forum acusticum 2026
8-12 September · Graz, Austria

DYNAMICS, COMPRESSION, AND PERCEIVED GROOVE IN SNARE-DRUM PATTERNS

Ulysse Lefeuvre[1,2,3], Axel Marmoret[3], Nicolas Epain[1], Nicolas Farrugia[3], Luc Arnal[2]

[1] *L-Acoustics, 13 Rue Levacher Cintrat, 91460, Marcoussis, France*
[2] *Université Paris Cité, Institut Pasteur, AP-HP, INSERM, CNRS, Fondation Pour l'Audition, Institut de l'Audition, IHU reConnect, 75012, Paris, France*
[3] *IMT Atlantique, Lab-STICC, UMR CNRS 6285, Brest, France*

Abstract

Groove, defined as the pleasurable urge to move in response to music, is a central dimension of rhythmic perception. While previous research has examined the contributions of syncopation, microtiming, and tempo to groove, the role of sound dynamics remains poorly understood. This study investigates how variations in sound level across drum patterns influence perceived groove, and whether dynamic range compression can be used to enhance it. Twenty participants rated the groove, pleasure, and complexity of 159 short snare-drum sequences designed to systematically vary dynamic structure. A linear mixed-effects model revealed significant effects of dynamic pattern on groove ratings, with rankings that remained stable across different levels of gain variability. To predict groove from sound dynamics, we introduce a fuzzy representation of gain levels combined with a ridge regression model. A transition-based feature representation, encoding sequential dependencies between successive gain events, outperformed a distribution-based approach, suggesting that the temporal ordering of dynamic levels is a key perceptual cue. Using the trained model in a generative framework, we further show that modifying sound dynamics with dynamic compression can significantly increase predicted groove across a wide range of dynamic configurations. These findings suggest that compression could be used not only for loudness control, but also as a perceptually motivated tool for shaping groove in live music contexts.

Keywords: *Groove, sound dynamics, dynamic compression*

Corresponding author: *ulysse.lefeuvre@l-acoustics.com.*

1. Introduction

Music often induces a pleasurable urge to move in time with the rhythm, an experience commonly referred to as groove [1]. Groove is consistently associated with sensorimotor synchronization and positive affect, and is considered a central mechanism through which listeners physically and emotionally engage with music. As such, it has become a key construct for understanding the relationship between rhythmic perception, movement, and pleasure. Previous studies have examined several musical properties that contribute to groove, including syncopation [2], microtiming [3] or tempo [4]. Because percussion often carries the most salient rhythmic information in groove-based music, many studies have used drum-based stimuli to investigate groove perception. Drummers use a wide range of musical parameters to create groove, including sound dynamics, that is variations in sound level [5]. The relevance of sound dynamics is further suggested by beat-perception research showing that intensity accents influence how easily listeners infer a beat [6]. In live music contexts, sound dynamics may be further modified by mixing engineers, as drum channels are often processed with dynamic range compression to control peaks and shape sound dynamics [7]. Despite the recognized importance of sound dynamics in performance practice and the widespread use of tools to manipulate it, its contribution to groove remains poorly understood. The present study investigates whether sound dynamics affect groove and whether compression can be used to improve groove.

2. Experimental setup

2.1. Participants

Twenty participants were recruited via the online data collection platform Prolific. All participants provided informed consent prior to taking part in the experiment. We only recruited adult participants. Upon completion, participants were financially compensated in accordance with Prolific's recommended rates.

