COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 47 (2023) e13363

ISSN: 1551-6709 online
DOI: 10.1111/cogs.13363

Recurrence Quantification Analysis of Crowd Sound Dynamics

Shannon Proksch,[a,b] Majerle Reeves,[c] Kent Gee,[d] Mark Transtrum,[d] Chris Kello,[b] Ramesh Balasubramaniam[b]

[a]*Department of Psychology, Augustana University*
[b]*Department of Cognitive and Information Sciences, University of California, Merced*
[c]*Department of Applied Mathematics, University of California, Merced*
[d]*Department of Physics and Astronomy, Brigham Young University*

Received 4 November 2022; received in revised form 17 August 2023; accepted 22 September 2023

Abstract

When multiple individuals interact in a conversation or as part of a large crowd, emergent structures and dynamics arise that are behavioral properties of the interacting group rather than of any individual member of that group. Recent work using traditional signal processing techniques and machine learning has demonstrated that global acoustic data recorded from a crowd at a basketball game can be used to classify emergent crowd behavior in terms of the crowd's purported emotional state. We propose that the description of crowd behavior from such global acoustic data could benefit from nonlinear analysis methods derived from dynamical systems theory. Such methods have been used in recent research applying nonlinear methods to audio data extracted from music and group musical interactions. In this work, we used nonlinear analyses to extract features that are relevant to the behavioral interactions that underlie acoustic signals produced by a crowd attending a sporting event. We propose that recurrence dynamics measured from these audio signals via recurrence quantification analysis (RQA) reflect information about the behavioral dynamics of the crowd itself. We analyze these dynamics from acoustic signals recorded from crowds attending basketball games, and that were manually labeled according to the crowds' emotional state across six categories: angry noise, applause, cheer, distraction noise, positive chant, and negative chant. We show that RQA measures are useful to differentiate the emergent

Correspondence should be sent to Shannon Proksch, Department of Psychology, Augustana University, Sioux Falls, SD 57197, USA. E-mail: shannon.proksch@augie.edu

acoustic behavioral dynamics between these categories, and can provide insight into the recurrence patterns that underlie crowd interactions.

Keywords: Recurrence quantification analysis; Acoustical analysis; Crowd behavior; Dynamical systems; Emergence

1. Background

Thousands of people attending a basketball game are clapping and shouting, some chatting with their neighbors, some yelling toward the court, when the chant of "Let's Go, [Team Name]" seems to emerge out of the crowd without any particular warning or leader, spreading through the fans until everyone is participating in the chant. The spreading of this synchronized acoustical behavior is not unlike the "wave" that physically spreads through groups of fans raising their arms and standing in succession (Farkas, Helbing, & Vicsek, 2002), and is one example of collective social interaction that may spontaneously emerge during a sporting event through a form of "social contagion" (Mann, Faria, Sumpter, & Krause, 2013). The emergence of this chant, or repetitions of words and rhythmic claps (e.g., "De-fense!" *clap, clap*), are exemplars of collective acoustical behavior among large crowds. These simple chants and rhythmic clapping behavior can serve as an auditory signal to enhance synchrony and coordination among crowd members who are spatially separated or otherwise outside of visual contact (Mann et al., 2013).

The question of synchrony or other forms of coordination during large group interactions has been investigated in the relatively structured social interactions of musical participation, and increasingly, it is being recognized that the behavior of a large interacting group may not be wholly explained by description of local interactions alone, but rather by description of the emergent dynamics of the group itself (Demos & Palmer, 2022; Høffding et al., 2023; Schiavio, Maes, & van der Schyff, 2022). We argue that the methods used to study these musical social interactions (from coordination dynamics and dynamical systems, c.f. Schiavio, Maes, and van der Schyff, 2022) can be expanded to study the dynamics of less scripted behavior, such as the dynamics of crowd behavior at sporting events. Farrera and Ramos-Fernández (2022) describe a phenomenon of emergent collective rhythm (like the collective "wave") as the emergence of "group-level rhythmic patterns" that result from the self-organization of social behavior through entrainment and synchronization (Farrera & Ramos-Fernández, 2022). This social behavior is characterized by "context-specific interactions" and mutual "fast adaptation to other group members" actions' (Demos & Palmer, 2022). This coordinated crowd behavior at sporting events can include both the physical movement of the crowd and variations of coordinated or rhythmic auditory signals generated by their cheers and chants. But of particular importance is the *reciprocal interaction* between members of the crowd, with local patterns of signals nested into larger dynamics that reflect the behavior of the crowd as a whole. In this sense, the basketball crowd with its collective rhythms and emergent chants can be thought of as an interaction-dominant system.

Interaction-dominant systems—such as a collection of individuals interacting within a crowd—can be described by emergent structures and dynamics that are behavioral properties of the system itself, rather than of any individual component (Riley, Richardson, Shockley, & Ramenzoni, 2011). The emergent dynamics of crowd behavior have been fruitfully modeled according to biological phenomena such as swarm behavior (Kok, Lim, & Chan, 2016). Classification of emergent crowd dynamics, often using computer vision technology, has typically relied on analysis of video data for features, such as crowd density estimation, motion detection, and movement/behavior tracking of individual signals or group behavior (Kok et al., 2016; Swathi, Shivakumar, & Mohana, 2017). However, it is not always feasible to obtain high-quality image, video, or speech data of a crowd in action, nor is it always feasible to obtain signals measured from each individual in an interacting crowd. We seek to extend the study of emergent crowd behavior to include analysis of the global acoustic output of a crowd as a whole. This acoustical analysis of crowd behavior can augment current video-based crowd behavior analysis, and can also mediate in cases where video data are incomplete or unclear.

Recent work using traditional signal processing techniques (e.g., spectral analysis) and machine learning has demonstrated that global acoustic data recorded from a crowd at a basketball game can be used to classify crowd behavior in terms of the crowd's purported emotional state (Butler et al., 2018). Importantly, these data were not a collection of individual acoustic signals from individual members of the crowd, but rather a global acoustic signal measured from the crowd as a whole. Common acoustic analyses used for classifying human speech, as well as crowd noise, include spectral and mel frequency cepstral coefficients (Reddy, Sinha, & Seshadri, 2013), where the latter measure is based on an approximation of human hearing (Singh & Rani, 2014). These measures assume that at short enough time scales important features of an audio signal are reasonably stationary. Nonlinear analysis techniques, such as recurrence quantification analysis (RQA), are adept at capturing exactly that nonstationarity that characterizes audio signals at longer time scales (Wallot & Leonardi, 2018). RQA has been used to quantify the coordinated and uncoordinated acoustical activity of a large interacting group in a scripted musical interaction (Proksch, Reeves, Spivey, & Balasubramaniam, 2022). We build on Proksch et al. (2022), and propose that analysis of crowd behavior from global acoustic data in the ecological and nonscripted environment of a basketball game could similarly benefit from taking a dynamical systems approach that embraces the nonlinearity and nonstationarity present in the sounds generated by an interacting crowd.

Our paper presents a case study analysis of two data sets containing labeled instances of crowd sound recorded from the student section of two basketball games. Specific aims of our project are expanded in Section 1.3. Briefly, our first aim is to describe the nonlinear dynamics present in the sound of the crowds using RQA (more details in Section 2.1). We test whether there are differences in RQA metrics for specific labels of crowd sound based on a theory-driven hypothesis that some categories of crowd interactions (e.g., chanting) require higher coordination than other categories (e.g., noise-specific hypotheses are outlined in Section 1.3). Our second aim concerns the practical use of RQA metrics in the prediction of crowd sound behavior in a data-driven approach. We train a support vector machine (SVM) classifier on crowd sound labels and RQA results from a subset of crowd sound samples, and

test whether the model can predict the crowd sound label based on RQA results of unseen data. To motivate each of these aims, we first provide background into the Dynamical Systems framework with further explanation of RQA (Section 1.1), the Dynamics of Collective Interaction, particularly involving behavioral and physiological synchrony during group interpersonal interaction (Section 1.2), and the concept of Acoustic Social Coordination and its relevance to the context of sporting events (Section 1.2.2). We then introduce the two data sets which make up this case study and our methods of applying RQA to describe (Aim 1) and predict (Aim 2) the crowd's acoustical behavior in each basketball game. We end with a discussion of results of this case study and the potential application to further research on large group interaction.

