻平滑权重全为零的原因所在。为了解决这一问题，可以利用 $t_l = t-l$ 时刻的前向滤波分布和 $t_u = t+u$ 时刻的反向滤波分布来计算平滑分布 $\pi_{t|T}(\mathbf{s}; \boldsymbol{\theta}^c)$ ，只要满足 $l+u \geq L$ ，即使 t_l 时刻的粒子中任何元素都不与 t_u 时刻的粒子中元素一致，也不会导致平滑分布的权重全为零。公式 (5.19) 展示了如何利用 t_l 时刻的前向滤波分布和 t_u 时刻的反向滤波分布来计算 t 时刻的平滑分布。

$$\begin{aligned}
 & P(\mathbf{s}_t; \boldsymbol{\theta}^c | y_{1:T}) \\
 & \propto p(y_{1:T}, \mathbf{s}_t = \mathbf{s}; \boldsymbol{\theta}^c) \\
 & = \sum_{\mathbf{s}_{t_u}} \sum_{\mathbf{s}_{t_l}} p(y_{t_u:T}, \mathbf{s}_{t_u}, y_{t:t_u-1}, \mathbf{s}_t, y_{t_l+1:t-1}, \mathbf{s}_{t_l}, y_{1:t_l}; \boldsymbol{\theta}^c) \\
 & \propto \sum_{\mathbf{s}_{t_u}} \sum_{\mathbf{s}_{t_l}} (P(\mathbf{s}_{t_u}; \boldsymbol{\theta}^c | y_{t_u:T}) p(\mathbf{s}_{t_u}, y_{t:t_u-1}; \boldsymbol{\theta}^c | \mathbf{s}_t) \\
 & \quad \times p(\mathbf{s}_t, y_{t_l+1:t-1}; \boldsymbol{\theta}^c | \mathbf{s}_{t_l}) P(\mathbf{s}_{t_l}; \boldsymbol{\theta}^c | y_{1:t_l})) \quad \dots (5.19) \\
 & = \sum_{\mathbf{s}_{t+L_u}} \sum_{\mathbf{s}_{t_l+1}} (P(\mathbf{s}_{t_u}; \boldsymbol{\theta}^c | y_{t_u:T}) p(\mathbf{s}_{t+L_u}, y_{t:t_u-1}; \boldsymbol{\theta}^c | \mathbf{s}_t) \\
 & \quad \times p(\mathbf{s}_{t_l+1:t-1}, \mathbf{s}_t, y_{t_l+1:t-1}; \boldsymbol{\theta}^c | \mathbf{s}_{t_l}) P(\mathbf{s}_{t_l}; \boldsymbol{\theta}^c | y_{1:t_l})) \\
 & = \sum_{\mathbf{s}_{t+L_u}} \sum_{\mathbf{s}_{t_l+1}} (P(\mathbf{s}_{t_u}; \boldsymbol{\theta}^c | y_{t_u:T}) \prod_{\tau=t_l+1}^{t_u-1} p(y_\tau; \boldsymbol{\theta}^c | \mathbf{s}_\tau) P(\mathbf{s}_{\tau+1} | \mathbf{s}_\tau) \\
 & \quad \times P(\mathbf{s}_{t_l+1} | \mathbf{s}_{t_l}) P(\mathbf{s}_{t_l}; \boldsymbol{\theta}^c | y_{1:t_l}))
 \end{aligned}$$

当 $l+u = L$ 时， \mathbf{s}_{t_l} 和 \mathbf{s}_{t_u} 唯一确定 \mathbf{s}_t ， \mathbf{s}_t 的后 $L-l$ 个元素与 \mathbf{s}_{t_l} 的前 $L-l$ 个元素一致， \mathbf{s}_t 的前 $L-u$ 个元素与 \mathbf{s}_{t_u} 的后 $L-u$ 个元素一致。所以在 t 时刻，可由 t_l 时刻的 N_F 个粒子和 t_u 时刻的 N_B 个粒子组合形成 $N_F \times N_B$ 个粒子。虽然利用公式 (5.19) 的分解可以解决某些时刻平滑分布中粒子权重全部为零的问题，但是其在计算权重时的算法复杂度为 $O(N_F \times N_B)$ ，极大地增加了运算量。

5.2.3 JTFS 算法

在前文提及的两种近似平滑分布的算法中, 近似平滑分布的粒子要么来源于前向滤波所得粒子, 要么来源于反向滤波所得粒子。当信道类型是最小相位信道时, 利用前向滤波所得的粒子能够较好地近似平滑分布; 当信道类型是最大相位信道时, 利用反向滤波所得的粒子能够较好地近似平滑分布。如果提前知道信道类型, 可以选择采用前向滤波或者反向滤波, 但是在实际通信过程中, 无法预先知道确切的信道类型。JTFS 算法尝试解决这一问题, 该算法的核心思想是将前向滤波所得粒子与反向滤波所得粒子结合, 更新前向滤波分布和反向滤波分布的近似值, 然后利用 TFS 算法计算近似平滑分布。通过将前向滤波粒子和反向滤波粒子的集合, 不论当信道类型是最小或最大相位时, 利用更新后的前向滤波分布或反向滤波分布中的粒子都能较好的近似平滑分布。

为了描述方便, 用前向滤波分布中的粒子和其对应的未归一化的权重 (由公式 (4.11) 计算而得) 集合 $\{\xi_{t|t}^i, w_{t|t}^i\}_{i < N_F}$ 来表示 $\hat{\pi}_{t|t}'(\mathbf{s}; \boldsymbol{\theta}^c) = \sum_{i=0}^{N_F-1} w_{t|t}^i \delta_{\xi_{t|t}^i}(\mathbf{s})$ 。 $\hat{\pi}_{t|t}'(\mathbf{s}; \boldsymbol{\theta}^c)$ 是公式 (5.20) 中 $\pi_{t|t}'(\mathbf{s}; \boldsymbol{\theta}^c)$ 的近似值。

$$\pi_{t|t}'(\mathbf{s}; \boldsymbol{\theta}^c) = \sum_{\mathbf{s}} \pi_{t-1|t-1}'(\mathbf{s}; \boldsymbol{\theta}^c) P(\mathbf{s}_t = \mathbf{s} | \mathbf{s}_{t-1} = \mathbf{s}') p(y_t; \boldsymbol{\theta}^c | \mathbf{s}_t = \mathbf{s}) \quad \dots (5.20)$$

其中 $\pi_{t-1|t-1}'(\cdot; \boldsymbol{\theta}^c)$ 表示前向滤波分布。这里再定义 $\Omega_t \triangleq \sum_{i=0}^{N_F-1} w_{t|t}^i$, 很明显分布

$\hat{\pi}_{t|t}'(\mathbf{s}; \boldsymbol{\theta}^c) / \Omega_t$ 是前向滤波分布 $\pi_{t|t}'(\mathbf{s}; \boldsymbol{\theta}^c)$ 的近似值。

同样地, 用反向滤波分布中的粒子和其对应的未归一化的权重集合 (基于公式 (5.14) 的分解计算而得) $\{\xi_{t|t:T}^i, w_{t|t:T}^i\}_{i < N_B}$ 来表示 $\hat{\pi}_{t|t:T}'(\mathbf{s}; \boldsymbol{\theta}^c) = \sum_{i=0}^{N_B-1} w_{t|t:T}^i \delta_{\xi_{t|t:T}^i}(\mathbf{s})$ 。 $\hat{\pi}_{t|t:T}'(\mathbf{s}; \boldsymbol{\theta}^c)$ 是公式 (5.21) 中 $\pi_{t|t:T}'(\mathbf{s}; \boldsymbol{\theta}^c)$ 的近似值。

$$\pi_{t|t:T}'(\mathbf{s}; \boldsymbol{\theta}^c) = p(y_t; \boldsymbol{\theta}^c | \mathbf{s}_t = \mathbf{s}) \sum_{\mathbf{s}'} P(\mathbf{s}_{t+1} = \mathbf{s}' | \mathbf{s}_t = \mathbf{s}) \pi_{t+1|t+1:T}'(\mathbf{s}'; \boldsymbol{\theta}^c) \quad \dots (5.21)$$

其中 $\pi_{t+1|t+1:T}'(\cdot; \boldsymbol{\theta}^c)$ 表示反向滤波分布。定义 $\Omega_{t|t:T} \triangleq \sum_{i=0}^{N_B-1} w_{t|t:T}^i$, 分布 $\hat{\pi}_{t|t:T}'(\mathbf{s}; \boldsymbol{\theta}^c) / \Omega_{t|t:T}$ 是反向滤波分布 $\pi_{t|t:T}'(\mathbf{s}; \boldsymbol{\theta}^c)$ 的近似值。

首先说明如何利用 $\{\xi_{t|t}^i, w_{t|t}^i\}_{i < N_F}$ 和 $\{\xi_{t|t:T}^i, w_{t|t:T}^i\}_{i < N_B}$ 更新 $\pi_{t|t}'(\mathbf{s}; \boldsymbol{\theta}^c)$ 的近似值

$\hat{\pi}_{t|t}'(\mathbf{s}; \boldsymbol{\theta}^c)$, 更新后的近似值记为 $\tilde{\pi}_{t|t}'(\mathbf{s}; \boldsymbol{\theta}^c)$ 。假设已经得到了 $t-1$ 时刻 $\pi_{t-1|t-1}'(\mathbf{s}; \boldsymbol{\theta}^c)$

更新后的近似值 $\tilde{\pi}'_{t-1|t-1}(\mathbf{s}; \boldsymbol{\theta}^c)$ ，其由粒子和权重集合 $\{\tilde{\xi}_{t-1|t-1}^i, \tilde{w}_{t-1|t-1}^i\}_{i < \tilde{N}_{t-1}^F}$ 组成，定义 $\tilde{\Omega}_{t-1}^F \triangleq \sum_{i=0}^{\tilde{N}_{t-1}^F-1} \tilde{w}_{t-1|t-1}^i$ 。利用 t 时刻前向、反向滤波所得粒子集合确定更新后近似前向滤波分布的粒子集合 $\{\tilde{\xi}_{t|t}^i\}_{i < \tilde{N}_t^F} = \{\xi_{t|t}^i\}_{i < N_F} \cup \{\xi_{t|t:T}^i\}_{i < N_B}$ ，对 $\{\tilde{\xi}_{t|t}^i\}_{i < \tilde{N}_t^F}$ 进行排序使得对于所有 $i < N_F$ ，都有 $\tilde{\xi}_{t|t}^i = \xi_{t|t}^i$ 成立，也就是说集合 $\{\tilde{\xi}_{t|t}^i\}_{i < \tilde{N}_t^F}$ 的前 N_F 个元素与集合 $\{\xi_{t|t}^i\}_{i < N_F}$ 一致，剩下的元素为集合 $\{\xi_{t|t:T}^i\}_{i < N_B}$ 中不同于集合 $\{\xi_{t|t}^i\}_{i < N_F}$ 的元素。更新后， $\pi'_{t|t}(\mathbf{s}; \boldsymbol{\theta}^c)$ 的近似值如公式 (5.22) 所示。

$$\begin{aligned} \tilde{\pi}'_{t|t}(\mathbf{s}; \boldsymbol{\theta}^c) &= \sum_{i=0}^{\tilde{N}_t^F-1} \tilde{w}_{t|t}^i \delta_{\tilde{\xi}_{t|t}^i}(\mathbf{s}) \\ &= \tilde{\pi}'_{t|t}(\mathbf{s}; \boldsymbol{\theta}^c) + \sum_{i=N_F}^{\tilde{N}_t^F-1} \tilde{w}_{t|t}^i \delta_{\tilde{\xi}_{t|t}^i}(\mathbf{s}) \end{aligned} \quad \dots (5.22)$$

这里保持了前向滤波中所得粒子的权重值不变，也就是对所有的 $i < N_F$ ， $\tilde{w}_{t|t}^i = w_{t|t}^i$ 。

利用公式 (5.20) 以及 $\tilde{\pi}'_{t-1|t-1}(\mathbf{s}; \boldsymbol{\theta}^c)$ 计算新加入的粒子的权重值，所以对于

$i = N_F, \dots, \tilde{N}_t^F - 1$ ，粒子 $\tilde{\xi}_{t|t}^i$ 所对应的未归一化权重 $\tilde{w}_{t|t}^i$ 计算方式如公式 (5.23) 所示。

$$\tilde{w}_{t|t}^i = \sum_{j=0}^{\tilde{N}_{t-1}^F-1} \frac{\tilde{w}_{t-1|t-1}^j}{\tilde{\Omega}_{t-1}^F} P(\mathbf{s}_t = \tilde{\xi}_{t|t}^i | \mathbf{s}_{t-1} = \tilde{\xi}_{t-1|t-1}^j) p(y_t; \boldsymbol{\theta}^c | \mathbf{s}_t = \tilde{\xi}_{t|t}^i) \quad \dots (5.23)$$

