A Geometric Framework for Analyzing Drum Groove: Coupled Timing-Amplitude Vector Representations

Tyler Garcia *tgarcia2@uco.edu*, Dr. Erin Williams *ewilliams50@uco.edu*, Dr. Britney Hopkins *bhopkins3@uco.edu*
University of Central Oklahoma Department of Mathematics and Statistics

Research Question

What quantifiable short-range patterns in timing and amplitude give drum performances their characteristic feel?

Prior work analyzes timing and amplitude separately or only measures their statistical correlation. We propose treating each drum hit as a 2D vector, enabling pattern discovery involving both dimensions of the performance.

Background

Groove is often described as the subjective urge to move when listening to music, and emerges from micro-level timing and amplitude deviations in drum performance [1, 2].

Previous research has focused on **long-range correlations** in timing, finding $1/f$-noise patterns across entire performances [1]. Short-range patterns remain understudied. We hypothesize:

- Groove may be characterized by measureable patterns that exist within microtiming deviations
- Timing deviations cannot be meaningfully analyzed in isolation from their associated amplitudes
- These patterns are present in remarkably brief periods of music

The song featured in this analysis, *I Keep Forgettin'* by Michael McDonald, was chosen for being widely regarded as a drum performance with excellent groove, and for its inclusion and study in the Räsänen et. al. paper [2]. To simplify matters while building out this framework, we are looking only at the hi hat hits in this performance.

The Hit Vector

Each drum hit is represented as a 2D vector:

$$\mathbf{h}_n = (\Delta t_n,\ v_n) \tag{1}$$

where Δt_n is the timing deviation from the nearest grid position (in ms or normalized by beat period), and v_n is the amplitude (MIDI velocity, normalized 0–1).

A performance of N hits becomes a matrix:

$$H = \begin{bmatrix} \Delta t_1 & v_1 \\ \Delta t_2 & v_2 \\ \vdots & \vdots \\ \Delta t_N & v_N \end{bmatrix} \tag{2}$$