1.1. Dynamical systems

Dynamical systems theory seeks to describe the nonlinear behavior of large-scale systems that emerges from interacting components/individuals (Connell, DiMercurio, & Corbetta, 2017). Such emergent behavior arises due to the soft-assembly of individual components into metastable patterns of behavior (Kello & Van Orden, 2009). When large groups of people gather together, they consciously and unconsciously coordinate their behavior in a number of ways, from cheering with the same chants to spontaneously synchronizing in their applause at concert. The patterns of synchronicity in the sounds generated by crowds demonstrate a process of social self-organization (Néda, Ravasz, Bréchet, Vicsek, & Barabási, 2000).

One tool in the dynamical systems toolbox is RQA. RQA is used to quantify structures that can be visualized in recurrence plots generated from the nonlinear behavior of a time series that has been subject to state space reconstruction (Marwan, Wessel, Meyerfeld, Schirdewan, & Kurths, 2002; Marwan, Romano, Thiel, & Kurths, 2007; Takens, 1981; Vlachos & Kugiumtzis, 2010). Traditional RQA, as well as multivariate approaches like cross RQA and multidimensional RQA (mdRQA), has proven useful in describing the behavioral aspects of joint action in dyadic and group interaction. These analyses are robust to the nonlinearity and nonstationarity of time-dependent signals, and can be used to evaluate relative coordination dynamics as well as transitions between order and chaos in such systems (c.f. Wallot and Leonardi, 2018 for a detailed review and tutorial).

In a recurrence plot, time series data are plotted on axes of time by time. A point (i, j) is plotted if the value at time i and time j are recurrent within a specified neighborhood size of an N-dimensional state-space after state space reconstruction. The line of incidence (LOI) along the main diagonal shows the time series at a time lag of 0. Each step away from the LOI represents the trajectory of the system at a time lag, depicting self-similarity of the system over time.

Information from a recurrence plot is quantified into a variety of metrics via RQA. The recurrence rate quantifies the percentage of points on a recurrence plot and represents patterns of behavior that persist over time. Determinism quantifies the percentage of points that fall on any diagonal line in the plot (except the LOI), and represents behaviors that belong to a particular pattern of behavior over time. Entropy is the variability in these line lengths, representing disorder of these sequences. Finally, laminarity quantifies the percentage of points

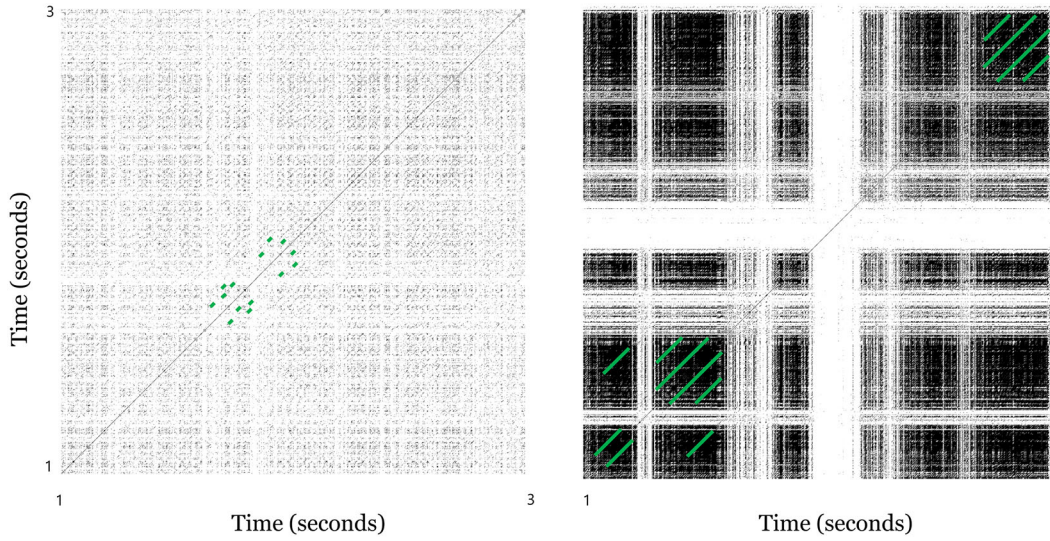

Fig. 1. Green lines depict trajectories of consecutive recurrent points over time. Longer trajectories are quantified in higher values of determinism. The plot on the left has many short diagonal trajectories consisting of only a few consecutive recurrent points, while the plot on the right contains many long diagonal trajectories consisting of many consecutive recurrent points over time (even more than we have highlighted). These plots are zoomed in on 2 s of data from two different 5-s samples of crowd sound. At left: A 2-s sample of Distraction Noise. At right: A 2-s sample of positive chant. For zoomed out plots from full 5-s samples, refer to Fig. 3.

that fall on a vertical line on a recurrence plot, and represents clusters of behavior over a short period of time to which the system may temporarily visit, leave, and return. Examples of determinism and laminarity depicted from two samples of basketball crowd sound data are shown in Figs. 1 and 2.

RQA has additionally proven useful in describing and classifying acoustic data. Zhang et al. (2011) made use of recurrence plots and RQA to classify audio signals into noise-like, transient, harmonic-like, and mixed signals. Proksch et al. (2022) further justified the use of RQA to describe differences in acoustical signals generated by multiagent behavior of a performing musical ensemble who were either uncoordinated with each other (the asignal they collectively generated was noise-like), or were coordinated with each other (the audio signal they collectively generated was harmonic-like).

1.2. Dynamics of collective interaction

The context of a social event affects the emergence of collective synergies, synchronicities, and multistable dynamics in the acoustical behavior of collective interactions. To understand how analysis of the acoustical behavior of a crowd might expand research on joint action and crowd dynamics, it is necessary to provide context on the growing body of research evaluating coordination dynamics that arise during group interactions, much of which has focused on the behavioral and physiological modalities.

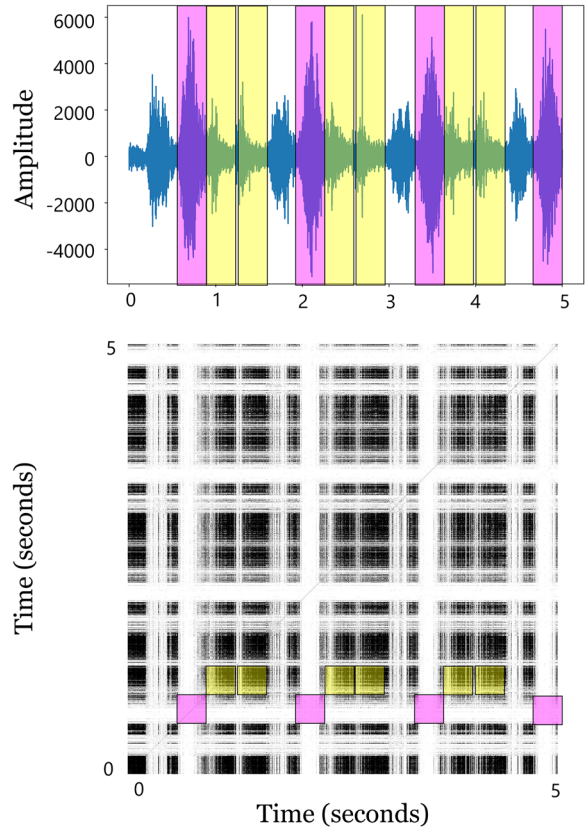

Fig. 2. Highlighted instances of laminar states in a time series reflected in a recurrence plot. The bursts in the time series appear as white space with few recurrent points in the recurrence plot. This is because this bursty state is revisited only three additional times after the instance highlighted in pink along the main diagonal. Meanwhile, the state highlighted in yellow repeats an additional five times for the duration of the behavior. This plot is one 5-s sample of crowd sound.

