中国科学 : 信息科学 2026 年 第 56 卷 第 11 期: 2879–2894
SCIENTIA SINICA Informationis
Agentic AI 赋能的 6G 融合通信专题 • 论文

面向 6G 高动态场景的基于条件扩散模型的 OTFS 信道估计方法

郑殿鑫[1], 杨禹志[2], 郭天昊[1*]

1. 山西大学物理电子工程学院, 无线通信与检测山西省重点实验室, 太原 030006, 中国
2. Digital Future Institute, Khalifa University of Science and Technology, Abu Dhabi 127788, UAE
* 通信作者. E-mail: tianhao_guo@sxu.edu.cn

收稿日期: 2026–04–07; 修回日期: 2026–08–10; 接受日期: 2026–10–23; 网络出版日期: 2026–11–09

山西省基础研究计划 (批准号: 202303021212014) 和山西省回国留学人员科研 (批准号: 2023-005) 资助项目

摘要 正交时频空间调制适用于 6G 高动态场景, 但在低信噪比和高多普勒条件下, 时延 – 多普勒域信道估计仍面临严重的噪声放大与跨场景泛化困难等问题. 本文提出一种基于条件扩散模型的正交时频空间信道估计方法. 该方法以嵌入式导频提取的最小二乘估计为初始锚点, 采用复数值 U-Net 保留相位信息, 引入速度 – 信噪比联合条件嵌入刻画环境变化, 并结合 v-prediction 与去噪扩散隐式模型采样降低推理开销. 基于 DeepMIMO 射线追踪数据的仿真结果表明, 所提方法在 0∼10 dB 信噪比和 10∼120 km/h 速度范围内均优于最小二乘、最小均方误差、正交匹配追踪、Transformer 和 ChannelNet 等代表性基线; 在 10 和 0 dB 时, 全速度平均归一化均方误差分别为 −23.17 和 −20.02 dB. 多个未见城市场景测试表明, 所提方法在不同城市传播环境下具有初步的跨场景泛化能力. 端到端误码率评估进一步表明, 在当前理想同步与无编码 QPSK 的统一检测链路中, 所提方法达到 10^{-2} 目标误码率所需的信噪比较最小二乘估计低约 3.20 dB.

关键词 OTFS 调制, 信道估计, 条件扩散模型, 时延 – 多普勒域, 6G 高动态场景

1 引言

随着未来 6G 网络的持续演进, 高铁、车联网以及低轨卫星等高移动性场景已成为新一代无线通信系统的核心业务形态 [1,2]. 在这些超高动态场景下, 极高的相对运动速度会引发显著的多普勒频移, 导致传统正交频分复用 (orthogonal frequency division multiplexing, OFDM) 技术的子载波正交性遭到破坏, 引发严重的子载波间干扰, 使得系统性能明显下降 [3,4]. 为应对这一挑战, 研究者提出了正交时频空间 (orthogonal time frequency space, OTFS) 调制技术 [5,6]. OTFS 通过二维辛有限傅里叶变换 (symplectic finite Fourier transform, SFFT), 将信号映射至时延 – 多普勒 (delay-Doppler, DD) 域. 这种变换将原本在时频域快速变化的双选信道, 转化为 DD 域中相对稀疏且较为稳定的二维信道结构, 从而为提取全时频分集增益提供了物理可能 [7,8]. 然而, OTFS 系统理论优势的释放, 高度依赖于接收端能否获取准确的 DD 域信道状态信息 (channel state information, CSI). 在恶劣环境下若难以实现准确的信道估计, 后续的信号检测与均衡将受到影响 [9,10].

引用格式: 郑殿鑫, 杨禹志, 郭天昊. 面向 6G 高动态场景的基于条件扩散模型的 OTFS 信道估计方法. 中国科学: 信息科学, 2026, 56: 2879–2894, doi: 10.1360/SSI-2026-0135
Zheng D X, Yang Y Z, Guo T H. Conditional diffusion-based OTFS channel estimation for 6G high-mobility scenarios. Sci Sin Inform, 2026, 56: 2879–2894, doi: 10.1360/SSI-2026-0135

© 2026《中国科学》杂志社

目前, 针对 OTFS 系统的信道估计方法主要分为传统数理统计算法与数据驱动的深度学习算法两大类. 在传统算法中, 最小二乘 (least squares, LS) 算法因其较低的计算复杂度被广泛采用, 但理论与实践均表明, 在低信噪比 (signal-to-noise ratio, SNR) 条件下, LS 估计容易受到噪声放大效应的影响, 导致估计精度明显下降 [11]; 同时, 为了对抗严重干扰, 传统估计方案往往需要给 OTFS 系统带来较高的导频开销. 最小均方误差 (minimum mean square error, MMSE) 算法虽然能够利用信道的二阶统计特性提升精度, 但其涉及复杂的高维矩阵求逆运算, 且难以在高速时变场景下实时获取准确的协方差矩阵, 因而难以在实际高动态网络中高效部署 [12]. 此外, 基于正交匹配追踪 (orthogonal matching pursuit, OMP) [13] 与稀疏贝叶斯学习 (sparse bayesian learning, SBL) [14] 等压缩感知算法虽然有效利用了 DD 域的信道稀疏性, 但在实际复杂电磁环境下却面临两大问题: 其一, 对基扩展模型的网格失配极为敏感 [15], 易出现迭代发散 [16,17]; 其二, 为了保证信道恢复的精确度, 此类算法往往会给系统带来较大的导频开销 [18].

近年来, 基于深度学习的信道估计逐渐成为重要研究方向 [19,20]. 然而, 现有的深度学习估计算法多关注于传统 OFDM 系统的空 – 频域或角度 – 时延域信道, 其信道结构与 OTFS 独特的 DD 域二维稀疏流形存在本质差异. 此外, OTFS 系统为了充分获取分集增益, 其二维网格规模通常远大于传统 OFDM 系统, 这为高维信道特征的提取与模型的稳定收敛带来了额外的计算负担. 另一方面, 现有判别式信道估计方法通常采用 ChannelNet 等卷积架构 [21] 或 Transformer 架构 [22], 并以均方误差作为训练目标. 在极低信噪比下, 此类方法往往难以保持稳定估计性能, 容易破坏 DD 域中的多径尖峰结构 [23,24]. 近期, 已有研究将分数模型或扩散先验用于常规多输入多输出 (multiple-input multiple-output, MIMO) 信道矩阵的后验采样与迭代恢复 [25,26], 验证了生成式先验用于高维 CSI 估计的可行性. 然而, 这些方法主要围绕常规 MIMO 信道矩阵或角度域稀疏展开, 尚未解决 OTFS DD 域中的二维循环卷积结构、嵌入式导频 LS 观测以及高动态物理条件联合表征问题.

基于上述分析, 本文将条件扩散模型 (conditional diffusion models, CDM) [27,28] 引入 OTFS 信道估计. 依托其逐步去噪与分数匹配机制, 本文提出了一种面向复杂时变场景的 OTFS 信道估计方法. 该方法将初始的 LS 估计结果作为物理锚点, 将速度与 SNR 作为联合条件, 面向复杂时变场景恢复 DD 域信道.

本文的主要贡献总结如下.

• 本文提出了一种基于条件扩散模型的 OTFS 信道估计方法. 该方法面向 MIMO-OTFS DD 域 CSI, 以嵌入式导频 (embedded pilot, EP) 和保护间隔 (guard interval, GI) 提取的 LS 粗估计为物理锚点, 将信道恢复构造为相对于该锚点的条件残差生成过程, 并由物理锚点与所学残差叠加得到最终估计.

• 本文设计了复数值去噪架构与多物理条件显式嵌入机制. 针对 OTFS 信号的复数物理特征, 定制了复数值 U-Net 去噪网络, 较好地保留了底层物理相位信息. 同时, 引入速度和 SNR 联合显式条件嵌入机制, 用于调节模型对不同环境条件的响应并自适应调整去噪过程.

• 本文进一步降低了扩散模型的推理开销, 并在受控链路下评估了信道估计误差对检测性能的影响. 引入 v-prediction 目标重构与去噪扩散隐式模型 (denoising diffusion implicit models, DDIM) 少步采样, 以较少的网络调用完成信道恢复. 在理想同步、无编码的统一检测链路中, 本文方法达到目标误码率 (bit error rate, BER) 所需的 SNR 明显低于 LS.

为便于实验复现和后续研究, 本文所提方法的相关实现代码已公开[1)].

2　相关工作

现有信道估计研究可按照物理先验的使用方式与估计目标分为传统模型驱动方法、判别式深度学习方法和生成式方法. 传统模型驱动方法利用导频结构、信道统计或 DD 域稀疏性求解信道参数. 判别式深度学习方法学习从含噪观测到目标信道的确定性映射. 生成式方法则进一步学习信道先验或观测条件下的概率分布.

1) 代码仓库: https://github.com/tellsgo/OTFS-Diffusion-CE.

2.1　传统 OTFS 信道估计方法

传统 OTFS 信道估计首先依赖导频结构与解析观测模型 [29]. 嵌入式导频方法通过在导频周围设置保护间隔, 将局部接收网格中的数据干扰与目标信道响应分离, 从而能够利用阈值判决或 LS 除法直接形成初始估计 [11]. 这类方法计算过程清晰、无需离线训练, 但 LS 估计会将观测噪声直接传递至信道结果; 当信噪比较低或保护区不能充分覆盖时延和多普勒扩展时, 估计误差与数据泄漏会更加明显. MMSE 类方法可结合二阶信道统计抑制噪声, 但性能依赖协方差估计的准确性, 并涉及矩阵求逆或相应的近似计算 [12].

另一类模型驱动方法利用 DD 域信道由少量物理路径构成的结构稀疏性, 将估计转化为稀疏恢复问题. 正交匹配追踪能够以较低存储开销逐步选择候选路径 [13], 稀疏贝叶斯学习则通过层次概率模型联合估计稀疏系数及其统计参数 [14,17]. 对于连续时延或分数多普勒造成的网格失配, 已有研究进一步采用离网格稀疏贝叶斯学习或基扩展模型进行修正 [15,16]. 这些改进增强了物理可解释性, 但通常需要迭代求解, 并对稀疏字典、停止准则和导频配置较为敏感; 数据辅助或叠加导频设计能够在估计精度与资源开销之间提供新的折中 [18], 但仍需结合具体收发结构选择参数.

综合而言, 传统方法的优势在于观测关系明确且无需训练, 其局限主要来自噪声放大、统计失配和迭代求解开销. 本文保留低复杂度 LS 结果作为接收观测提供的确定性锚点, 并利用学习模型恢复固定解析规则难以刻画的误差成分, 形成由物理观测到残差精化的估计过程.

