

Motivation

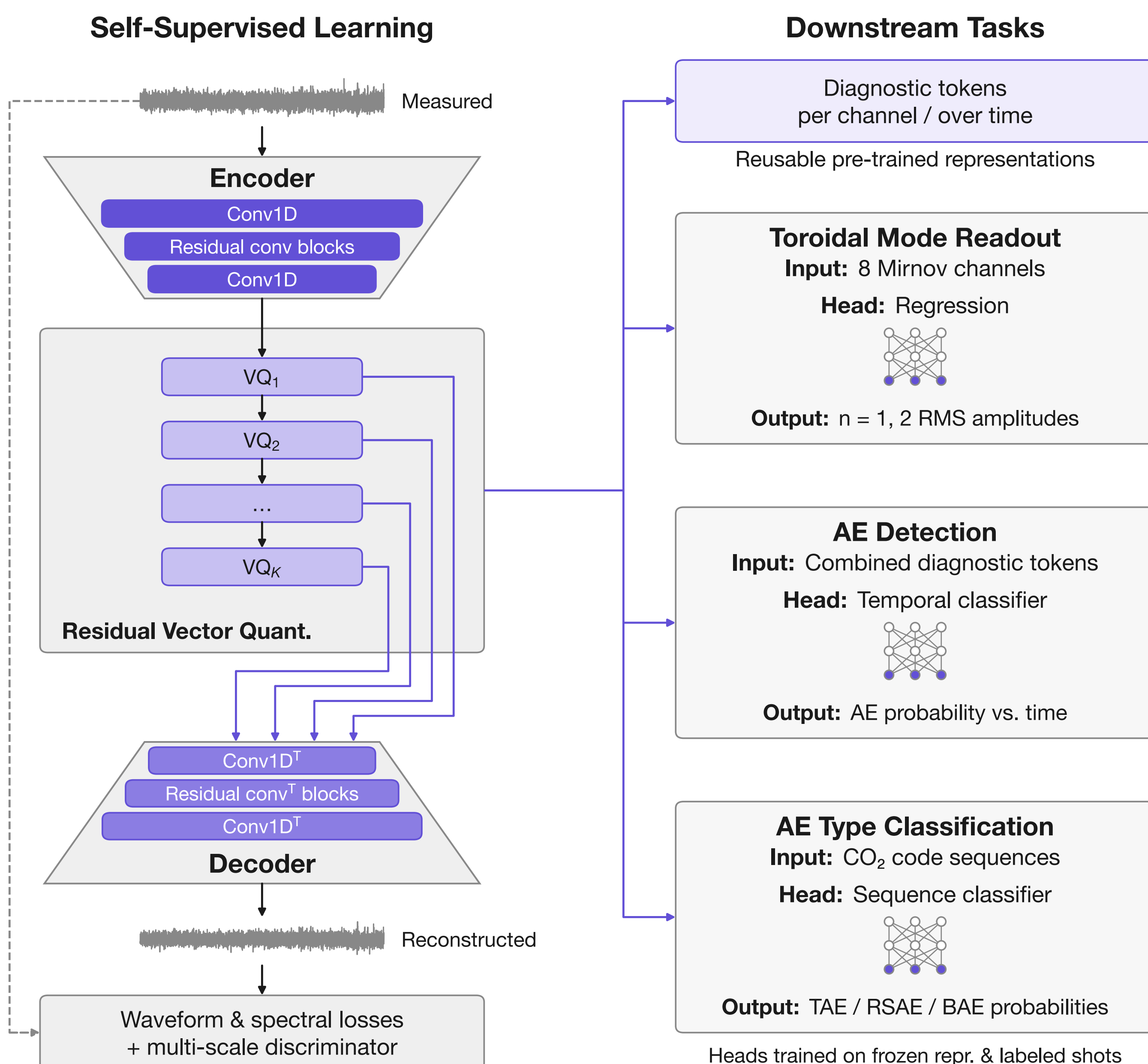
- **Foundation models:** Large neural nets pre-trained on diverse data to learn general-purpose representations adaptable to multiple downstream tasks.
- **Diagnostic fusion:** Building on the success of *Diag2Diag* [1], we seek to integrate multiple diagnostics to build a foundation model for fusion.
- **Self-supervised learning** is able to leverage large-scale multi-modal data to learn a universal representation of plasma dynamics.

