MULTI-VIEW SUBBAND AUTOREGRESSIVE POLE HARVESTING FOR MODAL PLATE IDENTIFICATION

Facundo Franchino, Jatin Chowdhury, Jin Woo Lee

Music Technology
Massachusetts Institute of Technology
Cambridge, MA
{jnlee | chowdsp | facundof}
@mit.edu

Soohyun Kim

CCRMA
Stanford University
Stanford, CA
soohyun@ccrma.stanford.edu

Mark Rau

Music Technology
Massachusetts Institute of Technology
Cambridge, MA
mrau@mit.edu

ABSTRACT

We describe a Task B submission for the 1st DAFx Parameter Estimation Challenge. A matching-pursuit anchor stage seeds a multi-view subband autoregressive (AR) pole harvester on the raw IR and its first two finite differences; since linear filtering preserves pole locations while changing residues, the three views expose complementary subsets of the same pole set. Bands in which the AR order saturates are recursively split, and any remaining under-resolved region is completed from an IR-derived saturation indicator. Gains are assigned by a global ridge least-squares (LS) fit in the damped-biquad atom convention. The pipeline uses only the unnormalized IR; so it does not use plate parameters, analytical modal-frequency or decay laws, .wav files, or Task A information.

1. INTRODUCTION

Identifying modal parameters from frequency response functions has long been a fundamental task in diverse fields such as acoustics or structural dynamics. Extracting a system's intrinsic dynamic characteristics, namely natural frequencies, damping ratios, and mode shapes, from the frequency response is crucial for many applications such as realistically emulating acoustic behavior, diagnosing structural integrity, and optimizing designs [1–3].

Despite its long history and extensive application, recovering modal parameters from frequency response remains a challenging task, fraught with technical hurdles such as measurement noise [4], limited frequency resolution [5], and transducer loading effects [6], making it difficult to discern the underlying physical parameters. Furthermore, complex systems exhibit high modal density and severe modal overlap particularly in the high-frequency range [5, 7]. Due to this modal overlap, conventional curve-fitting algorithms, such as the rational fraction polynomial method, struggle to uniquely isolate and estimate individual modes [8].

Task B of the 1st DAFx Parameter Estimation Challenge [9] specifically addresses this issue. The challenge poses the problem of estimating modal parameters for the Kirchhoff–Love plate model's reverberation [10] for given an unnormalized displacement impulse response (IR). A single plate carries 10^3–10^4 modes packed densely enough that FFT peak picking and matching pursuit alike leave plenty unresolved, and backfilling with placeholder frequencies injects meaningless decay and gain values into the output.

Copyright: © 2026 Facundo Franchino et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, adaptation, and reproduction in any medium, provided the original author and source are credited.

In this regard, we propose a method that recovers the missing modes directly from the IR. Every mode corresponds to a pole, and pole locations do not move when the IR is passed through any linear filter. The raw IR and its first two finite differences therefore share the same poles, but each view makes a slightly different subset of those poles easier to see [11]. A subband autoregressive (AR) estimator runs on each view, bands that run out of capacity are split into smaller bands, and the merged pole list is de-duplicated using cross-view agreement. A saturation indicator computed from the IR fills any remaining gap, and a global ridge least-squares fit assigns gains.

Our three contributions are as follows. *(i)* Multi-view formulation of modal analysis: We treat the raw IR together with its first two finite-differences as three views of the same mode set, leveraging the insight that linear filtering alters modal residues while leaving pole locations invariant; *(ii)* Recursive band-splitting scheme: We introduce a recursive method that splits any frequency band whenever the AR estimator runs out of capacity; and *(iii)* Residual-driven mode enumeration metric: We propose using the capacity shortfall of the AR esimator as the signal for quantifying missing modes. Fig. 1 summarizes the pipeline.