2.2　判别式深度学习的信道估计方法

为降低对精确统计先验和迭代优化的依赖, 数据驱动的深度学习 (deep learning, DL) 信道估计逐渐受到关注 [30,31]. 在 OFDM 等系统中, ChannelNet 和 ReEsNet 分别利用卷积特征提取与残差学习, 将粗糙估计或插值结果映射为精化信道 [21,32]; 这些架构也为 OTFS 信道二维特征处理提供了可借鉴的判别式基线. 针对 OTFS 任务, 已有 Transformer 方法先由阈值策略获得初始信道, 再利用跨时刻相关性和自注意力机制进行信道响应预测与精化 [22,33]. 模型驱动学习则将通信机理嵌入网络结构, 以降低对训练分布的依赖 [34,35].

判别式网络通常只需一次前向传播, 因而具有较低推理开销; 在训练与测试分布一致时, 其监督映射能够获得较好的估计精度. 但其性能依赖训练数据覆盖范围和损失函数定义. 当含噪观测对应多个可能信道时, 或测试场景偏离训练分布时, 以均方误差为目标的单点回归倾向于学习条件均值, 可能平滑弱多径尖峰或复数相位细节 [23,24]. 这表明, 在强噪声和高动态分布变化下, 仅依赖固定的单步映射可能难以显式表达观测不确定性.

因此, 判别式方法为本文提供两方面启示. 一方面, 利用粗估计与目标信道之间的残差降低学习难度; 另一方面, OTFS 信道实部与虚部之间的耦合关系需要在特征提取中显式表征. 本文以 LS 观测为锚点, 采用复数值网络提取信道特征, 并将确定性单步回归扩展为条件残差的分布建模与渐进恢复.

2.3　生成式信道估计与本文定位

生成式模型通过学习信道样本的概率结构, 以不同于单点回归的方式对观测不确定性建模. 早期研究采用生成对抗网络 (generative adversarial network, GAN) 和变分自编码器 (variational autoencoder, VAE) 进行随机信道建模、端到端通信或信道去噪 [36~39]. GAN 通过对抗训练学习复杂样本分布, 但其训练过程需要协调生成器与判别器; VAE 具有显式潜变量结构, 但变分近似可能限制复杂分布的表达. 这些研究验证了数据驱动信道分布建模的可行性, 并推动了扩散与分数模型在物理层中的进一步应用.

扩散模型 (diffusion models, DM) 通过逐步加噪和逆向去噪学习数据分布 [40]. 在信道估计领域, Arvinte 等 [25] 针对常规 MIMO 信道矩阵学习对数先验梯度, 并将训练后的分数模型与导频测量似然结合, 在测试阶段通过后验采样迭代恢复信道. Fesl 等 [26] 则在稀疏角度域学习轻量级扩散先验, 对 SNR 进行位置编码, 并通过避免重采样、截断低 SNR 逆扩散阶段来降低复杂度. 这些工作表明, 生成式先验可与当前测量结合完成 MIMO 信道估计. 本文进一步面向 MIMO-OTFS DD 域任务, 针对该域的观测结构与高动态条件重新建模.

在建模对象与观测组织方面, 已有代表性方法主要学习常规 MIMO 信道矩阵或角度域稀疏先验, 本文则直接处理 MIMO-OTFS DD 域复数 CSI 张量, 使网络输入对应二维循环卷积形成的多径结构. 已有分数先验方法

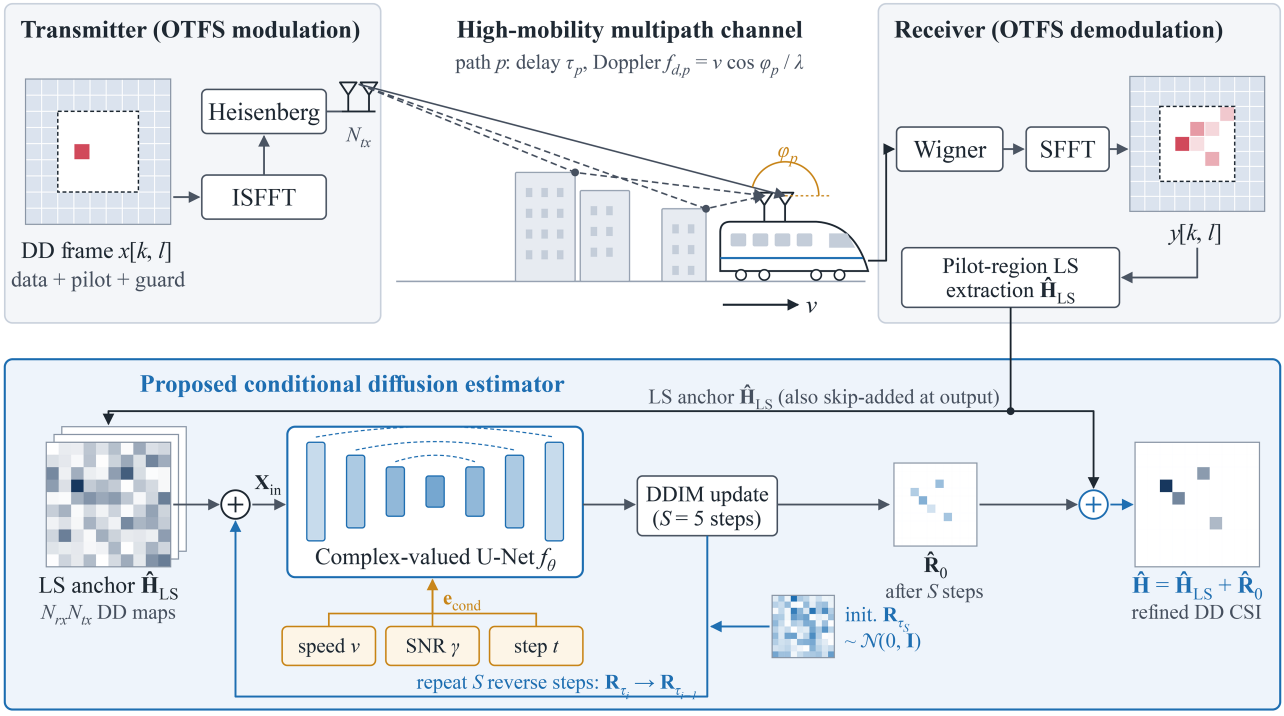

图 1　(网络版彩图) 基于条件扩散模型的 MIMO-OTFS 系统与信道估计整体架构.

Figure 1　(Color online) Conditional-diffusion MIMO-OTFS system and channel-estimation architecture.

在采样过程中显式使用导频测量似然, 本文先由嵌入式导频形成逐链路 LS 结果, 再将堆叠后的完整 MIMO LS 张量作为物理锚点, 学习真实信道相对于该锚点的条件残差. 在条件信息与推理机制方面, 已有轻量级扩散估计器利用 SNR 调节推理过程. 本文面向高速移动场景联合输入速度和 SNR, 以分别表征移动性与噪声水平, 并结合 v-prediction 训练目标和 DDIM 少步采样, 调节估计精度与迭代开销之间的折中. 综上, 本文以 MIMO-OTFS DD 域 LS 张量为物理锚点, 通过复数条件残差建模与 DDIM 少步采样, 形成面向高动态场景的信道估计框架.

3　系统模型

3.1　MIMO-OTFS 收发信号模型

本文考虑采用 N_{tx} 根发射 (transmitter, Tx) 天线与 N_{rx} 根接收 (receiver, Rx) 天线的 MIMO-OTFS 通信系统 [41], 其包含条件扩散信道估计模块的整体收发架构, 如图 1 所示. 本文采用 $M \times N$ 的离散 DD 域表示, 其中 M, N 分别为多普勒与时延维度的采样数. 令 $x[k,l]$ 表示网格点 (k,l) 处的发送数据符号, 其中 $k \in \{0,\ldots,M-1\}$, $l \in \{0,\ldots,N-1\}$ 分别为多普勒与时延索引.

本文采用基于 OFDM 的 OTFS 收发机 [6]. 发送端通过二维逆辛有限傅里叶变换 (inverse symplectic finite Fourier transform, ISFFT) 和海森堡变换将 DD 域符号依次映射至时频 (time-frequency, TF) 域和时域; 接收端则通过维格纳变换与二维辛有限傅里叶变换完成逆向映射. 在理想成型脉冲假设下, 固定一组 Tx-Rx 链路的 DD 域接收信号 $y[k,l]$ 可表示为发送信号与 DD 域信道响应的二维离散周期循环卷积 [9]

$$y[k,l] = \sum_{k'=0}^{M-1} \sum_{l'=0}^{N-1} h(k',l') x[(k-k')_M, (l-l')_N] + n[k,l], \tag{1}$$

其中, $(\cdot)_M$ 与 $(\cdot)_N$ 分别表示模 M 与模 N 运算, $n[k,l] \sim \mathcal{CN}(0,\sigma_w^2)$ 为加性复高斯白噪声. 相比于时频域显著的

衰落起伏, DD 域信道具有固有的结构稀疏性, 可表示为少量离散物理传播路径的叠加

$$h(k,l)=\sum_{p=1}^{P}\alpha_p\delta[k-k_p]\delta[l-l_p], \tag{2}$$

式中, P 为多径总数; 对第 p 条路径, α_p, k_p 与 l_p 分别为其复增益、离散多普勒频移索引与时延索引; $\delta[\cdot]$ 为离散克罗内克冲激函数. 在扩展至 MIMO 场景时, 系统整体的 DD 域 CSI 可表示为高维复数张量 $\mathbf{H}\in\mathbb{C}^{N_{rx}\times N_{tx}\times M\times N}$. 对于完整的 MIMO 数据传输, 第 r 根接收天线的 DD 域信号是所有发射天线贡献的叠加, 即

$$y_r[k,l]=\sum_{t=1}^{N_{tx}}\sum_{k'=0}^{M-1}\sum_{l'=0}^{N-1}h_{r,t}(k',l')x_t[(k-k')_M,(l-l')_N]+n_r[k,l], \tag{3}$$

其中 $h_{r,t}$ 表示第 t 根发射天线到第 r 根接收天线的 DD 域信道. 式 (1) 是固定一组 Tx-Rx 索引后用于说明 OTFS 二维卷积结构的单链路形式, 式 (3) 则给出了其在完整 MIMO 数据传输阶段的叠加关系.