Dataset

- Currently **~17,000 DIII-D shots** used to train with 3 high-res. diagnostics:
 - **Electron Cyclotron Emission:** 48 channels @ 500 kHz, edge-to-core.
 - **CO₂ Interferometry:** 4 chords @ 1.6 MHz, 3 vertical and 1 radial.
 - **High-Speed Magnetic Probes:** 8 channels @ 2 MHz.
- **Raw waveforms** are fed directly into the encoders, channel by channel.
 - The encoders learn representations from the time-domain signals.

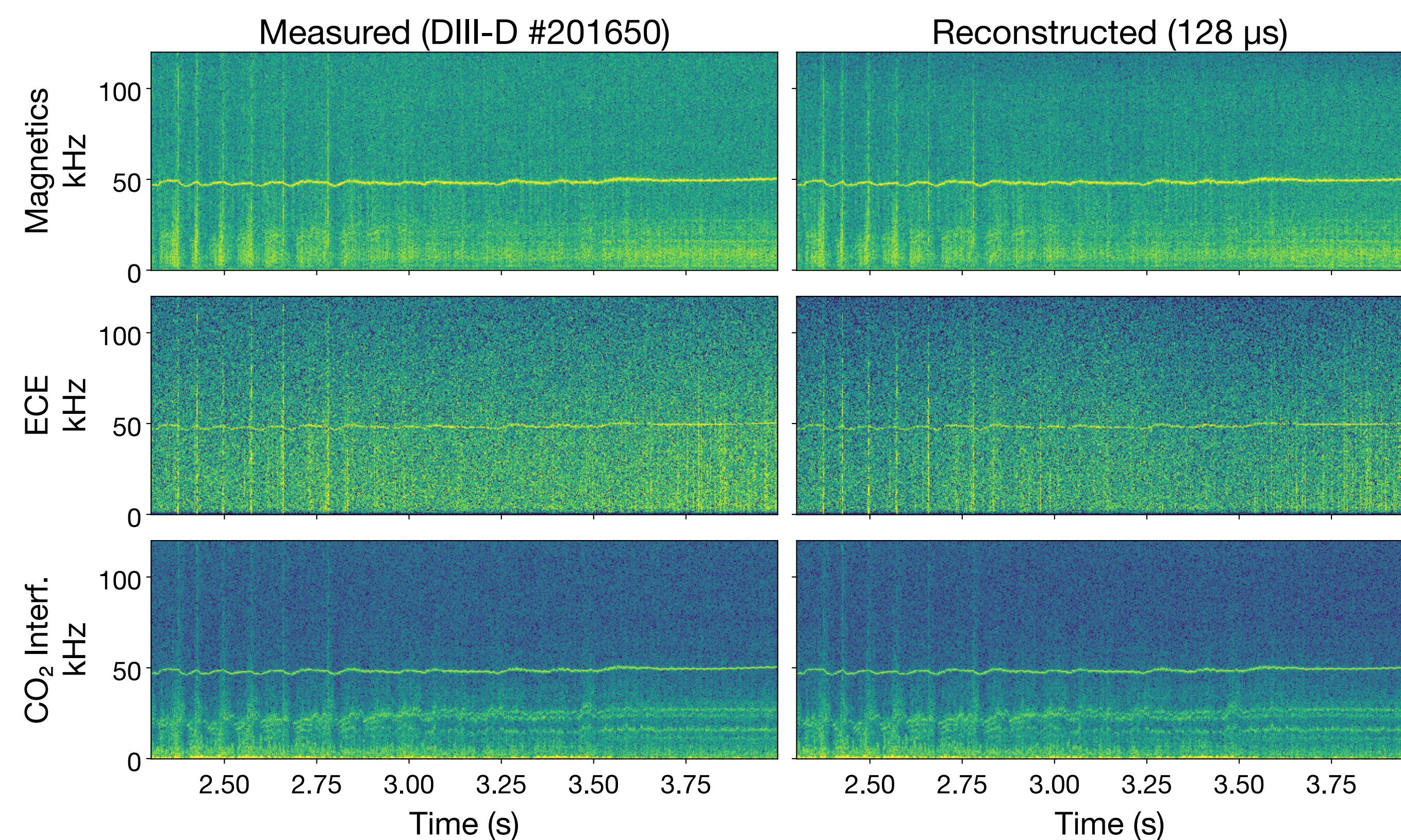
Architecture

- Inspired by Neural Audio Codecs (NACs): **EnCodec** [2] and **DAC** [3].