最后定义 $\tilde{\Omega}_t^F \triangleq \sum_{i=0}^{\tilde{N}_t^F-1} \tilde{w}_{t|t}^i$ 。至此便可以得到 $\pi'_{t|t}(\mathbf{s}; \boldsymbol{\theta}^c)$ 更新后的近似值 $\tilde{\pi}'_{t|t}(\mathbf{s}; \boldsymbol{\theta}^c)$ ， $\tilde{\pi}'_{t|t}(\mathbf{s}; \boldsymbol{\theta}^c) / \tilde{\Omega}_t^F$ 即为前向滤波分布 $\pi_{t|t}(\mathbf{s}; \boldsymbol{\theta}^c)$ 的更新后的近似值。

类似地，利用 $\{\xi_{t|t}^i, w_{t|t}^i\}_{i < N_F}$ 和 $\{\xi_{t|t:T}^i, w_{t|t:T}^i\}_{i < N_B}$ 更新 $\pi'_{t|t:T}(\mathbf{s}; \boldsymbol{\theta}^c)$ 近似值 $\tilde{\pi}'_{t|t:T}(\mathbf{s}; \boldsymbol{\theta}^c)$ ，更新后的近似值记为 $\tilde{\pi}'_{t|t:T}(\mathbf{s}; \boldsymbol{\theta}^c)$ 。假设已经得到 $t+1$ 时刻的 $\tilde{\pi}'_{t+1|t+1:T}(\mathbf{s}; \boldsymbol{\theta}^c)$ ，其由粒子和权重集合 $\{\tilde{\xi}_{t+1|t+1:T}^i, \tilde{w}_{t+1|t+1:T}^i\}_{i < \tilde{N}_{t+1}^B}$ 组成，定义 $\tilde{\Omega}_{t+1}^B \triangleq \sum_{i=0}^{\tilde{N}_{t+1}^B-1} \tilde{w}_{t+1|t+1:T}^i$ 。这时， t 时刻的 $\tilde{\pi}'_{t|t:T}(\mathbf{s}; \boldsymbol{\theta}^c)$ 中的粒子可由 $t+1$ 时刻前向、反向粒子集合更新为

$\{\tilde{\xi}_{t|t:T}^i\}_{i < \tilde{N}_t^B} = \{\xi_{t|t:T}^i\}_{i < N_B} \cup \{\xi_{t|t}^i\}_{i < N_F}$ ，同样对 $\{\tilde{\xi}_{t|t}^i\}_{i < \tilde{N}_t^B}$ 中元素进行排序使得对所有

$i < N_B$, $\tilde{\xi}_{t|t:T}^i = \xi_{t|t:T}^i$ 。基于更新后的粒子集合, $\tilde{\pi}_{t|t:T}'(\mathbf{s}; \boldsymbol{\theta}^c)$ 可表示为公式 (5.24) 的形式。

$$\begin{aligned}\tilde{\pi}_{t|t:T}'(\mathbf{s}; \boldsymbol{\theta}^c) &= \sum_{i=0}^{\tilde{N}_t^B-1} \tilde{w}_{t|t:T}^i \delta_{\tilde{\xi}_{t|t:T}^i}(\mathbf{s}) \\ &= \tilde{\pi}_{t|t:T}'(\mathbf{s}; \boldsymbol{\theta}^c) + \sum_{i=N_B}^{\tilde{N}_t^B-1} \tilde{w}_{t|t:T}^i \delta_{\tilde{\xi}_{t|t:T}^i}(\mathbf{s})\end{aligned}\quad \dots (5.24)$$

这里保持了反向滤波所得的粒子对应的权重不变, 也就是对所有 $i < N_B$,

$\tilde{w}_{t|t:T}^i = w_{t|t:T}^i$ 。利用公式 (5.21) 以及 $t+1$ 时刻的分布 $\tilde{\pi}_{t+1|t+1:T}'(\mathbf{s}; \boldsymbol{\theta}^c)$ 计算新加入粒子的权重值, 所以对于 $i = N_B, \dots, \tilde{N}_t^B - 1$, 粒子 $\tilde{\xi}_{t|t:T}^i$ 的权重计算方式如公式 (5.25) 所示。

$$\begin{aligned}\tilde{w}_{t|t:T}^i &= p(y_t; \boldsymbol{\theta}^c | \mathbf{s}_t = \tilde{\xi}_{t|t:T}^i) \sum_{j=0}^{\tilde{N}_{t+1}^B-1} (P(\mathbf{s}_{t+1} = \tilde{\xi}_{t+1|t+1:T}^j | \mathbf{s}_t = \tilde{\xi}_{t|t:T}^i) \\ &\quad \times \frac{\tilde{\pi}_{t+1|t+1:T}'(\tilde{\xi}_{t+1|t+1:T}^j; \boldsymbol{\theta}^c)}{\tilde{\Omega}_{t+1|t+1:T}^B})\end{aligned}\quad \dots (5.25)$$

定义 $\tilde{\Omega}_{t|t:T}^B \triangleq \sum_{i=0}^{\tilde{N}_t^B-1} \tilde{w}_{t|t:T}^i$, 所以 $\tilde{\pi}_{t|t:T}'(\mathbf{s}; \boldsymbol{\theta}^c) / \tilde{\Omega}_{t|t:T}^B$ 为反向滤波分布 $\pi_{t|t:T}(\mathbf{s}; \boldsymbol{\theta}^c)$ 的更新后的近似值。

在获得前向、反向滤波分布拓展粒子集合后的近似值后, 可以利用公式 (5.15) 计算平滑分布的近似值。JTFS 算法步骤总结如下:

- 1) 进行粒子数为 N_F 的前向粒子滤波得到每个时刻的粒子及未归一化权重集合

$$\{\xi_{t|t}^i, w_{t|t}^i\}_{i < N_F}, \text{ 其中 } t = 1, \dots, T。$$

- 2) 进行粒子数为 N_B 的反向粒子滤波得到每个时刻的粒子及未归一化权重集合

$$\{\xi_{t|t:T}^i, w_{t|t:T}^i\}_{i < N_B}, \text{ 其中 } t = T, \dots, 1。$$

- 3) 利用公式 (5.22) 和 (5.23) 更新前向滤波近似值, 得到粒子及未归一化权重集合 $\{\tilde{\xi}_{t|t}^i, \tilde{w}_{t|t}^i\}_{i < \tilde{N}_t^F}$, 计算 $\tilde{\Omega}_t^F = \sum_{i=0}^{\tilde{N}_t^F-1} \tilde{w}_{t|t}^i$, 其中 $t = 1, \dots, T$ 。

- 4) 利用公式 (5.24) 和 (5.25) 更新反向滤波近似值, 得到粒子及未归一化权重集合 $\{\tilde{\xi}_{t|t:T}^i, \tilde{w}_{t|t:T}^i\}_{i < \tilde{N}_t^B}$, 计算 $\tilde{\Omega}_{t|t:T}^B = \sum_{i=0}^{\tilde{N}_t^B-1} \tilde{w}_{t|t:T}^i$, 其中 $t = T, \dots, 1$ 。

5) 根据公式 (5.15) 的平滑分布分解方式, 平滑分布的近似值如公式 (5.26) 所示。

$$\hat{\pi}_{t|1:T}(\mathbf{s}; \boldsymbol{\theta}^c) = \sum_{i=0}^{\tilde{N}_t^F-1} w_{t|1:T}^i \delta_{\tilde{\xi}_{t|t}^i}(\mathbf{s}) \quad \dots (5.26)$$

其中权重 $w_{t|1:T}^i \propto \tilde{w}_{t|t}^i \sum_{j=0}^{\tilde{N}_t^B-1} P(\mathbf{s}_{t+1} = \tilde{\xi}_{t+1|t+1:T}^j | \mathbf{s}_t = \tilde{\xi}_{t|t}^i) \tilde{w}_{t+1|t+1:T}^j$ 。

步骤 1) 和 2) 的算法复杂度同前文分析的前向、反向粒子滤波; 步骤 3) 和 4) 在计算某一时刻新添加的粒子权重时, 需要去寻找其先代或后代粒子, 所以需要前一时刻或后一时刻粒子集合进行排序, 所以每一时刻的算法复杂度为 $O((N_F + N_B) \log_2(N_F + N_B))$; 步骤 5) 同样需要对某一时刻前向滤波粒子集合及其下一时刻反向滤波粒子集合进行排序, 所以每一时刻的算法复杂度同样为 $O((N_F + N_B) \log_2(N_F + N_B))$ 。综上分析可见, JTFS 算法的复杂度主要集中在步骤 1) 和 2) 中, 其算法复杂度为 $O(m(N_F + N_B)T)$ 。

5.3 基于 RGSE 的近似算法

在上述介绍的基于粒子平滑的算法中, 用来近似平滑分布的粒子并非是从平滑分布中采样得来, 而是从滤波分布中采样获得。在最大相位信道下, 从前向滤波分布中很难采样到具有较大平滑概率的粒子; 在最小相位信道下, 从反向滤波分布中很难采样到具有较大平滑概率的粒子, 所以利用基于粒子平滑算法的盲信道估计的准确度受到信道类型的影响较大。同时, 粒子平滑算法的算法复杂度受调制阶数的影响依然较大。为了有效地解决上述问题, 提出了基于 RGSE 的信道盲估计算法。下面从 RGSE 的基本原理、实现方法以及方差降低方案三个方面展开介绍, 并提出了 BRGSE 算法。

5.3.1 RGSE 基本原理

RGSE 是第一个无偏平滑估计器, 易于并行化的特性使得其更适用于现代的计算平台。通过构造出满足如下所述的三个条件的两条马尔科夫链 $(X^{(n)})_{n \geq 0}$ 和 $(\tilde{X}^{(n)})_{n \geq 0}$, 可以实现 RGSE。

(1) $(X^{(n)})_{n \geq 0}$ 和 $(\tilde{X}^{(n)})_{n \geq 0}$ 均以平滑分布 $P(\mathbf{s}_{1:T}; \boldsymbol{\theta} | y_{1:T})$ 为稳态分布;

(2) 对于所有的 $n \geq 0$, $X^{(n)}$ 和 $\tilde{X}^{(n)}$ 具有相同的边缘分布;

(3) 存在常数 τ , 使得当 $n \geq \tau$ 时, $X^{(n)} = \tilde{X}^{(n-1)}$ 基本上恒成立。这里称 τ 为相遇时间, 指两条马尔科夫链在时刻 τ 相遇, 并且之后基本上保持一致。

RGSE 利用构造出满足上述条件的 $(X^{(n)})_{n \geq 0}$ 和 $(\tilde{X}^{(n)})_{n \geq 0}$ 可以估计任何关于 $\mathbf{s}_{1:T}$ 的函数 $h(\mathbf{s}_{1:T})$ 的期望值 $E[h(\mathbf{s}_{1:T})] \triangleq \sum_{\mathbf{s}_{1:T}} h(\mathbf{s}_{1:T}) P(\mathbf{s}_{1:T}; \boldsymbol{\theta} | y_{1:T})$ 。首先, 因为 $(X^{(n)})_{n \geq 0}$ 的稳态分布为 $P(\mathbf{s}_{1:T}; \boldsymbol{\theta} | y_{1:T})$, 所以 $E[h(\mathbf{s}_{1:T})] = \lim_{n \rightarrow \infty} E[h(X^{(n)})]$; 其次, 利用条件 (2), 对于所有 $n \geq 0$, $X^{(n)}$ 和 $\tilde{X}^{(n)}$ 具有相同的边缘分布, $\lim_{n \rightarrow \infty} E[h(X^{(n)})] = E[h(X^{(0)})] + \sum_{n=1}^{\infty} E[h(X^{(n)})] - E[h(\tilde{X}^{(n-1)})]$, 交换求和与求期望的顺序可得 $E[h(\mathbf{s}_{1:T})] = E[h(X^{(0)})] + \sum_{n=1}^{\infty} E[h(X^{(n)}) - h(\tilde{X}^{(n-1)})]$; 最后根据条件 (3), 当 $n \geq \tau$ 时, $h(X^{(n)}) = h(\tilde{X}^{(n-1)})$, 所以可以得到公式 (5.27)。

$$E[h(\mathbf{s}_{1:T})] = E[h(X^{(0)})] + \sum_{n=1}^{\tau-1} E[h(X^{(n)}) - h(\tilde{X}^{(n-1)})] \quad \dots (5.27)$$

于是, 定义随机变量 $H \triangleq h(X^{(0)}) + \sum_{n=1}^{\tau-1} h(X^{(n)}) - h(\tilde{X}^{(n-1)})$, 这里 H 便是 $E[h(\mathbf{s}_{1:T})]$ 的 RGSE。

5.3.2 RGSE 实现方法

下面详细介绍如何产生符合上述三个条件的两条马尔科夫链 $(X^{(n)})_{n \geq 0}$ 和 $(\tilde{X}^{(n)})_{n \geq 0}$ 。