2. PLATE MODEL AND EVALUATION METRIC

2.1. Problem Statement

The solution of a Kirchhoff–Love plate equation for plate reverberation [10] is expressed in the z-domain [9] as a sum of M damped modes, where each mode is in an exact form as a discrete biquad

$$H_m(z) = \frac{b_m z^{-1}}{1 - 2r_m \cos(\Omega_m T)\, z^{-1} + r_m^2\, z^{-2}}, \tag{1}$$

with $r_m = e^{-\sigma_m T}$ and sampling period T. This gives a time-domain impulse response's m-th *atom*:

$$h_m[k] = b_m\, r_m^{k-1} \frac{\sin(k\, \Omega_m T)}{\sin(\Omega_m T)}, \quad k \geq 1. \tag{2}$$

The goal is to estimate $\{(f_m, \sigma_m, b_m)\}_{m=1}^{M}$ with $f_m = \Omega_m/(2\pi)$ for each given impulse response. The total number of modes M is assumed to be unknown.

2.2. Evaluation Metric

Identified and reference modes are matched by Hungarian assignment [12] on log-frequency distance with a half-octave reject threshold; per-mode errors are clipped to 1 and unmatched true modes contribute 1 to each component, giving the total error metric as a combination of the per-plate relative error (RE) in three modal parameters and mode number error [9] as

$$\mathrm{RE} = \tfrac{1}{3}(\mathrm{RE}_\Omega + \mathrm{RE}_\sigma + \mathrm{RE}_b) + \min\Big(1, \tfrac{|M-\tilde{M}|}{M}\Big), \quad (3)$$

where

$$\mathrm{RE}_\chi = \frac{1}{M}\sum_{m=1}^{M} \min\left(1, \frac{|\chi_m^{(\mathrm{est})} - \chi_m^{(\mathrm{ref})}|}{\chi_m^{(\mathrm{ref})}}\right) \quad (4)$$

is the per-plate relative error for each parameter $\chi \in \{\Omega, \sigma, b\}$. Our presented design is based on insights gained from the following two observations. First, missing a true mode costs much more than emitting a spurious one. A missed mode adds the maximum penalty to all three component errors *and* to the count term, while a spurious mode only adds to the count term. The output count should therefore lean high when in doubt. Second, when many modes lie within a fraction of a hertz of each other, their atoms in Eq. (2) become nearly linearly dependent, so the gain LS can recover their joint amplitude but not the individual gains, regardless of window length or regularisation.

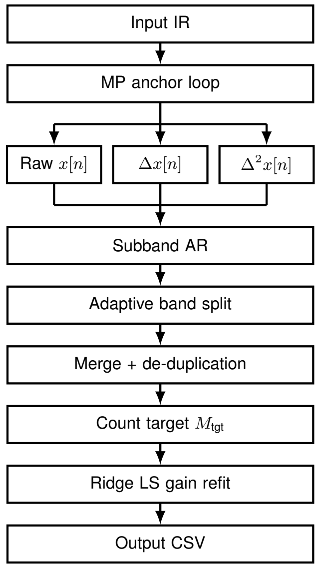

Figure 1: Pipeline overview. MP anchors seed three pole-preserving views of the IR; per-view subband AR with adaptive splitting harvests poles; merge, count completion, gain refit produce the output.

3. PIPELINE

The presented algorithm densely extracts the poles of the reverberant signal by combining Matching Pursuit (MP) loop-based FFT peak picking (subsection 3.1) with sub-band AR modeling (subsection 3.2) and adaptive band partitioning (subsection 3.3) applied to multiple versions of the signal processed through a derivative filter. Subsequently, a density supplementation process fills in missing components with synthetic frequencies (subsection 3.4), and finally, the precise magnitude of each mode is calculated using Ridge least squares with full-band energy balancing to determine the final parameters (subsection 3.5).

3.1. Anchor extraction

The MP loop [13] runs eight residual-subtraction iterations of an FFT peak picker. After each pass the joint set of anchor atoms (2) is refit by ridge LS and subtracted, so the next pass sees lower-prominence modes. The output is an anchor set $\mathcal{A} = \{(f_a, \sigma_a)\}$ of several hundred entries per plate.

