SPAR-RT: EVOLVING REPRESENTATIONS OF MUSICAL TIMBRE, PITCHES, AND TUNING IN REAL-TIME USING SYMMETRIC PROJECTION ATTRACTOR RECONSTRUCTION

Wentao Hao, Merlin Chao, Mateusz Soliński, Elaine Chew
Department of Engineering, Faculty of Natural, Mathematical & Engineering Sciences
School of Biomedical Engineering & Imaging Sciences, Faculty of Life Sciences & Medicine
King's College London, London, UK
{wentao.hao,merlin.chao,mateusz.solinski,elaine.chew}@kcl.ac.uk

ABSTRACT

The Symmetric Projection Attractor Reconstruction (SPAR) method was proposed to visualise and analyse nonlinear characteristics of physiological signals. It has recently been applied to music signals to visualise instrument timbre. Here we evaluate its use for characterising pitch sets, timbre, and tuning in music signals, and create a real-time visualisation tool based on SPAR to analyse live music. We show how SPAR integrates phase space reconstruction with the Discrete Fourier Transform (DFT) to generate attractors whose shapes reflect the signal's dynamic structures. Using synthetic sounds, the attractors are shown to quantify the interference of different harmonics, music intervals, chords or tuning systems. The aliasing effect is found to significantly affect the shape of the attractor, and perceptual properties like dissonance and commensurability are linked to the attractor's entropy. Using sampled audio in the NSynth dataset, the attractors discriminate between timbres through geometric density patterns over a wide range of pitches. Cluster analysis achieved perfect classification accuracy between flute and brass instruments. The real-time SPAR visualisation enables interactive analysis and visualisation of pitch, timbral, and tuning characteristics in live music sounds evolving dynamically over time, offering novel ways to represent and process music audio. Applications include music signal analysis and learning, sound and instrument design, and live multimedia performance.

1. INTRODUCTION

Music signals exhibit rich acoustic characteristics such as timbre, pitch, and tuning variations, which are encoded in the harmonic structure, amplitude envelope, and waveform morphology. The analysis of these elements is mainly based on traditional time domain or frequency-based representations such as spectrograms and Mel-frequency cepstral coefficients (MFCCs). However, these methods may fail to capture morphological trajectories or nonlinear phenomena that give rise to perceptual differences across musical signals. To address such limitations, recent studies of nonlinear dynamics have provided promising approaches such as the Variable Markov Oracle(VMO) to analysing complex audio signals beyond traditional methods using low-level acoustic representations [1]. Applications such as pitch perception [2] and affect recognition in speech [3] have explored the use of nonlinear dynamic models. One approach is phase space reconstruction, grounded in Takens' embedding theorem, which provides a geometric framework for reconstructing dynamical states in time series by embedding them into a higher-dimensional phase space. An earlier study [4] compared phase space reconstruction analysis to MFCC features in speech recognition, suggesting its potential in audio processing. However, there is a lack of tools in the literature for systematic analysis of the influence of various signal features on its shape and phase space characteristics.

In this study, we developed an implementation of the Symmetric Projection Attractor Reconstruction (SPAR), called SPAR-RT, to represent a signal's phase space as attractors in two-dimensions. The SPAR method was initially introduced to analyse physiological signals [5], then applied to musical signals by Solinski et al. [6] for the visualisation of instrument timbres. SPAR-RT offers animated visualisations of attractors generated from uploaded music files or real-time audio captured via a microphone. We use this framework to create attractors based on synthetic and real acoustic sounds to investigate how musical features such as timbre, pitch, and tuning shape dynamical and nonlinear characteristics of sound signals.

The remainder of the paper is organised as follows: Section 2 introduces the theoretical background and principles of the SPAR method; Section 3 describes the datasets and experiment design; Section 4 presents the results including both quantitative measurements and visualisations of the SPAR attractors; Section 5 discusses the findings in relation to the theoretical framework; Section 6 concludes the paper and outlines the future work.

2. METHODS

This section details the SPAR methodology, covering the mathematical framework, quantification measures, and se-

lection of parameter used for generating attractors, as well as the real-time implementation (SPAR-RT) developed for interactive analysis.

2.1 Symmetric Projection Attractor Reconstruction

We adopt the SPAR method as a new representation of audio signals. The introduction and mathematical interpretation of the SPAR method are described in [7]. The SPAR method embeds time series into N-dimensional phase space delay coordinates with a specified time delay $\tau = T/N$ (T is the average period of the raw signal) and projects the results onto a 2D representation of an attractor. The attractor's trajectories capture the local waveform morphology and variability between the signal's cycles.