首先, 独立进行两次 BPF 产生 $X^{(0)}$ 和 $\tilde{X}^{(0)}$ 。具体 BPF 过程为: 初始化, $t=1$ 时刻, 从状态空间 S 中均匀采样出 N 个状态 $\{\mathbf{s}_i^n\}_{i=1}^N$, 同时初始化权重 $w_i^n \propto p(y_1; \boldsymbol{\theta} | \mathbf{s}_i^n)$, 并归一化使其满足 $\sum_{i=1}^N w_i^n = 1$ 。由于从离散概率分布中进行采样可等效为某一确定函数作用在服从均匀分布 $U(0,1)$ 的随机变量的过程, 初始化的过程中产生的随机变量记为 $\mathbf{U}_1 \triangleq \{U_1^n\}_{n=1}^N$, 因此 $U_1^n \sim U(0,1)$ 。完成初始化之后, BPF 在时刻 t , $1 \leq t \leq T-1$, 主要包含粒子序号重采样和粒子传递两个步骤。粒子序号重采样步骤产生序号向量 $\mathbf{a}_t = [a_t^1, \dots, a_t^N]^T \in \{1, \dots, N\}^N$, 其依据 t 时刻的权重

值进行重采样, 满足 $P(a_t^n = j) = w_t^j$ 对所有的 $n, j \in \{1, \dots, N\}$ 都成立。粒子传递步骤主要依据 t 时刻的粒子集合 $\{\mathbf{s}_t^n\}_{n=1}^N$ 和序号向量 \mathbf{a}_t 产生 $t+1$ 时刻的粒子集合 $\{\mathbf{s}_{t+1}^n\}_{n=1}^N$ 和权重集合 $\{w_{t+1}^n\}_{n=1}^N$ 。以 \mathbf{s}_{t+1}^n 的产生过程为例, 首先从整数集合 $\{1, \dots, m\}$ 中等概率采样随机变量 U_{t+1}^n , 其中 m 表示调制假设 H_i 所对应的调制星座点集合 \mathbf{M} 中的点数。根据状态转移概率, 可以得到 $\mathbf{s}_{t+1}^n = [\mathbf{M}(U_{t+1}^n), \mathbf{s}_t^{a_t^n}(1:L-1)^T]^T$, 其中 $\mathbf{M}(i)$ 表示 \mathbf{M} 中序号为 i 的星座点, $\mathbf{s}_t^i(j:k)$ 表示状态 \mathbf{s}_t^i 中第 j 到第 k 个元素组成的向量。 $t+1$ 时刻的权重 $w_{t+1}^n \propto p(y_{t+1}; \boldsymbol{\theta} | \mathbf{s}_{t+1}^n)$, 并将权重归一化。 t 时刻的粒子传递步骤中产生的随机变量记为 $\mathbf{U}_{t+1} = \{U_{t+1}^n\}_{n=1}^N$, 其中 U_{t+1}^n 为集合 $\{1, \dots, m\}$ 中均匀分布的点。重复上述粒子序号重采样步骤和粒子传递步骤直到 $t = T-1$ 。在最后的时刻 $t = T$, 我们可以保存生成的状态路径集合 $\{\mathbf{s}_{1:T}^n\}_{n=1}^N$ 。从 T 时刻的粒子 \mathbf{s}_T^n 开始, 反向寻找前一时刻用来产生 \mathbf{s}_T^n 的粒子, 依次往前追溯便可得到状态路径 $\mathbf{s}_{1:T}^n$ 。具体产生路径 $\mathbf{s}_{1:T}^n$ 的方式为: 首先定义 $b_T^n = n$, 然后当 $t = T-1, \dots, 1$ 时, $b_t^n = a_t^{b_{t+1}^n}$ 。这样便可知道路径 $\mathbf{s}_{1:T}^n$ 在时刻 t 的成分为 $\mathbf{s}_t^{b_t^n}$ 。最后依据 T 时刻的权重集合 $\{w_T^n\}_{n=1}^N$ 采样状态路径集合 $\{\mathbf{s}_{1:T}^n\}_{n=1}^N$, 满足 $P(X = \mathbf{s}_{1:T}^n) = w_T^n$ 。这里我们称随机变量集合 $\mathbf{U} = \{\mathbf{U}_t\}_{t=1}^T$ 为 BPF 过程生成变量, 定义其联合分布为 φ 。我们用 $X \sim \text{BPF}(\mathbf{U})$ 表示利用 BPF 采样状态路径 X 。所以独立进行两次 BPF 产生 $X^{(0)}$ 和 $\tilde{X}^{(0)}$ 的过程可简化为: 首先采样 $\mathbf{U}^{(0)} \sim \varphi$ 和 $\tilde{\mathbf{U}}^{(0)} \sim \varphi$, 然后采样 $X^{(0)} \sim \text{BPF}(\mathbf{U}^{(0)})$ 和 $\tilde{X}^{(0)} \sim \text{BPF}(\tilde{\mathbf{U}}^{(0)})$ 。因为过程生成变量 $\mathbf{U}^{(0)}$ 和 $\tilde{\mathbf{U}}^{(0)}$ 同分布, 所以可以得到 $X^{(0)}$ 和 $\tilde{X}^{(0)}$ 具有相同的边缘分布。

其次, 在给定 $X^{(0)}$ 的条件下, 利用 CPF 产生 $X^{(1)}$ 。CPF 可以构建出以平滑分布为稳态分布的马尔科夫链, 所以利用 CPF 可以得到满足条件 (1) 的马尔科夫链。一般地, 在给定参考状态路径 $S' = \mathbf{s}_{1:T}'$ 的条件下, 利用 CPF 产生新的状态路径 $S = \mathbf{s}_{1:T}$ 的具体过程为: 初始化, $t=1$ 时, 从状态空间 S 中均匀采样出 $N-1$ 个状态 $\{\mathbf{s}_1^n\}_{n=1}^{N-1}$, 将参考状态路径 $t=1$ 时刻的成分 \mathbf{s}_1' 赋值给 \mathbf{s}_0^N 。初始化权重

$w_1^n \propto p(y_1; \boldsymbol{\theta} | \mathbf{s}_1^n)$, 并归一化使其满足 $\sum_{n=1}^N w_1^n = 1$ 。在初始化过程中产生的随机变量记为 $\mathbf{U}_1 \triangleq \{U_1^n\}_{n=1}^{N-1}$, 同样地, $U_1^n \sim U(0,1)$ 。CPF 在初始化之后每一时刻 t , $1 \leq t \leq T-1$, 同样包含粒子序号重采样和粒子传递两个步骤。对于 t 时刻的前 $N-1$ 个粒子序号向量的采样方式同 BPF, 只需对所有 $n=1, \dots, N-1$, $j=1, \dots, N$ 满足 $P(a_t^n = j) = w_t^j$ 。对于第 N 个粒子序号的采样方案有两种: 第一种直接使 $a_t^N = N$; 第二种采用 AS 方案, 基于状态 \mathbf{s}_{t+1}' 和观察序列 $y_{1:t+1}$, 采样 a_t^N : 因为 $P(\mathbf{s}_t; \boldsymbol{\theta} | \mathbf{s}_{t+1}', y_{1:t+1}) \propto P(\mathbf{s}_{t+1}' | \mathbf{s}_t) P(\mathbf{s}_t; \boldsymbol{\theta} | y_{1:t})$, 所以 $P(a_t^N = j) \propto P(\mathbf{s}_{t+1}' | \mathbf{s}_t^j) w_t^j$ 。这里我们选择采用 AS 方案, 原因可参考文献[94]。对于 t 时刻的粒子传递步骤, 当 $n=1, \dots, N-1$ 时, 产生 \mathbf{s}_{t+1}^n 的机制同 BPF: 首先从整数集合 $\{1, \dots, m\}$ 中等概率采样随机变量 U_{t+1}^n , 根据状态转移概率, 可以得到 $\mathbf{s}_{t+1}^n = [\mathbf{M}(U_{t+1}^n), \mathbf{s}_t^{a_t^n}(1:L-1)^T]^T$ 。将参考状态路径 $t+1$ 时刻的成分 \mathbf{s}_{t+1}' 赋值给 \mathbf{s}_{t+1}^N 。同样地, $t+1$ 时刻的权重 $w_{t+1}^n \propto p(y_{t+1}; \boldsymbol{\theta} | \mathbf{s}_{t+1}^n)$, 并将权重归一化。 t 时刻的粒子传递步骤中产生的随机变量记为 $\mathbf{U}_{t+1} = \{U_{t+1}^n\}_{n=1}^{N-1}$, 其中 U_{t+1}^n 为集合 $\{1, \dots, m\}$ 中均匀分布的点。重复上述粒子序号重采样步骤和粒子传递步骤直到 $t=T-1$ 。在最后的时刻 $t=T$, 我们可以利用与 BPF 相同的方法采样出状态路径 $S = \mathbf{s}_{1:T}$ 。由此可知, CPF 的过程生成变量 $\mathbf{U} = \{\mathbf{U}_t\}_{t=1}^T$ 同样分布于 φ 。我们用 $S \sim \text{CPF-AS}(S', \mathbf{U})$ 表示给定参考路径 S' 的条件下, 利用 CPF 结合 AS 采样出新的路径 S 。至此, 在给定 $X^{(0)}$ 的条件下, 利用 CPF 产生 $X^{(1)}$ 的过程可简化表示为: 首先采样 $\mathbf{U}^{(1)} \sim \varphi$, 然后采样 $X^{(1)} \sim \text{CPF-AS}(X^{(0)}, \mathbf{U}^{(1)})$ 。

最后, $n \geq 2$ 时, 在给定 $X^{(n-1)}$ 和 $\tilde{X}^{(n-2)}$ 的条件下, 利用 CCPF 采样得到 $X^{(n)}$ 和 $\tilde{X}^{(n-1)}$, 直到 $\tilde{X}^{(n)} = \tilde{X}^{(n-1)}$ 为止。这里主要利用 CRN [95][96] 和 ICR [97][98] 技术对两个 CPF 进行耦合, 以产生满足条件 (3) 的两条马尔科夫链。一般地, 在给定参考状态路径 $S' = \mathbf{s}_{1:T}'$ 和 $\tilde{S}' = \tilde{\mathbf{s}}_{1:T}'$ 的条件下, 利用 CCPF 产生新的状态路径

$S = \mathbf{s}_{1:T}$ 和 $\tilde{S} = \tilde{\mathbf{s}}_{1:T}$ 的过程为: 首先, 利用 CRN 技术初始化, 对于所有 $n=1, \dots, N-1$, 采样随机变量 $U_1^n \sim U(0,1)$, 利用 U_1^n 分别从状态空间 S 中均匀采样出初始状态粒子 \mathbf{s}_1^n 和 $\tilde{\mathbf{s}}_1^n$ 。将 \mathbf{s}_1^n 和 $\tilde{\mathbf{s}}_1^n$ 分别赋值给 \mathbf{s}_1^N 和 $\tilde{\mathbf{s}}_1^N$ 。对于所有 $n=1, \dots, N$, 初始化权重 $w_1^n \propto p(y_1; \boldsymbol{\theta} | \mathbf{s}_1^n)$ 和 $\tilde{w}_1^n \propto p(y_1; \boldsymbol{\theta} | \tilde{\mathbf{s}}_1^n)$, 并分别对其归一化。同样地, 之后每一个时刻 t , $1 \leq t \leq T-1$, 主要包含粒子序号重采样和粒子传递两个步骤。这里我们采用 ICR 技术进行粒子序号重采样: 对于 $n \in \{1, \dots, N-1\}$, 根据概率矩阵 \mathbf{P}_t 采样粒子序号对 (a_t^n, \tilde{a}_t^n) , 其中概率矩阵 \mathbf{P}_t 的计算方式如公式 (5.28) 所示。

$$\mathbf{P}_t = \alpha_t \text{diag}(\boldsymbol{\mu}_t) + (1 - \alpha_t) \mathbf{r}_t \tilde{\mathbf{r}}_t^T \quad \dots (5.28)$$

其中 $\boldsymbol{\mu}_t = \mathbf{v}_t / \alpha_t$, \mathbf{v}_t 是 N 维列向量, 第 n 个元素 $v_t^n = \min(w_t^n, \tilde{w}_t^n)$, $\alpha_t = \sum_{n=1}^N v_t^n$, \mathbf{r}_t 和 $\tilde{\mathbf{r}}_t$ 同样也是 N 维列向量, 它们第 n 个元素分别是 $r_t^n = (w_t^n - v_t^n) / (1 - \alpha_t)$ 和 $\tilde{r}_t^n = (\tilde{w}_t^n - v_t^n) / (1 - \alpha_t)$ 。根据 \mathbf{P}_t 采样粒子序号对 (a_t^n, \tilde{a}_t^n) 的过程可以算法复杂度正比于 N 的方式完成: 产生随机数 $\rho \sim U(0,1)$, 如果 $\rho \leq \alpha_t$, 根据离散分布 $\boldsymbol{\mu}_t$ 采样 a_t^n , 同时使 $\tilde{a}_t^n = a_t^n$; 如果 $\rho > \alpha_t$, 根据离散分布 \mathbf{r}_t 和 $\tilde{\mathbf{r}}_t$, 利用 CRN 技术或者独立采样 a_t^n 和 \tilde{a}_t^n 。关于 a_t^N 和 \tilde{a}_t^N 的采样方法则是采用 CPF 中的 AS 采样方案,