2.2. Stimuli

The stimuli were 5 s snare drum sequences, which differed only in rhythmic structure and sound dynamics. To this end, sequences were quantized onto a fixed temporal grid with a resolution of 125 ms, eliminating any microtiming variation. Each stimulus thus consisted of 40 discrete time steps. The construction of the stimuli proceeded in two stages. First, the rhythmic structure of each sequence was defined by a *binary pattern* $\mathbf{b} \in \{0,1\}^L$, where each element indicates either the absence (0) or presence (1) of a snare onset at a given time step, and $L \in \{1,2,3,4\}$ is the pattern length (e.g., $\mathbf{b} = \mathtt{1010}$). Second, each onset was assigned to a discrete gain level category, yielding a *dynamic pattern* $\mathbf{d} \in \{0,1,2,3\}^L$, where 0 denotes silence, 1 soft, 2 medium, and 3 loud (e.g., $\mathbf{d} = \mathtt{1210}$). The full 40-step sequence was then obtained by periodically repeating $\mathbf{d}$. Note that $\mathbf{b}$ is implicitly encoded in $\mathbf{d}$: it is recovered by mapping any non-zero element to 1. All possible dynamic patterns of length $L \in \{1,2,3,4\}$ were enumerated, then filtered to remove redundant patterns according to four equivalence criteria: (i) *rotational equivalence* ($\mathtt{1203} \equiv \mathtt{2031}$), (ii) *reducible periodicity* ($\mathtt{1212} \equiv \mathtt{12}$), (iii) *gain translation* ($\mathtt{2323} \equiv \mathtt{1212}$), and (iv) *tempo scaling* ($\mathtt{1020} \equiv \mathtt{12}$). This yielded a set of 6 unique binary patterns and 53 unique dynamic patterns.

For each dynamic pattern, the gain values (in dB) associated with soft, medium, and loud levels were instantiated under three variability conditions. In the no-variance condition (NV), gain values were fixed at -20, -10, and 0 dB for soft, medium, and loud events, respectively. In the mid-variance condition (MV), gain values were drawn independently from uniform distributions over $[-30,-20]$ dB, $[-20,-10]$ dB, and $[-10,0]$ dB for soft, medium, and loud events. In the high-variance condition (HV), the same procedure was applied with wider ranges: $[-40,-20]$ dB, $[-30,-10]$ dB, and $[-20,0]$ dB, respectively. This resulted in a total of $53 \times 3 = 159$ stimuli. All stimuli were RMS-normalized to 0 dB. After normalization, $d_{\min}$ and $d_{\max}$ denote the minimum and maximum gain values observed across all generated stimuli. An example of stimuli under the three variability conditions is shown in Fig. 1.

2.3. Procedure

Participants completed the online experiment on a computer using headphones. Each trial consisted of the presentation of a stimulus, followed by a rating task where participants rated rhythmic complexity ("How complex did this rhythm feel?"), pleasure ("How pleasurable was it to listen to this rhythm?"), and groove ("To what extent did this rhythm give you an urge to move?") using a 7-point Likert scale. Prior to the experimental trials, participants completed a brief familiarization phase during which they listened to a subset of stimuli. No data from this phase was included in the analyses. For the MV and HV conditions, gain values were sampled anew immediately before each stimulus was presented. Consequently, although the underlying stimulus conditions were identical, the instantiated gain values for a given sequence varied across participants. Stimuli were played in a random order.