Given an audio signal $x(t)$ with average period T, the (N, k) attractor is defined as the k-th projection of an N-dimensional phase space representation of the signal $x(t)$. The delay vector is defined as $x_{N,j}(t) = x(t - j\tau), j = 0, ..., N-1$. The coordinates of the (N, k) attractor $a(t)$ and $b(t)$ are as follows:

$$a_{N,k}(t) = \mathrm{Re}\left\{\frac{1}{\sqrt{N}}\sum_{j=0}^{N-1} e^{-i\frac{2\pi jk}{N}} x_{N,j}(t)\right\}, \quad (1)$$

$$b_{N,k}(t) = \mathrm{Im}\left\{\frac{1}{\sqrt{N}}\sum_{j=0}^{N-1} e^{-i\frac{2\pi jk}{N}} x_{N,j}(t)\right\}. \quad (2)$$

2.2 Analysis and Parameter Selection

Next, we describe the quantitative measures used to analyse the generated attractors and detail the approach for selecting proper parameters N and k of the SPAR method.

2.2.1 Quantification and Analysis of Attractors

The attractors can be quantified and analysed by four measures based on polar coordinates:

- **r density:** the density distribution in function of the distance (radius) from the attractor's centre;
- θ **density:** the density distribution with angle $(0 - 2\pi)$;
- **outline:** the maximum r of an attractor for a given θ angle; and,
- **joint entropy:** the joint entropy distribution over the (r, θ) space.

2.2.2 Selection of Parameters

The coordinates of attractors $(a_{N,k}(t), b_{N,k}(t))$ are the real and imaginary parts of the k-th coefficients of the Discrete Fourier Transform for the vector $x_{N,j}(t)$. This fundamental linkage to spectral representations of signals guides the selection of the attractor's embedding dimension N and projection index k.

The SPAR method defines $\lfloor \frac{N-1}{2} \rfloor$ distinct projections for a given N based on the properties of the circulant matrix. Consequently, the selection of projection index is $k \in \{1, \ldots, \lfloor \frac{N-1}{2} \rfloor\}$.

The average period T determines the frequency resolution of the SPAR method as $1/T$. The frequency of th k-th projection is consequently given by $f_k = k/T$. For extracted N points, the effective sampling rate is $f_{eff} = N/T$, so the Nyquist frequency is:

$$f_{Nyquist} = \frac{N}{2T}. \quad (3)$$

According to the Nyquist–Shannon sampling theorem, any signal component with a frequency higher than $f_{Nyquist}$ cannot be independently resolved by the DFT system of N points, which results in aliasing. Therefore, the selection of N is crucial to a proper analysis using SPAR as it relates to the maximum number of harmonics that can be represented without aliasing. However, increasing N indefinitely does not always yield better results. It may introduce unnecessary computational effort and complexity in the analysis.

2.3 Real-time Implementation of SPAR for Music

We present a dedicated implementation of the SPAR method referred to as SPAR-RT as shown in Figure 1.

Figure 1: The user interface of the SPAR-RT tool

Built using Python, SPAR-RT is designed to enable the interactive analysis and intuitive observation with two modes:

- **Real-time:** The tool captures the surrounding sound and generates corresponding real-time attractors.
- **Animation generation:** This mode allows users to upload audio files and export corresponding attractor animation videos synchronised with the audio.

Both operational modes support the adjustment of key SPAR parameters (N, k) and frame settings (window and hop size) to optimise the visualisation for specific analysis goals. The user may also choose to view the visualisations on a white or dark background.

Example videos of visualisations demonstrating the dynamic visualisation capabilities of SPAR-RT are provided in the Supplementary Material.

3. DATA

This section describes the synthetic and real audio datasets used to investigate how musical characteristics like timbre, pitches and tuning are represented by the SPAR method.

3.1 Synthetic Sinusoid Harmonic Waves

We generated synthetic sinusoid harmonic waves by combining sinusoidal waveforms having different frequencies, where each frequency was an integer multiple of the fundamental frequency. The mathematical formulation of the synthetic sinusoid harmonic wave is as follows:

$$x(t) = \sum_{m=1}^{M} A_m \cdot \sin(2\pi m f_0 t + \phi_m), \tag{4}$$

where f_0 is the fundamental frequency, and A_m and ϕ_m are the amplitude and phase of the m-th harmonic, respectively.

To explore how different harmonic components impact the shape of attractors, we used two settings of composite sinusoidal signals, containing three and six harmonic components, respectively. Different embedding dimensions ($N = 5, 10$) were selected for attractor generation. The fundamental frequency was set to $f_0 = 440$ Hz and the average period T was $1/f_0$. Signals were generated at the sampling rate (fs) of 44000 Hz, ensuring that $T * fs/N$ was an integer. This prevented misalignment between the sampling rate and the SPAR delay ($\tau = T/N$) caused by rounding errors.

3.2 Intervals

Using SuperCollider, we synthesise all dyads in the chromatic scale as a pair of pure sine waves. For each dyad, we generate its attractor with parameters $N = 4$ and $k = 1$ to avoid aliasing. We then compute the joint entropy and plot it alongside a dissonance curve modelled on pure sine wave interactions, as proposed in [8].