$P(a_t^N = j) \propto P(\mathbf{s}_{t+1}^j | \mathbf{s}_t^j) w_t^j$ 和 $P(\tilde{a}_t^N = k) \propto P(\tilde{\mathbf{s}}_{t+1}^k | \tilde{\mathbf{s}}_t^k) \tilde{w}_t^k$ 。在粒子传递步骤中, 采用 CRN 技术产生 $t+1$ 时刻的粒子集合 $\{\mathbf{s}_{t+1}^n\}_{n=1}^{N-1}$ 和 $\{\tilde{\mathbf{s}}_{t+1}^n\}_{n=1}^{N-1}$ 。对于任意 $n \in \{1, \dots, N-1\}$, 从整数集合 $\{1, \dots, m\}$ 中等概率采样随机变量 U_{t+1}^n , 根据状态转移概率, 可以得到 $\mathbf{s}_{t+1}^n = [\mathbf{M}(U_{t+1}^n), \mathbf{s}_t^{a_t^n} (1:L-1)^T]^T$ 和 $\tilde{\mathbf{s}}_{t+1}^n = [\mathbf{M}(U_{t+1}^n), \tilde{\mathbf{s}}_t^{\tilde{a}_t^n} (1:L-1)^T]^T$ 。同时, 设置 $\mathbf{s}_{t+1}^N = \mathbf{s}_{t+1}^1$ 和 $\tilde{\mathbf{s}}_{t+1}^N = \tilde{\mathbf{s}}_{t+1}^1$ 。 $t+1$ 时刻的权重 $w_{t+1}^n \propto p(y_{t+1}; \boldsymbol{\theta} | \mathbf{s}_{t+1}^n)$, $\tilde{w}_{t+1}^n \propto p(y_{t+1}; \boldsymbol{\theta} | \tilde{\mathbf{s}}_{t+1}^n)$, 并分别进行归一化。在最后的时刻 $t=T$, 利用与 BPF 中相同的方法分别构建出粒子路径集合 $\{\mathbf{s}_{1:T}^n\}_{n=1}^N$ 和 $\{\tilde{\mathbf{s}}_{1:T}^n\}_{n=1}^N$, 同时根据 $t=T$ 时刻的权重集合 $\{w_T^n\}_{n=1}^N$ 和 $\{\tilde{w}_T^n\}_{n=1}^N$, 利用 ICR 技术采样出新的粒子路径 $S = \mathbf{s}_{1:T}$ 和 $\tilde{S} = \tilde{\mathbf{s}}_{1:T}$ 。在 CCPF 中过程生成变量为

$\mathbf{U} = \{\mathbf{U}_t\}_{t=1}^T$ 同样分布于 φ 。将 $S = \mathbf{s}_{1:T}$ 和 $\tilde{S} = \tilde{\mathbf{s}}_{1:T}$ 的产生过程分开来看，其分别是基于参考状态路径 $S' = \mathbf{s}'_{1:T}$ 和 $\tilde{S}' = \tilde{\mathbf{s}}'_{1:T}$ ，利用 CPF-AS 采样而来，也就是 $S \sim \text{CPF-AS}(S', \mathbf{U})$ 和 $\tilde{S} \sim \text{CPF-AS}(\tilde{S}', \mathbf{U})$ 。所以利用 CCPF 迭代产生的两条状态路径都可以构成以平滑分布为稳态分布的马尔科夫链，也就是满足上述的条件(1)。将其产生过程联合起来看，可以表示为 $(S, \tilde{S}) \sim \text{CCPF-AS}(\tilde{S}', S', \mathbf{U})$ ，由于利用 CRN 和 ICR 技术对两个 CPF 进行了耦合，所以 CCPF 的输出路径 S 和 \tilde{S} 将会有一定概率相等，从而达到满足上述条件(3)的目的。当 $n \geq 2$ 时，给定 $X^{(n-1)}$ 和 $\tilde{X}^{(n-2)}$ 的条件下，利用 CCPF 产生 $X^{(n)}$ 和 $\tilde{X}^{(n-1)}$ 的过程可简化为：首先采样 $\mathbf{U}^{(n)} \sim \varphi$ ，然后采样 $(X^{(n)}, \tilde{X}^{(n-1)}) \sim \text{CCPF-AS}(X^{(n-1)}, \tilde{X}^{(n-2)}, \mathbf{U}^{(n)})$ 。

由上述过程，可以产生两条马尔科夫链 $(X^{(n)})_{n \geq 0}$ 和 $(\tilde{X}^{(n)})_{n \geq 0}$ 。首先由于 $(X^{(n)})_{n \geq 1}$ 和 $(\tilde{X}^{(n)})_{n \geq 1}$ 都是由 CPF 迭代采样出来，所以 $(X^{(n)})_{n \geq 0}$ 和 $(\tilde{X}^{(n)})_{n \geq 0}$ 都是以平滑分布为稳态分布的马尔科夫链，也就是这两条链满足上述条件(1)；其次，因为 $(\mathbf{U}^{(n)})_{n \geq 0}$ 独立同分布，所以对于每一个 n ，其中 $n \geq 0$ ， $X^{(n)}$ 和 $\tilde{X}^{(n)}$ 具有相同的边缘分布，也就是这两条链满足上述条件(2)；最后，因为采用了 CRN 和 ICR 技术对两个 CPF 进行耦合，当 $n \geq 2$ 时，存在一定概率会出现 $X^{(n)} = \tilde{X}^{(n-1)}$ ，所以这两条链满足上述条件(3)。

综上所述，基本的 RGSE 的算法流程如算法 5.1 所示。

算法 5.1 基本的 RGSE

- 1: 采样 $\mathbf{U}^{(0)} \sim \varphi$ ，利用 BPF 采样 $X^{(0)} \sim \text{BPF}(\mathbf{U}^{(0)})$ ；
- 2: 采样 $\tilde{\mathbf{U}}^{(0)} \sim \varphi$ ，利用 BPF 采样 $\tilde{X}^{(0)} \sim \text{BPF}(\tilde{\mathbf{U}}^{(0)})$ ；
- 3: 采样 $\mathbf{U}^{(1)} \sim \varphi$ ，利用 CPF-AS 采样 $X^{(1)} \sim \text{CPF-AS}(X^{(0)}, \mathbf{U}^{(1)})$ ；
- 4: 计算 $\Delta^{(0)} = h(X^{(0)})$ 和 $\Delta^{(1)} = h(X^{(1)}) - h(\tilde{X}^{(0)})$ ，设置 $H = \Delta^{(0)} + \Delta^{(1)}$ ；
- 5: 赋值 $n = 2$ ；
- 6: 采样 $\mathbf{U}^{(n)} \sim \varphi$ ，利用 CCPF-AS 采样 $(X^{(n)}, \tilde{X}^{(n-1)}) \sim \text{CCPF-AS}(X^{(n-1)}, \tilde{X}^{(n-2)}, \mathbf{U}^{(n)})$ ；

- 7: 判断 $X^{(n)}$ 是否等于 $\tilde{X}^{(n-1)}$, 如果两者不相等, 则计算 $\Delta^{(n)} = h(X^{(n)}) - h(\tilde{X}^{(n-1)})$, 并更新 $H = H + \Delta^{(n)}$, 同时更新 $n = n + 1$, 重复步骤 6 和 7; 如果两者相等, 则停止循环, 返回 H 。

5.3.3 RGSE 方差降低方案

下面介绍降低基本的 RGSE 方差的方案。

首先, 当 M 和 m 都是整数且 $M > m \geq 0$ 时, 定义

$H_{m,M} \triangleq h(X^{(m)}) + \sum_{n=m+1}^M h(X^{(n)}) - h(\tilde{X}^{(n-1)})$, 其中 $(X^{(n)})_{n \geq 0}$ 和 $(\tilde{X}^{(n)})_{n \geq 0}$ 是满足上述三个条件的两条马尔科夫链。由于 $H_{m,M} = h(X^{(M)}) + \sum_{n=m}^{M-1} h(X^{(n)}) - h(\tilde{X}^{(n)})$, 利用 $X^{(n)}$ 和 $\tilde{X}^{(n)}$ 具有相同的边缘分布可得 $E[H_{m,M}] = E[h(X^{(M)})]$ 。同时, 因为 $E[h(\mathbf{s}_{1:T})] = \lim_{M \rightarrow \infty} E[h(X^{(M)})]$, 所以当 $M = \infty$ 时, $H_{m,\infty}$ 滤波器满足 $E[h(\mathbf{s}_{1:T})] = E[H_{m,\infty}]$ 。根据定义可知, $H_{m,\infty} = h(X^{(m)}) + \sum_{n=m+1}^{\infty} h(X^{(n)}) - h(\tilde{X}^{(n-1)})$, 利用 $(X^{(n)})_{n \geq 0}$ 和 $(\tilde{X}^{(n)})_{n \geq 0}$ 满足条件 (3), $H_{m,\infty}$ 其可在如下有限步骤内实现该滤波器。

循环运行算法 5.1 的步骤 6 直到 $n = \max(\tau, m)$ 。如果相遇时间 $\tau \leq m+1$, 说明 $X^{(n)}$ 和 $\tilde{X}^{(n-1)}$ 在 $n = m+1$ 之前已经相等, 所以此时 $H_{m,\infty} = h(X^{(m)})$; 如果 $\tau > m+1$, 则 $H_{m,\infty} = h(X^{(m)}) + \sum_{n=m+1}^{\tau-1} h(X^{(n)}) - h(\tilde{X}^{(n-1)})$ 。直观上来看, 参与计算 $H_{m,\infty}$ 的项越少, 估计器方差将会越小, 所以当 m 取相对较大值时, $H_{m,\infty}$ 将会有很大概率是由单一项 $h(X^{(m)})$ 计算得来。 $H_{m,\infty}$ 滤波器的具体实现过程如算法 5.2 所示。

算法 5.2 $H_{m,\infty}$ 平滑估计器

- 1: 采样 $\mathbf{U}^{(0)} \sim \varphi$, 利用 BPF 采样 $X^{(0)} \sim \text{BPF}(\mathbf{U}^{(0)})$;
- 2: 采样 $\tilde{\mathbf{U}}^{(0)} \sim \varphi$, 利用 BPF 采样 $\tilde{X}^{(0)} \sim \text{BPF}(\tilde{\mathbf{U}}^{(0)})$;
- 3: 采样 $\mathbf{U}^{(1)} \sim \varphi$, 利用 CPF-AS 采样 $X^{(1)} \sim \text{CPF-AS}(X^{(0)}, \mathbf{U}^{(1)})$;
- 4: 赋值 $n = 2$;

- 5: 采样 $\mathbf{U}^{(n)} \sim \varphi$, 利用 CCPF-AS 采样 $(X^{(n)}, \tilde{X}^{(n-1)}) \sim \text{CCPF-AS}(X^{(n-1)}, \tilde{X}^{(n-2)}, \mathbf{U}^{(n)})$,
更新 $n = n + 1$;
- 6: 重复步骤 5 直到 $n = m$ 为止;
- 7: 计算 $H = h(X^{(m)})$;
- 8: 如果 $X^{(n)}$ 不等于 $\tilde{X}^{(n-1)}$, 更新 $n = n + 1$, 采样 $\mathbf{U}^{(n)} \sim \varphi$, 利用 CCPF-AS 采样
 $(X^{(n)}, \tilde{X}^{(n-1)}) \sim \text{CCPF-AS}(X^{(n-1)}, \tilde{X}^{(n-2)}, \mathbf{U}^{(n)})$, 并更新 $H = H + h(X^{(n)}) - h(\tilde{X}^{(n-1)})$;
- 9: 重复步骤 8 直到 $X^{(n)}$ 等于 $\tilde{X}^{(n-1)}$ 为止。

其次, 可以利用 RB 方法进一步降低估计器方差。在上述算法中, 我们从状态路径集合 $\{\mathbf{s}_{1:T}^n\}_{n=1}^N$ 中以概率 $\{w_T^n\}_{n=1}^N$ 采样出 $\mathbf{s}_{1:T}^k$ 来计算随机变量 $h(X^{(n)}) = h(\mathbf{s}_{1:T}^k)$ 。

因为随机变量 $\sum_{k=1}^N w_T^k h(\mathbf{s}_{1:T}^k)$ 与 $h(X^{(n)})$ 具有相同的期望值, 并且其充分利用了 CPF 产生的状态路径, 所以利用该方法替换平滑估计中的 $h(X^{(n)})$ 或 $h(\tilde{X}^{(n)})$ 可以进一步降低估计器方差。当粒子滤波采样点数 N 较大时, 采用所有状态路径集合进行计算会增加算法的计算复杂度, 这时可以从状态路径集合中按照概率 $\{w_T^n\}_{n=1}^N$ 采样出 $N_s (1 \leq N_s \leq N)$ 条路径, 记为 $\{\mathbf{s}_{1:T}^k\}_{k=1}^{N_s}$, 然后用随机变量 $1/N_s \sum_{k=1}^{N_s} h(\mathbf{s}_{1:T}^k)$ 代替 $h(X^{(n)})$ 以达到降低方差的目的, 我们称其为部分 RB 方法。