3.2　时变信道物理建模

为描述高速移动场景下的时变传播特性, 本文采用基于几何的多径信道模型. 对于第 t 根发射天线到第 r 根接收天线的链路, 第 p 条传播路径由复增益 $\alpha_{r,t,p}$、到达时延 τ_p 以及路径到达方向与移动方向之间的夹角 ϕ_p 表征. 当移动终端以速度 v 运动时, 第 p 条路径对应的多普勒频移为

$$f_{d,p}=\frac{v}{\lambda}\cos(\phi_p), \tag{4}$$

其中 $\lambda=c/f_c$ 为载波波长. 由各路径的时延和多普勒相位叠加可得相应的时频域信道响应

$$H_{\mathrm{TF}}^{(r,t)}[m,n]=\sum_{p=1}^{P}\alpha_{r,t,p}\mathrm{e}^{\mathrm{j}2\pi f_{d,p}mT_{\mathrm{sym}}}\mathrm{e}^{-\mathrm{j}2\pi n\Delta f\tau_p}, \tag{5}$$

其中 T_{sym} 为符号周期, Δf 为子载波间隔, m 和 n 分别表示时频域中的时间与频率索引. 接收端通过二维辛有限傅里叶变换将其映射至 DD 域,

$$H_{\mathrm{DD}}^{(r,t)}[k,l]=\mathrm{SFFT}\left\{H_{\mathrm{TF}}^{(r,t)}[m,n]\right\}. \tag{6}$$

由式 (4)∼(6) 可见, 速度通过各路径的多普勒相位共同影响 DD 域信道的能量分布与相位结构. 信道样本构造采用三维射线追踪获得的连续路径增益、时延与角度参数, 并依据上述关系映射为 $M\times N$ 的 DD 域张量. 式 (2) 是路径时延与多普勒恰好落在整数网格点时的理想形式; 实际映射保留分数多普勒分量及其在相邻网格中的泄漏.

3.3　信道估计目标与初始重构

在实际的通信接收端, 系统的主要目标是从叠加了高斯白噪声的观测张量 $\mathbf{Y}$ 中, 准确重构出未知的时变信道复数张量 $\mathbf{H}$ [9, 26, 42]. 传统 OFDM 的梳状或块状导频因 DD 域二维周期循环卷积会与数据符号相互干扰, 难以直接用于 OTFS, 因此本文采用广泛使用的嵌入式导频方案 [9].

下述公式首先以单条 Tx-Rx 链路说明嵌入式导频观测与 LS 提取过程. 在 MIMO 信道训练阶段, 本文采用发射天线间的时分正交嵌入式导频. 一个训练周期包含 N_{tx} 个相互正交的导频时隙, 在第 t 个时隙内仅第 t 根发射天线发送嵌入式导频帧, 其余发射天线保持静默, 全部 N_{rx} 根接收天线同时获取对应的导频观测. 因此, 不同发射天线的导频响应在时间资源上相互分离, 各接收天线能够分别提取相应 Tx-Rx 链路的 LS 估计. 时分正交仅用于信道训练阶段, 完整 MIMO 数据传输仍采用式 (3) 所示的多发射天线信号叠加模型.

具体而言, 在 DD 域预设位置 (k_p,l_p) 发送高能量孤立导频脉冲 x_p, 并在其周围设置矩形无数据保护间隔以降低二维周期卷积的数据干扰. 结合信道因果性与双选扩展, 保护间隔按最大时延扩展 $l_{\max}$ 与最大多普勒扩展 $k_{\max}$ 非对称设计, 网格结构如图 2 所示. 保护区网格索引集合 $\mathcal{G}$ 定义为

$$\mathcal{G}=\{(k,l)\mid k\in[k_p-2k_{\max},k_p+2k_{\max}],l\in[l_p-l_{\max},l_p+l_{\max}]\}. \tag{7}$$

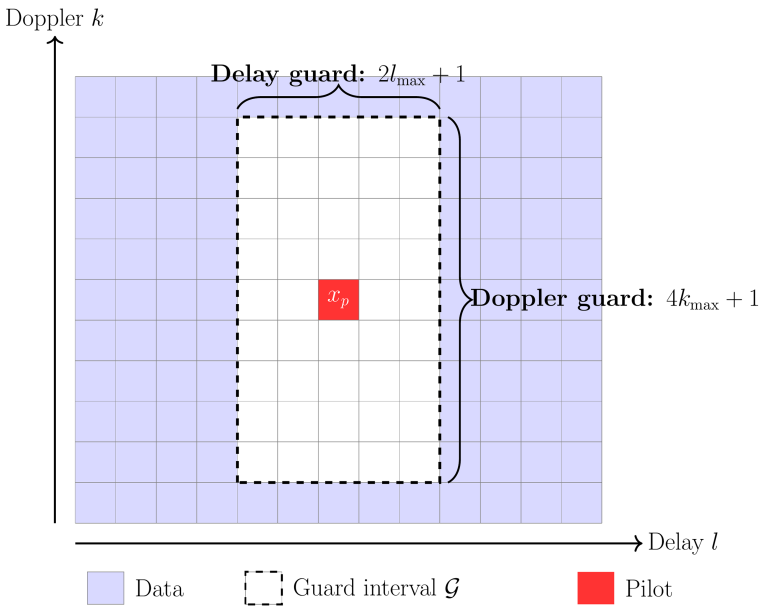

图 2　(网络版彩图) OTFS 嵌入式导频与二维保护间隔 DD 域网格结构示意图.

Figure 2　(Color online) Schematic diagram of OTFS embedded pilot and 2D guard interval structure in DD domain.

在此区域内, 发送符号被严格置零 (即 $x[k,l]=0, \forall (k,l)\in \mathcal{G}\setminus\{(k_p,l_p)\}$). 由式 (7) 可得, 对于 $M\times N$ DD 域张量, 单个嵌入式导频及其保护区的归一化占用率为

$$\eta_{\mathrm{p}}(k_{\max})=\frac{|\mathcal{G}|}{MN}=\frac{(4k_{\max}+1)(2l_{\max}+1)}{MN}. \tag{8}$$

相应的数据网格比例为 $1-\eta_{\mathrm{p}}$. 该定义表示一个发射天线导频时隙内的嵌入式导频及保护区占用. 完整 MIMO 训练周期包含 N_{tx} 个时分正交导频时隙, 因而系统总训练开销还取决于导频周期与数据帧的组织方式.

在保护间隔充分覆盖信道支撑且满足理想同步条件时, 有效信道支撑区内的数据干扰可以忽略, 相应观测可表示为导频脉冲与对应信道响应的乘积叠加背景噪声

$$y[(k_p+k)_M,(l_p+l)_N]=h[k,l]\cdot x_p+n[(k_p+k)_M,(l_p+l)_N],\quad k\in[-k_{\max},k_{\max}],l\in[0,l_{\max}]. \tag{9}$$

基于上述导频观测, 接收端可直接截取相应信号, 并利用标量除法进行低复杂度的粗略初始信道提取 $\hat{\mathbf{H}}_{\mathrm{LS}}$,

$$\hat{h}_{\mathrm{LS}}[k,l]=\frac{y[(k_p+k)_M,(l_p+l)_N]}{x_p}=h[k,l]+\tilde{n}[k,l], \tag{10}$$

其中 $\tilde{n}[k,l]=n[\cdot]/x_p$ 为初始提取区域的等效加性高斯白噪声. 对于 MIMO 设置, 记 $y_r^{(t)}$ 为第 r 根接收天线在第 t 根发射天线对应的正交导频时隙中获得的观测, 则相应的逐链路导频观测与 LS 估计可写为

$$\begin{aligned} y_r^{(t)}[(k_p+k)_M,(l_p+l)_N]&=h_{r,t}[k,l]x_{p,t}+n_r^{(t)}[(k_p+k)_M,(l_p+l)_N],\\ \hat{h}_{\mathrm{LS}}^{(r,t)}[k,l]&=\frac{y_r^{(t)}[(k_p+k)_M,(l_p+l)_N]}{x_{p,t}}=h_{r,t}[k,l]+\tilde{n}_r^{(t)}[k,l]. \end{aligned} \tag{11}$$

对全部 $r=1,\ldots,N_{rx}$ 和 $t=1,\ldots,N_{tx}$ 重复上述提取并按天线索引堆叠, 得到

$$\hat{\mathbf{H}}_{\mathrm{LS}}=\left[\hat{h}_{\mathrm{LS}}^{(r,t)}[k,l]\right]_{r,t,k,l}\in\mathbb{C}^{N_{rx}\times N_{tx}\times M\times N}. \tag{12}$$

全部链路堆叠后的完整 MIMO LS 张量统一输入扩散网络, 由网络联合各链路特征, 一次输出完整 CSI 张量.

上述嵌入式导频机制避免了高维矩阵求逆与全网格插值, 但在高速移动和低信噪比环境下, 直接提取的物理锚点仍受噪声污染, 有效多径可能被噪声掩盖. 另一方面, 保护间隔需覆盖信道的时延与多普勒扩展, 其资源占用随保护范围增大. 本文利用生成式先验对 LS 初始估计进行条件残差恢复, 以改善 OTFS 信道的估计.

4　基于条件扩散模型的信道估计算法

4.1　条件残差概率建模

传统方法及判别式网络通常求解确定性点估计 $\hat{\mathbf{H}} = f(\mathbf{Y})$, 即从观测到目标信道的单点映射. 然而, 含噪观测下, 真实信道可由后验分布 $p(\mathbf{H}|\mathbf{Y})$ 描述; 在低信噪比或观测不确定性较强的条件下, 仅以均方误差优化单点输出可能难以充分表征这一不确定性. 因此, 本文基于生成式概率建模, 将信道恢复表述为条件分布拟合问题.

具体而言, 本文以低复杂度初始重构 $\hat{\mathbf{H}}_{\mathrm{LS}}$ 为物理锚点, 构建基于深度残差学习 [43] 的生成框架. 模型不直接生成完整信道 $\mathbf{H}$, 而是拟合真实信道与粗估计之间的条件残差分布 $p(\mathbf{R}|\hat{\mathbf{H}}_{\mathrm{LS}})$, 其中 $\mathbf{R} = \mathbf{H} - \hat{\mathbf{H}}_{\mathrm{LS}}$. 当 $\hat{\mathbf{H}}_{\mathrm{LS}} = \mathbf{H} + \mathbf{N}$ 时, 目标残差对应 LS 估计误差的负值, 即 $\mathbf{R} = -\mathbf{N}$. 在常规信噪比下, 残差能量小于信道本身, 使网络能够围绕当前观测集中学习需要修正的部分, 并将点估计扩展为条件残差分布的逐步恢复. 本框架在推理阶段仅以 LS 初估 $\hat{\mathbf{H}}_{\mathrm{LS}}$、速度 v 和 SNR γ 为条件输入, 真实信道 $\mathbf{H}$ 仅在离线训练时作为监督标签.