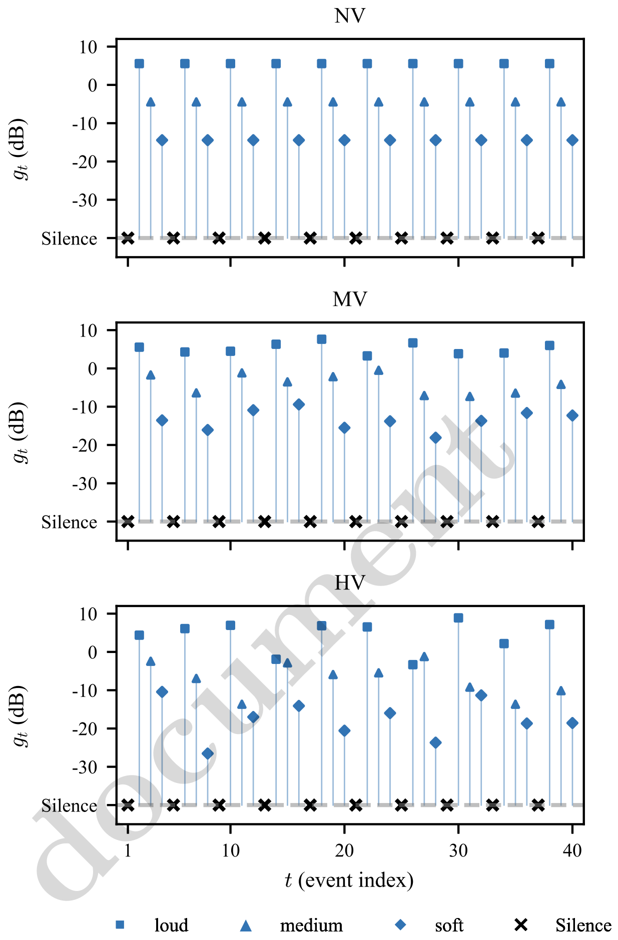

Figure 1: Examples of sequences for dynamic pattern $\mathbf{d} = \mathtt{0321}$ in the three probability distribution conditions. The top, middle and bottom plots correspond to the no-variance (NV), mid-variance (MV) and high-variance (HV) conditions, respectively.

3. Experimental data analysis

Because pleasure, groove and complexity ratings are highly correlated (see Fig. 2), only results for groove ratings are reported. A linear mixed-effects model was fitted to predict groove ratings, including fixed effects of dynamic pattern, probability distribution and their interaction, with random intercepts for participants. The analysis revealed a significant main effect of dynamic pattern ($F(52, 2844) = 5.69, p < 0.001$), indicating that groove ratings varied across dynamic patterns. Post hoc pairwise comparisons with Benjamini–Hochberg false discovery rate (FDR-BH) correction showed significant differences between some

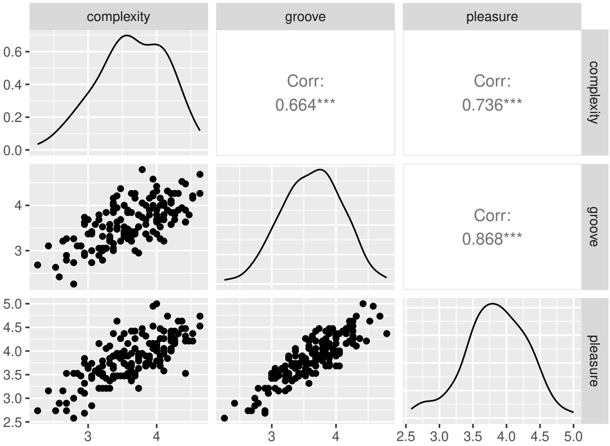

Figure 2: Pairwise scatterplots and distributions of complexity, groove, and pleasure ratings across stimuli, showing a strong correlation between all ratings.

dynamic patterns within each binary pattern (see Fig. 3). The main effect of probability distribution was non-significant ($F(2, 2844) = 2.25, p = .11$), suggesting that overall groove ratings were not influenced by probability distribution. The absence of a significant effect of probability distribution condition suggests that groove perception is not driven by an exact repetition of dynamic pattern, implying a degree of perceptual invariance to fine-grained amplitude fluctuations. Importantly, the interaction between dynamic pattern and probability distribution was non-significant ($F(104, 2844) = 1.14, p = .16$), indicating that the relative ordering of rhythmic structures was stable across probability distributions. This result was further supported by a strong and significant Kendall's coefficient of concordance across probability distributions ($W = 0.72$, $\chi^2(52) = 112, p < .001$), confirming a high level of agreement in the ranking of dynamic patterns. The variance associated with the participant effect ($\sigma^2 = 2.09$) was higher than the residual variance ($\sigma^2 = 1.59$), corresponding to an intraclass correlation of approximately 0.57.