最后, 可以通过独立进行 R 次 RGSE, 每一次估计结果记为 $H^{(r)}$, 最终用 $\bar{H} = 1/R \sum_{r=1}^R H^{(r)}$ 作为平滑估计的结果, 以达到降低估计方差的目的。由于是独立计算, 所以该降低方差的方案易于并行实现。

5.3.4 BRGSE 算法

从[92][93]中得知, 随着观察信号长度 T 的增加, 需要增加粒子滤波采样点数 N 才能使 RGSE 的平均相遇时间 τ 和方差保持相对稳定, 导致了算法复杂度的增加。在我们所研究的盲信道估计的问题中, 随着信噪比的增加, 观察信号的似然函数值提供的信息量随之增加, t 时刻的隐藏状态与该时刻的观察信号之间的关系越来越紧密。所以在信噪比相对较高的情况下, 我们提出将观察信号 $y_{1:T}$ 分成 B 块, 每一块的长度为 $l = T/B$ (这里选择 B 使得 l 为整数), 那么第 i 块观察

信号的序号区间为 $\Delta_i = (i-1) \times l + 1 : i \times l$ 。根据第 i 块观察信号 y_{Δ_i} ，采用 RGSE 估计 $E_{s_{\Delta_i}; \theta | y_{\Delta_i}}[h(s_{\Delta_i})]$ ，估计的结果记为 H_i 。利用每一小块的估计值，我们提出近似计算 $E_{s_{1:T}; \theta | y_{1:T}}[h(s_{1:T})]$ 的估计值 H 的方案如公式 (5.29) 所示。

$$H \approx \sum_{i=1}^B H_i \quad \dots (5.29)$$

这样做的优势有两点：其一，在每一个小块中，由于观察信号的长度 l 相对较小，所以 RGSE 需要较少的粒子数 N 便可使得平均相遇时间 τ 较小，从而使得每一小块 RGSE 的算法复杂度较低；其二，由于每一小块的 RGSE 相互独立，所以便于并行运行，进一步加快盲信道估计的速度。我们称该方案为 BRGSE 算法。针对每一小块中的平滑估计，同样可以采用 5.3.3 节介绍的方案降低估计器的方差。采用降低方差方案后的每一小块平滑估计器 $H_i^{(r)}$ 的算法复杂度为

$O(\max(m, \tau)Nl)$ ，其中 m 表示 $H_{m, \infty}$ 滤波器中的输入参数。由于在该小块中独立进行 R 次平滑估计可以通过并行化实现，所以该小块中的平滑估计的运算时间只与 $O(\max(m, \tau)Nl)$ 有关。同时，由于 B 块平滑估计器同样可以并行实现，所以估计 $E_{s_{1:T}; \theta | y_{1:T}}[h(s_{1:T})]$ 的运算时间也只与 $O(\max(m, \tau)Nl)$ 有关。

5.4 仿真及分析

首先，我们研究 RGSE 的算法复杂度与信道类型（最小/最大相位信道）、信道阶数以及调制阶数之间的关系。由于 RGSE 的算法复杂度与平均相遇时间 τ 和粒子滤波采样点数 N 有关，而 τ 会随着 N 的增大而减小，所以这里为了研究不同信道类型、信道阶数以及调制阶数对于 RGSE 算法复杂度的影响，我们使得每组对比实验的 τ 保持相近，比较 N 的大小来说明各种情形下算法复杂度的差异。

在研究信道类型对 RGSE 算法复杂度的影响时，设置最小相位信道时的信道系数为 $\mathbf{h} = [0.63, 0.05, -0.3]^T$ ，最大相位信道时的信道系数为 $\mathbf{h} = [-0.3, 0.05, 0.8]^T$ 。除此之外，实验相关参数保持一致， $\text{SNR} = 12\text{dB}$ ， $T = 10$ ，发射机选择的调制类型为 64QAM。表 5.1 显示了这两种信道类型下的平均相遇时间 τ 以及相应的粒子滤波采样点数 N 。从表 5.1 可以看出，在最大相位信道下，想要维持 $E(\tau)$ 与

最小相位时相近，所需的 N 大于最小相位信道下的值。所以 RGSE 的算法复杂度受信道类型的影响较大，最大相位信道类型下的算法的复杂度高于最小相位下的算法复杂度。

表 5.1 不同信道类型下的 $E(\tau)$ 及其所对应的 N

Table 5.1 $E(\tau)$ and N under different channel type

信道类型	最小相位信道	最大相位信道
N	500	30000
$E(\tau)$	3.226	3.186

在研究信道阶数对于 RGSE 算法复杂度的影响时，选择信道类型为最小相位信道，信道系数分别为 $\mathbf{h} = [0.63, -0.3]^T$ 、 $\mathbf{h} = [0.63, 0.05, -0.3]^T$ 和 $\mathbf{h} = [0.63, 0.05, 0.05, -0.3]^T$ 。除此之外，实验相关参数保持一致， $\text{SNR} = 12\text{dB}$ ， $T = 10$ ，发射机选择的调制类型为 64QAM。表 5.2 显示了不同信道阶数下的 $E(\tau)$ 及其所对应的 N 。从表 5.2 中可以看出信道阶数对于 RGSE 算法复杂度影响相对较小。

表 5.2 不同信道阶数下的 $E(\tau)$ 及其所对应的 N

Table 5.2 $E(\tau)$ and N under different channel order

信道阶数	$L = 2$	$L = 3$	$L = 4$
N	500	500	500
$E(\tau)$	2.661	3.226	3.853

在研究调制阶数对于 RGSE 算法复杂度的影响时，我们选择信道类型为最小相位信道，信道系数为 $\mathbf{h} = [0.63, 0.05, -0.3]^T$ ，发射机分别采用 64QAM、128QAM 和 256QAM。除此之外，实验相关参数保持一致， $\text{SNR} = 12\text{dB}$ ， $T = 10$ 。表 5.3 显示了不同调制阶数下的 $E(\tau)$ 及其所对应的 N 。从表 5.3 中可以看出调制阶数对于 RGSE 算法复杂度影响相对较小。

表 5.3 不同调制阶数下的 $E(\tau)$ 及其所对应的 N Table 5.3 $E(\tau)$ and N under different modulation order

调制类型	64QAM	128QAM	256QAM
N	500	500	500
$E(\tau)$	3.226	3.273	3.304

其次，我们研究了 RGSE 的估计性能。因为公式 (5.5) 中的 $E[d(\mathbf{s}_{1:T})]$ 是标量，所以选择其作为研究对象来展示 RGSE 的性能。在 $T=100$ ， $\text{SNR}=12\text{dB}$ ，64QAM 选做发射调制类型，选择最小相位信道系数为 $\mathbf{h}=[0.63, 0.05, -0.3]^T$ 的情形下，首先随机生成一个观察信号序列，利用公式 (5.5) 对 $E[d(\mathbf{s}_{1:T})]$ 的定义，准确地计算出给定该观察信号序列条件下的 $E[d(\mathbf{s}_{1:T})]$ 。其次，多次采用上述 $H_{m,\infty}$ 估计器对 $E[d(\mathbf{s}_{1:T})]$ 进行估计，将估计结果与准确值进行比较来说明估计器的性能，其中 $H_{m,\infty}$ 的相关参数设置为： $m=5$ ， $N=2000$ 。图 5.1 展示了 $H_{m,\infty}$ 估计器在使用 RB 技术和未使用 RB 技术两种情形下， $R=10, 50, 100$ 时估计结果的分布情况。图 5.1 中左边三个箱线图表示使用了 RB 技术的 $H_{m,\infty}$ 估计器的估计结果分布情况，右边三个箱线图则表示未使用 RB 技术的 $H_{m,\infty}$ 估计器的估计结果分布情况，图中红色虚线表示 $E[d(\mathbf{s}_{1:T})]$ 的准确值。从图 5.1 中可以看出估计结果对称分布在准确值的两侧，并且采用 RB 技术和增加 R 值可以降低估计器的方差。在相同的实验条件下，利用上述随机生成的接收信号，我们比较了 RGSE 与 BRGSE 估计器的性能。两者均采用 RB 技术以及独立进行 $R=10$ 次估计取平均值的方案来降低估计器方差。BRGSE 的分块数 B 取 2, 5, 10, 20 进行研究，每一小块的 $H_{m,\infty}$ 估计器参数同 RGSE： $m=5, N=2000$ 。图 5.2 显示了 RGSE 与不同分块数下 BRGSE 对于 $E[d(\mathbf{s}_{1:T})]$ 的估计值分布情况。从图 5.2 中可以发现所有分块数大于 1 的 BRGSE 均为有偏估计，随着分块数的减小，BRGSE 估计值的分布越接近 RGSE，也就是越接近无偏估计。

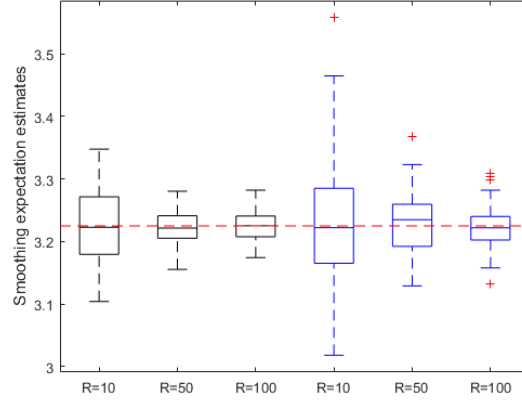


图 5.1 最小相位信道下，RGSE 对 $E[d(s_{l:T})]$ 的估计结果分布图

Figure 5.1 Distributions of estimates of $E[d(s_{l:T})]$ by RGSE under Minimum phase channel

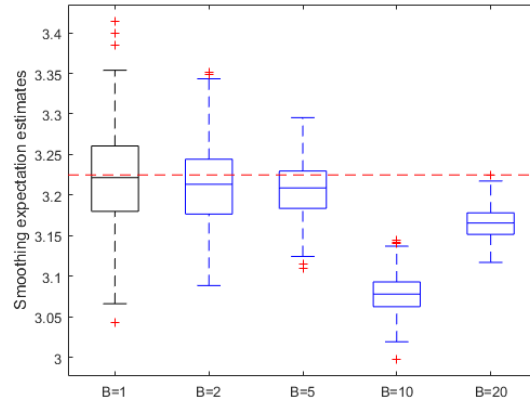


图 5.2 不同分块数下的 BRGSE 对 $E[d(s_{l:T})]$ 的估计结果分布图

Figure 5.2 Distributions of estimates of $E[d(s_{l:T})]$ by BRGSE when B is different

在研究最大相位信道下 RGSE 性能时，设置相关参数： $T = 10$ ， $\text{SNR} = 12\text{dB}$ ，64QAM 选做发射调制类型，信道系数 $\mathbf{h} = [-0.3, 0.05, 0.8]^T$ 。同样地，首先随机生成一个观察信号序列，计算出 $E[d(s_{l:T})]$ 作为参考值。然后多次采用上述 $H_{m,\infty}$ 估计器对 $E[d(s_{l:T})]$ 进行估计，将估计结果与准确值进行比较来说明估计器的性能，其中 $H_{m,\infty}$ 的相关参数设置为： $m = 5$ ， $N = 30000$ 。图 5.3 展示了 $H_{m,\infty}$ 估计器分别在使用 RB 技术，未使用 RB 技术和使用部分 RB 技术三种情形下，

$R=10,50,100$ 时估计结果的分布情况。图 5.3 中左边三个箱线图表示使用了 RB 技术的 $H_{m,\infty}$ 估计器的估计结果分布情况,中间三个箱线图表示未使用 RB 技术的 $H_{m,\infty}$ 估计器的估计结果分布情况,右边三个箱线图表示使用部分 RB 技术 ($N_s=100$) 的 $H_{m,\infty}$ 估计器的估计结果分布情况,图中红色虚线表示 $E[d(s_{l,T})]$ 的准确值。从图 5.3 中可以看出估计结果对称分布在准确值的两侧,采用 RB 技术和增加 R 值可以降低估计器的方差,并且采用部分 RB 技术同样可以降低估计器方差,且性能接近 RB 技术。

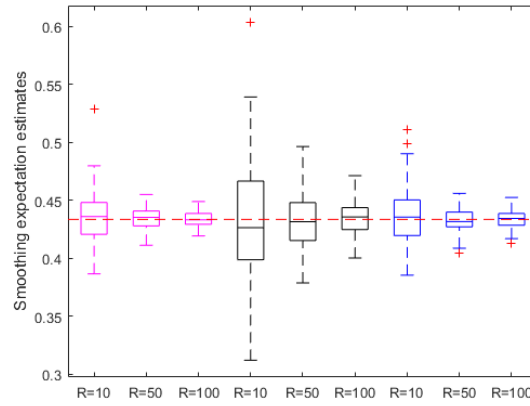


图 5.3 最大相位信道下, RGSE 对 $E[d(s_{l,T})]$ 的估计结果分布图

Figure 5.3 Distributions of estimates of $E[d(s_{l,T})]$ by RGSE under Maximum phase channel