在低信噪比下残差能量可能与信道相当, 残差学习虽不再具备 “小能量目标” 的优势, 但仍以 LS 观测为参照, 将学习限定于观测误差的条件修正, 并借相对中心化减小样本间绝对衰落尺度的差异. 结合后续的速度和 SNR 条件, 模型学习分布族 $p(\mathbf{R}|\hat{\mathbf{H}}_{\mathrm{LS}}, v, \gamma)$, 从而根据当前观测及环境状态调整残差恢复过程.

为学习该残差分布, 本文引入马尔可夫扩散过程 [27]. 前向扩散过程向目标残差张量 $x_0 = \mathbf{R}$ 中逐步注入高斯噪声, 以构建平滑的数据流形. 在任意时间步 $t \in \{1, \ldots, T\}$, 含噪残差变量 $x_t = \mathbf{R}_t$ 的转移概率分布表示为

$$q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{\alpha}_t}x_0, (1-\bar{\alpha}_t)\mathbf{I}), \tag{13}$$

其中, $\bar{\alpha}_t = \prod_{i=1}^{t}(1-\beta_i)$, 而 $\{\beta_t \in (0,1)\}_{t=1}^{T}$ 为控制每次注入噪声方差大小的预设调度参数. 本文采用余弦噪声调度策略 [44], 使前向扩散早期的噪声方差平缓增加, 从而减缓有效多径结构在早期时间步的衰减.

4.2　物理条件联合嵌入

在高度时变的无线通信环境中, 移动速度 v 影响信道的时变特征和多普勒扩展, 以 γ 表示的 SNR 表征接收观测的噪声水平. 为使扩散模型根据移动状态和噪声条件调节恢复过程, 本文采用速度 – SNR 联合条件嵌入机制. 模型学习由多物理变量索引的条件概率分布 $p(\mathbf{R}|\hat{\mathbf{H}}_{\mathrm{LS}}, v, \gamma)$. 其中, v 和 γ 由模型外部提供.

在实际系统中, 速度状态可由全球导航卫星系统等定位与导航信息 [45] 获得, 也可根据接收信号的多普勒特征进行估计 [46]. SNR 及相关链路质量信息可由导频或参考信号的物理层测量获得 [47]. 实际估计结果可能受到测量误差、反馈延迟和时间失配的影响, 因此条件信息的获取精度是模型实际应用中需要考虑的因素.

为实现这一目标, 本文首先将物理标量映射为高维特征空间. 受 Transformer 位置编码机制启发 [33], 采用正弦位置编码向量 $\mathbf{p}(\cdot)$,

$$\mathbf{p}(x)_{2k} = \sin\left(\frac{x}{10000^{2k/d}}\right), \quad \mathbf{p}(x)_{2k+1} = \cos\left(\frac{x}{10000^{2k/d}}\right). \tag{14}$$

式中 d 为所编码特征的维度. 设定全局特征隐藏层维度为 d_t. 扩散时间步 t 经正弦位置编码与多层感知机映射为特征 $\mathbf{e}_t \in \mathbb{R}^{d_t}$. 对于物理条件, 速度 v 归一化后经线性特征映射得到 $\mathbf{e}_v \in \mathbb{R}^{d_t/2}$; 以 dB 表示的 SNR γ 完成尺度变换后, 经正弦编码与多层感知机得到 $\mathbf{e}_\gamma \in \mathbb{R}^{d_t/2}$. 最终, 将 $\mathbf{e}_v$ 与 $\mathbf{e}_\gamma$ 在特征维度拼接, 再与时间步特征 $\mathbf{e}_t$ 相加, 形成全局条件向量 $\mathbf{e}_{\mathrm{cond}} \in \mathbb{R}^{d_t}$,

$$\mathbf{e}_{\mathrm{cond}} = \mathbf{e}_t + [\mathbf{e}_v \parallel \mathbf{e}_\gamma]. \tag{15}$$

4.3　复数网络与目标重构

本文根据 DD 域 CSI 的复数特性, 采用复数值 U-Net (complex-valued U-Net) [48] 作为去噪网络. 为使网络输入与 DD 域信道张量维度一致, 首先对仅在信道支撑区内具有有效提取值的初始锚点 $\hat{\mathbf{H}}_{\mathrm{LS}}$ 进行外围补零. 对于维度对齐后的输入张量 $\mathbf{X}_{\mathrm{in}} = \hat{\mathbf{H}}_{\mathrm{LS}} + \mathbf{R}_t \in \mathbb{C}^{N_{rx}\times N_{tx}\times M\times N}$, 网络分别学习实部和虚部卷积核, 并按照复数乘法规则对实部与虚部特征进行交叉耦合, 从而显式保留二者的关联并表征相位结构. 随后, 采用加性空间广播 $\mathbf{F}_{\mathrm{out}} = \mathbf{F} + \mathrm{Broadcast}(\mathbf{W}_{\mathrm{proj}}\mathbf{e}_{\mathrm{cond}})$, 以较低的额外计算开销将条件信息注入各尺度特征图 $\mathbf{F}$ 中.

在网络训练目标的设计上, 扩散模型可采用噪声 $\boldsymbol{\epsilon}$ 预测或目标残差 x_0 预测, 二者在不同噪声水平下误差敏感性不同. 当 $\bar{\alpha}_t$ 较小时, 由噪声预测结果反解 x_0 需要除以较小的 $\sqrt{\bar{\alpha}_t}$; 当 $\bar{\alpha}_t$ 接近 1 时, 由 x_0 预测结果反解噪声需要除以较小的 $\sqrt{1-\bar{\alpha}_t}$. 相应预测误差可能在上述反解过程中被放大. 为在不同噪声水平下统一表征目标残差与噪声, 本文采用 v-prediction 目标重构机制 [49], 将网络预测向量定义为二者的线性组合

$$\mathbf{v}_t = \sqrt{\bar{\alpha}_t}\boldsymbol{\epsilon} - \sqrt{1-\bar{\alpha}_t}x_0. \tag{16}$$

$\mathbf{v}_t$ 同时包含目标残差 x_0 与噪声 $\boldsymbol{\epsilon}$ 的信息, 由 x_t 与 $\mathbf{v}_t$ 反解二者时仅涉及 $\sqrt{\bar{\alpha}_t}$ 和 $\sqrt{1-\bar{\alpha}_t}$ 的线性组合, 无需除以接近零的系数. 当 $\bar{\alpha}_t \to 0$ 时, $\mathbf{v}_t \approx -x_0$, 预测目标主要保留信号结构; 当 $\bar{\alpha}_t \to 1$ 时, $\mathbf{v}_t \approx \boldsymbol{\epsilon}$, 其形式接近标准噪声预测. 因此, v-prediction 为不同扩散时间步提供了统一的预测目标, 有助于平衡高、低噪声阶段的监督信息. 基于上述网络架构与目标重构, 算法 1 总结了本框架离线阶段的完整训练流程.

算法 1 基于 v-prediction 的时延 – 多普勒域条件扩散模型离线训练过程.

输入: 训练数据集 $\mathcal{D} = \{(\hat{\mathbf{H}}_{\mathrm{LS}}^{(i)}, \mathbf{H}^{(i)}, v^{(i)}, \gamma^{(i)})\}_{i=1}^{N}$; 预设噪声参数 $\{\bar{\alpha}_t\}_{t=1}^{T}$; 学习率 η.
输出: 优化后的复数值去噪网络参数 θ.

随机初始化网络参数 θ;
while 未达到最大训练轮数 **do**
  for 从 $\mathcal{D}$ 中采样 Mini-Batch $(\hat{\mathbf{H}}_{\mathrm{LS}}, \mathbf{H}, v, \gamma)$ **do**
    计算目标残差 $\mathbf{R} \leftarrow \mathbf{H} - \hat{\mathbf{H}}_{\mathrm{LS}}$, 并采样时间步 $t \sim \mathrm{Uniform}(1, T)$, 噪声 $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$;
    依式 (13) 计算 $\mathbf{R}_t$, 依式 (16) 计算预测目标 $\mathbf{v}_t$;
    依式 (15) 融合 t, v, γ 得到引导向量 $\mathbf{e}_{\mathrm{cond}}$;
    构建显式条件组合输入: $\mathbf{X}_{\mathrm{in}} \leftarrow \hat{\mathbf{H}}_{\mathrm{LS}} + \mathbf{R}_t$;
    网络前向预测 (复数域交叉耦合计算与加性条件注入): $\hat{\mathbf{v}}_t \leftarrow f_\theta(\mathbf{X}_{\mathrm{in}}, \mathbf{e}_{\mathrm{cond}})$;
    损失计算与参数更新: $\mathcal{L} \leftarrow \|\hat{\mathbf{v}}_t - \mathbf{v}_t\|_2^2$, $\theta \leftarrow \theta - \eta\nabla_\theta\mathcal{L}$;
  end for
end while
return 网络参数 θ.

4.4 DDIM 加速推理机制

扩散模型的反向恢复通常需要在多个时间步重复调用去噪网络, 由此带来较高的迭代开销. 为减少网络调用次数, 本文采用去噪扩散隐式模型采样方法 [50]. DDIM 通过非马尔可夫的跳步更新构造确定性逆向轨迹

$$x_{\tau_{i-1}} = \sqrt{\bar{\alpha}_{\tau_{i-1}}}\,\hat{x}_0(x_{\tau_i}) + \sqrt{1-\bar{\alpha}_{\tau_{i-1}}}\,\hat{\boldsymbol{\epsilon}}(x_{\tau_i}), \tag{17}$$

其中, $\{\tau_S > \cdots > \tau_1 > \tau_0 = 0\}$ 表示从完整 T 步扩散序列中选取的递减采样时间步, S 为逆向过程的网络调用次数. 根据网络在 τ_i 时刻输出的 $\hat{\mathbf{v}}_{\tau_i}$, 目标残差与噪声预测分别由 $\hat{x}_0(x_{\tau_i}) = \sqrt{\bar{\alpha}_{\tau_i}}x_{\tau_i} - \sqrt{1-\bar{\alpha}_{\tau_i}}\hat{\mathbf{v}}_{\tau_i}$ 和 $\hat{\boldsymbol{\epsilon}}(x_{\tau_i}) = \sqrt{1-\bar{\alpha}_{\tau_i}}x_{\tau_i} + \sqrt{\bar{\alpha}_{\tau_i}}\hat{\mathbf{v}}_{\tau_i}$ 联立得到. DDIM 仅在所选时间步上执行逆向更新, 从而将去噪网络调用次数由 T 次减少至 S 次. 完整推理流程如算法 2 所示.