1.2.1. Modalities of synchrony and coordination in groups

Recently, MdRQA was used to measure physiological synchrony in the heart rates of fans who attended a live basketball game, and fans who gathered in small groups to attend a live screen of a basketball game on television (Baranowski-Pinto, Profeta, Newson, Whitehouse, & Xygalatas, 2022). MdRQA is able to compute recurrence measures across multiple signals (i.e., the heart rates of multiple individuals), in contrast to RQA that evaluates recurrence measures across the length of a single signal. Increased interdependence in heart rate was found for fans who attended the live game, indicating that there is an enhanced social effect of the interpersonal dynamics inherent in attending this sporting event live and in-person compared to virtually engaging with the game over a television screen. This was demonstrated through increases in both determinism (DET: indicating stability in the system and the ability to predict future states from past states) and average diagonal line length

(ADL: indicating the length of time, or persistence, of recurrent states within the system). Further, Baranowski-Pinto et al. found that individuals who attended the live game exhibited stronger social cohesion, as reported by stronger feelings of transformativeness, or the sense that their individual identity has "fused" to the identity of the group. These self-report measures were correlated with recurrence measures for fans who attended the game in person. A second study evaluated behavioral synchrony of an audience attending a live or prerecorded rock concert (Swarbrick et al., 2019). Where basketball fans in Baranoiwski-Pinto et al.'s study were both watching a live game, which differed only by being in-person or screened on television, Swarbrick et al.'s study maintained the interpersonal dynamic between concert attendees by having concert-goers in each condition be physically present in the concert venue (Swarbrick et al., 2019). In the live concert, the rock band performed on the stage, while in the nonlive condition, a recording of that performance was projected onto the stage. Faster head movements, a measure of vigor and engagement, were found during the live performance than the nonlive performance. No effect was found between performances for entrainment with the music—however, it was not analyzed whether there was enhanced movement synchrony *between* audience members during either performance.

In a third study, both physiological and behavioral synchrony were evaluated in groups of three people engaged in a joint drumming task (Gordon, Gilboa, Cohen, & Kleinfeld, 2020). Synchronous or asynchronous drumming interaction was achieved by asking participants to drum along to an auditory beat with a predictable or unpredictable tempo, respectively. Gordon et al. (2020) found that the drumming task itself led to increased synchrony of heart-beat inter-beat-intervals (IBIs) between group members compared to baseline (non-interaction). Groups with higher physiological synchrony during the initial drumming task were more coordinated during a subsequent free-improvisation drumming task. Interestingly, this increase in heartbeat IBI synchrony was not related to whether the initial drumming task was synchronous or asynchronous, indicating that the enhanced heartbeat IBI synchrony may stem from the effect of the interpersonal interaction itself, rather than the behavioral synchronization of the drumming itself. This may be similar to interaction at a sporting event, where individuals are not always synchronizing directly with other fans in attendance. Further, even during a coordinated cheer, individuals may be "in sync" with a global signal without being "in sync" to other individuals directly nearby.

These studies highlight the importance of interpersonal interaction—as well as a live, in-person interaction context—in facilitating physiological and behavioral synchrony and the emergence and maintenance of shared social bonds. Although the specific modality of focus differed in each study (from physiological measures of heart-rate, to behavioral measures of movement), all of the interactions in these studies occurred in a shared acoustic and auditory environment. Anthropologist and ethnomusicologist, Blake and Cross (2015), state that it is precisely this environment that is "one of the most powerful and flexible tools that humans use to manage and mediate relationships with each other and with the environments that they construct or modify." In the next section, we situate our interest in acoustic social coordination and in the shared social scripts that underlie the coordination of acoustical behavior during the interactions of a crowd.

1.2.2. Acoustic social coordination

The emergence of different joint action dynamics can be analyzed in terms of social scripts—implicit or explicit norms for organizing behavior in social contexts that are "underwritten by culturally specific narratives" (Albarracin, Constant, Friston, & Ramstead, 2021). The acoustical behavior of the musical ensemble described in Section 1.1 was carried out according to an *explicitly* social script—a musical score—governing the transition from uncoordinated action of individuals to coordinated interaction of a multi-agent group. These two coordination modes are reflected in the emergence of structured recurrence patterns over time (Proksch et al., 2022). Applause after concerts can also follow certain *implicit* social scripts—spreading by initial social contagion (Mann et al., 2013) and perhaps persisting while slowly dying down, or ending abruptly as soon as a loudness threshold is passed (Michard & Bouchaud, 2005). Fluctuations in the relative synchrony of sound generated by the applause of a crowd attending a classical music concert have been shown to display an emergent periodic signal. Initial applause is fast and asynchronous, and as synchrony increases, the overall signal of the sound behavior increases, while the average noise of the sound decreases. This decrease in average noise is a result of a slower clapping period that emerges as individuals clap in unison (Néda et al., 2000). Such audiences fluctuate between asynchronous and synchronous behavior before ultimately fading out as the event draws to a close.

Basketball games are another social context that affords the emergence of coordinated acoustical behavior among a group of interacting people. Rather than an explicitly written script, fans at sporting events follow an at times *explicit* or *implicit* social script, where events in the game and prompts from the announcer or cheerleaders, or other fans, govern the behavior of fans gathered in the arena. Patterns of social self-organization emerge and dissipate according to local interaction among fans and global interactions associated with the game. When your team scores, the social script affords a cheer, when the other team is attempting a free throw, the social script affords generating raucous noises in attempt to distract the player on the court, and when the cheer leaders or a group of fans begin a rehearsed chant ("Defense";"B-Y-U Cougars"), the social script requires that you chant along. These rehearsed chants are an example of joint speech, a collective phenomenon where multiple individuals repeat the same words simultaneously with the purpose of engaging in group expression (Cummins, 2013).

Synchrony demonstrated in this acoustical behavior, including the synchrony of joint speaking during group chants, is an important characteristic of interpersonal interaction. The repetition of chants or short rhythmic utterances in sporting events enables "synchronized activity…an extreme form of coordination," whereby individuals enact a collective "we" and establish a coordinated group identity for as long as the behavior persists (Cummins, 2020). Joint action research has demonstrated that higher levels of synchronous behavior are associated with various aspects of prosocial cognition including: increased affiliation (Hove & Risen, 2009; Wiltermuth & Heath, 2009), social cohesion (Marsh, Richardson, & Schmidt, 2009), group identity (McNeill, 2022), and cooperation (Kirschner & Tomasello, 2010). A recent meta-analysis has shown that the effects of synchrony on prosocial behaviors and positive affect are larger for larger groups (Mogan, Fischer, & Bulbulia, 2017). However, it can be difficult to measure synchronous activity or joint speech from very large groups of people

engaged in naturalistic social interactions. It may not always be possible to obtain one signal from each member of a large group to evaluate correlations and synchrony between those signals. What may be more feasible in such situations, is to record a global acoustic signal generated by the group as a whole.