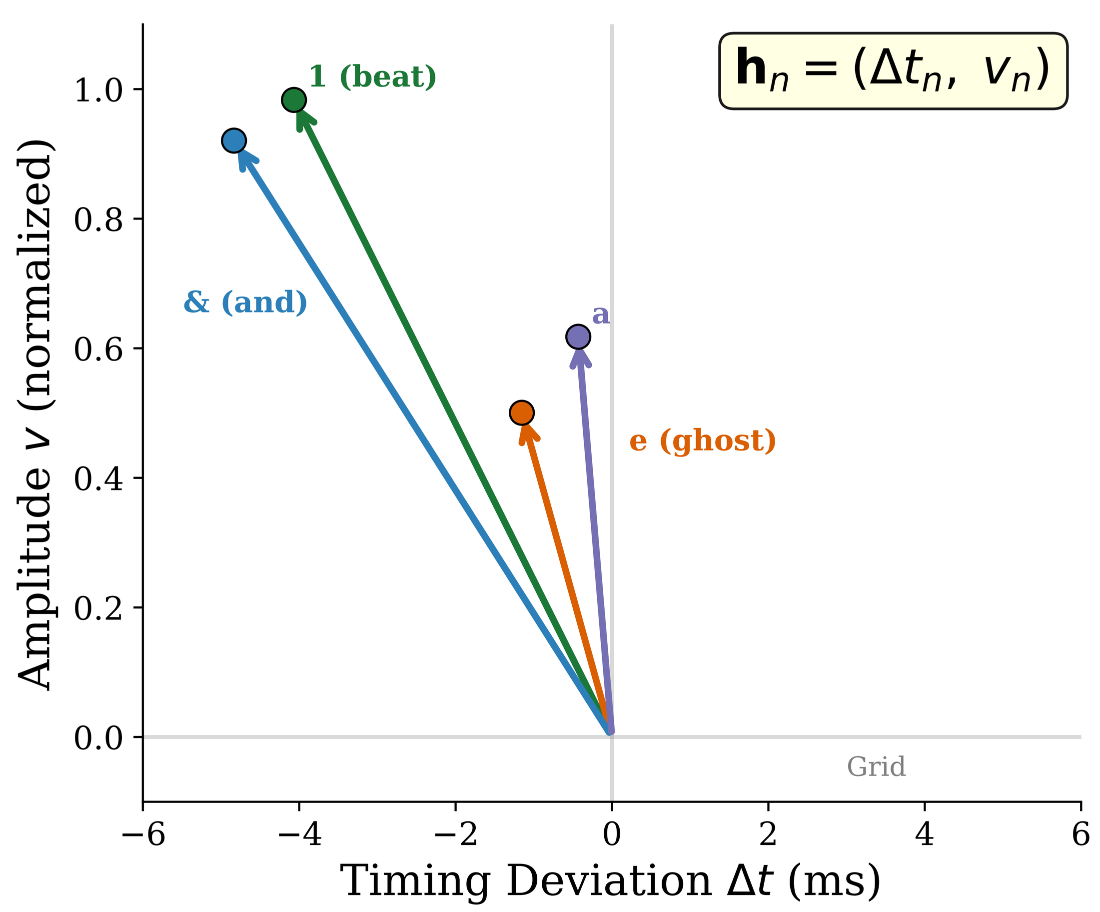

Figure 1: Drum hits in timing-amplitude space. Each arrow represents one hit as a displacement from the ideal grid.

Onset Detection

Pipeline:

1. **Drum separation**: Stem splitter software isolates the drum track from the original studio mix
2. **Timing normalization**: Since the source recording was performed without a metronome, the individual phrases are all slightly different lengths. Each 2-bar phrase was manually aligned to the grid at the first downbeat and time-normalized to a common tempo by stretching to match the following downbeat. This preserves the internal microtiming structure while enabling comparison across sections.
3. **MIDI extraction**: Drum hit detection plugin inside of digital audio workstation software converts the isolated audio to MIDI note events with precise timing and velocity
4. **Timing deviation detection**: Each hit is assigned to the nearest 16th-note grid position; the deviation Δt is the signed distance from that grid point

At 480 PPQ (Pulses per quarter note effective resolution of this MIDI file) and 97 BPM, one MIDI tick $\approx$ **1.3 ms**—well below the $\sim$5 ms threshold of human timing perception. Velocity values (0–127) are recorded exactly as performed, with no amplitude estimation required.

Profile Stacking Methodology

Improving signal to noise ratio through epoch averaging: a technique utilized in analysis of periodic noisy signals like EEG and radar, we align and stack multiple iterations of the same drum pattern to average out noise and reveal faint structure. In this analysis, the performance is divided into 2-bar phrases, with 8 of these sections chosen due to their high degree of stationarity by ear.

Procedure:

1. Export k instances of the same 2-bar pattern as separate MIDI files from the source recording
2. For each grid position p, collect velocity and timing values across all k iterations. Timing deviation determined by comparing each hit to the 16th-note grid (32 hits per 2-bar phrase)
3. Compute the sample means $\bar{v}_p$ and $\bar{\Delta t}_p$ at each position

Expected behavior:

- Random performance variation averages out as $\sigma/\sqrt{k}$
- Systematic patterns (the "true groove") are preserved and sharpened
- Consistency (coefficient of variation) reveals which positions are more tightly controlled