4. Groove prediction from sound dynamics

4.1. Fuzzy representation of gain values

To predict listeners' ratings with sound dynamics, each sequence was represented as a time-series of gain values. A fuzzy representation was used to map each event onto a set of fuzzy gain categories. For each time step $t \in 1, \ldots, T$, the gain value g_t was mapped to a vector of membership coefficients $(\mu_k(g_t))_{0 \leq k \leq K}$, where μ_k is the fuzzy membership function of the fuzzy gain category G_k, with $k = 0$ corresponding to the silent category, and $k \in 1, \ldots, K$ corresponding to ordered gain categories ranging from the softest to the loudest. For silent events, the membership function was defined as:

$$\mu_0(g_t) = \begin{cases} 1 & \text{for} \quad g_t = -\infty \\ 0 & \text{for} \quad g_t \neq -\infty \end{cases} \tag{1}$$

For non-silent events, membership functions were defined as piecewise linear functions spanning the range $[g_{min}, g_{max}]$. These functions were constructed to satisfy:

$$\begin{cases} \displaystyle\sum_{k=0}^{K} \mu_k(x) & = & 1 \quad \forall x \\ \mu_k & \geq & 0 \\ \displaystyle\int \mu_k(x)\, dx & = & \Delta, \quad \Delta = \frac{g_{max} - g_{min}}{K} \end{cases} \tag{2}$$

An illustration of membership functions is shown in Fig. 4.

4.2. Feature representations

Two feature representations were considered to characterize the sound dynamics of a sequence based on the fuzzy representation introduced above. In both cases, each sequence was ultimately described by a feature vector $\mathbf{x} \in \mathbb{R}^d$, allowing a unified treatment within the prediction framework.

4.2.1. Distribution-based representation

In a first approach, each sequence was described by the distribution of fuzzy gain categories, independently of temporal order. The occupancy of fuzzy gain category G_k was defined as:

$$x_k = \frac{1}{T} \sum_{t=1}^{T} \mu_k(g_t) \tag{3}$$

The resulting feature vector is:

$$\mathbf{x} = (x_k)_{0 \leq k \leq K} \in \mathbb{R}^{K+1} \tag{4}$$

This representation captures the relative prevalence of each fuzzy gain category over time.

4.2.2. Transition-based representation

In a second approach, temporal dependencies between successive events were encoded through a transition-based representation. Transitions between fuzzy gain categories G_i and G_j were defined as:

$$X_{ij} = \frac{1}{T-1} \sum_{t=1}^{T-1} \mu_i(g_t)\, \mu_j(g_{t+1}) \tag{5}$$

The resulting matrix $\mathbf{X} = (X_{ij})_{0 \leq i,j \leq K}$ describes the occupancy of transitions between fuzzy gain categories. The transition matrix was vectorized into a feature vector:

$$\mathbf{x} = \mathrm{vec}(\mathbf{X}) \in \mathbb{R}^{(K+1)^2} \tag{6}$$

An example of transition-based representations is shown in Fig. 5.