1.3. Project aims

Previously, Proksch et al. (2022) applied RQA analysis to a global acoustic recording of a performing musical ensemble. Whether individuals within that ensemble were coordinating their behavior, or not, was dictated by a musical score. These two patterns of behavioral dynamics (uncoordinated vs. coordinated) were reflected in RQA metrics derived from analyzing recurrence plots generated from the resampled audio signals. Here, we present case studies of two crowds at two Brigham Young University (BYU) basketball games. We analyze the recorded audio signal of the crowds of students engaging in various forms of acoustical behavior at each game. Following from both Zhang et al. (2011) and Proksch et al. (2022), we used nonlinear analyses to extract features that are relevant to the behavioral interaction and coordination of the crowd who produced these audio signals. We propose that recurrence dynamics measured from this global audio signal reflect information about the behavioral dynamics of the crowd itself. We calculate recurrence features using RQA to evaluate the emergent acoustical behavioral dynamics of the interacting crowd.

This paper has two objectives. The first objective is a theory-driven description of crowd sound dynamics using specific RQA metrics relevant for describing system-level collective behavior: Recurrence Rate, Determinism, Entropy, and Laminarity (described in more detail in the methods below). We predict that the coordinated acoustical behavior of the crowd will exhibit higher stability and recurrence (measured by determinism and recurrence rate) during pseudo-rhythmically organized joint speech such as rhythmic chants. Meanwhile, the less structured nature of acoustic events such as distraction noise will exhibit lower measures of stability and recurrence. We argue that chant requires whole-crowd coordination to produce specific words, that is, acoustic patterns as a whole. Cheering and applause are behaviors involving intermittent coordination, engendering individual patterns of clapping, hooting, and so on that may be locally coordinated within earshot and eyeshot, but specific acoustic patterns will vary across the crowd. Distraction or angry noise is an aim for uniform noise with minimal structured variability in the acoustic signal, which is noise that may hinder or perturb performance of the players on the court. We predict that the strongest coordination (i.e., chanting) will result in higher RQA values due to a higher percentage of recurring behavior (measured through recurrence rate), higher stability in sequences of behavior (determinism), increased variability in the possible sequences of behavior (entropy), and increased clusters of behavioral states that are revisited over time (laminarity). Intermittent or local coordination (i.e., cheer/applause) will result in lower values across these RQA metrics. Finally, inhibited synchronization and minimal coordination toward structured variation (i.e., noise) will result in the lowest values. Thus, we hypothesize that RQA will reflect these three distinct modes of acoustic patterning that vary in recurrence structure from most to some to least.

Table 1
Crowd sound categories and descriptions adapted from Butler et al. (2018)

Crowd Sound Category	Description
Angry Noise	Crowd shouting in anger.
Applause	Crowd clapping that can include crowd vocalization.
Cheer	Loud, positive crowd vocalization.
Distraction Noise	Attempts by crowd to draw an opposing team member's attention away from the game, most commonly when the opposing team possesses the ball or is about to shoot a free throw.
Negative Chant	Crowd shouting in anger or distress, usually directed toward referees after a less than ideal call or toward a player from the opposite team.
Positive Chant	Rhythmic crowd shouting, usually directed toward the home team, for example, "De-fense" or "B-Y-U- Cougars."

The second objective is a data-driven, machine learning approach to classify each crowd sound based on the full suite of metrics available from the PyRQA package (a total of 19 RQA metrics, listed in the methods). A combination of RQA metrics and SVM classification has proven effective at discriminating between nonlinear (and nonstationary) dynamical systems that exhibit similar dynamics. dos Santos, Barroso, Godoy, Macau, & Freitas (2014) showed that RQA plus an SVM classifier showed successful classification of time series data generated from the Logistic map—a canonical example of a nonlinear dynamical system—as well as classification of real biological data describing the (nonlinear and nonstationary) dynamics of human heart rate variability across different age groups and health contexts (dos Santos et al., 2014). We apply an SVM classifier on RQA metrics of samples of crowd data that were labeled according to differing classes of acoustical behavior. We predict that the machine learning analysis will show that RQA features can clearly distinguish between the acoustic patterns that correspond to chanting (strongest coordination/synchronization) versus cheering (intermittent coordination/synchronization) versus distracting (weakest coordination/inhibited synchronization).

2. Methods

2.0.1. Crowd sound data sets

Our data sets are previously collected audio recordings and associated text files with labeled crowd events. Each audio recording was recorded from the student section of two Men's BYU basketball games (Butler et al., 2018). The data sets were previously labeled by BYU undergraduates—who were unaware of the hypotheses of the current study—into different classes of acoustical behavior along with labels of associated game events, shown in Table 1. The specific labels for crowd events were agreed upon by consensus and with reference to specific events in accompanying video recordings of the basketball game (e.g., distraction noise associated with free-throws by the opposing team, positive chant directed toward the home team, negative chant directed toward the opposing team). Elements of crowd sound

Table 2
Nonoverlapping 5-s samples used for RQA and linear regression

Crowd Sound Category	Game 1 # of samples	Game 2 # of samples
Angry Noise	16	17
Applause	15	25
Cheer	47	79
Distraction Noise	116	77
Negative Chant	14	0
Positive Chant	63	58

could have multiple labels (e.g., applause often accompanied by cheer). In the current paper, we discarded any samples with multiple labels.

For Game 1, we analyze six classes of acoustical behavior that were observed in the crowd during this basketball game: Angry Noise, Applause, Cheer, Distraction Noise, Negative Chant, and Positive Chant. For Game 2, we analyze the same classes except for Negative Chant, of which no events were labeled for this game. We did not analyze sound events labeled as Singing (which was defined as "Harmonic crowd vocalization accompanied by the pep band or PA system") or Silence. The raw acoustic data were recorded at a sampling rate of 50 kHz. We resampled by a factor of 10 for nonlinear analysis at 5 kHz. As described in Proksch et al. (2022), resampling to a lower sampling rate focuses on the higher-order rhythmic properties and aggregate amplitude of the acoustic signal, essentially filtering out much of the pitch-information from the acoustical behavior of the audience as well as semi-pitched signals from shoes across the basketball court that may have been picked up by the microphones.

2.1. Part 1: Recurrence quantification analysis

For Part 1, 5-s samples were created using nonoverlapping windows (such that a 12-s event will have two 5-s samples, e.g., 0–5 s, and 5–10 s, and the remaining 2 s in the event are discarded). Any samples shorter than 5 s were discarded, and residual data longer than multiples of 5 s were also discarded. Five-second samples were chosen as the smallest sample that could capture relevant behavioral information about the crowd. For example, a 5-s sample allows for at least one repetition of a simple chant if performed by a crowd at roughly 100–120 beats per minute. This same 5-s sample size was also applied to detect recurrence information in Proksch et al. (2022). The number of samples for each crowd sound category is reported in Table 2.

Based on Proksch et al. (2022), we chose four RQA metrics to evaluate, which are each indicative of different aspects of behavior in nonlinear dynamical systems: Recurrence Rate, Determinism, Entropy, and Laminarity.

- Recurrence rate (the percentage of recurrence points on the recurrence plot) represents patterns of behavior that repeat over time.

- Determinism (the percentage of points that fall on any diagonal line in the recurrence plot) represents behaviors that belong to a longer sequence of behavior.
- Entropy (the variability in lengths of these diagonal lines) represents the amount of disorder there is in these sequences.
- Laminarity (the percentage of points that fall on a vertical line in the recurrence plot) represents clusters of behavior for a length of time, that is, when a system visits a behavior for a period of time, leaves, and then returns to that behavioral state.

