

Eigenform: A Neural-Feedback Hybrid System for Improvisational Electronic Music Performance

Nicola Casetta
Conservatory of Teramo
n.casetta@conservatoriobraga.it

Abstract

This paper presents *Eigenform*, a hybrid feedback-neural system for solo improvisational electronic music performance. Built on a closed eight-channel nonlinear processing network implemented in Max/MSP using `mc.gen~`, the system is coupled with a 16-channel RAVE model trained on recordings of its predecessor system, *Quiet Catastrophe Unleashed*, and deployed via `nn~`. The model operates in two combinable modes: as an independent synthesis voice sounding alongside the live process, and as a signal injected into the feedback loop, where the deterministic network processes and re-encodes its output. Real-time latent navigation via LFO modulation and MIDI controller allows the performer to move continuously between them. The recursion that gives the work its name is located inside the performance session rather than across training: a live deterministic process circulating repeatedly through a fixed model of an earlier state of itself.

1 Introduction

Feedback-based sound synthesis occupies a productive tension between control and emergence. A prior system, *Quiet Catastrophe Unleashed* (QCU) (Casetta, 2026), introduced a closed multichannel feedback architecture for solo improvisational performance, implementing a formally navigable dynamical space through a bank of eight nonlinear processing modules (Sanfilippo and Valle, 2013; Kiefer et al., 2025). Its behavior, however, was governed by relationships fixed in advance by the designer, and so remained external to the sonic history the system itself produced.

Eigenform integrates a RAVE model (Caillon and Esling, 2021a) trained exclusively on approximately three hours of recordings of QCU, so that a learned distribution over that history now conditions how the network behaves. The term *eigenform*—developed by von Foerster (1981) to describe a form stabilized through recursive self-application—names what occurs when such a model is placed back inside a live loop.

2 Practice Field

Feedback systems built on internal signal recirculation, rather than a conventional path from source to output, have a long lineage in experimental electronic music. No-input mixing board practice, explored since at least the 1960s and strongly associated with Nakamura (Nakamura, 2017; Mudd and Brown, 2023), is the foundational analog reference; digital counterparts include Sanfilippo and Valle’s SD/OS (Sanfilippo and Valle, 2013), Manousakis’s *Fantasia on a Single Number* (Manousakis, 2009), and Elia and Overholt’s digitally-controlled analog no-input mixer (Elia and Overholt, 2024).

The role of machine learning in such instruments is only beginning to be explored. The Proto-Langspil (Armitage et al., 2022) is a feedback string instrument built in a lab explicitly oriented toward embedded intelligence, though

the learned component does not sit inside its loop. Kiefer’s Xiasri (Kiefer, 2022) couples physical string feedback to a software loop that resynthesizes the instrument’s recent past through self-concatenative synthesis; Kotowski’s *Broken Forecasts* (Kotowski, 2025) routes a latent generator recursively through delayed feedback, treating the resulting instability as compositional material; and Shaheed and Wang’s *I Am Sitting in a (Latent) Room* (Shaheed and Wang, 2024) repeatedly re-encodes a clip through two parallel variational autoencoders while performers steer the latent embeddings. These are the closest precedents I have found. What separates them from the far larger body of work using neural audio models is position: the model is not an endpoint receiving a finished signal but a stage the signal passes through and returns from. *Eigenform* differs in the provenance and persistence of its memory, which is neither a clip degraded within one performance nor an external corpus, but the recorded output of a distinct predecessor, carried between performances.

3 Methods

3.1 Architecture

Eigenform is implemented in Max/MSP using `mc.gen~` for sample-level multichannel processing, with a 16-channel RAVE model deployed via `nn~` (ACIDS/IRCAM, 2022). As shown in Figure 1, the signal passes through eight nonlinear modules in series—Multi-Effect Delay, Chebyshev Shaper, Quantizer, Glideverb, Exciter, Frequency Shifter, Chaotic Attractor, and Limiter—whose output feeds the RAVE model and the system output. The encoder maps incoming audio to a 16-dimensional latent vector; the decoder maps it back to audio, with the latent state shaped in real time by LFO modulation and MIDI controller. The final limiter bounds the loop, guaranteeing that the network remains stable under arbitrary injection depth.

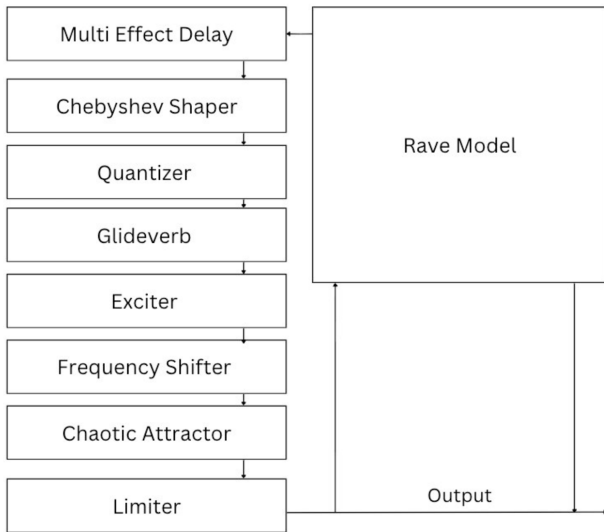


Figure 1: Signal flow of *Eigenform*. The nonlinear processing chain feeds the RAVE model, whose output recirculates to the top of the chain (feedback injection) and to the system output (parallel voice).