紧接着,我们研究 EM-BRGSE 算法在信道参数盲估计方面的性能。由于参数初始化值对 EM 算法收敛结果影响较大,在研究中采用文献[78]提供的初始化方案。EM 算法的迭代次数均设置为 80。主要通过信道系数估计值与真实信道系数值之间的均方误差来衡量所提出算法对信道估计的准确度情况,该性能指标的定义如公式 (5.30) 所示。

$$\text{MSE}(\hat{\mathbf{h}}) = E\left[\sum_{l=1}^L |\hat{h}_l - h_l|^2\right] \quad \dots (5.30)$$

其中 \hat{h}_l 表示第 l 径信道系数估计值, h_l 表示第 l 径信道系数真实值。实验的比较对象为 EM-FB 算法 (也称为 BW 算法) 和 EM-JTFS 算法,在信道系数估计值

的均方误差以及信道估计所需时间两个方面，将提出算法与它们做比较，以说明提出算法的优越性。

在最小相位信道下，选择信道系数为 $\mathbf{h} = [0.63, 0.05, -0.3]^T$ ， $\text{SNR} = 15\text{dB}$ ，64QAM 作为发射信号采用的调制类型，研究信号长度 $T = 300$ 时，三种算法信道系数估计值的均方误差情况。设置 EM-BRGSE 算法的分块数 $B = 6, 3$ 来研究不同分块数下该算法对信道估计的准确度情况。当 $B = 6$ 时，设置粒子滤波采样点数 $N = 2000$ ，利用 $H_{m,\infty}$ 平滑估计器（设置 $m = 5$ ）、RB 技术以及独立进行 $R = 10$ 次 BRGSE 取平均值这三种方案来降低估计器的方差；当 $B = 3$ 时，设置粒子滤波采样点数 $N = 20000$ ，利用 $H_{m,\infty}$ 平滑估计器（设置 $m = 5$ ）、部分 RB 技术（设置 $N_s = 100$ ）以及独立进行 $R = 10$ 次 BRGSE 取平均值这三种方案来降低估计器的方差；对于 EM-JTFS 算法，根据文献[91]，利用 4.2.3 节介绍的随机选择方法 1 对每一次迭代产生的粒子进行采样；同时设置前向滤波粒子数和反向滤波粒子数 $N_F = N_B = 500$ 。图 5.4 展示了这三种算法对信道系数估计的均方误差随着迭代次数的变化情况图。图 5.5 显示了在第 80 次迭代时，三种算法信道系数估计值与真实值之间的误差的平方和的分布情况。表 5.4 展示了这三种算法对信道盲估计所需时间，仿真实验采用 Matlab 软件，运行在 Intel(R) Xeon(R) Platinum 8163 主频为 2.50GHz 的 CPU 上。由于 BRGSE 算法易于并行实现的特性，这里利用 Matlab 模拟统计该算法并行运行所需的时间（在每一次迭代中，针对接收信号的每一小块所独立运行的 R 次 RGSE 可以并行实现，同时不同信号块间的 RGSE 也可独立进行运算，所以在仿真中，取这些独立并行小块中运行时间最长的作为该次迭代所需时间）。从图 5.4 中可以看出随着迭代次数的增加，三种算法信道系数估计的 MSE 相应地趋于稳定，达到收敛的状态。从图 5.5 可以看出在最小相位信道下，三种算法对信道系数估计的准确度相近。对于 EM-BRGSE 算法，随着分块数 B 的减小，信道系数估计准确度会相应地提高；结合表 5.4，随着分块数 B 的减小，信道系数估计所需的时间会相应地增加。从表 5.4 中可以看出 EM-BRGSE 算法可以大幅度地提升信道盲估计的速度。

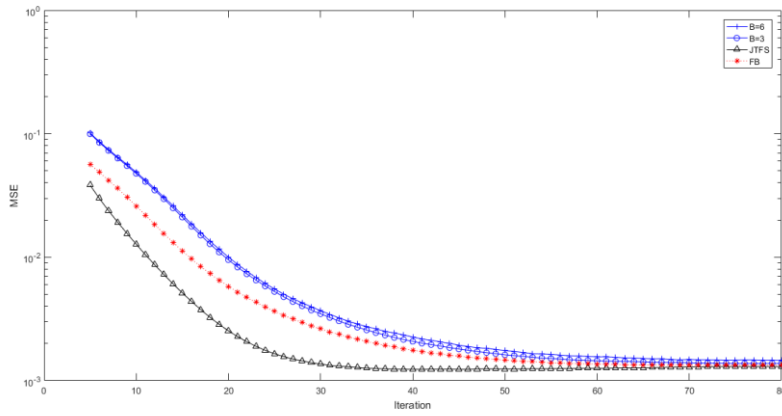


图 5.4 最小相位信道下，三种算法对信道系数估计的 MSE 随着迭代次数的变化情况图

Figure 5.4 The MSE of channel coefficients changes with the number of iterations under Minimum phase channel

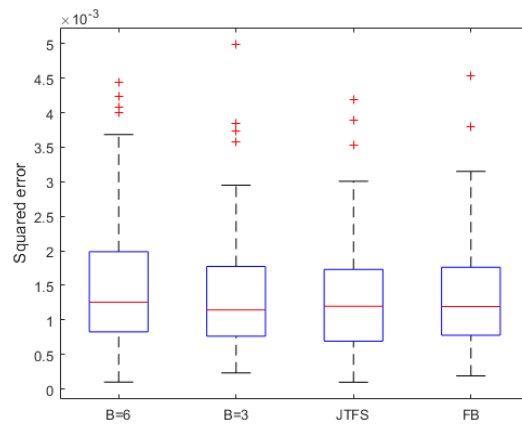


图 5.5 最小相位信道下，三种算法信道系数估计值与真实值之间的误差的平方和的分布情况

Figure 5.5 Distributions of squared error of channel coefficients under Minimum phase channel

表 5.4 最小相位信道下，三种算法对信道盲估计所需的时间

Table 5.4 The time required for blind channel estimation by the three algorithms under Minimum phase channel

算法	EM-BRGSE		EM-JTFS	EM-FB
	B=6	B=3		
时间（秒）	84.81	750.66	2236.42	112975.18

在最大相位信道条件下，选择信道系数 $\mathbf{h} = [-0.3, 0.05, 0.8]^T$ ， $\text{SNR} = 15\text{dB}$ ，64QAM 作为发射信号采用的调制类型，研究信号长度 $T = 300$ 时，三种算法信道系数估计值的均方误差情况。对于 EM-BRGSE 算法，设置分块数 $B = 60$ ，粒子滤波采样点数 $N = 30000$ ，同样采用 $H_{m,\infty}$ 平滑估计器（设置 $m = 5$ ）、部分 RB（设置 $N_s = 100$ ）以及独立进行 $R = 10$ 次 BRGSE 取平均值这三种方案来降低估计器的方差；对于 EM-JTFS 算法，研究了前向滤波和反向滤波粒子数 $N_F = N_B = 2000$ 以及 $N_F = N_B = 3000$ 这两种情形。图 5.6 展示了这三种算法对信道系数估计的均方误差随着迭代次数的变化情况图。图 5.7 显示了在第 80 次迭代时，三种算法信道系数估计值与真实值之间的误差的平方和的分布情况。表 5.5 展示了这三种算法对信道盲估计所需时间。从图 5.6 中可以看出随着迭代次数的增加，三种算法信道系数估计的 MSE 相应地趋于稳定，达到收敛的状态。从图 5.7 可以看出，在最大相位信道下，EM-BRGSE 算法对信道估计的准确度高于 EM-JTFS 算法，接近 EM-FB 算法。表 5.5 说明了 EM-BRGSE 算法对信道盲估计的速度快于 EM-JTFS 以及 EM-FB 算法。因此，在最大相位信道下，EM-BRGSE 算法在速度和信道系数估计准确度方面均优于 EM-JTFS 算法。

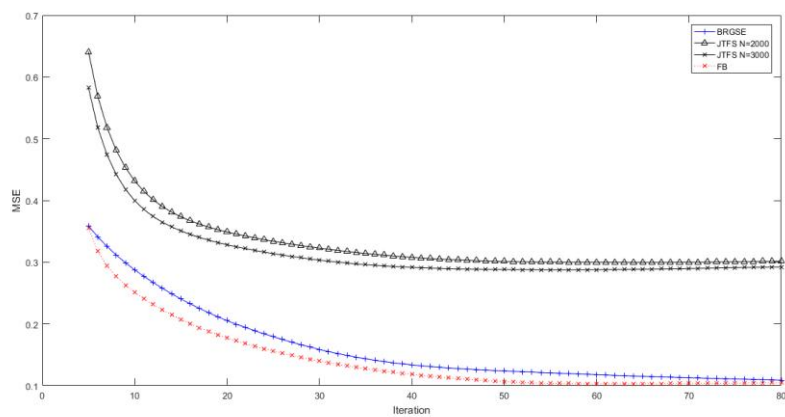


图 5.6 最大相位信道下，三种算法对信道系数估计的 MSE 随迭代次数变化情况图

Figure 5.6 The MSE of channel coefficients changes with the number of iterations under Maximum phase channel

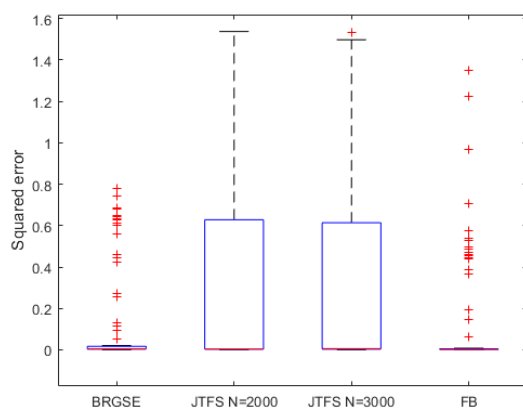


图 5.7 最大相位信道下，三种算法信道系数估计值与真实值之间的误差的平方和的分布情况

Figure 5.7 Distributions of squared error of channel coefficients under Maximum phase channel

表 5.5 最大相位信道下，三种算法对信道盲估计所需的时间

Table 5.5 The time required for blind channel estimation by the three algorithms under Maximum phase channel

算法	EM-BRGSE	EM-JTFS		EM-FB
		$N_F = N_B = 2000$ $N_F = N_B = 3000$		
时间（秒）	82.19	18354.86	45918.86	93012.46

最后，在最小、最大相位信道下，利用最大似然调制分类的框架对 64QPSK 与 64QAM 进行分类。这里信道参数盲估计部分分别采用 EM-BRGSE 算法与 EM-JTFS 算法，似然函数计算部分都采用第四章中介绍的前向粒子滤波算法（利用确定选择方法对每一次迭代产生的粒子进行采样）进行近似计算。

最小相位信道下，信道系数 $\mathbf{h} = [0.63, 0.05, -0.3]^T$ ， $\text{SNR} = 15\text{dB}$ ， $T = 200$ 。BRGSE 算法参数设置为： $B = 20, N = 10000$ ，采用 $H_{m,\infty}$ 平滑估计器（设置 $m = 5$ ）、部分 RB（设置 $N_s = 100$ ）以及独立进行 $R = 10$ 次 BRGSE 取平均值这三种方案来降低估计器的方差；JTFS 算法参数设置为： $N_F = N_B = 1000$ 。前向粒子滤波粒子采样点数 $N = 10000$ 。表 5.6 展示了两种算法下平均分类准确度以及平均分类所需时间。从表 5.6 中可以看出，在最小相位信道下，基于 EM-BRGSE 的最大似然调制分类器的分类准确度与基于 EM-JTFS 算法一致，但是前者的速度明显快于后者。

表 5.6 最小相位信道下, 基于 EM-BRGSE 和 EM-JTFS 算法的调制分类器平均分类准确度与时间

Table 5.6 The average accuracy and time of classifier based on EM-BRGSE and EM-JTFS under Minimum phase channel

算法	EM-BRGSE	EM-JTFS
平均分类准确度	100%	100%
平均分类时间 (秒)	111.47	6162.07

最大相位信道下, 信道系数 $\mathbf{h} = [-0.3, 0.05, 0.8]^T$, $\text{SNR} = 15\text{dB}$, $T = 200$ 。