RQA was run on the time series data extracted from the resampled audio for each of these 5-s samples using PyRQA version 8.0.0 (Rawald, Sips, & Marwan, 2017), with an embedding dimension of 5, a delay of 5, and a neighborhood value fixed radius of 1*standard deviation (SD), using maximum norm to calculate neighbors of the phase space trajectory. Parameters for the time delay and embedding dimension were chosen based on Average Mutual Information (AMI) and Fale Nearest Neighbors (FNN), respectively, using a custom MATLAB GUI provided from the 2019 APA Advanced Training Institute in Nonlinear Methods for Psychological Science.

There are a variety of approaches to setting the neighborhood threshold value. A standard approach, which we use, is setting this threshold value according to a fixed amount of nearest neighbors at some ratio of the standard deviation of the data. This holds constant the number of neighbors within a neighborhood and also holds the number of recurrence points constant in a column of the recurrence plot (Eckmann, Kamphorst, & Ruelle, 1987). It has been suggested that 5*SD be used to accurately detect a signal in the presence of significant observational noise (Thiel et al., 2002). However,"this approach fails for signals of very low signal to noise ratio, or when the amount of noise is unknown" (Schinkel, Dimigen, & Marwan, 2008). Additionally, the so-called "noise" is the signal in our data, therefore, we settle on a similar approach to Zhang et al. (2011), setting this threshold based on a Fixed Amount of Nearest neighbors value of 1*SD. (Note: Zhang et al. (2011) used 1*standard error—we chose standard deviation because the value given by SD of the mean is always larger than the SE of the mean, assuring that our radius is large enough to sufficiently capture the recurrence structures in the recurrence plots.)

2.1.1. Statistical analyses

A separate linear regression model with planned contrasts was fit for each RQA metric of interest (Recurrence Rate, Determinism, Entropy, and Laminarity) as a function of the level of Crowd Sound Category:

- Recurrence Rate ~ CrowdSoundCategory
- Determinism ~ CrowdSoundCategory
- Entropy ~ CrowdSoundCategory
- Laminarity ~ CrowdSoundCategory

There were six levels of Crowd Sound Category for Game 1 (Angry Noise, Applause, Cheer, Distraction Noise, Negative Chant, and Positive Chant), and five levels for Game 2 (Angry Noise, Applause, Cheer, Distraction Noise, and Positive Chant). Sum-to-zero contrasts were used to specify a specific linear combination for each predictor in a priori planned comparisons. Since we are not testing a specific treatment or change from any initial baseline

Table 3
Contrast matrix for a priori sum-to-zero contrasts

Game 1					
Angry Noise	1	0	0	0	0
Applause	0	1	0	0	0
Cheer	0	0	1	0	0
Distraction Noise	0	0	0	1	0
Negative Chant	0	0	0	0	1
Positive Chant	−1	−1	−1	−1	−1
Game 2					
Angry Noise	1	0	0	0	
Applause	0	1	0	0	
Cheer	0	0	1	0	
Distraction Noise	0	0	0	1	
Positive Chant	−1	−1	−1	−1	

of crowd sound behavior, sum contrasts provide the advantage of comparing RQA metrics in each category to the average value across all categories, rather than to a baseline or control category (Schad, Vasishth, Hohenstein, & Kliegl, 2020). That is, we are comparing each contrast to the mean of all means (the grand mean) of the RQA metric of interest. The contrast matrix is shown in Table 3.

Following regression analysis, we implemented post hoc pairwise comparisons of estimated marginal means to compare the relative RQA metrics between each pair of crowd sound categories. Pairwise comparisons were implemented in the R package emmeans, version 1.5.4 (Lenth, 2022). The linear regression and pairwise comparisons of estimated marginal means were implemented separately for each of the four RQA measures of interest. This is because each RQA measure addresses a different aspect of the crowd's behavior over time, as described previously.

2.2. Results

2.2.1. RQA results

Fig. 3 shows recurrence plots generated from the time series data of a representative 5-s audio sample from two crowd sound categories. These recurrence plots visualize characteristic patterns of recurrence that are quantified through RQA. Qualitatively, the recurrence plot generated from 5 s of distraction noise resembles recurrence plots of uncoordinated group behavior (Proksch et al., 2022) or noisy-like audio signals (Zhang, Liu, Zhang, & Bu, 2011), while the recurrence plot generated from 5 s of labeled positive chant resembles a recurrence plot generated from coordinated group behavior (Proksch et al., 2022) or mixed audio signal (Zhang et al., 2011). Distraction noise shows low levels of stability and recurrence, while positive chant shows high levels of stability and recurrence, as quantified by RQA. Further statistical analysis on the distribution of RQA metrics in each crowd sound category is described below.

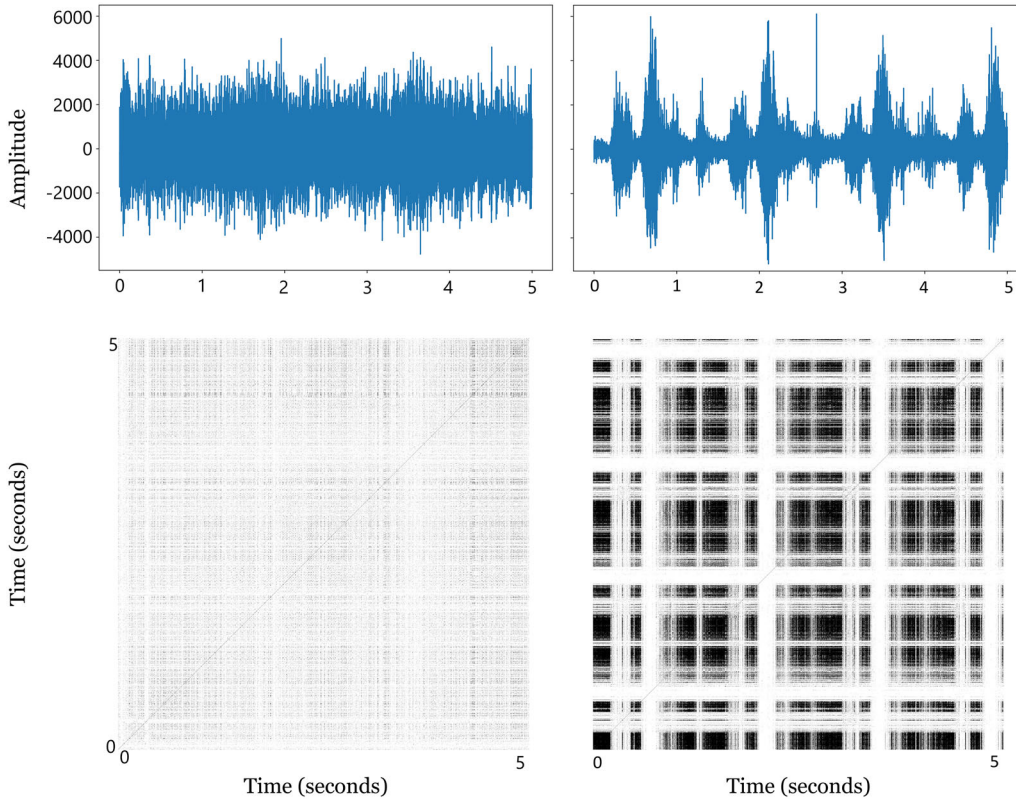

Fig. 3. Representative time series and recurrence plots from 5-s samples of two categories of crowd sound: Distraction Noise (left) and Positive Chant (right). During this sample of chant, the audience is repeating "De-fense (clap clap), De-fense (clap clap)."