3.3 Tuning and Chords

To evaluate how different tuning systems are represented by SPAR, we synthesised both C major and C minor chords using pure sine waves. For each chord, one version was tuned according to the just intonation tuning system, in which pitch relationships follow simple integer ratios (e.g., $5:4$ for a major third, $6:5$ for a minor third), and another in equal temperament system, where frequency ratios are defined by dividing the octave into twelve equal parts on a logarithmic scale. Table 1 provides an example of the frequency ratios in each tuning system for a major triad and a minor triad, and Table 2 gives the actual frequencies for a C major triad and a C minor triad.

Chord	Just Int. Ratios	Equal Temp. Ratios
major triad	$\left[1,\ \frac{5}{4},\ \frac{3}{2}\right]$	$\left[1,\ 2^{4/12},\ 2^{7/12}\right]$
minor triad	$\left[1,\ \frac{6}{5},\ \frac{3}{2}\right]$	$\left[1,\ 2^{3/12},\ 2^{7/12}\right]$

Table 1: Frequency ratios of a major and a minor triad in Just Intonation and Equal Temperament tuning systems

Pitches	Just Int. Freq. (Hz)	Eq. Temp. Freq. (Hz)
C, E, G	[261.7, 327.0, 392.4]	[261.7, 329.6, 392.0]
C, E♭, G	[261.7, 314.0, 392.4]	[261.7, 311.1, 392.0]

Table 2: Rounded frequencies of generated C major and C minor chords in Just Intonation and Equal Temperament tuning systems (base note: C = 261.7 Hz)

3.4 Real Acoustic Instrument Audio

We applied the SPAR method to real acoustic signals to show differences in generated attractors regarding the timbre of the instruments. We used 22 samples of flute sounds and a brass instrument (44 samples in total) from the NSynth database by Google Inc [9]. Each audio sample consists of a single note with a fixed pitch, ranging from G4 to E6 across samples, and a duration of 4 seconds.

We applied the SPAR method to these samples using two embedding dimensions, $N = 5$ and $N = 10$. The generated attractors were compared and analysed in terms of both pitch and instrument timbre. To quantify attractors, we extracted three measure profiles (r density, θ density, and outline) and concatenated them into a single profile vector as the numerical representation for each attractor.

Finally, we clustered the vectors using hierarchical clustering, with the results demonstrated in a dendrogram.

4. RESULTS

This section presents the results for both synthetic and real acoustic audio signals, demonstrating how the SPAR method effectively characterises and distinguishes musical features such as harmonics, dissonance, tuning, pitch and timbre.

4.1 Synthetic Sinusoid Harmonic Waves

Figure 2 shows the attractors generated from a sinusoidal harmonic signal consisting of three components based on different dimensions N and projections k. We observed:

When $N = 5$, $k \in \{1, 2\}$, $f_{Nyquist} = 1100$ Hz, and

- $k = 1$ (440 Hz), the attractor appears as a perfect circle.
- $k = 2$ (880 Hz), the attractor is a star-shaped structure with a five-fold rotational symmetry due to aliasing from the harmonic with 1320 Hz.

When $N = 10$, $k \in \{1, 2, 3\}$, $f_{Nyquist} = 2200$ Hz,

- All three attractors show perfect circle shapes, which indicates that three sinusoidal components have been fully resolved by the SPAR with a high dimension $N = 10$.

Figure 3 illustrates the attractors for the sinusoidal harmonic signals consisting of six components based on different dimensions N and projections k.

When $N = 5$, $k \in \{1, 2\}$, $f_{Nyquist} = 1100$ Hz, and

- $k = 1$ (440 Hz), the attractor exhibits a unique petal-like structure with five-fold rotational symmetry. Five

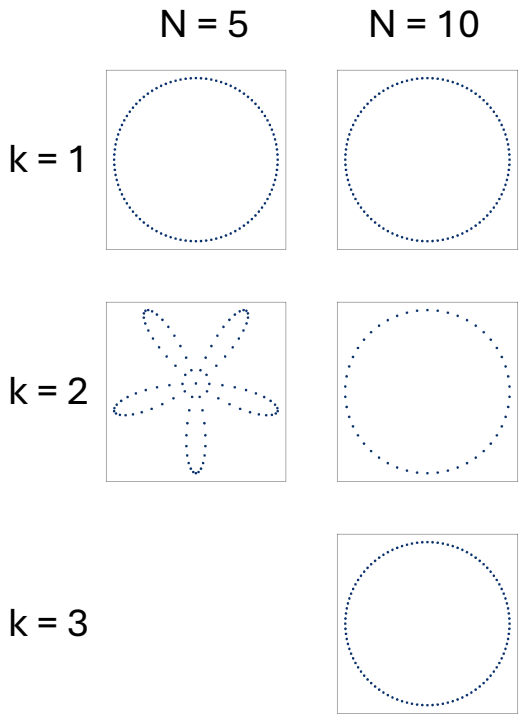


Figure 2: The shapes of attractors for audio signals consisting of three sinusoid functions with $N = 5, 10$ and $k = 1, 2, 3$. The non-circular attractor $(5, 2)$ is a manifestation of aliasing from higher harmonics.

distinct petals are symmetrically distributed around the centre point.