Stacked Velocity Profile: 8 Iterations of a 2-Bar Pattern

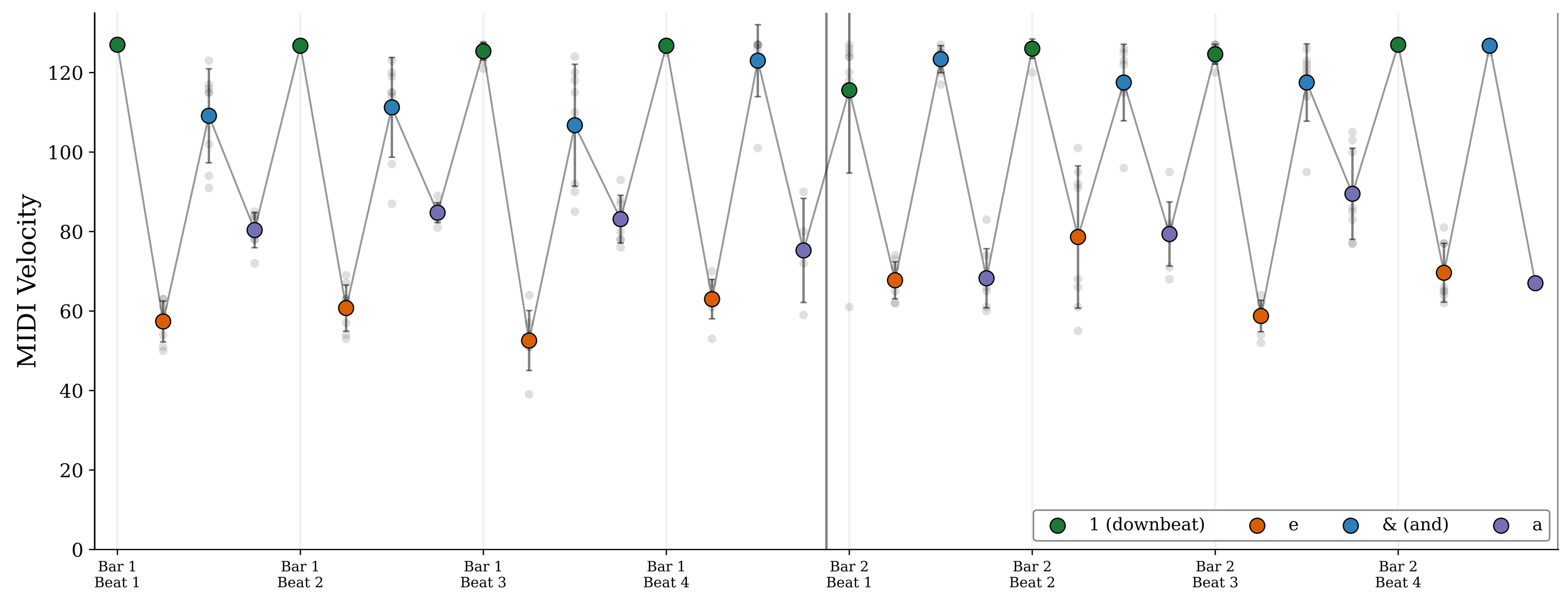

Figure 2: Stacked velocity profile from 8 iterations of the same 2-bar 16th-note hi-hat pattern ("Keep Forgettin'" Michael McDonald, 97 BPM). Gray dots show individual iterations; colored markers show the stacked mean $\pm\sigma$ at each of the 32 grid positions. A clear **4-level velocity hierarchy** emerges: downbeats (green), and's (blue), a's (purple), and e's (orange).