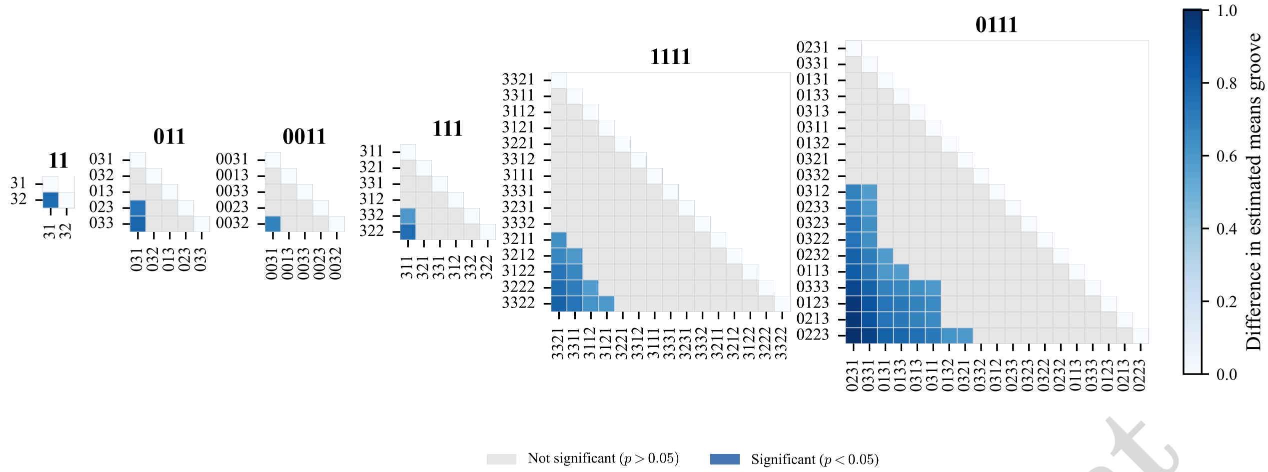

Figure 3: Pairwise comparisons between dynamic patterns within each binary pattern. Each panel corresponds to one binary pattern, with dynamic patterns ordered by increasing estimated mean groove in x-axis. Only the lower triangular part of each matrix is displayed. Grey cells indicate non-significant differences, whereas blue cells indicate significant differences between pairs of dynamic patterns. Color intensity reflects the magnitude of the difference in estimated means between dynamic pattern in y-axis and dynamic pattern in x-axis. Pairwise comparisons were corrected for multiple testing using the Benjamini–Hochberg false discovery rate (FDR-BH) procedure.

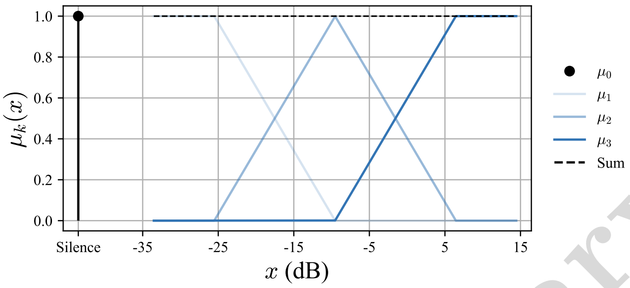

Figure 4: Fuzzy membership functions for $K = 3$. The x-axis shows gain values in dB between d_{min} and d_{max} with a specific gain category for silence, and the y-axis shows the membership degree of gain values for all fuzzy gain categories.

4.3. Ridge regression model

Both feature representations were used as explanatory variables in a regression framework. For each stimulus, features were computed at the participant level and subsequently averaged across participants. Groove ratings were likewise averaged, resulting in a single mean rating per stimulus. We model the relationship between features and ratings using ridge regression. The number of fuzzy gain categories K was treated as a hyperparameter and evaluated separately. For each value of K, model performance was estimated using a nested 5-fold cross-validation procedure. In each outer fold, a subset of stimuli was held out as test data. The regularization parameter λ was selected exclusively on the corresponding training data using an inner 5-fold cross-validation. A ridge regression model was then trained with the selected λ and evaluated on the held-out test fold. Model performance was quantified using the coefficient of determination:

$$R^2 = 1 - \frac{\sum_n (y_n - \hat{y}_n)^2}{\sum_n (y_n - \bar{y})^2} \tag{7}$$

For each outer fold, R^2 was computed on both the training and test subsets, yielding R^2_{train} and R^2_{test}. Mean values and standard deviations were then computed across folds. Following cross-validation, a final model was fitted on the complete dataset. The regularization parameter λ for this final model was determined through 5-fold cross-validation on the full dataset. To further characterize prediction accuracy, the maximum absolute error M was also computed:

$$M = \max_n |y_n - \hat{y}_n| \tag{8}$$

This measure was calculated for both the training and test subsets within each outer cross-validation fold, yielding M_{train} and M_{test}. Mean values and standard deviations across folds were reported.