DDIM 的主要计算开销来自在每个采样状态调用一次去噪网络, 因此其网络调用次数随采样步数 S 增加. 较小的 S 能够降低迭代开销, 但相邻采样状态之间的跨度随之增大; 较大的 S 提供更细致的逆向轨迹离散化, 同时增加计算量. 因此, 采样步数 S 是控制逆向恢复粒度与计算开销的关键参数.

5 仿真结果与分析

5.1 数据集与仿真设置

本文使用 DeepMIMO v3 提供的 Remcom Wireless InSite 射线追踪场景构建 3.5 GHz 物理信道. 洛杉矶、芝加哥、休斯顿和凤凰城用于训练、验证及已见场景测试, 纽约、圣迭戈、达拉斯和丹佛仅用于未见场景测试. 每个用户最多保留 $P = 25$ 条有效传播路径, 并生成 4×4 MIMO 信道及相应路径参数.

算法 2 基于 DDIM 采样的 MIMO-OTFS 信道在线推理过程.

输入: MIMO 训练阶段的逐 Tx-Rx 导频观测 $\{\mathbf{Y}_p^{(r,t)}\}_{r,t}$, 导频符号 x_p, 速度 v, SNR γ; 预训练网络 f_θ; 推理步数 S.
输出: 时延 – 多普勒域高精度信道估计结果 $\hat{\mathbf{H}}$.

1: 依式 (11) 计算逐链路 LS 估计, 并依式 (12) 堆叠为完整 MIMO 初始张量 $\hat{\mathbf{H}}_{\mathrm{LS}}$, 再通过外围补零与 DD 域张量维度对齐;
2: 初始化与残差张量同维的高斯噪声 $\mathbf{R}_{\tau_S}\sim\mathcal{N}(\mathbf{0},\mathbf{I})$, 并选取递减时间步序列 $\{\tau_S,\ldots,\tau_1,\tau_0=0\}$;
3: **for** $i=S$ **to** 1 **step** -1 **do**
4:　令 $t\leftarrow\tau_i$, $t_{\mathrm{prev}}\leftarrow\tau_{i-1}$; 依式 (15) 融合 t,v,γ 得到 $\mathbf{e}_{\mathrm{cond}}$, 并构建 $\mathbf{X}_{\mathrm{in}}\leftarrow\hat{\mathbf{H}}_{\mathrm{LS}}+\mathbf{R}_t$;
5:　网络预测 $\hat{\mathbf{v}}_t\leftarrow f_\theta(\mathbf{X}_{\mathrm{in}},\mathbf{e}_{\mathrm{cond}})$, 并反解得到 $\hat{\mathbf{R}}_0$ 与 $\hat{\boldsymbol{\epsilon}}$;
6:　依式 (17) 确定性更新 DDIM 状态 $\mathbf{R}_{t_{\mathrm{prev}}}$;
7: **end for**
8: **return** 最终重构信道 $\hat{\mathbf{H}}=\hat{\mathbf{H}}_{\mathrm{LS}}+\hat{\mathbf{R}}_0$.

表 1　MIMO-OTFS 信道采样与训练参数.

Table 1　MIMO-OTFS channel sampling and training parameters.

Parameter	Symbol	Value	Parameter	Symbol	Value
Channel frequency span	B_{ch}	38.4 MHz	Time observation window	T_{obs}	25 ms
Tx/Rx antennas	$N_{tx}\times N_{rx}$	4×4	Nominal Doppler resolution	$\Delta\nu_{\mathrm{nom}}$	40 Hz
Frequency/delay samples	N	128	Batch size	B_s	64
Time/Doppler samples	M	64	Training epochs	E	100

数据覆盖 12 档速度 $v\in\{10,20,\ldots,120\}$ km/h 和 5 档 SNR $\gamma\in\{0,3,5,8,10\}$ dB. 受射线追踪场景中建筑遮挡影响, 部分用户位置不存在有效传播路径; 本文在四个已见城市约 30 万原始样本中剔除功率近似为零的无效样本, 并将有效样本按各 “城市 – 速度 – SNR” 组合以 8:1:1 划分为训练集、验证集与已见测试集, 其中训练集约含 78567 条有效样本; 未见城市仅用于独立测试. 所有方法采用相同的有效性判据与固定数据划分.

训练、验证和测试阶段, 速度和 SNR 作为数据生成参数与各 CSI 样本一一对应, 直接读取当前样本的仿真标签作为条件输入, 而不由网络估计; 各样本信道先按平均元素功率归一化, 再提取 LS 初估, 模型以真实信道与 LS 初估之差 $\mathbf{H}_{\mathrm{DD}}-\hat{\mathbf{H}}_{\mathrm{LS}}$ 为目标残差, 并使用训练集统计量完成标准化. 其余参数见表 1.

评价指标采用通信领域标准的端到端误码率与归一化均方误差 (normalized mean square error, NMSE). NMSE 定义为

$$\mathrm{NMSE}=10\log_{10}\left(\mathbb{E}\left[\frac{\|\mathbf{H}-\hat{\mathbf{H}}\|_F^2}{\|\mathbf{H}\|_F^2}\right]\right),\tag{18}$$

其中 $\mathbb{E}[\cdot]$ 表示统计均值, $\|\cdot\|_F$ 表示 Frobenius 范数. 深度学习模型均在 NVIDIA RTX PRO 6000 GPU 上训练, 采用 AdamW 优化器, 初始学习率为 10^{-4}、权重衰减系数为 10^{-6}, 并使用余弦退火调度.

本文选取的对比方法包括传统 LS 与对角近似 MMSE、压缩感知类 OMP 与 SBL 以及判别式学习 ChannelNet [21], ReEsNet 与 Transformer, 与本文方法一并评测; 端到端 BER 评估另以理想信道 (Perfect CSI) 结果作为不参与排名的理想性能参考.

5.2　信道估计性能比较

本节在相同测试样本与统一评测口径下, 比较不同信道估计方法随 SNR 和移动速度变化的 NMSE 性能.

图 3 展示了在 10∼120 km/h 全速度平均下, 各方法 NMSE 随 SNR 的变化. 传统 LS 与 MMSE 受限于线性滤波机制, 难以有效抑制强背景噪声, NMSE 仅徘徊在约 −10 dB 水平; OMP, SBL 等压缩感知方法在低 SNR 下仍受稀疏恢复误差制约; 判别式基线 ChannelNet, ReEsNet 与 Transformer 整体优于传统方法, 却在高 SNR 区间出现明显饱和. 相比之下, 本文方法在全部测试 SNR 下均保持最低 NMSE 且持续下降, 未出现饱和瓶颈: 0 dB 下全速度平均 NMSE 仍稳固于约 −20.02 dB, 10 dB 达到 −23.17 dB. 且相对各基线的增益在低 SNR 区间更为显著, 与生成式先验对高噪声后验分布的刻画能力相符.

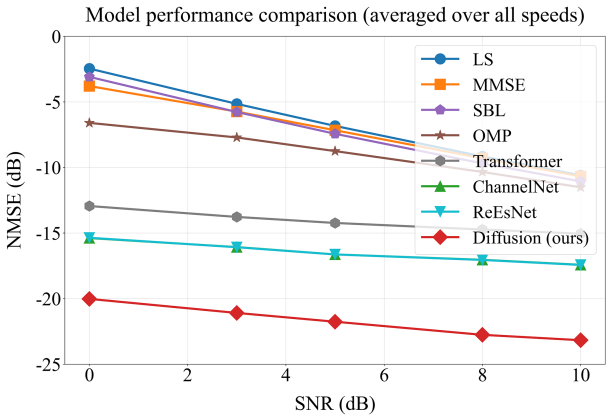

图 3　(网络版彩图) 八种信道估计方法统一评测后的 NMSE 性能对比 (全速度平均).

Figure 3　(Color online) Speed-averaged NMSE comparison after unified evaluation of eight channel estimators.

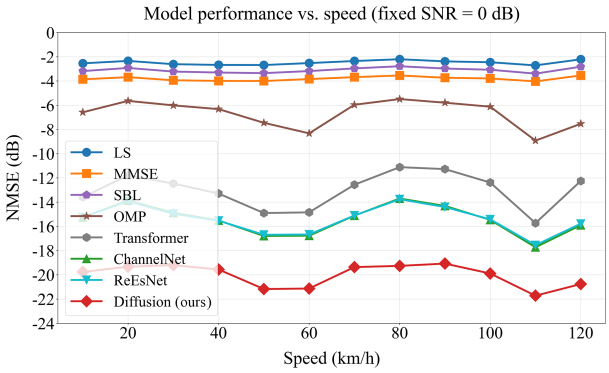

图 4　(网络版彩图) 八种信道估计方法统一评测后在 0 dB 下的分速度 NMSE 对比.

Figure 4　(Color online) Speed-wise NMSE comparison after unified evaluation of eight channel estimators at 0 dB.

图 4 固定 SNR 为 0 dB, 比较 10∼120 km/h 内各方法的估计精度. 传统 LS 与 MMSE 随速度基本恒定但整体偏低, 判别式基线虽有提升却波动受限. 本文方法则凭借速度 – SNR 联合条件嵌入与多步去噪的协同, 在全速度范围内均超越所有基线, NMSE 均优于 −19 dB 并在部分速度逼近 −21.7 dB; 即使在 120 km/h 最大多普勒频移下也未出现明显衰退. 此外, 各方法普遍存在的 NMSE 速度波动源于 OTFS 固有的分数多普勒效应: 当实际多普勒频移未精确对齐离散网格时, 频谱泄漏会降低 DD 域信道的稀疏性. 本文的速度适应性在 10∼120 km/h 典型高动态范围内得到验证; 对更宽的速度区间, 可通过将相应速度纳入训练进一步扩展.

5.3　DDIM 采样步数与计算开销

训练阶段的前向扩散总步数设为 $T = 1000$, 推理阶段默认采用 $S = 5$ 步 DDIM 采样. 本节结合采样步数 – NMSE 曲线和统一测量的模型复杂度, 分析默认设置下的精度与计算开销折中.

如图 5 所示, 本文从估计精度与推理耗时两个维度评估 DDIM 采样步数 S 的影响: 图 5(a) 给出精度随步数的演变, 图 5(b) 直观呈现精度与耗时的折中. 结合两图可见, 5 步采样已达到 50 步参考收敛精度的约 93.5%; 在此拐点之下降至 3 步时 NMSE 回退约 1.3 dB, 在其之上增至 10 步则耗时翻倍并突破 100 ms, 精度增益却不足 1 dB. 因此 5 步是以较少耗时换取较高去噪精度的折中点. 此外, OTFS 调制将剧烈变化的时变信道转化为 DD 域中准静态的二维网格, 去噪网络面对的始终是一张相对稳定的二维特征图, 其特征畸变复杂度不随终端速度剧烈起伏, 从而保证固定 5 步在各速度区间均能稳健收敛.