3.2 Corpus and Training

The model was trained on seven recordings of QCU, totalling approximately 2 hours 59 minutes and spanning that system’s full behavioral range, from sparse resonant states to dense chaotic eruptions. The set includes QCU’s first public performance, structured around a prepared score mapping known stable and unstable parameter regions. Training used the v2 configuration of the official RAVE implementation (Caillon and Esling, 2021b), following the architecture of Caillon and Esling (2021a). The corpus is closed: the model has heard the predecessor and nothing else.

3.3 Operational Modes

Two independent gains route the decoded output: to the system output (**parallel voice**) and into the loop (**feedback injection**), where the modules process it as any other signal and the result is re-encoded on the following cycle. Both run simultaneously, with continuously variable injection depth.

4 Performance and Listening

In parallel voice mode the two layers remain separable by ear. The decoded output is recognizably of the same family as the live signal—same grain, same resonances—but it does not answer what the live process is doing at that instant. It arrives slightly out of register with the situation: a plausible version of the present that is not the present one. This is what *temporal displacement* means here—not delay, but a second voice with the right vocabulary and the wrong tense.

Injection collapses that separation. Once decoded material enters the chain it is no longer a voice but a source: shaped by the modules, bounded by the limiter, re-encoded on the next cycle. It stops being audible as such and becomes a bias in the network’s behavior, pulling the loop toward

states the model finds probable. *Turbulent fusion* names this condition: the performer hears only the composite.

The continuum between the modes is the substance of the piece. Playing alongside the model is playing with a document of what the lineage has already done: the past is quoted, audible, and answerable. Playing through it means the live signal cannot reach the output without first traversing a compressed representation of that past—every gesture filtered by what the model considers likely. Injection depth is therefore a control over how much of the present is permitted to be unprecedented. At low depth the performer improvises with a history; at high depth, that history improvises with the performer. Moving audibly along this axis is the performative claim of the work.

5 Recursion and its Limits

Two clarifications are owed to the framing. The model is trained once and frozen: its memory does not update during or between performances. The recursion is therefore not the ongoing self-production of the cybernetic ideal, in which the operator stabilizing a form is itself continually re-derived. It is signal-level recursion within a single session—the loop passing repeatedly through a fixed operator, its stable textures the fixed points of that composite process. The eigenform is earned in the session, not in the training. Online adaptation, in which the model’s distribution shifts with what it is currently hearing, remains the open frontier, alongside characterization of latent regions against behavioral regimes and multi-performer configurations sharing one network.

Author Declarations

This work was developed during a research residency at CNMAT, University of California, Berkeley, from September to December 2025, hosted by Carmine Emanuele Cella. The RAVE model was trained by Jeremy Wagner.

References

- ACIDS/IRCAM (2022). nn~: Real-time neural network inference in max/msp. https://github.com/acids-ircam/nn_tilde.
- Armitage, J., Magnusson, T., Shepardson, V., and Ulfarson, H. (2022). The proto-langspil: Launching an icelandic NIME research lab with the help of a marginalised instrument. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*. <https://nime.pubpub.org/pub/langspil>.
- Caillon, A. and Esling, P. (2021a). RAVE: A variational autoencoder for fast and high-quality neural audio synthesis. *arXiv preprint arXiv:2111.05011*.
- Caillon, A. and Esling, P. (2021b). RAVE: Official implementation. <https://github.com/acids-ircam/RAVE>.

- Casetta, N. (2026). Design and formalization of a multichannel feedback network: Quiet Catastrophe Unleashed. In *Proceedings of SpaMec 2026*. Forthcoming.
- Elia, G. and Overholt, D. (2024). Musicking with dynamical systems: Introducing a digitally-controlled analog no-input mixer. In *Proceedings of the 19th International Audio Mostly Conference*.
- Kiefer, C. (2022). Xiasri. Feedback string instrument. Performance documentation, Moving Strings, Intelligent Instruments Lab, Reykjavik.
- Kiefer, C., Eldridge, A., and Overholt, D. (2025). Future research challenges in feedback musicianship.
- Kotowski, B. (2025). Broken forecasts: Feedbacking latent generators for sonic instability. Music submission, 6th Conference on AI Music Creativity (AIMC 2025), Brussels, Belgium, 10–12 September.
- Manousakis, S. (2009). Fantasia on a single number. Composition for digital feedback (live electronics). Released on *Primeval Sonic Atoms*, Inkilino Records, 2020.
- Mudd, T. and Brown, A. (2023). Musical pathways through the no-input mixer. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*. https://nime.org/proceedings/2023/nime2023_56.pdf. VERIFY page range before submission.
- Nakamura, T. (2017). The art of noise: Toshimaru nakamura interview. MusicOff, online music magazine.
- Sanfilippo, D. and Valle, A. (2013). Feedback systems: An analytical framework. *Computer Music Journal*, 37(2):12–27.
- Shaheed, N. and Wang, G. (2024). I am sitting in a (latent) room. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, pages 333–338, Utrecht, Netherlands.
- von Foerster, H. (1981). Objects: Tokens for (eigen-)behaviors. In *Observing Systems*, pages 274–285. Intersystems Publications, Seaside, CA. Originally presented at the University of Geneva, 29 June 1976, on the occasion of Jean Piaget’s 80th birthday. Reprinted in *Understanding Understanding*, Springer, 2003, pp. 261–271.