5. Results and discussion

5.1. Model comparisons

Model performance for both feature representations is summarized in Fig. 6. Across the full range of K, the transition-based model consistently outperformed the distribution-based model regarding all scores. Regarding mean R^2_{test}, the transition-based model achieved its best performance for $K = 5$ ($R^2_{\text{test}} = 0.49 \pm 0.05$). While the highest mean R^2_{test} was obtained for $K = 5$, the lowest mean M_{test} was observed for $K = 8$ (0.71 ± 0.09), suggesting a trade-off between overall predictive accuracy and worst-case prediction error. Since the dimensionality of the transition-based representation grows as $(K + 1)^2$, we selected $K = 5$ as

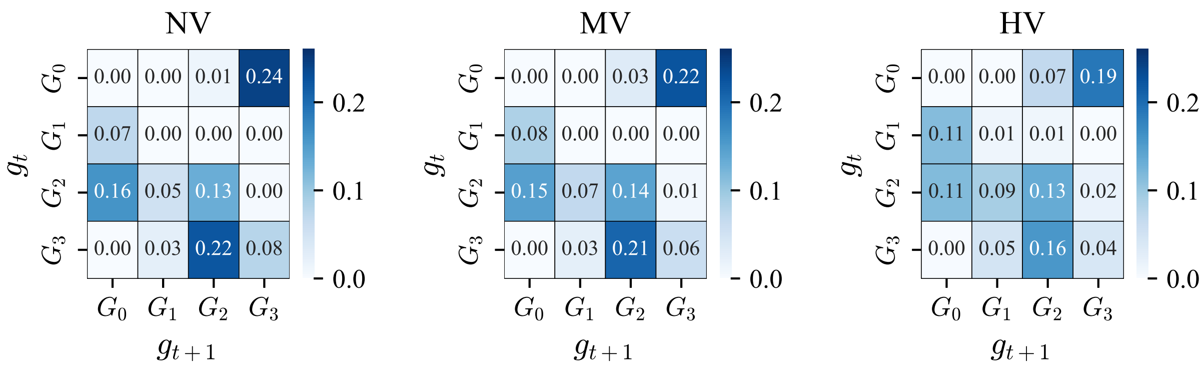

Figure 5: Examples of transition-based representation for the dynamic pattern $\mathbf{d} = \texttt{0321}$ in various probability distribution conditions. The left figure is the no-variation condition (NV), the middle figure is the mid-variance condition (MV) and the right figure is the high-variance (HV). Each cell represents the relative prevalence of transitions between fuzzy gain categories at time t (y-axis) and time $t+1$ (x-axis).

a compromise between predictive performance and model complexity. Moreover, the superior performance of the transition-based representation suggests that temporal dependencies between successive events play a central role in groove perception. This finding is consistent with predictive coding theories of perception, which posit that the brain continuously generates and updates predictions about incoming sensory input based on its temporal structure [8]. In this framework, sensitivity to transitions between sound levels may reflect the encoding of predictive relationships, whereby groove emerges from the alignment between expected and actual dynamic patterns over time.