表 2 对比了各方法的计算复杂度与推理延迟. 本文 5 步 DDIM 与单步 direct regression 共享相同的 32.42M 参数主干, 迭代采样使其计算量与延迟约为单步模型的 5 倍, 但峰值显存几乎不增, 说明额外开销主要来自采样步数而非模型规模. 与各类基线相比, 本文方法推理延迟虽相对更高但仍处于毫秒量级. 在上述 GPU 上单样本预热 20 次后同步计时并重复 100 次取平均, 本文 5 步 DDIM 信道估计的单样本推理约为 50.96 ms, 采用 16 样本批处理时每样本可进一步摊薄至约 15 ms; 该量级相对车联万物等低时延业务约 100 ms 的端到端预算 [51] 仍有充裕空间. 因此, 本文以适度采样开销换取了更高估计精度. 完整端到端时延仍取决于系统实现.

5.4　跨场景泛化与检测可靠性

本节通过多城市未见场景测试、统一端到端 BER 链路和逐子载波误码分布, 评价模型的初步跨场景泛化能力及信道估计误差对数据检测的影响.

图 6 比较了四个已见城市的聚合结果与四个未见城市的泛化性能. 面对完全陌生的城市建筑排布与多径散

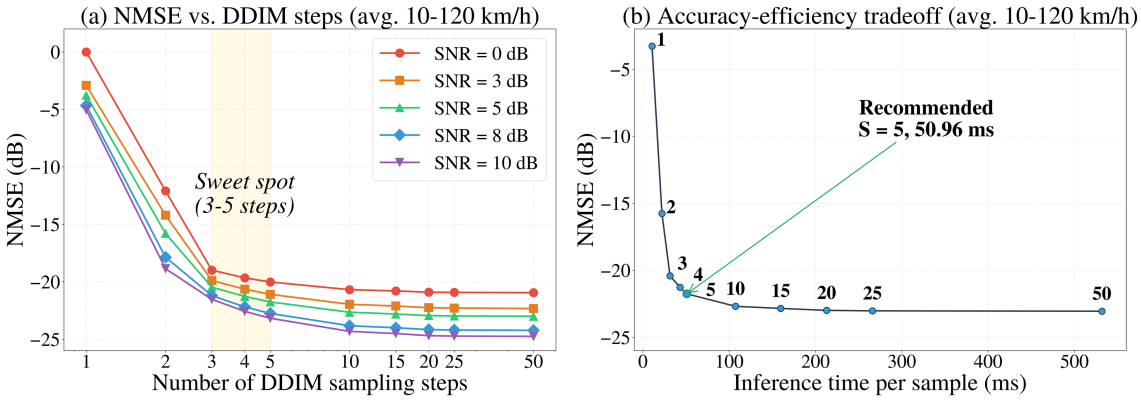

图 5　(网络版彩图) 扩散模型 NMSE 性能随 DDIM 采样步数 S 的演变趋势.

Figure 5　(Color online) Evolution of NMSE performance with DDIM sampling steps S.

表 2　代表性信道估计方法的计算复杂度与单样本推理开销. N_p, M, N 分别表示导频数、多普勒采样点数与时延采样点数; 传统方法给出渐近复杂度, 神经网络方法给出浮点运算量 (FLOPs), 二者不具备直接可比性. 时延与显存均在批大小为 1、单精度浮点 (FP32) 条件下测得, 显存为单样本峰值占用.

Table 2　Computational complexity and single-sample inference overhead of representative channel estimation methods. N_p, M, and N denote the numbers of pilots, Doppler samples, and delay samples, respectively; traditional methods report asymptotic complexity while neural networks report FLOPs, which are not directly comparable. Latency and memory are measured at batch size 1 with FP32; memory is the single-sample peak.

Method	Params	FLOPs/complexity	Latency (ms)	Peak memory (MB)	Training	Statistical prior
LS	–	$O(N_pMN)$	0.15	1.0	No	No
MMSE	–	$O(MN)$	0.16	2.1	No	Yes
OMP	–	$O(MN\log MN)$	0.32	7.7	No	No
ChannelNet	0.56 (M)	145.56 G	2.91	172.2	Yes	No
ReEsNet	0.63 (M)	164.89 G	3.44	206.1	Yes	No
Transformer	5.97 (M)	5.17 G	0.92	49.5	Yes	No
Direct regression (single-step)	32.42 (M)	141.07 G	10.42	234.6	Yes	No
Proposed (5-step DDIM)	32.42 (M)	705.34 G	50.96	241.7	Yes	No

射, 本文模型表现出初步的跨场景泛化能力: 四个未见城市的 NMSE 均随 SNR 单调改善, 且全程较 LS 基线优约 10∼15 dB; 其平均 NMSE 约 −19.4 dB, 相对已见基准 −21.76 dB 的平均泛化差仅约 2.4 dB. 各城市按传播复杂度呈现难易分层, 其中纽约建筑最为密集、传播环境最复杂, 泛化差约 4.5 dB, 是泛化难度最大的场景; 其余城市则明显更接近已见性能, 丹佛泛化差不足 0.5 dB. 这表明以条件残差方式建模有助于降低模型对特定训练场景的依赖, 使其在未见城市仍能保持可用的估计精度.

图 7 采用统一的无编码正交相移键控 (quadrature phase shift keying, QPSK)、逐子载波线性最小均方误差 (linear minimum mean square error, LMMSE) 均衡和最大比合并链路; 各方法共享信道、噪声与有效样本, 并以 Perfect CSI 作为理想参考. 随 SNR 提高, 本文方法的 BER 曲线持续下降并接近 Perfect CSI, 表明 NMSE 改善能够进一步转化为检测性能增益. 在 10^{-2} 目标 BER 处, 本文方法所需 SNR 较 LS 低约 3.20 dB, 与 Perfect CSI 相差约 0.18 dB. 总体而言, 本文方法在理想链路下将信道估计精度充分转化为了检测增益, 逼近理想信道性能, 显示出在 6G 高可靠通信场景中的应用潜力.

为直观剖析信道恢复的空间特性, 图 8 给出扩散模型在不同 SNR 下 DD 域全网格的逐子载波误码分布. 图中中心区域对应无数据传输的导频与保护间隔. 随 SNR 由 0 升至 10 dB, 数据区误码点逐步消退、零误码区域向外延展并覆盖全网格, 仅残存零星错误. 这一可视化结果表明, 本文方法在 DD 域未产生局部误差聚集, 而是在整个网格上都实现了对双选信道的稳健恢复, 与图 7 整体 BER 的下降趋势相互印证.

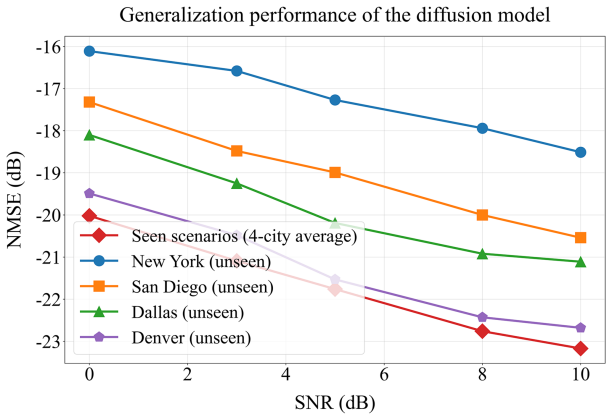

图 6　(网络版彩图) 四个已见城市聚合结果与四个未见城市的跨场景泛化性能.

Figure 6　(Color online) Cross-scenario generalization from the aggregate of four seen cities to four unseen cities.

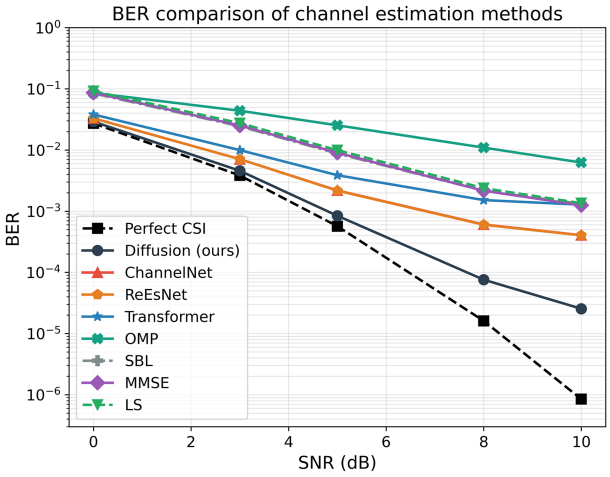

图 7　(网络版彩图) 八种信道估计方法统一评测后的端到端 BER 对比.

Figure 7　(Color online) End-to-end BER comparison after unified evaluation of eight channel estimators.

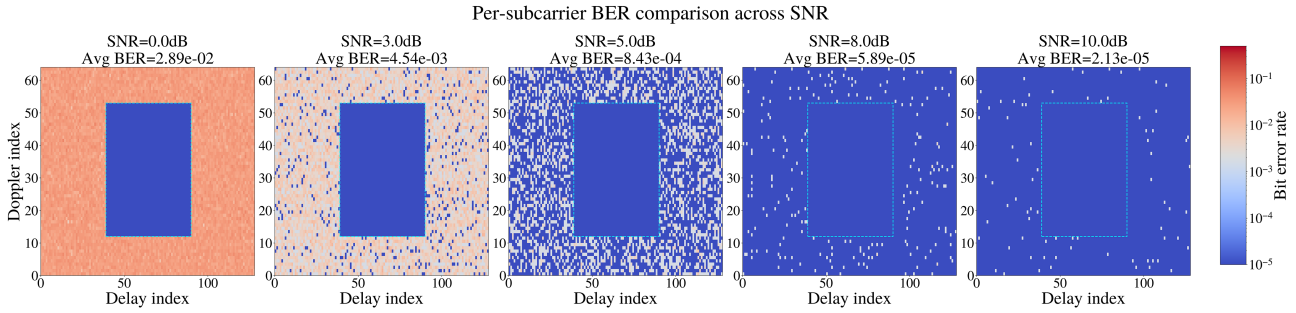

图 8　(网络版彩图) 扩散模型逐子载波误码率热力图.

Figure 8　(Color online) Per-subcarrier BER heatmaps of the proposed diffusion model.

5.5　资源效率与消融分析

本节首先分析导频保护区配置对有效数据传输效率的影响, 随后通过保护窗、网络结构、训练目标和条件失配实验评价各项设计的作用.