2.2.2. *Summary statistics*

Table 4 lists the associated descriptive means, median, and standard deviation of RQA metrics, and Fig. 4 shows the smoothed density distributions and quartile lines for the RQA data from each crowd sound category. For Game 1, the two chant categories (Positive and Negative chant) display the highest values of recurrence rate, determinism, entropy, and laminarity. The two noise categories (Angry and Distraction Noise) display consistently low values of these RQA metrics, with Angry Noise having particularly low values of Entropy and Laminarity. Cheer and Applause are the most variable, with multimodal or nearly flat distributions exhibited by Cheer. Game 2 follows a similar pattern, with the chant category (Positive Chant) displaying the highest values of all four RQA measures. Recurrence rate values are low for all other categories. The distribution of determinism, entropy, and laminarity values for cheer is relatively long-tailed, while the two noise categories (Angry and Distraction) and the Applause category show a more peaked distribution in each of these categories, albeit lower than Positive Chant.

Table 4
Summary statistics for each RQA measure across six Crowd Sound Categories: Mean, median, and standard deviation as calculated from original, nonbootstrapped data

Game 1 Recurrence Rate	mean	median	sd
Angry Noise	0.06	0.05	0.02
Applause	0.08	0.08	0.03
Cheer	0.10	0.08	0.06
Distraction Noise	0.07	0.06	0.04
Negative Chant	0.12	0.11	0.05
Positive Chant	0.17	0.16	0.07
Determinism	**mean**	**median**	**sd**
Angry Noise	0.25	0.20	0.10
Applause	0.37	0.36	0.11
Cheer	0.39	0.38	0.23
Distraction Noise	0.34	0.29	0.16
Negative Chant	0.52	0.50	0.14
Positive Chant	0.64	0.62	0.13
Entropy	**mean**	**median**	**sd**
Angry Noise	0.72	0.66	0.24
Applause	0.94	0.90	0.26
Cheer	0.96	1.00	0.32
Distraction Noise	0.92	0.86	0.27
Negative Chant	1.17	1.13	0.25
Positive Chant	1.32	1.33	0.21
Laminarity	**mean**	**median**	**sd**
Angry Noise	0.40	0.34	0.11
Applause	0.52	0.51	0.11
Cheer	0.51	0.53	0.22
Distraction Noise	0.45	0.40	0.16
Negative Chant	0.65	0.65	0.13
Positive Chant	0.75	0.74	0.10
Game 2 Recurrence Rate	mean	median	sd
Angry Noise	0.04	0.04	0.00
Applause	0.06	0.05	0.03
Cheer	0.04	0.04	0.01
Distraction Noise	0.06	0.05	0.02
Positive Chant	0.10	0.09	0.04

(*Continued*)

Table 4
(Continued)

Determinism	mean	median	sd
Angry Noise	0.15	0.15	0.04
Applause	0.28	0.25	0.12
Cheer	0.22	0.16	0.14
Distraction Noise	0.31	0.28	0.12
Positive Chant	0.60	0.59	0.20
Entropy	**mean**	**median**	**sd**
Angry Noise	0.56	0.52	0.13
Applause	0.74	0.69	0.22
Cheer	0.77	0.67	0.30
Distraction Noise	0.98	0.96	0.22
Positive Chant	1.39	1.23	0.50
Laminarity	**mean**	**median**	**sd**
Angry Noise	0.24	0.24	0.07
Applause	0.42	0.40	0.13
Cheer	0.32	0.25	0.18
Distraction Noise	0.28	0.25	0.07
Positive Chant	0.72	0.71	0.17

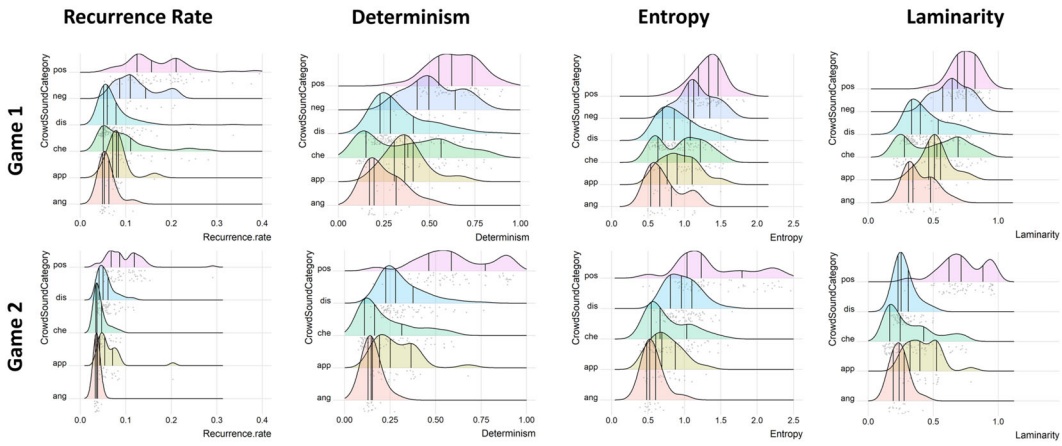

Fig. 4. Ridgeline plots show the smoothed distribution and individual data points along with quantile lines for four RQA measures across six Crowd Sound Categories: Angry Noise, Applause, Cheer, Distraction Noise, Negative Chant ,and Positive Chant.

2.2.3. *Linear regression with sum-to-zero contrasts and pairwise comparisons*

Results of the linear regression with sum-to-zero contrasts are shown in Table 5. The first row shows the grand mean of the RQA metric averaged across all categories of crowd sound. The beta value for each subsequent category shows the deviations from this grand mean for

Table 5
Results from the linear regression with sum-to-zero contrasts reported for each RQA measure across six crowd sound categories (Game 1) and five crowd sound categories (Game 2)

Game 1 Recurrence Rate	Estimate	Std. Error	*t* value	Pr(>|t|)
GrandMean(Int)	0.10	0.00	23.96	< 2e−16***
Angry Noise	−0.04	0.01	−3.72	0.000244***
Applause	−0.02	0.01	−1.69	0.091698
Cheer	−0.00	0.01	−0.56	0.576278
Distraction Noise	−0.03	0.01	−4.88	<1.9e−06***
Negative Chant	0.02	0.01	1.92	0.055668
Positive Chant	0.07			

Note. Residual standard error: 0.05056 on 265 degrees of freedom. Multiple R^2 0.3922, Adjusted R^2 0.3808. *F*-statistic: 34.2 on 5 and 265 DF, *p*-value: < 2.2e-16

Determinism	Estimate	Std. Error	*t* value	Pr(>|t|)
GrandMean(Int)	0.42	0.01	31.34	< 2e−16 ***
Angry Noise	−0.17	0.04	−4.72	3.89e−06 ***
Applause	−0.05	0.04	−1.29	0.19917
Cheer	−0.028	0.02	−1.21	0.22654
Distraction Noise	−0.08	0.02	−4.22	3.41e−05 ***
Negative Chant	0.10	0.04	2.7	0.00722 **
Positive Chant	0.22			

Note. Residual standard error: 0.1615 on 265 degrees of freedom. Multiple *R*-squared: 0.3845, Adjusted *R*-squared: 0.3729. *F*-statistic: 33.11 on 5 and 265 DF, *p*-value: < 2.2e-16

Entropy	Estimate	Std. Error	*t* value	Pr(>|t|)
GrandMean(Int)	1.01	0.02	45.84	< 2e−16 ***
Angry Noise	−0.29	0.06	−4.89	1.78e−06 ***
Applause	−0.07	0.06	−1.16	0.24685
Cheer	−0.04	0.04	−1.06	0.29013
Distraction Noise	−0.09	0.03	−2.95	0.00344 **
Negative Chant	0.17	0.06	2.68	0.00787 **
Positive Chant	0.31			

Note. Residual standard error: 0.265 on 265 degrees of freedom. Multiple *R*-squared: 0.3263, Adjusted *R*-squared: 0.3136. *F*-statistic: 25.67 on 5 and 265 DF, *p*-value: < 2.2e-16.

