SCAD: A Robust Framework for Spectral Correction via Adaptive Deconvolution

Ivan Rodionov

November 28, 2025

Abstract

This paper presents SCAD (Spectral Correction via Adaptive Deconvolution), a robust framework for estimating stable inverse filters from noisy audio observations. The method addresses common linear distortion artifacts in poor-quality microphone recordings, including resonant frequencies, short echoes, and temporal smearing. SCAD combines regularized least-squares estimation with automatic model selection and multi-criteria noise-to-signal ratio optimization to produce stabilized Wiener filters. The implementation emphasizes practical deployment with diagnostic variants and computational efficiency. Experimental results demonstrate effective correction of microphone deficiencies while avoiding common inverse-filtering artifacts.

1 Introduction

This work presents SCAD (Spectral Correction via Adaptive Deconvolution), a robust framework for estimating stable inverse filters from noisy observations. The method is particularly effective for correcting audio recorded on poor-quality microphones suffering from linear distortion artifacts including short echoes, resonant frequencies, and temporal smearing.

The implementation emphasizes practical deployment, with algorithm descriptions closely aligned with the accompanying MATLAB code. Key design choices, mathematical operations, and considerations for reproducibility are highlighted to provide both theoretical understanding and practical implementation guidance.

2 Algorithm Overview

Algorithm 1 SCAD Implementation Pipeline

```
Extract initial segment y_seg from distorted audio
Initialize via blind deconvolution with parametric seed
Compute autocorrelations for candidate filter lengths
for each candidate L_h do
    Solve regularized normal equations
    Compute RSS, NSR estimate, and BIC
end for
Select optimal L_h via BIC with tie-breaking
Iteratively refine NSR using multi-criteria objective
Construct stabilized Wiener filter W(ω)
Apply via overlap-add with sample-wise normalization
Generate diagnostic outputs and variant comparisons
```

The SCAD framework excels at correcting common microphone deficiencies including:

- **Resonant peaks/dips** from poor frequency response
- **Short echoes** from internal reflections
- **Temporal smearing** ("flabby" bass response)
- **Comb filtering** from multipath interference

3 Forward Model and Problem Formulation

We consider the standard finite-length linear time-invariant (LTI) model with additive noise:

$$y[n] = (x * h)[n] + \eta[n], \tag{1}$$

where $h[n]$ is an FIR filter of unknown length L_h, $x[n]$ is the source signal, and $\eta[n]$ represents additive noise (wide-sense stationary and uncorrelated with $x[n]$).

Practical Objectives:

- Automatically estimate stable finite-length filter h without manual specification of L_h

- Derive a time-invariant frequency-domain inverse filter $W(\omega)$ with appropriate stabilization
- Apply $W(\omega)$ via overlap-add convolution with correct normalization
- Provide diagnostic variants ("safe/best/aggressive") for rapid evaluation

4 Implementation Methodology

4.1 Short-Segment Estimation

The algorithm processes an initial segment y_{seg} (default: 1.0 s) under local stationarity assumptions. This approach reduces computational load during blind initialization and minimizes sensitivity to non-stationary artifacts. Segment duration is controlled by the `seg_seconds` parameter.

4.2 Blind Deconvolution Initialization

A parametric "swingy exponential" seed initializes the blind deconvolution:

$$h_{\text{seed}}[n] = e^{-n/\tau} \left(1 + \alpha \sin\left(2\pi f_0 n/F_s\right) e^{-n/(\kappa\tau)}\right), \quad n = 0, \ldots, L_h^{\text{init}} - 1. \tag{2}$$

This causal, normalized seed includes weak sinusoidal perturbation to avoid degenerate solutions. The implementation falls back to using y_{seg} directly if blind deconvolution fails.

4.3 Regularized Least Squares via Autocorrelations

The efficient Toeplitz-based formulation solves:

$$\left(\mathbf{R}_{xx}^{(L_h)} + \lambda I + \mu D^\top D\right) h = r_{xy}^{(L_h)}, \tag{3}$$

where $\mathbf{R}_{xx}^{(L_h)}$ is the Toeplitz autocorrelation matrix and $r_{xy}^{(L_h)}$ contains cross-correlation lags. Regularization includes:

- **Tikhonov:** $\lambda = \lambda_{\text{rel}} \cdot \mathrm{Var}(y_{\text{seg}})$ (eigenvalue stabilization)
- **Smoothness:** $\mu\|Dh\|^2$ with D as second-derivative operator (prevents oscillatory solutions)

4.4 Automatic Model Order Selection

The algorithm evaluates candidate lengths `Lh_coarse` $= [64, 128, 256, 512]$ using the Bayesian Information Criterion:

$$\mathrm{BIC}(L_h) = L_h \log(M) + M \log\left(\mathrm{RSS}(L_h)/M\right). \tag{4}$$

Among candidates within 1% of the minimum BIC, the longest filter is selected to prevent undermodeling.

4.5 Iterative Noise-to-Signal Ratio Refinement

The implementation alternates between NSR selection and filter refinement. For each NSR candidate, three metrics are computed:

1. Normalized residual variance
2. Spectral flatness (windowed geometric/arithmetic mean ratio)
3. Peak inverse gain magnitude

These are normalized and combined into a weighted objective function. The selected NSR balances reconstruction quality, perceptual characteristics, and stability.

4.6 Global Wiener Filter Construction

The frequency-domain inverse filter is constructed as:

$$W(\omega) = \frac{H^*(\omega)}{|H(\omega)|^2 + \Gamma}, \tag{5}$$

with practical stabilization via:

- Gaussian smoothing of magnitude response
- Hard capping at `maxInverseGain` to prevent noise amplification

4.7 Overlap-Add Application

The overlap-add process uses block length $B = 65536$ with critical sample-wise normalization by the accumulated contribution counter. This preserves local amplitude without global renormalization artifacts.

5 Experimental Results

Table 1: Implementation Default Parameters

Parameter	Variable	Default
Segment duration	`seg_seconds`	1.0 s
Coarse lengths	`Lh_coarse`	[64,128,256,512]
Tikhonov regularization	λ_{rel}	5×10^{-6}
Smoothness penalty	`smoothness_mu`	0.05
NSR floor	`NSR_floor`	1×10^{-3}
Maximum inverse gain	`maxInverseGain`	12 dB
Block length	`block_len`	65536
FFT length	`Lfft`	$2^{\lceil \log_2(B+L_h-1) \rceil}$

6 Conclusion and Future Work

The SCAD framework provides a robust solution for spectral correction of microphone-distorted audio. Key contributions include automatic model order selection, multi-criteria NSR optimization, and practical stabilization techniques.

Future enhancements:

- Perceptual weighting (A-weighting or critical bands) in NSR selection
- Optional synthesis windows with proper normalization
- Unit test suite for core numerical routines
- Subjective listening tests for perceptual validation