5.2. Generative exploration of sound dynamics and implications for dynamic compression

Because the proposed model predicts groove ratings from sound dynamics, it can be used not only as an explanatory tool but also in a generative framework to explore how different dynamic configurations shape groove. To illustrate this approach, we focus on one representative binary pattern ($\mathbf{b} = \texttt{0111}$), for which a large set of admissible dynamic patterns was generated. These dynamic patterns were constructed under two constraints: RMS normalization at 0 dB and a minimum gain value of -25 dB for all non-silent events. For each admissible dynamic pattern, the model first estimated the predicted groove of the original, uncompressed pattern. Dynamic compression was then applied to each pattern using a static hard-knee compressor. For a gain value x (in dB) exceeding a threshold T, the compressed gain y was computed as:

$$y = \begin{cases} x & \text{if } x \leq T \\ T + \dfrac{x - T}{\rho} & \text{if } x > T \end{cases} \tag{9}$$

where ρ denotes the compression ratio. For each dynamic pattern, several combinations of threshold and ratio values were tested. After each compression setting, the resulting pattern was RMS-renormalized before being passed back through the model. For each dynamic pattern, the maximum predicted groove value across all tested compression settings was retained, yielding the optimal post-compression groove estimate. As shown in Fig. 7, the predicted groove values prior to compression vary substantially across dynamic patterns, reflecting the sensitivity of groove perception to sound dynamics. Crucially, applying optimal compression led to significant increases in predicted groove for a large proportion of these configurations. This suggests that compression does not merely homogenize dynamic contrasts: it can reshape the transition structure of gain values in ways that align more closely with perceptually preferred dynamic profiles. These findings carry practical implications for live music mixing choice. Audio engineers typically apply dynamic range compression to control peak levels or ensure loudness consistency. The present results suggest an additional, perceptually motivated use case: compression as a tool for enhancing groove.

6. Conclusion

Results demonstrate that dynamic pattern significantly affects groove perception, and that the temporal ordering of gain values is a stronger predictor of groove than the overall distribution of sound levels. A ridge regression model trained on these features achieved moderate predictive accuracy, providing evidence that groove perception can, to a meaningful extent, be predicted from sound dynamics. Beyond its explanatory value, the proposed model was used generatively to explore how dynamic range compression modifies the predicted groove of drum patterns. The results suggest that compression, when applied with appropriate dynamic compression settings, can substantially increase perceived groove. This points to a novel, perceptually grounded perspective on compression in music production: rather than treating it solely as a tool for

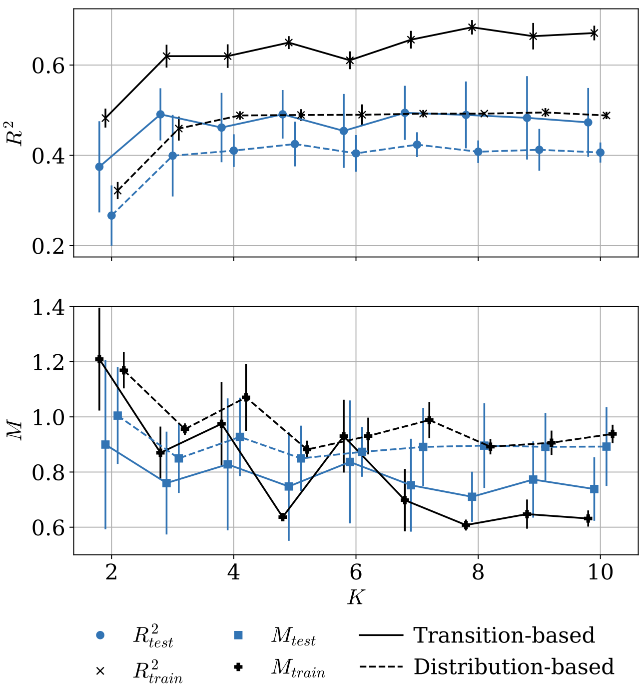

Figure 6: Comparisons of ridge regression models performance for distribution-based and transition-based representations for different number of fuzzy membership functions K. Top figure represents R^2 while bottom figure represents M scores.

loudness control or dynamic smoothing, practitioners may consider optimizing compression parameters with respect to groove perception. Several limitations of the present work should be acknowledged. The stimuli were restricted to isolated snare-drum sequences, and it remains to be determined whether the findings generalize to full drum patterns or multi-instrument contexts. Future work should validate the generative compression analysis through dedicated perceptual experiments, and extend the modelling framework to richer rhythmic and timbral contexts.

