

Neural Source-Filter vs. Transformer Inference Latency in Cross-Domain Symbolic-to-Audio Generation

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: How does the inference latency of neural source-filter models compare to transformer-based architectures like CodeT5 when scaling sequence length in cross-domain symbolic-to-audio generation tasks. 7 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Length independent generalization bounds for deep SSM architectures via Rademacher contraction and stability constraints. Research question: How does the inference latency of neural source-filter models compare to transformer-based architectures like CodeT5 when scaling sequence length in cross-domain symbolic-to-audio generation tasks?.

2 Methodology

Systematic literature search across multiple databases yielded 15 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.3/10.

3 Results

15 papers retrieved. 7 claims extracted; 0 independently verified. Quality review score: 3.3/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
The Rademacher complexity of stable SSMs is bounded by their norm for any sequence length.	×	0.09
Rademacher Contraction is a universal framework that enables componentwise complexity estimation in deep models.	×	0.06
The set of compositions $\Phi_2 \Phi_1$ is $(\mu_1 \mu_2, \mu_2 c_1 + c_2)$ -RC if Φ_1 is (μ_1, c_1) -RC and Φ_2 is (μ_2, c_2) -RC.	×	0.01
For deep models composed of layers that each satisfy the RC property, the entire model class is RC as well.	×	0.02
Stable SSMs are robust to perturbations, i.e., their state and output are continuous in the initial state and input.	×	0.07
The induced norm of a stable SSM can be bounded uniformly in T.	×	0.11
An SSM block is a residual composition of the SSM with a non-linear function g applied element-wise.	×	0.05

References

- <http://arxiv.org/abs/2206.06811v2>
- <http://arxiv.org/abs/2405.20278v3>
- <http://arxiv.org/abs/2307.05639v2>