为量化估计精度与导频开销的权衡, 本文采用将估计误差等效为附加噪声的有效信噪比模型 [52]: 令第 i 个样本的线性 NMSE 为 $e_i = \|\hat{\mathbf{H}}_i - \mathbf{H}_i\|_F^2 / \|\mathbf{H}_i\|_F^2$, 接收 SNR 为 $\rho = 10^{\gamma/10}$, 则有效信噪比与净频谱效率为

$$\rho_{\mathrm{eff},i} = \frac{\rho}{1+\rho e_i}, \qquad R_{\mathrm{net}}(k_{\max},\gamma) = [1-\eta_{\mathrm{p}}(k_{\max})]\,\frac{1}{N_s}\sum_{i=1}^{N_s}\log_2(1+\rho_{\mathrm{eff},i})\,, \tag{19}$$

其中 N_s 为样本数, $T_{\mathrm{net}} = B_{\mathrm{ch}} R_{\mathrm{net}}$ ($B_{\mathrm{ch}} = 38.4$ MHz). 式 (19) 按单个发射天线导频时隙内的网格占用归一化, 未计入 N_{tx} 个时分正交导频时隙在训练周期中的调度比例. 该吞吐量为理想链路下按 B_{ch} 折算的结果.

图 9 显示, 净频谱效率并不随保护窗单调提高, 而在 $k_{\max} = 10$ 处取得峰值. 保护窗决定可覆盖的最大多普勒: $k_{\max} = 8$ 虽将导频占用降至约 20.5%, 但仅能覆盖约 99 km/h, 在 110~120 km/h 高速下 LS 前端因多普勒混叠而失效, 净频谱效率随之受限; 增大至 $k_{\max} = 12$, 14 虽能覆盖更高速度, 但估计精度已趋饱和, 新增开销反使净频谱效率下降. 因此在本文 120 km/h 移动性约束下, 能够覆盖该速度的 $k_{\max} \geqslant 10$ 配置中, $k_{\max} = 10$ 取得最高平均净频谱效率, 被选为折中点.

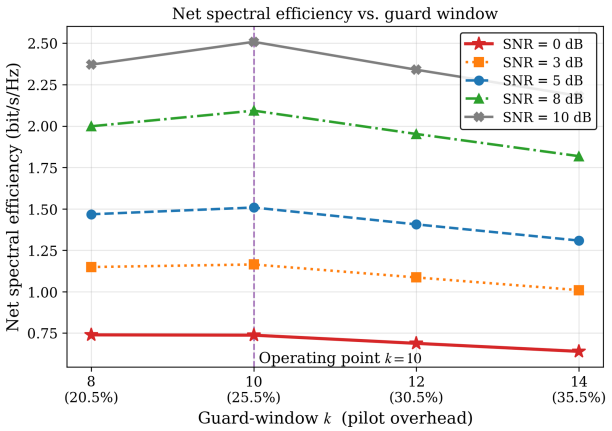

图 9　(网络版彩图) 净频谱效率随保护窗参数 $k_{\max}$ 的变化.

Figure 9　(Color online) Net spectral efficiency versus guard-window parameter $k_{\max}$.

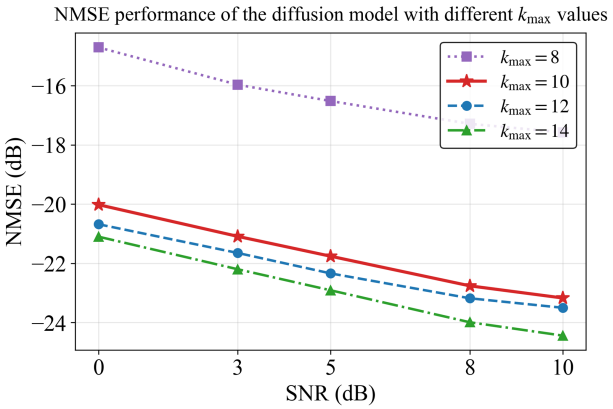

图 10　(网络版彩图) 不同 $k_{\max}$ 下扩散模型的 NMSE 比较.

Figure 10　(Color online) NMSE comparison of the diffusion model under different $k_{\max}$ values.

表 3　$k_{\max}=10$ 时不同信道信息质量下的净频谱效率与有效吞吐量.

Table 3　Net spectral efficiency and effective throughput under different channel-information qualities at $k_{\max}=10$.

Channel info	0 dB	3 dB	5 dB	8 dB	10 dB	Avg. net SE (bit/(s·Hz))	Throughput (Mbit/s)
LS (initial)	0.536	0.879	1.163	1.655	1.997	1.246	47.9
Proposed (diffusion)	0.737	1.164	1.508	2.093	2.509	1.602	61.5
Perfect CSI	0.745	1.179	1.532	2.137	2.576	1.634	62.7

表 3 在相同单时隙网格占用下隔离了信道信息质量的影响, 以 LS 初始估计为下界、理想 CSI 为上界. 本文扩散估计的平均净频谱效率达 1.602 bit/(s·Hz)、有效吞吐量达 61.5 Mbit/s, 已逼近理想 CSI 上界的 1.634 与 62.7 Mbit/s, 约为其 98%, 而优于 LS 下界的 1.246 与 47.9 Mbit/s, 吞吐量提升约 28%. 这说明本文估计质量已接近理想 CSI 所能达到的数据效率上限, 且该增益源于估计精度改善而非额外导频开销.

本节消融分析涵盖导频保护窗与网络及训练目标两方面, 均在 10∼120 km/h 全速度平均下进行. 保护窗方面比较不同保护窗 $k_{\max}$ 下的 NMSE, 见图 10; 网络与训练目标方面设定七种配置并与表 4 中各行一一对应: 完整模型 (proposed)、将复数值 U-Net 替换为实数双通道网络 (w/o complex U-Net)、同时移除速度与 SNR 两个条件 (w/o joint conditioning)、单独移除速度条件 (w/o velocity conditioning)、单独移除 SNR 条件 (w/o SNR conditioning)、以 ϵ-prediction 替代 v-prediction (ϵ-prediction), 以及保留相同复数 U-Net 主干的单步直接回归 (direct regression). 上述配置采用相同的训练与评价协议.

图 10 比较了不同保护窗 $k_{\max}$ 下的 NMSE. 其中 $k_{\max}=8$ 无法覆盖 120 km/h 的最大多普勒, 高速下 LS 前端混叠, 全速度平均 NMSE 明显偏高; 当 $k_{\max}\geqslant 10$ 时保护窗已覆盖目标移动范围, 三者估计精度趋于饱和, 各 SNR 下 NMSE 差异均在 1 dB 以内. 继续增大 $k_{\max}$ 仅带来微小精度提升, 导频开销却增加约 10 个百分点, 结合图 9 的净频谱效率, $k_{\max}=10$ 在满足 120 km/h 覆盖的前提下取得精度与资源的最佳平衡, 被选为默认配置.

表 4 给出各核心机制的消融结果. 本文完整架构在 0∼10 dB 全频段均保持最优精度, 移除任一机制均导致不同程度退化. 将复数值 U-Net 替换为实数双通道网络时, 五个 SNR 点平均 NMSE 退化 2.28 dB, 说明实虚部交叉耦合有助于保留 DD 域相位结构; 移除速度 – SNR 条件嵌入时全频段平均退化 0.93 dB, 单独移除速度或 SNR 条件分别退化 0.62, 0.63 dB, 均小于二者联合移除, 说明两路环境信息各自贡献有限且部分冗余, 由于 LS 锚点已隐含当前噪声与多普勒信息, 与显式条件部分重叠, 该增益虽小但稳定; 将 v-prediction 替换为 ϵ-prediction 后平均 NMSE 由 −21.76 骤降至 −6.43 dB, 源于 ϵ 目标与余弦噪声调度在少步采样下的不适配, 印证了 v-prediction 对少步恢复的必要性; 保留相同主干的单步 direct regression 退化 1.25 dB, 表明单步确定性映射易陷入局部最

表 4　核心机制消融实验的 NMSE 性能对比 (10∼120 km/h 全速度平均).

Table 4　NMSE comparison of core-mechanism ablations (averaged over 10–120 km/h).

Configuration	0 dB	3 dB	5 dB	8 dB	10 dB	Average	ΔNMSE
Proposed (full)	−20.02	−21.09	−21.76	−22.76	−23.17	−21.76	–
w/o complex U-Net	−17.65	−18.92	−19.72	−20.32	−20.79	−19.48	+2.28
w/o joint conditioning	−18.76	−20.14	−21.04	−21.86	−22.37	−20.83	+0.93
w/o velocity conditioning	−19.28	−20.52	−21.19	−22.20	−22.52	−21.14	+0.62
w/o SNR conditioning	−19.15	−20.48	−21.26	−22.45	−22.34	−21.13	+0.63
ϵ-prediction	−2.20	−4.69	−6.25	−8.81	−10.20	−6.43	+15.33
Direct regression	−18.13	−19.80	−20.70	−21.64	−22.26	−20.51	+1.25

表 5　SNR 条件偏差对扩散信道估计 NMSE 的影响.

Table 5　Effect of SNR-conditioning mismatch on the NMSE of diffusion-based channel estimation.

δ (dB)	−8	−5	−3	−2	−1	0	+1	+2	+3	+5	+8
NMSE	−22.00	−22.00	−21.89	−21.81	−21.78	−21.76	−21.61	−21.53	−21.51	−21.35	−21.20
ΔNMSE	−0.24	−0.24	−0.13	−0.05	−0.02	0	+0.15	+0.23	+0.24	+0.41	+0.56

优, 而多步残差细化提供了额外增益.

SNR 失配实验令条件值为 $\tilde{\gamma} = \gamma + \delta$, 保持接收观测、LS 锚点、速度条件和模型参数不变, 单独考察条件输入误差的影响, 结果见表 5. 以准确条件 $\delta = 0$ 为基准, 在 $|\delta| \leqslant 5$ dB 范围内最大 NMSE 退化为 0.41 dB; 即便将偏差扩大至 $|\delta| = 8$ dB, 退化也仅 0.56 dB, 表明模型对 SNR 条件误差具有较好的鲁棒性. 其原因在于 LS 锚点已经直接反映当前观测中的噪声和信道结构, 显式 SNR 标量主要提供辅助校准信息; 因此当条件值存在偏差时, 模型仍可依据物理观测完成稳定恢复.