Laminarity	Estimate	Std. Error	*t* value	Pr(>|t|)
GrandMean(Int)	0.55	0.01	42.53	< 2e−16 ***
Angry Noise	−0.15	0.03	−4.30	2.40e−05 ***
Applause	−0.03	0.04	−0.83	0.41
Cheer	−0.03	0.02	−1.44	0.15

(*Continued*)

Table 5
(Continued)

Laminarity	Estimate	Std. Error	*t* value	Pr(>|t|)
Distraction Noise	−0.09	0.02	−5.34	1.68e−07 ***
Negative Chant	0.10	0.04	2.81	0.00539 **
Positive Chant	0.20			

Note. Residual standard error: 0.1552 on 265 degrees of freedom. Multiple *R*-squared: 0.3911, Adjusted *R*-squared: 0.3796. *F*-statistic: 34.05 on 5 and 265 DF, *p*-value: < 2.2e-16.

Game 2 Recurrence Rate	Estimate	Std. Error	*t* value	Pr(>|t|)
GrandMean(Int)	0.06	0.00	30.93	< 2e−16 ***
Angry Noise	−0.02	0.01	−4.42	1.50e−05 ***
Applause	0.00	0.00	1.01	0.316
Cheer	−0.02	0.00	−5.45	1.19e−07 ***
Distraction Noise	−0.00	0.00	−1.07	0.286
Positive Chant	0.04			

Note. Residual standard error: 0.02522 on 251 degrees of freedom. Multiple *R*-squared: 0.4059, Adjusted *R*-squared: 0.3965. *F*-statistic: 42.88 on 4 and 251 DF, *p*-value: < 2.2e-16.

Determinism	Estimate	Std. Error	*t* value	Pr(>|t|)
GrandMean(Int)	0.31	0.01	28.21	< 2e−16 ***
Angry Noise	−0.17	0.03	−5.51	8.82e−08 ***
Applause	−0.03	0.03	−1.23	0.221
Cheer	−0.09	0.02	−5.27	2.98e−07 ***
Distraction Noise	−0.00	0.02	−0.07	0.947
Positive Chant	0.29			

Note. Residual standard error: 0.148 on 251 degrees of freedom. Multiple *R*-squared: 0.5124, Adjusted *R*-squared: 0.5046. *F*-statistic: 65.93 on 4 and 251 DF, *p*-value: < 2.2e-16

Entropy	Estimate	Std. Error	*t* value	Pr(>|t|)
GrandMean(Int)	0.89	0.02	36.29	< 2e−16 ***
Angry Noise	−0.33	0.07	−4.94	1.46e−06 ***
Applause	−0.15	0.06	−2.58	0.01049 *
Cheer	−0.12	0.04	−3.09	0.00224 **
Distraction Noise	0.09	0.04	2.30	0.02213 *
Positive Chant	0.50			

Note. Residual standard error: 0.3254 on 251 degrees of freedom. Multiple *R*-squared: 0.3928, Adjusted *R*-squared: 0.3831. *F*-statistic: 40.59 on 4 and 251 DF, *p*-value: < 2.2e-16. ***$p<0.001$, **$p<0.01$, *$p<0.05$

Laminarity	Estimate	Std. Error	t value	Pr(>|t|)
GrandMean(Int)	0.40	0.01	36.52	< 2e−16 ***
Angry Noise	−0.15	0.03	−5.30	2.51e−07 ***

(*Continued*)

Table 5
(Continued)

Laminarity	Estimate	Std. Error	t value	Pr(>|t|)
Applause	0.03	0.03	1.05	0.10
Cheer	−0.08	0.02	−4.81	2.57e−06 ***
Distraction Noise	−0.12	0.02	−7.04	1.81e−11 ***
Positive Chant	0.30			

Note. Residual standard error: 0.1438 on 251 degrees of freedom Multiple R-squared: 0.6103, Adjusted R-squared: 0.604 F-statistic: 98.25 on 4 and 251 DF, p-value: < 2.2e-16. ***$p<0.001$, **$p<0.01$, *$p<0.05$

Note: *Positive Chant values are not reported by the linear regression model, and were calculated independently by subtracting the GrandMean from the mean values of Positive Chant.*

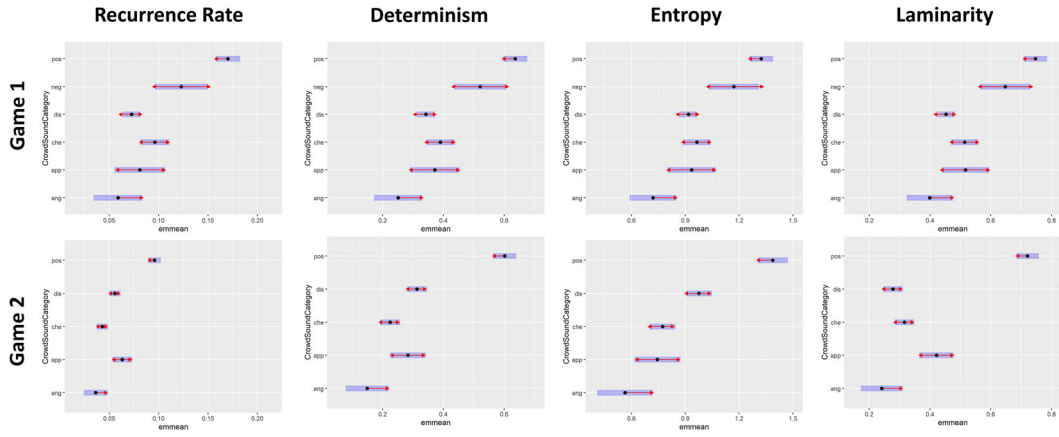

Fig. 5. Pairwise Comparisons of Estimated Marginal Means across crowd sound categories for four RQA metrics: (A) Recurrence Rate, (B) Determinism, (C) Entropy, and (D) Laminarity. Blue bars represent 95% confidence intervals. Red arrows represent comparisons among the means. Where a red arrow overlaps an arrow from another category, the difference between the overlapping categories is not significant.

that category. These results indicate that RQA metrics of some individual crowd sound categories can be differentiated from the grand average RQA metrics across all crowd sound categories.

At an alpha of 0.05, Cheer and Applause were not significantly different from the grand average for any RQA metrics in Game 1, and Distraction Noise is not significantly different from the grand average values of recurrence rate. For Game 2, Applause was not significantly different from the grand average values for all RQA metrics save for Laminarity, and Distraction Noise was not significantly different from the grand average values of Recurrence Rate and Determinism metrics.

Because positive chant was categorized as the final level for the sum-to-zero contrast coding scheme, its comparison to grand average RQA metrics and associated p-value is not directly reported from the model results. However, the change in RQA metrics can be calculated by subtracting the grand average value from the mean value for positive chant. With this in

mind, we can describe the general change from average RQA metrics for each Crowd Sound Category in each game. For Game 1, there was a general decrease from the grand average RQA values for most crowd sound categories, and an increase in values for both Negative and Positive Chant. For Game 2, there was a general decrease from the grand average for the nonchant RQA metrics, save for Entropy which saw an increase for Distraction Noise, and and increase in all RQA metrics for Positive Chant.