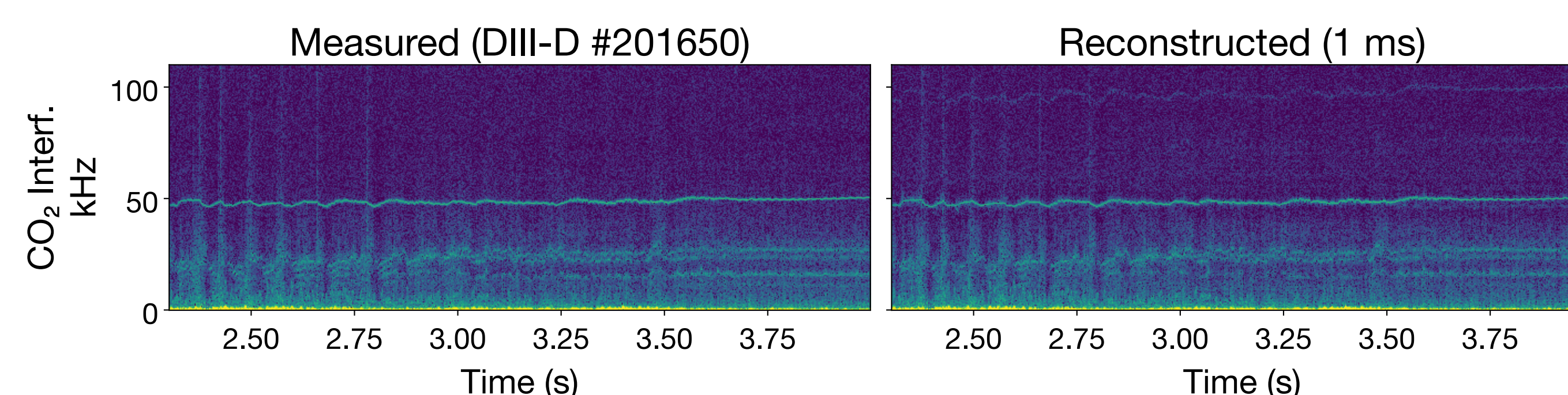


Example Reconstruction

- **Spectrograms** of measured (left) and codec-reconstructed (right) signals.



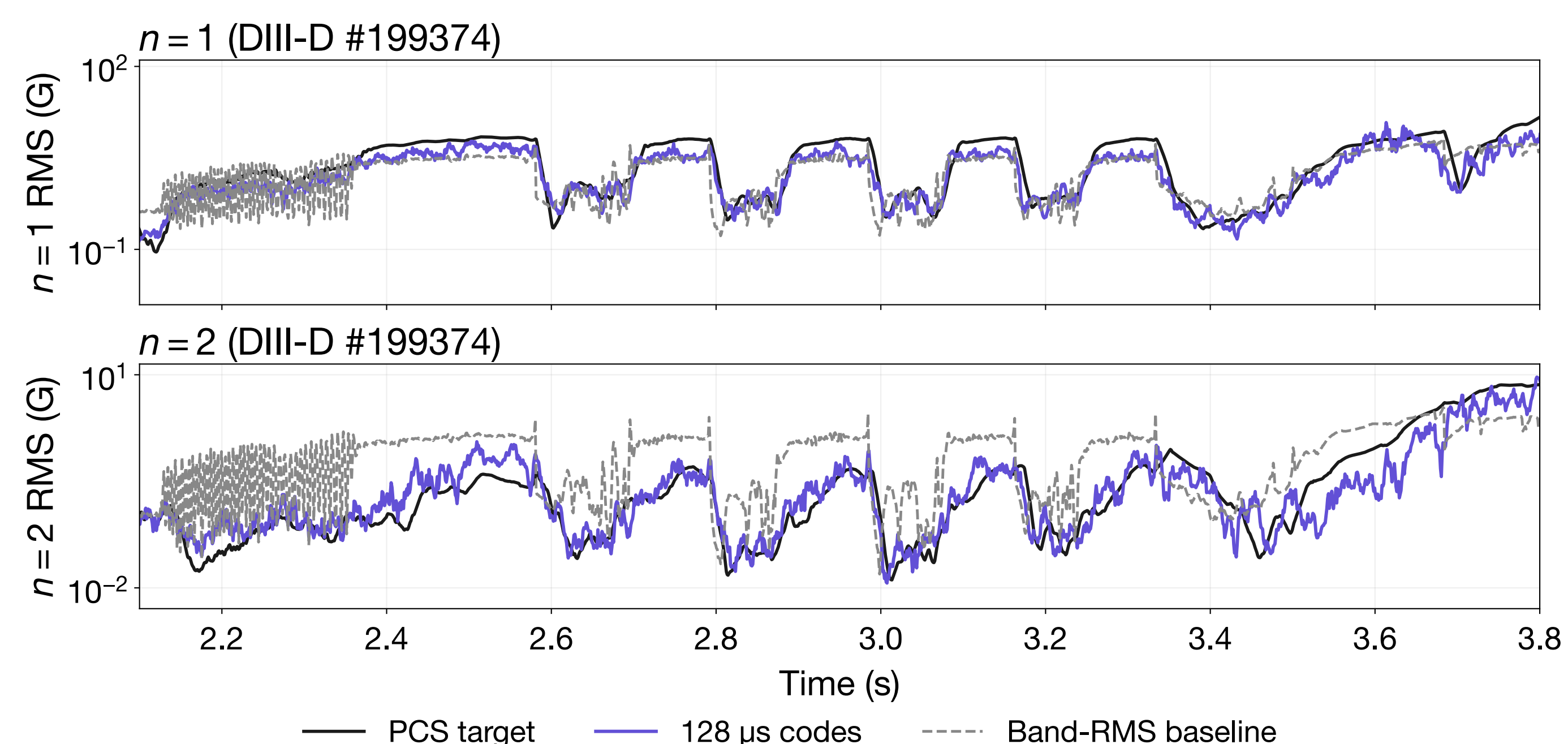
- **6.4x compression** vs. 16-bit PCM with 128 μs codes.