6　结论

本文面向高速移动 MIMO-OTFS 系统中的 DD 域信道估计问题, 提出了一种条件残差扩散估计方法. 在射线追踪信道数据上, 所提方法在 0∼10 dB SNR 和 10∼120 km/h 速度范围内均优于传统估计、压缩感知及学习型基线; 在理想链路中, 达到 10^{-2} 目标 BER 所需的 SNR 较 LS 低约 3.20 dB. 本文通过引入 DDIM, 在采样速度与恢复精度之间取得平衡, 在提升估计精度的同时保持了较快的采样速度. 尽管如此, 所提方法仍存在拓展空间: 在物理场景上, 可向高铁、低轨卫星等极高速场景与低于 0 dB 的极低信噪比延伸, 并引入多用户干扰与射频硬件非理想特性; 在推理效率上, 可结合知识蒸馏与一致性模型等进一步压缩采样步数.

参考文献

1 You X H, Wang C X, Huang J, et al. Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts. Sci China Inf Sci, 2021, 64: 110301

2 Saad W, Bennis M, Chen M. A vision of 6G wireless systems: applications, trends, technologies, and open research problems. IEEE Netw, 2020, 34: 134–142

3 Wang C X, Bian J, Sun J, et al. A survey of 5G channel measurements and models. IEEE Commun Surv Tutorials, 2018, 20: 3142–3168

4 Hwang T, Yang C, Wu G, et al. OFDM and its wireless applications: a survey. IEEE Trans Veh Technol, 2009, 58: 1673–1694

5 Hadani R, Rakib S, Tsatsanis M, et al. Orthogonal time frequency space modulation. In: Proceedings of IEEE Wireless Communications and Networking Conference, 2017. 1–6

6 Wei Z, Yuan W, Li S, et al. Orthogonal time-frequency space modulation: a promising next-generation waveform. IEEE Wireless Commun, 2021, 28: 136–144

7 Surabhi G D, Augustine R M, Chockalingam A. On the diversity of uncoded OTFS modulation in doubly-dispersive channels. IEEE Trans Wireless Commun, 2019, 18: 3049–3063
8 Li S, Yuan J, Yuan W, et al. Performance analysis of coded OTFS systems over high-mobility channels. IEEE Trans Wireless Commun, 2021, 20: 6033–6048
9 Raviteja P, Phan K T, Hong Y. Embedded pilot-aided channel estimation for OTFS in delay-Doppler channels. IEEE Trans Veh Technol, 2019, 68: 4906–4917
10 Shen W, Dai L, An J, et al. Channel estimation for orthogonal time frequency space (OTFS) massive MIMO. IEEE Trans Signal Process, 2019, 67: 4204–4217
11 Murali K R, Chockalingam A. On OTFS modulation for high-Doppler fading channels. In: Proceedings of Information Theory and Applications Workshop, 2018. 1–10
12 Surabhi G D, Chockalingam A. Low-complexity linear equalization for OTFS modulation. IEEE Commun Lett, 2020, 24: 330–334
13 Gómez-Cuba F. Compressed sensing channel estimation for OTFS modulation in non-integer delay-Doppler domain. In: Proceedings of IEEE Global Communications Conference (GLOBECOM), 2021. 1–6
14 Zhao L, Gao W J, Guo W. Sparse bayesian learning of delay-Doppler channel for OTFS system. IEEE Commun Lett, 2020, 24: 2766–2769
15 Wei Z, Yuan W, Li S, et al. Off-grid channel estimation with sparse bayesian learning for OTFS systems. IEEE Trans Wireless Commun, 2022, 21: 7407–7426
16 Shi D, Wang W, You L, et al. Deterministic pilot design and channel estimation for downlink massive MIMO-OTFS systems in presence of the fractional Doppler. IEEE Trans Wireless Commun, 2021, 20: 7151–7165
17 Srivastava S, Singh R K, Jagannatham A K, et al. Bayesian learning aided simultaneous row and group sparse channel estimation in orthogonal time frequency space modulated MIMO systems. IEEE Trans Commun, 2022, 70: 635–648
18 Yuan W, Li S, Wei Z, et al. Data-aided channel estimation for OTFS systems with a superimposed pilot and data transmission scheme. IEEE Wireless Commun Lett, 2021, 10: 1954–1958
19 Qin Z, Ye H, Li G Y, et al. Deep learning in physical layer communications. IEEE Wireless Commun, 2019, 26: 93–99
20 O'Shea T, Hoydis J. An introduction to deep learning for the physical layer. IEEE Trans Cogn Commun Netw, 2017, 3: 563–575
21 Soltani M, Pourahmadi V, Mirzaei A, et al. Deep learning-based channel estimation. IEEE Commun Lett, 2019, 23: 652–655
22 Sun T, Lv J, Zhou T. A Transformer-based channel estimation method for OTFS systems. Entropy, 2023, 25: 1423
23 Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2017. 105–114
24 Blau Y, Michaeli T. The perception-distortion tradeoff. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2018. 6228–6237
25 Arvinte M, Tamir J I. MIMO channel estimation using score-based generative models. IEEE Trans Wireless Commun, 2023, 22: 3698–3713
26 Fesl B, Baur M, Strasser F, et al. Diffusion-based generative prior for low-complexity MIMO channel estimation. IEEE Wireless Commun Lett, 2024, 13: 3493–3497
27 Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. Adv Neural Inf Process Syst, 2020, 33: 6840–6851
28 Song Y, Sohl-Dickstein J, Kingma D P, et al. Score-based generative modeling through stochastic differential equations. In: Proceedings of International Conference on Learning Representations, 2021
29 Mohammed S K. Derivation of OTFS modulation from first principles. IEEE Trans Veh Technol, 2021, 70: 7619–7636
30 LeCun Y, Bengio Y, Hinton G. Deep learning. Nature, 2015, 521: 436–444
31 Ye H, Li G Y, Juang B H. Power of deep learning for channel estimation and signal detection in OFDM systems. IEEE Wireless Commun Lett, 2018, 7: 114–117
32 Li L, Chen H, Chang H H, et al. Deep residual learning meets OFDM channel estimation. IEEE Wireless Commun Lett, 2020, 9: 615–618
33 Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. Adv Neural Inf Process Syst, 2017, 30: 5998–6008
34 Mashhadi M B, Gunduz D. Pruning the pilots: deep learning-based pilot design and channel estimation for MIMO-OFDM systems. IEEE Trans Wireless Commun, 2021, 20: 6315–6328
35 He H, Jin S, Wen C K, et al. Model-driven deep learning for physical layer communications. IEEE Wireless Commun, 2019, 26: 77–83
36 Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. Adv Neural Inf Process Syst, 2014, 27: 2672–2680
37 Kingma D P, Welling M. Auto-encoding variational Bayes. In: Proceedings of International Conference on Learning Representations, 2014
38 O'Shea T J, Roy T, West N. Approximating the void: learning stochastic channel models from observation with variational generative adversarial networks. In: Proceedings of International Conference on Computing, Networking and Communications,

2019. 681–686
39 Ye H, Li G Y, Juang B H F, et al. Channel agnostic end-to-end learning based communication systems with conditional GAN. In: Proceedings of IEEE GLOBECOM Workshops, 2018. 1–5
40 Yang L, Zhang Z L, Song Y, et al. Diffusion models: a comprehensive survey of methods and applications. ACM Comput Surv, 2023, 56: 1–39
41 Letaief K B, Chen W, Shi Y, et al. The roadmap to 6G: AI empowered wireless networks. IEEE Commun Mag, 2019, 57: 84–90
42 Hyvärinen A. Estimation of non-normalized statistical models by score matching. J Mach Learn Res, 2005, 6: 695–709
43 He K, Zhang X, Ren S, et al. Deep residual learning for image recognition. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2016. 770–778
44 Nichol A Q, Dhariwal P. Improved denoising diffusion probabilistic models. In: Proceedings of International Conference on Machine Learning, 2021. 8162–8171
45 3GPP. NG Radio Access Network (NG-RAN); Stage 2 Functional Specification of User Equipment Positioning in NG-RAN. 3GPP TS 38.305 V17.8.0. 2025
46 Tepedelenlioglu C, Giannakis G B. On velocity estimation and correlation properties of narrow-band mobile communication channels. IEEE Trans Veh Technol, 2001, 50: 1039–1052
47 3GPP. NR; Physical Layer Measurements. 3GPP TS 38.215 V17.5.0. 2024
48 Trabelsi C, Bilaniuk O, Zhang Y, et al. Deep complex networks. In: Proceedings of International Conference on Learning Representations, 2018
49 Salimans T, Ho J. Progressive distillation for fast sampling of diffusion models. In: Proceedings of International Conference on Learning Representations, 2022
50 Song J, Meng C, Ermon S. Denoising diffusion implicit models. In: Proceedings of International Conference on Learning Representations, 2021
51 3GPP. Service Requirements for V2X Services. 3GPP TS 22.185 V16.0.0. 2020
52 Yoo T, Goldsmith A. Capacity and power allocation for fading MIMO channels with channel estimation error. IEEE Trans Inform Theor, 2006, 52: 2203–2214

Conditional diffusion-based OTFS channel estimation for 6G high-mobility scenarios

Dianxin ZHENG[1], Yuzhi YANG[2] & Tianhao GUO[1*]

1. *Shanxi Key Laboratory of Wireless Communication and Detection, College of Physics and Electronic Engineering, Shanxi University, Taiyuan 030006, China*
2. *Digital Future Institute, Khalifa University of Science and Technology, Abu Dhabi 127788, UAE*
* Corresponding author. E-mail: tianhao_guo@sxu.edu.cn

Abstract OTFS modulation is suitable for 6G high-mobility scenarios, but delay-Doppler domain channel estimation still suffers from severe noise amplification and limited cross-scene generalization under low SNR and high-Doppler conditions. This paper proposes a conditional diffusion-based OTFS channel estimation method. The method uses the LS estimate extracted from embedded pilots as a physical anchor, employs a complex-valued U-Net to preserve phase information, introduces joint velocity-SNR conditioning to characterize environment variations, and combines v-prediction with DDIM sampling to reduce inference cost. Simulations based on the DeepMIMO ray-tracing dataset show that the proposed method outperforms representative baselines including LS, MMSE, OMP, Transformer, and ChannelNet over 0–10 dB SNR and 10–120 km/h. The average NMSE over the full speed range reaches -23.17 dB at 10 dB and -20.02 dB at 0 dB. Tests on multiple unseen urban scenes provide initial evidence that the proposed method generalizes across different urban propagation environments. End-to-end BER evaluation further shows that, under the unified detection chain with ideal synchronization and uncoded QPSK, the proposed method requires approximately 3.20 dB less SNR than LS to reach a target BER of 10^{-2}.

Keywords OTFS modulation, channel estimation, conditional diffusion model, delay-Doppler domain, 6G high-mobility scenarios