3.2. Multi-view subband AR

An autoregressive (AR) model, in the lineage of Prony and rotational-invariance methods [14, 15], is fit on each of three views of the IR: the raw signal $x[n]$, its first difference $\Delta x[n]$, and its second difference $\Delta^2 x[n]$. The differences attenuate low frequencies but do not move any modal pole, so each view exposes a slightly different subset of the modes. Each view is partitioned into eight log-spaced bands with 25% overlap. Within each band the signal is shifted down to baseband, lowpassed, and decimated to a sample rate just above the band width. A complex AR fit produces a polynomial whose roots are pole candidates; those inside the band with magnitude less than one and positive decay are kept. The AR order is set proportional to the expected mode count in the band, capped at 1800.

3.3. Adaptive band splitting

A band is flagged saturated when the AR order cap is reached and the core yields at least 50 stable poles. Each saturated band with log-frequency ratio greater than 1.20 is split log-evenly into two child bands and the AR estimator is rerun, to recursion depth 2. This concentrates capacity where mode density is highest without inflating cost on sparse low-frequency bands.

3.4. Saturation-aware modal-density completion

Anchors and AR roots are merged and de-duplicated by log-gap $\Delta \log_2 f < 0.0008$; each merge across views increments a per-pole support counter. Even after multi-view harvesting, the candidate count can stay too low in dense regions because the AR order is capped. To fill the gap we use two quantities computed from the IR: the anchor count and a measure of how badly the AR estimator saturated. An affine predictor

$$M_{\mathrm{lin}} = a\, n_a + b, \quad (a, b) = (10.10, -4437), \quad (5)$$

gives an initial target from the anchor count $n_a = |\mathcal{A}|$ with (a, b) fixed before final evaluation. With n_{AR} the de-duplicated AR pool size, the deficit ratio

$$\rho = \frac{M_{\mathrm{lin}} - n_{\mathrm{AR}}}{\max(M_{\mathrm{lin}}, 1)} \quad (6)$$

acts as a per-IR saturation indicator; large ρ signals that the finite-order estimator has run out of capacity. We raise the target conservatively by a monotone function of ρ,

$$\beta(\rho) = \begin{cases} 1.35 & \rho > 0.45 \\ 1.20 & 0.30 < \rho \leq 0.45 \\ 1.10 & 0.10 < \rho \leq 0.30 \\ 1.05 & \rho \leq 0.10, \end{cases} \tag{7}$$

$$M_{\text{tgt}} = \max(M_{\text{lin}},\ n_{\text{AR}} + \beta(\rho)\ \max(0, M_{\text{lin}} - n_{\text{AR}}))\,. \tag{8}$$

If the merged pool already exceeds $1.12\,M_{\text{tgt}}$, the candidates with the weakest cross-view support are dropped; otherwise the remaining slots are filled with log-spaced placeholder frequencies, and the final gain stage assigns their amplitudes.

3.5. Gain refit

Gains are assigned by a global ridge LS in atom (2), so the LS weight for mode m equals b_m directly. We use $n = 44100$ samples and a per-column ridge $\lambda_m = 10^{-8}\,\|a_m\|^2$ to balance low- and high-frequency modes whose column norms differ by orders of magnitude. The triples (f_m, σ_m, b_m) form the submission CSV.

4. RUNTIME AND COMPLIANCE

The pipeline is non-iterative at the top level. Per IR, one 8-pass MP loop, three AR views, eight bands per view with up to two levels of adaptive splitting, one de-duplication, one count-target decision, and one final LS. No outer search or random restarts are used. On an Apple M2 laptop with 16 GB unified memory and CPU-only compute, the 16 stripped-set IRs are processed in roughly 6 h of wall time using three-way parallelism over plate-density groups; per-plate cost scales with the anchor count. No part of the pipeline reads .wav files, plate-parameter files, or any ground-truth mode list at inference time; the constants (a, b) in Eq. (5) are fixed before final evaluation.