Timing-Velocity Space

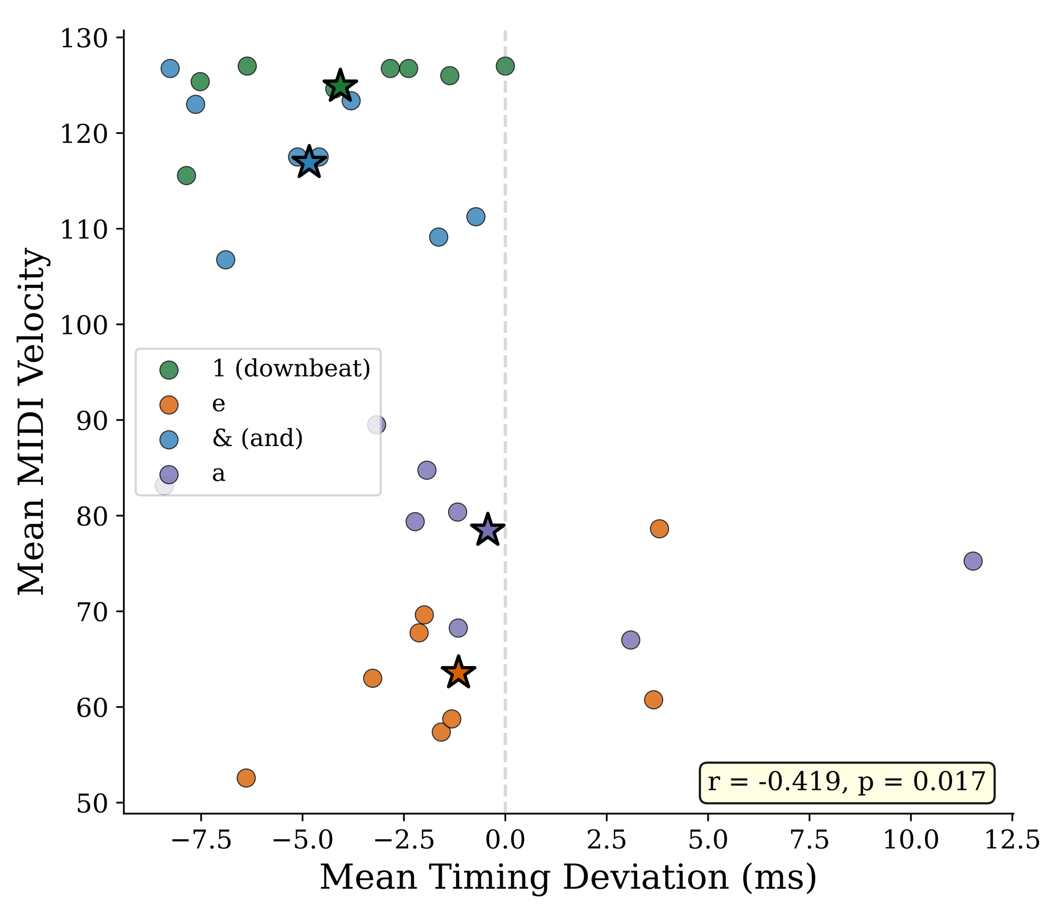

Figure 3: Stacked mean positions in $(\Delta t, v)$ space. Stars mark centroids for each beat-position type. The 4 hierarchical levels occupy distinct regions, confirming that timing and amplitude are structured jointly, not independently.