- **17x compression** vs. 16-bit PCM for CO₂ interferometry with 1 ms codes.

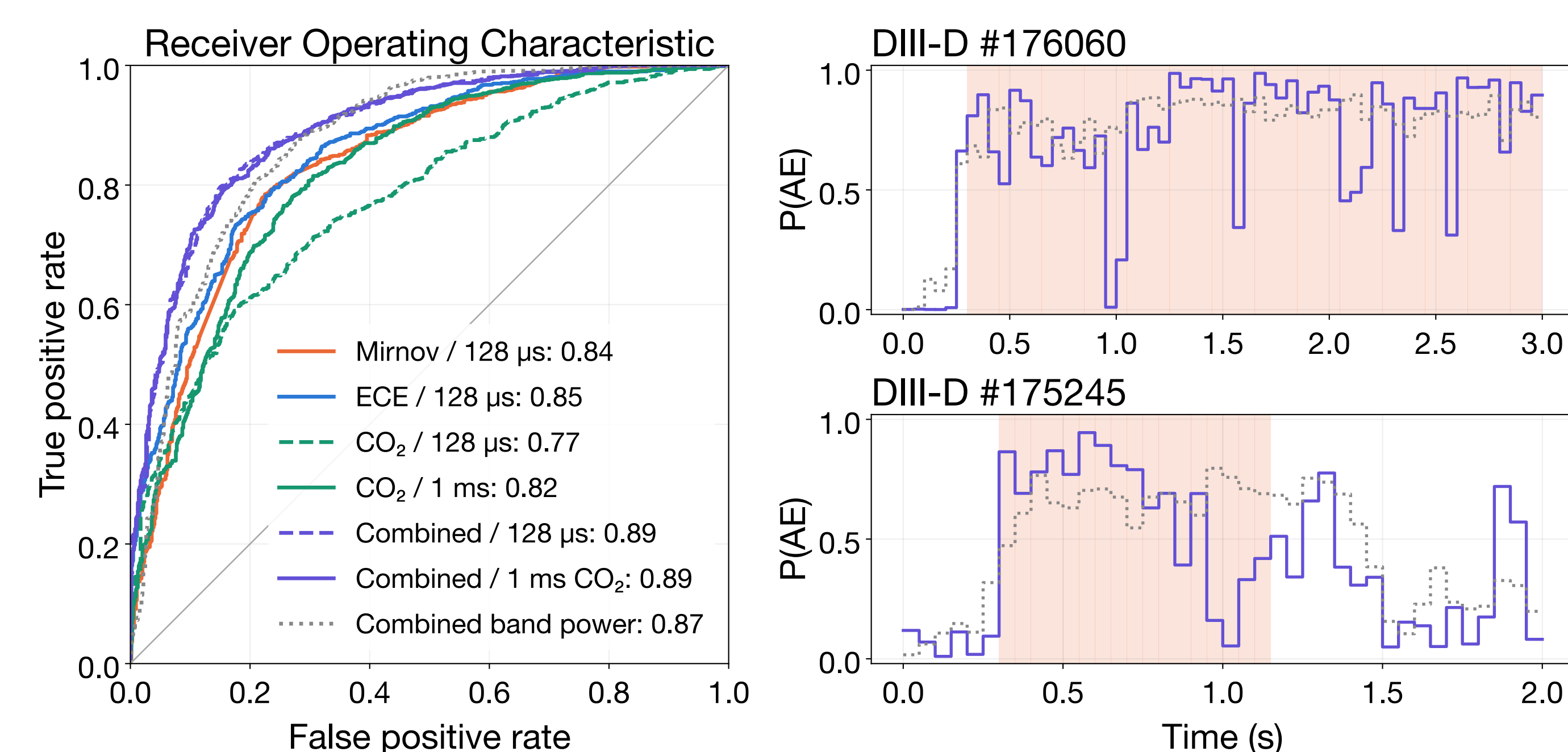
Toroidal Mode Amplitude Readout

- 8 Mirnov channels predicts DIII-D's $n = 1, n = 2$ **RMS amplitudes**.
 - **Input:** 128 μs codes. **Target:** PCS amplitudes smoothed over 20 ms.



- 58 held-out shots, $R^2 = \mathbf{0.845 / 0.939}$ vs. 0.760 / 0.797 for band-RMS.

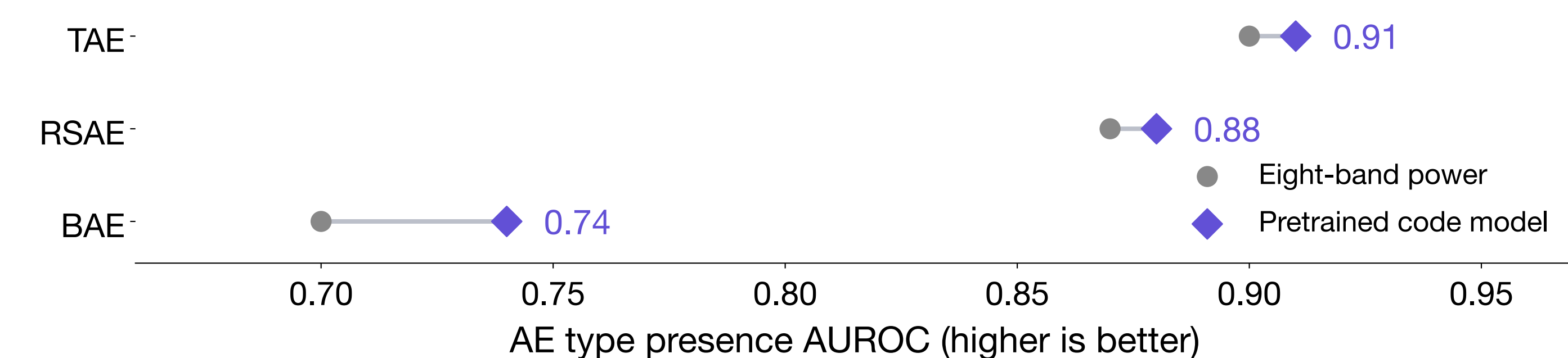
Alfvén-Eigenmode Detection



- **Input:** Diagnostic tokens. **Target:** Expert-labeled AE activity.
 - Compared against a spectral band power baseline.
- **Combined AUROC:** 0.900 → **0.916** with 1 ms CO₂ vs. band power: 0.889.
 - Three-seed means, with curves and legend scores on seed 0.
- **Right:** Per-bin AE probability on test shots.
 - Purple: combined model. Gray: band-power baseline. Shading: labels.

Alfvén-Eigenmode Type Classification

- **Input:** 128 μs CO₂ tokens. **Target:** Presence of each AE type.
 - Expert-labeled, compared against an eight-band power baseline.
- **AUROC:** TAE **0.91** vs. 0.90, RSAE **0.88** vs. 0.87, BAE **0.74** vs. 0.70.



Next Steps

- Scale the codec to more diagnostics and shots across the DIII-D archive.
- Train a shared model across diagnostics and actuators.
 - Test reconstruction of missing channels and prediction of future signals.

Conclusion

Per-channel neural audio codecs (NACs) are able to compress raw diagnostics, and their frozen codes beat band-power baselines for mode amplitudes and AE detection. We believe NACs are viable building blocks for future foundation models trained directly on raw diagnostic waveforms.