- $k = 2$ (880 Hz), the attractor presents a star-like structure with five-fold rotational symmetry.

When $N = 10$, $k \in \{1, 2, 3, 4\}$, $f_{Nyquist} = 2200$ Hz, and

- $k = 1, 2, 3$ (440 Hz, 880 Hz, 1320 Hz), the attractors show the perfect circles, indicating sinusoidal components are fully represented.
- $k = 4$ (1760 Hz), the attractor presents a star-like shape with five-fold rotational symmetry.

The results suggest that for the m-th sinusoidal component with a frequency $f_m > f_{Nyquist}$ we observe aliasing at $f_{alias} = |f_m - f_{eff}| = |f_m - 2 * f_{Nyquist}|$. The corresponding projection k of the attractor for resolving f_{alias} exhibits rotational symmetry.

4.2 Intervals

The shape of the attractor varies noticeably across dyads in the chromatic scale. We observe that its structure is primarily determined by the frequency ratio between the two tones, influenced by two key factors: beating and commensurability (the existence of a shared multiple among frequency periods). Both factors are known to impact auditory perception [10], and their effects may emerge at different time scales, either within a single analysis frame or across multiple frames. To illustrate effects that unfold

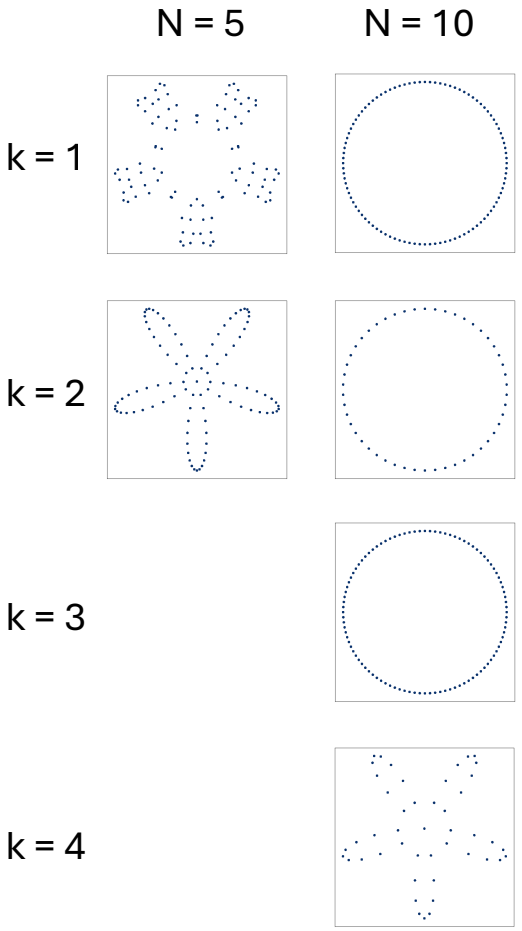
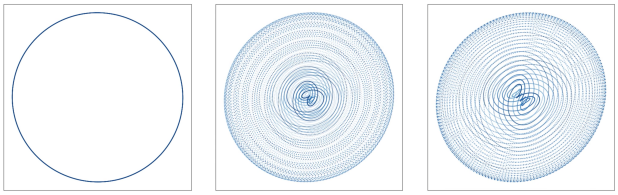


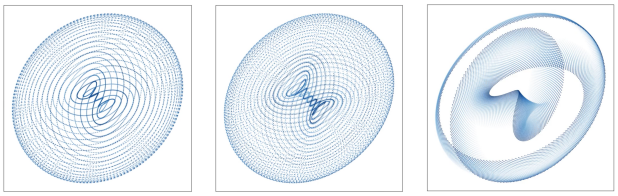
Figure 3: The shapes of attractors for audio signals consisting of six sinusoid functions with $N = 5, 10$ and $k = 1, 2, 3, 4$. The non-circular attractors $(5, 1)$, $(5, 2)$ and $(10, 4)$ are manifestations of aliasing from higher harmonics.

over time, a video showing the attractors for all dyads is provided in the Supplementary Material (Supplementary Animation 1). As expected, the attractor derived from the unison and octave forms a perfect circle, reflecting the fact that the sound frequencies were either identical or related by a factor of two, respectively. Notably, the intervals of the fourth and the fifth (see Figures 4f and 4h) produce attractors with shapes that are similar one to another but distinct from the shapes of other intervals.