BRGSE 算法参数设置为: $B = 40, N = 30000$, 采用 $H_{m,\infty}$ 平滑估计器 (设置 $m = 5$)、部分 RB (设置 $N_s = 100$) 以及独立进行 $R = 10$ 次 BRGSE 取平均值这三种方案来降低估计器的方差; JTFS 算法参数设置为: $N_F = N_B = 2000$ 。前向粒子滤波粒子采样点数 $N = 150000$ 。表 5.7 展示了两种算法下平均分类准确度以及平均分类所需时间。从表 5.7 中可以看出, 在最大相位信道下, 基于 EM-BRGSE 的最大似然调制分类器的分类准确度高于基于 EM-JTFS 算法, 并且前者的速度明显快于后者。

表 5.7 最大相位信道下, 基于 EM-BRGSE 和 EM-JTFS 算法的调制分类器平均分类准确度与时间

Table 5.7 The average accuracy and time of classifier based on EM-BRGSE and EM-JTFS under Maximum phase channel

算法	EM-BRGSE	EM-JTFS
平均分类准确度	96.09%	92.97%
平均分类时间 (秒)	343.51	20245.63

5.5 本章小结

本章首先介绍了现有的频率选择性衰落信道参数盲估计算法：EM-FB 和 EM-JTFS 算法，在详细介绍了这两个算法的原理基础上，重点分析了这两个算法存在的缺点。针对现有算法存在的问题，提出了 EM-BRGSE 算法，介绍了降低 RGSE 算法方差的方案以及 BRGSE 算法的优势。通过仿真实验，说明了 RGSE 的估计性能；从信道盲估计的准确度和所需时间对比了 EM-BRGSE 算法、EM-FB 算法和 EM-JTFS 算法；结合第四章提出的接收信号似然函数近似计算算法，比较了基于 EM-BRGSE 算法和 EM-JTFS 算法的调制分类器的分类准确度以及分类时间这两个方面的性能。

第六章 总结与展望

6.1 论文主要研究成果

本文主要研究了实际通信信道下的最大似然调制分类算法，信道环境为平坦衰落非高斯噪声信道和频率选择性衰落高斯噪声信道。最大似然调制分类的研究主要包括：信号建模、每个调制类型假设的似然函数值计算以及信号模型中相关参数值的盲估计。

在平坦衰落非高斯噪声信道环境下，利用 GMM 来建模非高斯噪声；结合信号模型，利用贝叶斯全概率公式计算每个调制类型假设的似然函数值；在信号模型中参数盲估计环节，由于利用 ECM 算法进行盲估计时存在收敛速度较慢的问题，我们基于 SQUAREM 算法提出了 SQUAREM-PC 算法来进行参数盲估计。研究结果表明 SQUAREM-PC 算法在进行参数盲估计时，收敛速度快于 ECM 算法，并且 SQUAREM-PC 算法对参数估计的准确度与 ECM 算法基本一致。

基于这一部分的相关研究，于 IEEE Communications Letters 发表学术论文一篇。

在频率选择性衰落高斯噪声信道环境下，根据多径接收信号产生机制，采用 HMM 作为接收信号的生成模型；在每个调制类型似然函数值计算时，由于前向迭代算法的计算复杂度与马尔科夫链的状态空间大小的平方成正比，所以当状态空间较大时（调制阶数较高或者信道阶数较高），前向迭代算法的复杂度将会非常高。在本文的研究中，我们采用基于前向粒子滤波近似计算接收信号似然函数值的方案来解决这一问题。

在频率选择性衰落高斯噪声信道参数盲估计环节，利用 FB 算法计算隐藏状态后验概率同样存在着算法复杂度随状态空间的增大而呈指数增长的问题。文献 [91] 提出了 JTFS 算法来近似计算隐藏状态的后验概率，但是该算法对信道估计的准确度受信道类型的影响较大，本文则基于 RGSE 算法，提出了利用 BRGSE 算法来近似计算隐藏状态的后验概率，可以有效地解决 FB 算法复杂度受状态空间影响较大以及 JTFS 算法信道估计准确度受信道类型影响较大的问题；同时，

BRGSE 算法具有易于并行计算的特性, 使其更加适用于现代计算平台, 可以进一步加快频率选择性衰落高斯噪声信道下, 参数盲估计的速度。

基于频率选择性衰落高斯噪声信道参数盲估计的相关研究, 于 IEEE Communications Letters 发表学术论文一篇。

6.2 工作展望

基于本文的研究, 发现在实际通信信道下的最大似然调制分类算法还存在以下问题需要进一步地研究解决:

- 1) 在平坦衰落非高斯噪声信道环境下, 通过仿真实验发现在 SNR 较高时, SQUAREM-PC 算法在信道参数估计准确度方面优于 ECM 算法, 关于这一现象的理论解释需要进一步研究;
- 2) 在频率选择性衰落高斯噪声信道环境下, BRGSE 算法的复杂度受信道类型的影响较大, 所以后续研究需要寻找一种受信道类型影响较小的信道参数盲估计算法;
- 3) 在频率选择性衰落高斯噪声信道环境下, 需要深入研究 EM-BRGSE 算法在理论方面的收敛特性;
- 4) 由于本文在研究频率选择性衰落信道环境下的自动调制识别算法时, 假设信道噪声为高斯噪声, 而在实际通信信道下, 信号可能同时经历频率选择性衰落和非高斯噪声的干扰, 所以后续可以假设信道噪声为非高斯噪声进行研究。

参考文献

- [1] <https://baike.baidu.com/item/电子对抗/3366613>.
- [2] <https://baike.baidu.com/item/链路自适应技术/16911566>.
- [3] Sills J A. Maximum-likelihood Modulation Classification for PSK/QAM[C]// Military Communications Conference Proceedings, 1999. MILCOM 1999. IEEE. IEEE, 1999.
- [4] Wei W, Mendel J M. Maximum-likelihood classification for digital amplitude-phase modulations[J]. IEEE Transactions on Communications, 2000, 48(2):189-193.
- [5] Kim K, Polydoros A. Digital modulation classification - The BPSK versus QPSK case[C]// Milcom 88, Century Military Communications-whats Possible? Conference Record Military Communications Conference. IEEE, 2002.
- [6] Polydoros A, Kim K. On the detection and classification of quadrature digital modulations in broad-band noise[J]. IEEE Transactions on Communications, 1990, 38(8):0-1211.
- [7] Huan C Y, Polydoros A. Likelihood methods for MPSK modulation classification[J]. IEEE Transactions on Communications, 1995, 43(234):0-1504.
- [8] Sapiiano P C, Martin J D. Maximum likelihood PSK classifier[C]// IEEE Military Communications Conference. 1996.
- [9] Long C S, Chugg K M, Polydoros A. Further results in likelihood classification of QAM signals[C]// Military Communications Conference. IEEE, 1994.
- [10] Hong L, Ho K C. Classification of BPSK and QPSK signals with unknown signal level using the Bayes technique[C]// International Symposium on Circuits & Systems. IEEE, 2003.
- [11] Beidas B F, Weber C L. Higher-order correlation-based approach to modulation classification of digitally frequency-modulated signals[J]. IEEE Journal on Selected Areas in Communications, 1995, 13(1):89-101.
- [12] Beidas B F, Weber C L. Asynchronous classification of MFSK signals using the higher order correlation domain[J]. IEEE Transactions on Communications, 1998, 46(4):480-493.
- [13] Beidas B F, Weber C L. General framework for the higher-order correlation domain[C]// IEEE Military Communications Conference. IEEE, 1995.

- [14] Beidas B F, Weber C L. Higher-order correlation-based classification of asynchronous MFSK signals[C]// IEEE Military Communications Conference. IEEE, 1996.
- [15] El-Mahdy A E, Namazi N M. Classification of multiple M-ary frequency-shift keying signals over a Rayleigh fading channel[J]. IEEE Transactions on Communications, 2002, 50(6):967-974.
- [16] Abdi A , Dobre O A , Choudhry R , et al. Modulation classification in fading channels using antenna arrays[C]// Military Communications Conference. IEEE, 2004.
- [17] Panagiotou P, Anastasopoulos A, Polydoros A. Likelihood ratio tests for modulation classification[C]// Milcom Century Military Communications Conference. IEEE, 2000.
- [18] Chugg K M, Long C S, Polydoros A. Combined likelihood power estimation and multiple hypothesis modulation classification[C]// Signals, Systems and Computers, 1995. 1995 Conference Record of the Twenty-Ninth Asilomar Conference on. IEEE, 1995.
- [19] Hong L, Ho K C. BPSK and QPSK Modulation classification with unknown signal level[C]// MILCOM 2000. 21st Century Military Communications Conference Proceedings. IEEE, 2000.
- [20] Hong L, Ho K C. Modulation classification of BPSK and QPSK signals using a two element antenna array receiver[C]// IEEE Military Communications Conference. IEEE, 2001.
- [21] Hong L, Ho K C. Antenna array likelihood modulation classifier for BPSK and QPSK signals[C]// Milcom. IEEE, 2002.
- [22] Lay N E, Polydoros A. Per-survivor processing for channel acquisition, data detection and modulation classification[C]// Proceedings of 1994 28th Asilomar Conference on Signals, Systems and Computers. IEEE Computer Society, 1994.
- [23] Lay N E, Polydoros A. Modulation classification of signals in unknown ISI environments[C]// IEEE Military Communications Conference. IEEE, 1995.
- [24] Dobre O A, Zarzoso J, Bar-Ness Y, et al. On the classification of linearly modulated signals in fading channel[C]//Proc. CISS. 2004, 26.
- [25] Abdi A, Dobre O A, Choudhry R, et al. Modulation classification in fading channels using antenna arrays[C]// Military Communications Conference. IEEE, 2004.
- [26] Hameed F, Dobre O, Popescu D. On the likelihood-based approach to modulation classification[J]. IEEE Transactions on Wireless Communications, 2009, 8(12):5884-5892.

- [27] Ramezani-Kebrya A, Kim I M, Kim D I, et al. Likelihood-Based Modulation Classification for Multiple-Antenna Receiver[J]. IEEE Transactions on Communications, 2013, 61(9):3816-3829.
- [28] 王明绪,张松泓,张成根.航空信道下基于假设检验的MQAM信号识别[J].国外电子测量技术,2018,37(07):15-18.
- [29] 王明绪,朱西方,黄国庆,张成根.基于分集接收技术和ALRT算法的MQAM信号识别[J].电子测量技术,2017,40(09):155-158.
- [30] 吴迪,葛临东,彭华.基于EM算法的联合调制识别与参数估计[J].四川大学学报(工程科学版),2014,46(01):114-120.
- [31] Conover W J. Practical Nonparametric Statistics, John Wiley & Sons[J]. INC, New York, 1999.
- [32] Wang F, Wang X. Fast and robust modulation classification via Kolmogorov-Smirnov test [J]. IEEE Transactions on Communications, 2010, 58(8): 2324-2332.
- [33] 付永明. 基于拟合优度检验的通信信号检测与识别技术[D].国防科学技术大学,2015.
- [34] Swami A, Sadler B M. Hierarchical digital modulation classification using cumulants[J]. IEEE Transactions on communications, 2000, 48(3): 416-429.
- [35] Urriza P, Rebeiz E, Pawelczak P, et al. Computationally efficient modulation level classification based on probability distribution distance functions[J]. IEEE Communications Letters, 2011, 15(5): 476-478.
- [36] Wang F, Dobre O A, Chan C, et al. Fold-based Kolmogorov-Smirnov modulation classifier[J]. IEEE Signal Processing Letters, 2016, 23(7): 1003-1007.
- [37] Yuan J, Zhao-Yang Z, Pei-Liang Q. Modulation classification of communication signals[C]// Military Communications Conference. IEEE Xplore, 2004.
- [38] Maliatsos K, Vassaki S, Constantinou P. Interclass and Intraclass Modulation Recognition using the Wavelet Transform[C]// IEEE International Symposium on Personal. IEEE, 2007.
- [39] Fu-Qing H, Jian-Li C, Yi-Tao X U, et al. Modulation Recognition of Symbol Shaped Digital Signals[J]. Electronic Information Warfare Technology, 2008:328-332.
- [40] De Vito L, Rapuano S, Villanacci M. Prototype of an Automatic Digital Modulation Classifier Embedded in a Real-Time Spectrum Analyzer[J]. IEEE Transactions on Instrumentation and Measurement, 2010, 59(10):2639-2651.