5. SUMMARY

We present a multi-view subband AR pole harvester for Task B of the 1st DAFx Parameter Estimation Challenge. Because linear filtering does not move modal poles, three filtered views of the same IR give three independent looks at the same mode set; saturated bands are split into smaller ones, and a saturation indicator computed from the IR completes any remaining gap in the modal density. The dominant remaining limitation is on the gains. When many modes sit within a fraction of a hertz of each other, the time-domain least-squares cannot fully separate their individual amplitudes. Tightening this would likely require fitting each gain locally in the frequency domain, or jointly refining the frequency, decay, and gain of each mode in a small neighbourhood.

6. REFERENCES

[1] M. Ducceschi and C. Touzé, “Modal approach for nonlinear vibrations of damped impacted plates: Application to sound synthesis of gongs and cymbals,” *Journal of Sound and Vibration*, vol. 344, pp. 313–331, 2015.

[2] G. Kerschen, K. Worden, A. F. Vakakis, and J.-C. Golinval, “Past, present and future of nonlinear system identification in structural dynamics,” *Mechanical systems and signal processing*, vol. 20, no. 3, pp. 505–592, 2006.

[3] E. Maestre, J. S. Abel, J. O. Smith, and G. P. Scavone, “Constrained pole optimization for modal reverberation,” in *Proc. of the 5th International Conference on Digital Audio Effects (DAFx)*, 2017, p. 155.

[4] M. Karjalainen, P. Ansalo, A. Mäkivirta, T. Peltonen, and V. Välimäki, “Estimation of modal decay parameters from noisy response measurements,” *Journal of the Audio Engineering Society*, vol. 50, no. 11, pp. 867–878, 2002.

[5] K. Ege, X. Boutillon, and B. David, “High-resolution modal analysis,” *Journal of sound and vibration*, vol. 325, no. 4-5, pp. 852–869, 2009.

[6] M. R. Ashory, “High quality modal testing methods,” Ph.D. dissertation, University of London, 1999.

[7] G. Rabbiolo, R. Bernhard, and F. Milner, “Definition of a high-frequency threshold for plates and acoustical spaces,” *Journal of Sound and Vibration*, vol. 277, no. 4-5, pp. 647–667, 2004.

[8] M. H. Richardson and D. L. Formenti, “Parameter estimation from frequency response measurements using rational fraction polynomials,” in *Proceedings of the 1st international modal analysis conference*, vol. 1. Union College Schenectady, NY, 1982, pp. 167–186.

[9] M. Ducceschi and L. Gabrielli, “1st DAFx parameter estimation challenge: Task and evaluation details,” *27th Int. Conf. on Digital Audio Effects (DAFx26)*, 2026.

[10] S. Bilbao, K. Arcas, and A. Chaigne, “A physical model for plate reverberation,” in *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, vol. 5. IEEE, 2006.

[11] C. Kereliuk, W. Herman, R. Wedelich, and D. J. Gillespie, “Modal analysis of room impulse responses using subband esprit,” in *Proceedings of the International Conference on Digital Audio Effects*, 2018.

[12] H. W. Kuhn, “The hungarian method for the assignment problem,” *Naval research logistics quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955.

[13] S. G. Mallat and Z. Zhang, “Matching pursuits with time-frequency dictionaries,” *IEEE Transactions on signal processing*, vol. 41, no. 12, pp. 3397–3415, 1993.

[14] S. L. Marple Jr, *Digital spectral analysis*. Courier Dover Publications, 2019.

[15] R. Roy and T. Kailath, “Esprit-estimation of signal parameters via rotational invariance techniques,” *IEEE Transactions on acoustics, speech, and signal processing*, vol. 37, no. 7, pp. 984–995, 1989.