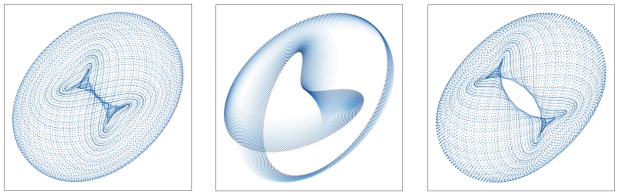
We compared the dissonance curve for two pure sine waves [8] with the joint entropy of the attractor; both measures show similar patterns as shown in Figure 5. The lowest values for both measures appeared at the unison and the octave (12 semitones) intervals, considered the most consonant. From the unison to the minor second, we observe a steep rise toward a maximum level for both entropy and dissonance, followed by a gradual, monotonic decrease toward the octave. This comparison shows an existing relationship between the SPAR representation and aspects of sound perception.



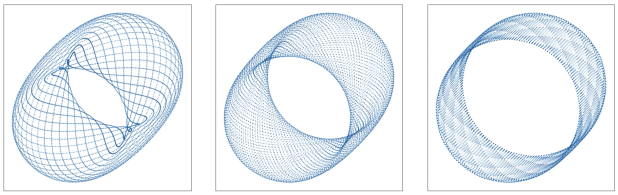
(a) Unison / Octave (b) Minor Second (c) Major Second



(d) Minor Third (e) Major Third (f) Fourth



(g) Tritone (h) Fifth (i) Minor Sixth



(j) Major Sixth (k) Minor Seventh (l) Major Seventh

Figure 4: Attractors for all dyads in the chromatic scale for N=4 and k=1

4.3 Tuning and Chords

The attractors exhibit significant differences between 12-tone equal temperament and just intonation tuning systems. This shows that the attractor's shape is highly sensitive to the commensurability of the signal's frequency components. Commensurability refers to how closely the individual signals' periods share a common multiple. To determine this common multiple, calculate the least common multiple (LCM) of the denominators of the frequency ratios [11]. For example, the major triad in just intonation, with frequency ratios $1 : \frac{5}{4} : \frac{3}{2}$, realigns every $\mathrm{LCM}(4, 2) = 4$ cycles of the fundamental frequency. The minor triad, with ratios $1 : \frac{6}{5} : \frac{3}{2}$, realigns every $\mathrm{LCM}(5, 2) = 10$ cycles, indicating a longer but still periodic alignment. These simple integer ratios produce bounded, low-entropy attractors, with the minor chord's attractor showing a slightly more intricate pattern (see Figures 6a and 6b). In contrast, equal temperament uses irrational frequency ratios that approximate just intervals logarithmically. These incommensurable components contribute to more complex attractor shapes with chaotic behaviour (without strict periodic loops; see Figures 6c and 6d).

Although the relationship between tuning systems and perceptual roughness remains debated, previous studies have linked commensurability to spectral interference and roughness perception [12], suggesting another connection between the attractor's shape and another aspect of sound perception.

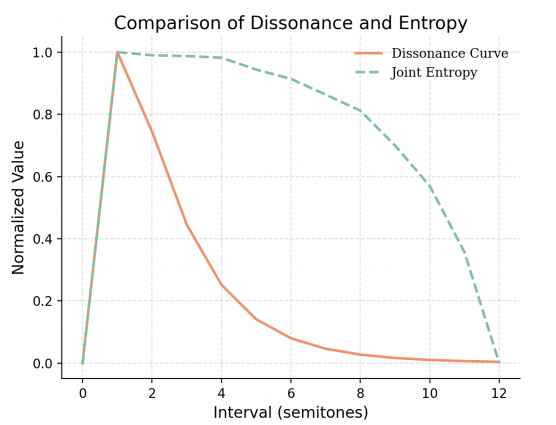
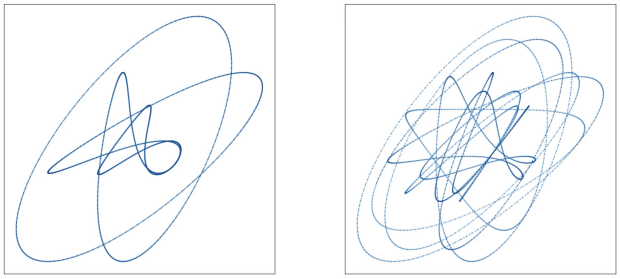


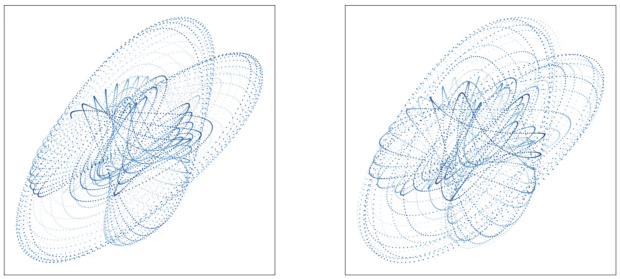
Figure 5: Normalised dissonance and joint entropy for pure sine dyads of the chromatic scale.