- [41] Liu Y, Liang G, Xu X, et al. The Methods of Recognition for Common Used M-ary Digital Modulations[C]// 2008 4th International Conference on Wireless Communications, Networking and Mobile Computing. IEEE, 2008.
- [42] Nandi A K, Azzouz E E. Algorithms for automatic modulation recognition of communication signals[J]. IEEE Transactions on Communications, 1998, 46(4):431-436.
- [43] Hazza A, Shoaib M, Saleh A, et al. Robustness of digitally modulated signal features against variation in HF noise model[J]. Eurasip Journal on Wireless Communications & Networking, 2011, 2011(1):24-0.
- [44] Muller F C B F, Cardoso C, Klautau A. A Front End for Discriminative Learning in Automatic Modulation Classification[J]. IEEE Communications Letters, 2011, 15(4):0-445.
- [45] Park C S, Choi J H, Nah S P, et al. Automatic Modulation Recognition of Digital Signals using Wavelet Features and SVM[C]// Advanced Communication Technology, 2008. ICACT 2008. 10th International Conference on. IEEE, 2008.
- [46] Dan W, Xuemai G, Qing G. A new scheme of automatic modulation classification using wavelet and WSVM[C]// International Conference on Mobile Technology. IEEE, 2009.
- [47] Prakasam P, Madheswaran M. Digital modulation identification model using wavelet transform and statistical parameters[J]. Journal of Computer Systems, Networks, and Communications, 2008, 2008.
- [48] Popoola J J, Olst R V. A Novel Modulation-Sensing Method[J]. IEEE Vehicular Technology Magazine, 2011, 6(3):60-69.
- [49] Li J, He C, Chen J, et al. Automatic digital modulation recognition based on euclidean distance in hyperspace[J]. IEICE transactions on communications, 2006, 89(8): 2245-2248.
- [50] 谭晓衡,褚国星,张雪静,杨扬.基于高阶累积量和小波变换的调制识别算法[J].系统工程与电子技术,2018,40(01):171-177.
- [51] 王东旭,陈殿仁,李兴广,陈磊.基于小波变换的调制识别算法[J].数字技术与应用,2017(01):120-122.
- [52] Orlic V D, Dukic M L. Multipath channel estimation algorithm for automatic modulation classification using sixth-order cumulants[J]. Electronics Letters, 2010, 46(19):1348.
- [53] 刘沛. 多累积量联合的调制方式识别方法[D].西安电子科技大学,2014.
- [54] 文永康,谭红芳,马勇.基于高阶累积量的 MPSK 信号的调制盲识别方法[J].硅

- 谷,2014,7(20):63-66.
- [55] Dobre O A , Abdi A , Bar-Ness Y, et al. Cyclostationarity-Based Modulation Classification of Linear Digital Modulations in Flat Fading Channels[J]. *Wireless Personal Communications*, 2010, 54(4):699-717.
- [56] Dobre O A, Rajan S, Inkol R. Joint Signal Detection and Classification Based on First-Order Cyclostationarity For Cognitive Radios[J]. *EURASIP Journal on Advances in Signal Processing*, 2009.
- [57] Pedzisz M, Mansour A. Automatic modulation recognition of MPSK signals using constellation rotation and its 4th order cumulant[J]. *Digital Signal Processing*, 2005, 15(3):295-304.
- [58] Gulati S, Bhattacharjee R. Automatic Blind Recognition of Noisy and Faded Digitally Modulated MQAM Signals[C]// *India Conference, 2006 Annual IEEE. IEEE*, 2006.
- [59] Hsue S Z, Soliman S S. Automatic modulation classification using zero crossing[C]//*IEEE Proceedings F (Radar and Signal Processing)*. IET Digital Library, 1990, 137(6): 459-464.
- [60] Hsue S Z, Soliman S S. Automatic modulation recognition of digitally modulated signals[C]//*IEEE Military Communications Conference, 'Bridging the Gap. Interoperability, Survivability, Security'*. IEEE, 1989: 645-649.
- [61] Vito L D, Rapuano S. Validating a method for classifying digitally modulated signals[J]. *Measurement*, 2009, 42(3):427-435.
- [62] O'Shea T J, Corgan J, Clancy T C . Convolutional Radio Modulation Recognition Networks[J]. 2016.
- [63] West N E, O'Shea T J. Deep Architectures for Modulation Recognition[J]. 2017.
- [64] Wu H, Wang Q, Zhou L, et al. VHF radio signal modulation classification based on convolution neural networks[C]//*Matec Web of Conferences*. EDP Sciences, 2018, 246: 03032.
- [65] Yashashwi K, Sethi A, Chaporkar P. A learnable distortion correction module for modulation recognition[J]. *IEEE Wireless Communications Letters*, 2019, 8(1): 77-80.
- [66] Nikias C L, Shao M. Signal processing with alpha-stable distributions and applications[M]. Wiley-Interscience, 1995.
- [67] Gulati K, Evans B L, Andrews J G, et al. Statistics of co-channel interference in a field of

- Poisson and Poisson-Poisson clustered interferers[J]. IEEE Transactions on Signal Processing, 2010, 58(12): 6207-6222.
- [68] Blackard K L, Rappaport T S, Bostian C W. Measurements and models of radio frequency impulsive noise for indoor wireless communications[J]. IEEE Journal on selected areas in communications, 1993, 11(7): 991-1001.
- [69] Kanemoto H, Miyamoto S, Morinaga N. A study on modeling of microwave oven interference and optimum reception[C]//1998 IEEE EMC Symposium. International Symposium on Electromagnetic Compatibility. Symposium Record (Cat. No. 98CH36253). IEEE, 1998, 1: 57-62.
- [70] Nassar M, Gulati K, DeYoung M R, et al. Mitigating near-field interference in laptop embedded wireless transceivers[J]. Journal of Signal Processing Systems, 2011, 63(1): 1-12.
- [71] Shi J, Bettner A, Chinn G, et al. A study of platform EMI from LCD panels-impact on wireless, root causes and mitigation methods[C]//2006 IEEE International Symposium on Electromagnetic Compatibility, 2006. EMC 2006. IEEE, 2006, 3: 626-631.
- [72] Erpek T, McHenry M A, Stirling A. DSA operational parameters with wireless microphones[C]//2010 IEEE Symposium on New Frontiers in Dynamic Spectrum (DySPAN). IEEE, 2010: 1-11.
- [73] Blankenship T K, Rappaport T S. Characteristics of impulsive noise in the 450-MHz band in hospitals and clinics[J]. IEEE Transactions on Antennas and Propagation, 1998, 46(2): 194-203.
- [74] Sanchez M G, De Haro L, Ramón M C, et al. Impulsive noise measurements and characterization in a UHF digital TV channel[J]. IEEE Transactions on electromagnetic compatibility, 1999, 41(2): 124-136.
- [75] Hall H M. A New Model for Impulsive Phenomena: Application to Atmospheric-Noise Communication Channels[R]. STANFORD UNIV CA STANFORD ELECTRONICS LABS, 1966.
- [76] El-Mahdy, A. E. Classification of MFSK signals over time-varying flat correlated fading channels under class-A impulsive noise environment[J]. IEE Proceedings-Communications, 2004, 151(6):619-0.
- [77] Chavali, Silva D. Maximum-Likelihood Classification of Digital Amplitude-Phase

- Modulated Signals in Flat Fading Non-Gaussian Channels[J]. IEEE Transactions on Communications, 2011, 59(8):2051-2056.
- [78] Puengnim A, Robert T, Thomas N, et al. Hidden Markov models for digital modulation classification in unknown ISI channels[C]//2007 15th European Signal Processing Conference. IEEE, 2007: 1882-1886.
- [79] Needham T. A Visual Explanation of Jensen's Inequality[J]. American Mathematical Monthly.
- [80] Anderson B D O, Moore J B. Optimal filtering[M]. Courier Corporation, 2012.
- [81] Middleton D. Statistical-Physical Models of Electromagnetic Interference[J]. IEEE Trans.Electromagn.Compat. 1977, 19.
- [82] Vastola, K. Threshold Detection in Narrow-Band Non-Gaussian Noise[J]. IEEE Transactions on Communications, 1984, 32(2):0-139.
- [83] Kozick R J, Sadler B M. Maximum-likelihood array processing in non-Gaussian noise with Gaussian mixtures[J]. IEEE Transactions on Signal Processing, 2000, 48(12):3520-3535.
- [84] Chavali V G, Da Silva C R C M. Detection of Digital Amplitude-Phase Modulated Signals in Symmetric Alpha-Stable Noise[J]. IEEE Transactions on Communications, 2012, 60(11):3365-3375.
- [85] Kay S M. Fundamentals of Statistical Signal Processing: Detection Theory Volume II[M]// false. 1998.
- [86] Roland R V. Simple and Globally Convergent Methods for Accelerating the Convergence of Any EM Algorithm[J]. Scandinavian Journal of Statistics, 2008, 35(2):335-353.
- [87] Raydan M, Svaiter B F. Relaxed steepest descent and Cauchy-Barzilai-Borwein method[M]. Kluwer Academic Publishers, 2002.
- [88] Rabiner L R. A tutorial on hidden Markov models and selected applications in speech recognition[J]. Proceedings of the IEEE, 1989, 77(2): 257-286.
- [89] <https://baike.baidu.com/item/中心极限定理/829451?fr=aladdin>
- [90] Kaleh G K, Vallet R. Joint parameter estimation and symbol detection for linear or nonlinear unknown channels[J]. IEEE Transactions on Communications, 1994, 42(7): 2406-2413.
- [91] Barembbruch S, Garivier A, Moulines E. On approximate maximum-likelihood methods for blind identification: how to cope with the curse of dimensionality[J]. IEEE Transactions on

- Signal Processing, 2009, 57(11): 4247-4259.
- [92] Jacob P E, Lindsten F, Schön T B. Smoothing with couplings of conditional particle filters[J]. Journal of the American Statistical Association, 2019 (just-accepted): 1-19.
- [93] Jacob P E, Lindsten F, Schön T B. Coupling of particle filters[J]. arXiv preprint arXiv:1606.01156, 2016.
- [94] Lindsten F, Jordan M I, Schön, Thomas B. Particle Gibbs with Ancestor Sampling[J]. Journal of Machine Learning Research, 2014, 15(15):2145–2184.
- [95] Asmussen, Søren, Glynn P W. Stochastic Simulation: Algorithms and Analysis[J]. Interfaces, 2007, 38(6):487-488.
- [96] Glasserman P, Yao D D . Some guidelines and guarantees for common random numbers[M]. INFORMS, 1992.
- [97] Chopin N, Singh S S. On particle Gibbs sampling[J]. Bernoulli, 2015, 21(3): 1855-1883.
- [98] Jasra A, Kamatani K, Law K J H, et al. Multilevel particle filters[J]. SIAM Journal on Numerical Analysis, 2017, 55(6): 3068-3096.

致 谢

时光荏苒，转眼在中科院上海微系统与信息技术研究所硕博连读的第六个年头就要结束了。回顾这六年的科研生涯，有亢奋、有苦楚、有激动，但更多的，还是感激！

感谢我的导师梁旭文研究员。初入实验室时，他教会我科研生涯中最重要的是制定短期与长期目标，秉承这一方法，我才可以一步一个脚印地攀登科研这座高峰；当我面对纷杂的科研课题而感到迷茫之时，他给我指明了前行的道路，让我能够在科研的道路上坚定地走下去；当一个一个的难题困扰着我时，他给予了我悉心地指导，培养了我科学的思维方式，让我能够突破一个又一个的困难；环境对于一个人的成长至关重要，非常感谢我的导师给我提供了优越的科研环境以及舒适的生活环境。

感谢我的师兄马陆副研究员和谢卓辰副研究员，给我提供了科研经费支持以及对于课题研究中的帮助。感谢钱玉璧师弟提供的 VPN，让我能够随时随地下载阅读各大数据库中的文献。感谢丁晟、胡伟胜、陈毅君、吴康师兄，侯绩玲师姐以及邹彦良师弟给予的科研以及生活中的帮助。

感谢中科院上海微系统与信息技术研究所人才教育处余翔波、邵天浩、罗琦、谢佳、沈玲燕老师，以及上海微小卫星创新研究院的贾冰清、陈家铭老师在工作 and 生活上不厌其烦的关心和帮助，让我们这些身在外地的学子能够感受到大家庭的温暖。

感谢余爱生同学给我提供的工作站，加快了 my 实验仿真的速度；感谢方晗同学在健身方面给予我的指导，以及在科研和生活上对我的鼓励和帮助；感谢孟祥、魏可友、游文斌、张身志等同窗好友给我生活带来的欢乐。因为有你们，科研生活不再枯燥、单调。

感谢我的未婚妻姚瑶女士一直以来的陪伴与支持；感谢我的父母对我的养育之恩，感谢他们在学习上给予我的全力支持，以及在生活中对我无微不至的照顾。他们的关心和支持，给了我无尽的力量来攻克一道道难关。

最后，由衷地感谢所有关心和帮助过我的人！

作者简历及在攻读博士学位期间发表的学术论文与研究成果

基本信息

姓名	陈文豪
性别	男
出生年月	1991 年 3 月
籍贯	江苏淮安

学习经历

2009.09 – 2013.06	华东理工大学	信息工程专业	工学学士
2013.09 – 2019.09	中科院上海微系统与信息技术研究所	通信与信息专业	工学博士

攻读博士学位期间发表的学术论文

- 1) Chen W, Ma L, Liang X. Blind Identification Based on Expectation-Maximization Algorithm Coupled With Blocked Rhee–Glynn Smoothing Estimator[J]. IEEE Communications Letters, 2018, 22(9): 1838-1841.
- 2) Chen W, Xie Z, Ma L, et al. A Faster Maximum-Likelihood Modulation Classification in Flat Fading Non-Gaussian Channels[J]. IEEE Communications Letters, 2019.