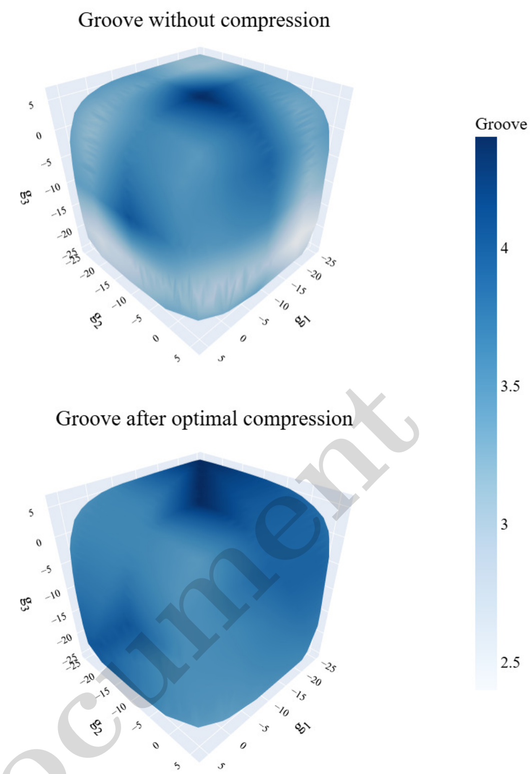

Figure 7: Groove predictions for $\mathbf{b} = \texttt{0111}$ in different dynamic patterns $\mathbf{d} = 0g_1g_2g_3$ with g_1,g_2,g_3 in dB, and its modification by optimal compression settings. The different gain values are organized along curved manifolds which are imposed by the RMS normalization constraint. Top figure shows groove prediction without dynamic compression. Bottom figure shows how groove change after applying optimal dynamic compression settings.

References

[1] G. Madison, "Experiencing Groove Induced by Music: Consistency and Phenomenology," en, *Music Perception*, vol. 24, no. 2, pp. 201–208, Dec. 2006.

[2] M. A. G. Witek, E. F. Clarke, M. Wallentin, M. L. Kringelbach, and P. Vuust, "Syncopation, Body-Movement and Pleasure in Groove Music," en, *PLoS ONE*, vol. 9, no. 4, R. Canal-Bruland, Ed., e94446, Apr. 2014.

[3] O. Senn, L. Kilchenmann, R. Von Georgi, and C. Bullerjahn, "The Effect of Expert Performance Microtiming on Listeners' Experience of Groove in Swing or Funk Music," en, *Frontiers in Psychology*, vol. 7, Oct. 2016.

[4] T. Etani, A. Marui, S. Kawase, and P. E. Keller, "Optimal Tempo for Groove: Its Relation to Directions of Body Movement and Japanese nori," en, *Frontiers in Psychology*, vol. 9, p. 462, Apr. 2018.

[5] A. Danielsen, M. Johansson, R. Brøvig, B. Sandvik, and K. K. Bøhler, "Shaping rhythm: Timing and sound in five groove-based genres," en, *Popular Music*, vol. 42, no. 1, pp. 20–41, Feb. 2023.

[6] F. L. Bouwer, J. A. Burgoyne, D. Odijk, H. Honing, and J. A. Grahn, "What makes a rhythm complex? The influence of musical training and accent type on beat perception," en, *PLOS ONE*, vol. 13, no. 1, S. Kotz, Ed., e0190322, Jan. 2018.

[7] D. Ronan, H. Gunes, and J. D. Reiss, "Analysis of the subgrouping practices of professional mix engineers," en, 2017.

[8] P. Vuust, O. A. Heggli, K. J. Friston, and M. L. Kringelbach, "Music in the brain," en, *Nature Reviews Neuroscience*, vol. 23, no. 5, pp. 287–305, May 2022.