4.4 Real Acoustic Instrument Audio

Figure 7 shows the 5-dimensional attractors ($N = 5$) extracted from the flute sound with pitches A4 (440.0 Hz) and C#6 (1108.73 Hz) and their corresponding power spectra. For A4, the power spectrum contains multiple harmonic frequency components, and the corresponding attractor exhibits a non-circular five-fold rotational symme-



(a) C Major triad in just intonation tuning system (b) C Minor triad in just intonation tuning system



(c) C Major triad in equal temperament tuning system (d) C Minor triad in equal temperament tuning system

Figure 6: Attractors of C Major and Minor triads with $N = 3$ and $k = 1$ in just intonation and equal temperament tuning systems.

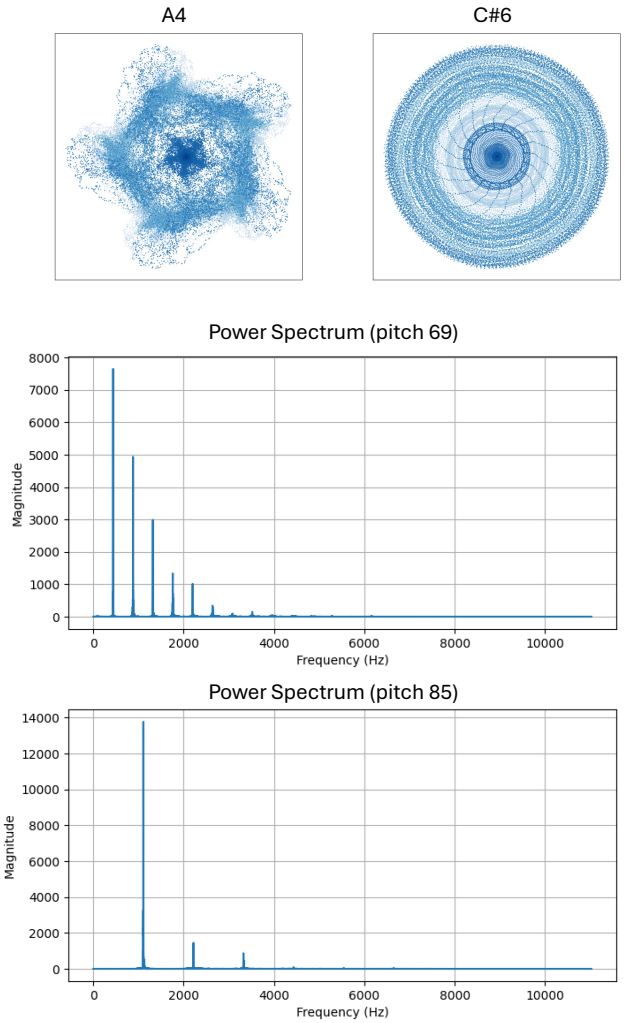


Figure 7: The attractors from the flute sounds on different pitches (A4 and C#6) and corresponding power spectra.

try. For pitch 85, the attractor has a circular outline, and the spectrum shows fewer harmonic components with a significantly dominant main fundamental frequency. A dynamic visualisation illustrating how the shape of the 5-dimensional attractor evolves as the pitch increases is provided in the Supplementary Material (Supplementary Animation 2)

Figure 8 shows the comparison between the 10-dimensional attractors extracted from a flute and a brass instrument at the same pitch of F#5 and their corresponding attractor measure profiles. Both attractors are circle-shaped, indicating that the selection of the dimension $N = 10$ resolved the audio signal well without significant aliasing. However, they are distinct in density distribution. The attractor representing a brass instrument reveals a higher density concentration in the centre and in the ring with a radius of approximately 0.15, compared to the attractor extracted from the sound of a flute. For the θ density profile, we observe an oscillatory-like pattern for the attractor linked with a brass instrument, while the profile for the flute is distinctly flatter. The outline profiles for flute and brass exhibit distinct geometric patterns across the angular domain. The flute profile shows higher and more fluctuating maxi-