Velocity Ratios

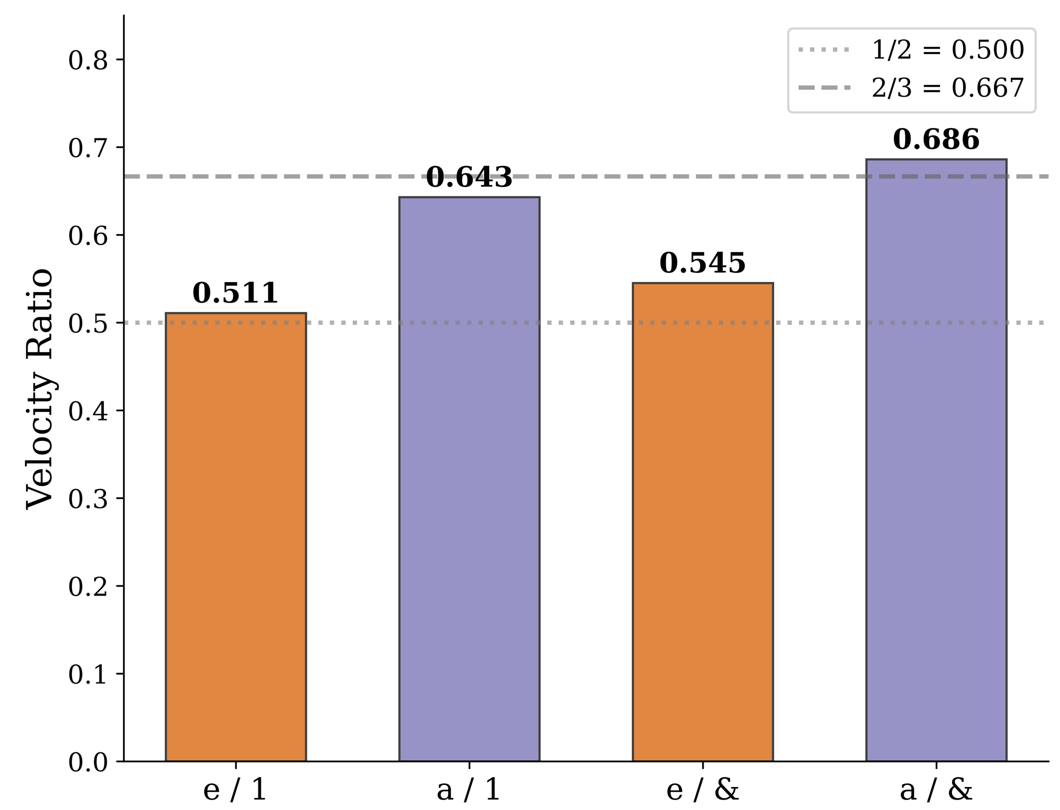

Figure 4: Pairwise velocity ratios between the 4 hierarchical levels. Dashed/dotted lines show simple-fraction reference values. The ghost/accent ratios cluster near $\frac{1}{2}$ and $\frac{2}{3}$.

Preliminary Findings

From 256 total hits (8 iterations $\times$ $\sim$32 hits):

1. Four-level velocity hierarchy

The 16th-note hi-hat pattern contains 4 distinct velocity levels, corresponding to the subdivisions within each beat. *The four 16th-note subdivisions within each beat are commonly spoken "1-e-&-a" by musicians.*

16th Note Position	Mean vel.	CV
1 (downbeat)	124.7	6.7%
e	63.7	17%
&	116.9	10.0%
a	80.2	12%

2. Accented beats are more consistent

The coefficient of variation increases as velocity decreases. The drummer's motor control is tightest on the accented beats and loosest on ghost notes.

3. Simple-fraction ratios

The ghost/accent velocity ratios are well-approximated by $\frac{1}{2}$ and $\frac{2}{3}$, suggesting the drummer maintains velocity relationships at simple integer ratios.

Ratio	Measured	95% CI
e / beat	0.511	[0.489, 0.535]
a / &	0.686	[0.659, 0.715]

Table 1: Ghost-to-accent ratios with confidence intervals. Simple fractions $\frac{1}{2}$ and $\frac{2}{3}$ fall within the CIs.

Future Directions

- **More data**: Stack 20+ iterations for tighter confidence intervals
- **Cross-song comparison**: Do other performances show the same hierarchy?
- **Perceptual testing**: Do listeners prefer patterns with these specific ratios?
- **Allan variance**: Borrow from clock stability analysis to characterize timing fluctuations at multiple scales

References

[1] H. Hennig, R. Fleischmann, A. Fredebohm, et al. (2011). The Nature and Perception of Fluctuations in Human Musical Rhythms. *PLoS ONE*, 6(10).

[2] E. Räsänen, O. Pulkkinen, T. Virtanen, M. Zollner, H. Hennig (2015). Fluctuations of Hi-Hat Timing and Dynamics in a Virtuoso Drum Track of a Popular Music Recording. *PLoS ONE*, 10(6).

Tools & Data

Analysis pipeline built with Python (`mido`, `numpy`, `matplotlib`). Drum stem extraction from mix via Apple Logic Pro stem splitter, Drum audio extracted via Massey DRT from the original studio recording. MIDI export and audio time stretching in Pro Tools.

Funding

This project is based on work supported by a grant from the Office of High-Impact Practices, University of Central Oklahoma.