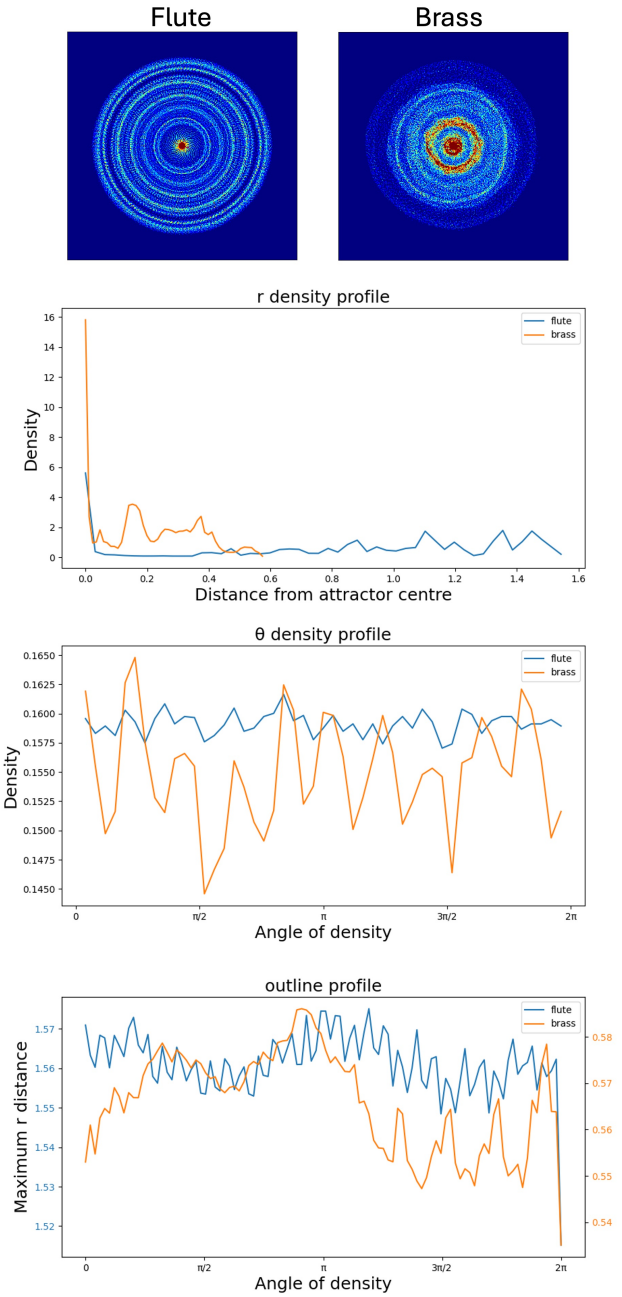


Figure 8: The attractors of flute and brass with $N = 10$, $k = 1$, pitch $= 78$, along with the corresponding profiles of the three attractor measures. Here, the attractors are shown as heat maps depicting the density of points in phase space used to calculate the density profiles.

mum r distances. In contrast, the brass profile is generally lower and more variable, suggesting a less symmetrical or more irregular attractor geometry.

Figure 9 illustrates the dendrogram generated from the hierarchical clustering of concatenated profile vectors. Using a distance threshold of 50, two distinct clusters emerge: one composed entirely of brass instruments and the other exclusively of flutes.

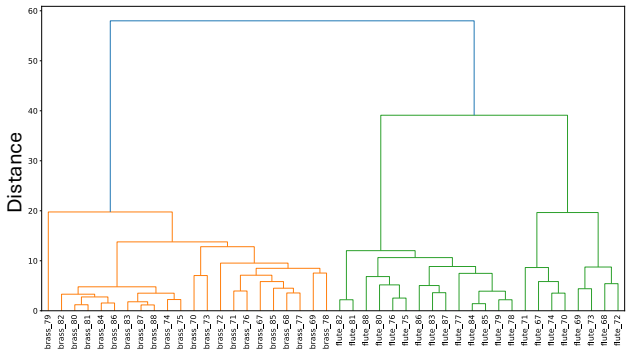


Figure 9: The dendrogram created from the distances between different 10-dimensional attractors.

5. DISCUSSION

In this work, we have applied the SPAR method to characterise dynamic changes in the synthetic and acoustic audio signals and presented the differences between attractors regarding timbre, pitch (including pitch sets) and tuning.

The analysis using harmonic sinusoid signals revealed the occurrence of aliasing, which deforms the complete reconstruction of the phase space. We showed a close relation between the embedding dimension N, projection index k, and the signals' frequency harmonics. These findings deepen our understanding of SPAR visualisation capabilities and enhance the interpretability of the geometric patterns, offering guidelines for parameter selection in future applications.

The differences between attractors for distinct dyads demonstrated the phenomenon of frequency intervals resulting in differential projections of phase space. Specifically, we observed that more dissonant intervals, typically associated with prominent beatings, correspond to attractors with higher joint entropy. This suggests a close relationship between the shape of the attractor and perceptual phenomena such as dissonance and beatings. The comparison between just intonation and equal temperament demonstrated that the attractor is sensitive to the commensurability of frequency components, which is also a measure of dissonance and roughness. These results highlight the potential of the SPAR method as a tool for linking signal structure to auditory perception.

The analysis of real acoustic signals confirmed the presence of a direct link between signal harmonic contents and attractor morphology, which is consistent with the observations for the synthetic harmonic sinusoids. The SPAR method was also able to distinguish between instrument families using geometric density profiles, even when the overall attractor shapes appear visually similar and circular. These examples demonstrated that SPAR can discriminate betwee subtle dynamics and inharmonicity within real acoustic audio signals, providing a powerful means to differentiate between sonic timbres. The classification using hierarchical clustering of the concatenated SPAR features suggests that the SPAR method may extract instrument-specific characteristics that complement traditional spectral features, thereby enhancing automatic instrument recognition. Future research involving supervised machine learning techniques could explore the effectiveness of this approach across a broader range of instruments and pitches.

This study offers a foundational analysis of the SPAR method as a music signal representation. There are some limitations to the approach. The analysis on synthetic data was conducted under controlled conditions, specifically by ensuring the sampling frequency was perfectly aligned with the SPAR time delay and dimension N. While the validation on a real acoustic dataset matched the findings on synthetic data, this perfect alignment may not occur in a real-world scenario, which may introduce distortions in the trajectories of attractors. Therefore, future work is needed to more thoroughly analyse how to quantify and interpret this real-world shift, which would enhance the SPAR method's robustness for practical applications.

The development of the real-time visualisation and animation generation tool provides a promising option for such future exploration, enabling interactive analysis and intuitive observation of evolving acoustic features over time, for instance in performance contexts where visualisation of attractors provides expressive feedback for sound and instrument designers, composers, performers, and the audience.

6. CONCLUSIONS

This study presents an analysis of the Symmetric Projection Attractor Reconstruction (SPAR) method as a novel representation for music signal analysis, focusing on its ability to encode musical characteristics such as timbre, pitch, and tuning. Through a series of simulated and real-world experiments, we demonstrate that SPAR attractors offer interpretable geometric representations that encode meaningful signal dynamics.

Our harmonic simulations reveal how attractor morphology reflects frequency content and aliasing effects, while analyses of musical intervals and tuning systems highlight SPAR's sensitivity to perceptual properties like dissonance and commensurability. The real acoustic music audio study further validates the method's ability to distinguish between instrument timbres through subtle geometric density patterns, achieving perfect classification in our cluster analysis. These results underscore SPAR's potential as a powerful and perceptually relevant tool for music signal analysis.

The real-time SPAR visualisation and animation-maker tools support not only analyses and learning but also creative applications in sound and instrument design, and live and multimedia performance. It will also support future explorations, enabling interactive analysis and development of new dynamic acoustic features. Future work will focus on further analyses of the discrepancies between controlled simulations and real-world scenarios, and on improving the robustness and interpretability of SPAR.

Acknowledgments

This result is part of the COSMOS project (`https://cosmos.isd.kcl.ac.uk`), which has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation

programme (Grant agreement No. 788960, https://cordis.europa.eu/project/id/788960). WH is supported by a King's College London China Scholarship Council (K-CSC) Scholarship.

7. REFERENCES

[1] P. Mouawad and S. Dubnov, "On modeling affect in audio with non-linear symbolic dynamics," *Advances in Science, Technology and Engineering Systems Journal*, vol. 2, no. 3, pp. 1727–1740, 2017.

[2] J. H. E. Cartwright, D. L. González, and O. Piro, "Pitch perception: A dynamical-systems perspective," *Proceedings of the National Academy of Sciences*, vol. 98, no. 9, pp. 4855–4859, 2001.

[3] A. Harimi, A. AhmadyFard, A. Shahzadi, and K. Yaghmaie, "Anger or joy? emotion recognition using nonlinear dynamics of speech," *Applied Artificial Intelligence*, vol. 29, no. 7, pp. 675–696, 2015.

[4] A. Lindgren, M. Johnson, and R. Povinelli, "Speech recognition using reconstructed phase space features," in *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03).*, vol. 1, 2003, pp. I–60.

[5] M. Nandi and P. J. Aston, "Extracting new information from old waveforms: Symmetric projection attractor reconstruction: Where maths meets medicine," *Experimental Physiology*, vol. 105, no. 9, p. 1444–1451, May 2020. [Online]. Available: http://dx.doi.org/10.1113/EP087873

[6] M. Solinski, W. Hao, and E. Chew, "Spar-timbre: Unlocking non-linear dynamics in music audio with symmetric projection attractor reconstruction," in *17th International Symposium on Computer Music Multidisciplinary Research*, 2025, accepted.

[7] J. V. Lyle and P. J. Aston, "Symmetric projection attractor reconstruction: Embedding in higher dimensions," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 31, p. 113135, 2021.

[8] W. A. Sethares, *Tuning, Timbre, Spectrum, Scale*, 2nd ed. London; New York: Springer Science & Business Media, 2005.

[9] J. Engel, C. Resnick, A. Roberts, S. Dieleman, D. Eck, K. Simonyan, and M. Norouzi, "Neural audio synthesis of musical notes with wavenet autoencoders," 2017.

[10] J. R. Pierce, "Periodicity and pitch perception," *The Journal of the Acoustical Society of America*, vol. 90, no. 4 Pt 1, pp. 1889–1893, Oct. 1991.

[11] F. Stolzenburg, "Harmony perception by periodicity detection," *Journal of Mathematics and Music*, vol. 9, no. 3, p. 215–238, Aug. 2015. [Online]. Available: http://dx.doi.org/10.1080/17459737.2015.1033024

[12] P. N. Vassilakis, "Perceptual and physical properties of amplitude fluctuation and their musical significance," Doctoral Dissertation, University of California, Los Angeles, 2001.