Marginal means and pairwise comparisons with associated confidence intervals were extracted from the linear regression, and computed for each crowd sound category using R package emmeans, version 1.5.4 and are plotted in Fig. 5 and listed in Table 6. Blue bars represent 95% confidence intervals, while red arrows represent comparisons among the means. Where a red arrow overlaps an arrow from another category, this means that the difference between the overlapping categories is not significant. Angry Noise shows the lowest values of each RQA metric, while Positive Chant (both games) and Negative Chant (Game 1) show the highest values of each RQA metric. There is some degree of overlap among the nonchant categories for all RQA metrics in both games, although the specific nonsignificant pairwise comparisons differ between games. All pairwise comparisons are listed in Tables S1 and S2. The pairwise comparison results indicate that among most RQA metrics, Positive Chant (and Negative Chant in Game 1) is significantly different from other nonchant Crowd Sound Categories in both games. In Game 2, Distraction Noise is also significantly different from the remaining nonchant categories on all recurrence metrics except for Applause for RR and DET. Remaining differences in crowd sound category differ by recurrence measure (see Tables S1 and S2).

2.2.4. Preliminary discussion for Part 1

In Part 1, we took a theory-driven approach to describe crowd sound dynamics using a set of four specific RQA metrics: Recurrence Rate, Determinism, Entropy, and Laminarity as in Proksch et al. (2022). These metrics were chosen to describe patterns of acoustical behavior from the crowd which repeat over time (RR), patterns that belong to a longer sequence of behavior (DET), the amount of disorder in these sequences (ENT), and clusters of acoustical behavior over a length of time (LAM). These four metrics are well studied in the literature describing the dynamics of interpersonal coordination and acoustical behavior (Fusaroli & Tylén, 2016; Paxton & Dale, 2013). We predicted that instances of whole-crowd coordination (such as the coordinated acoustical patterns of specific chants) versus locally coordinated activity (such as applause or cheer) versus noise-generating behavior represent three distinct modes of acoustic patterning which would be observable in distinct recurrence structures (from most to some to least, respectively). It was observed that the whole-crowd acoustical coordination required for generating chants resulted in the highest values overall across all RQA metrics. Pairwise comparisons demonstrate that these chant categories are significantly different from most other categories of crowd sound. When comparing each category of crowd sound to the grand mean RQA values (i.e., when comparing each category of crowd sound to the average acoustical behavior of the crowd throughout the game), there was some variability in which crowd sounds were differentiable from that average depending on RQA metric. Applause and Cheer were observed to *not* be significantly

Table 6
Estimated marginal means

Recurrence Rate	emmean	SE	df	lower.CL	upper.CL
Angry Noise	0.06	0.01	265	0.04	0.08
Applause	0.08	0.01	265	0.06	0.11
Cheer	0.10	0.01	265	0.08	0.11
Distraction Noise	0.07	0.01	265	0.06	0.08
Negative Chant	0.12	0.01	265	0.10	0.15
Positive Chant	0.17	0.01	265	0.16	0.18
Confidence level used: 0.95					
Determinism	**emmean**	**SE**	**df**	**lower.CL**	**upper.CL**
Angry Noise	0.25	0.04	265	0.17	0.33
Applause	0.37	0.04	265	0.29	0.45
Cheer	0.39	0.02	265	0.34	0.44
Distraction Noise	0.34	0.02	265	0.31	0.37
Negative Chant	0.52	0.04	265	0.44	0.61
Positive Chant	0.64	0.02	265	0.60	0.68
Confidence level used: 0.95					
Entropy	**emmean**	**SE**	**df**	**lower.CL**	**upper.CL**
Angry Noise	0.72	0.07	265	0.59	0.85
Applause	0.94	0.07	265	0.80	1.07
Cheer	0.96	0.04	265	0.89	1.04
Distraction Noise	0.92	0.03	265	0.87	0.97
Negative Chant	1.17	0.07	265	1.03	1.31
Positive Chant	1.32	0.03	265	1.26	1.39
Confidence level used: 0.95					
Laminarity	**emmean**	**SE**	**df**	**lower.CL**	**upper.CL**
Angry Noise	0.40	0.04	265	0.32	0.48
Applause	0.52	0.04	265	0.44	0.60
Cheer	0.51	0.02	265	0.47	0.56
Distraction Noise	0.45	0.01	265	0.42	0.48
Negative Chant	0.65	0.04	265	0.57	0.73
Positive Chant	0.75	0.02	265	0.71	0.79
Confidence level used: 0.95					
Game 2					
Recurrence Rate	emmean	SE	df	lower.CL	upper.CL
Angry Noise	0.04	0.01	251	0.02	0.05
Applause	0.06	0.01	251	0.05	0.0730
Cheer	0.04	0.00	251	0.04	0.05
Distraction Noise	0.06	0.00	251	0.05	0.06
Positive Chant	0.10	0.00	251	0.09	0.10
Confidence level used: 0.95					

(*Continued*)

Table 6
(Continued)

Determinism	emmean	SE	df	lower.CL	upper.CL
Angry Noise	0.15	0.04	251	0.08	0.22
Applause	0.28	0.03	251	0.23	0.34
Cheer	0.23	0.02	251	0.19	0.26
Distraction Noise	0.31	0.02	251	0.28	0.35
Positive Chant	0.60	0.02	251	0.56	0.64
Confidence level used: 0.95					
Entropy	**emmean**	**SE**	**df**	**lower.CL**	**upper.CL**
Angry Noise	0.56	0.08	251	0.41	0.72
Applause	0.74	0.07	251	0.62	0.87
Cheer	0.77	0.04	251	0.70	0.85
Distraction Noise	0.98	0.04	251	0.90	1.05
Positive Chant	1.39	0.04	251	1.30	1.47
Confidence level used: 0.95					
Laminarity	**emmean**	**SE**	**df**	**lower.CL**	**upper.CL**
Angry Noise	0.24	0.04	251	0.17	0.31
Applause	0.42	0.03	251	0.37	0.48
Cheer	0.32	0.02	251	0.28	0.35
Distraction Noise	0.28	0.02	251	0.25	0.31
Positive Chant	0.72	0.02	251	0.68	0.76
Confidence level used: 0.95					

differentiated from average crowd sound dynamics in Game 1 (as well as Negative Chant, but only for RR); Applause was observed to *not* be significantly differentiated from average crowd sound dynamics in Game 2 across all categories, as well as Distraction Noise for RR and DET.

Overall, the linear regression and pairwise comparison results support the hypothesis that chant categories will exhibit more coordination than non-chant categories, as observed by larger positive deviations from the grand average of each RQA metric for chant categories, and significant differences from other categories in pairwise comparisons. This is predicted because chanting requires coordination among the crowd as a whole to produce distinguishable acoustic patterns. However, the prediction that there may be differences in cheering and applause when compared to the noise categories is not strongly supported from these analyses. It was predicted that cheering/applause, which involve local coordination with individuals nearby, would show less recurrence than chant but more than distraction or angry noise, which produces noise with little structured variability in the acoustic signal. Indeed, applause and cheer appear to be near to the global average coordination dynamics of a crowd throughout a game. It appears from these analyses that, although there are strong differences between chant and other categories, there may not be a strong difference between cheering/applause and distraction/angry noise—distraction noise significantly differed from other nonchant categories only for Game 2. Thus, it remains unclear whether RQA metrics reflect *three* distinct modes

of acoustic patterning (from most coordination in chant categories, to some in applause/cheer, to least in noise categories), or perhaps *two* modes: chant categories with the most recurrent structures and nonchant categories with the least.

3. Part 2: SVM classification with RQA features

The objective in Part 2 was to use a data-driven approach to explore the usefulness of the full suite of all 19 RQA metrics computed by PyRQA as features for the classification of different crowd sounds. As stated previously, we predict that the machine learning analysis will show that RQA features can clearly distinguish between the acoustic patterns that correspond to chanting versus cheering versus distracting when applied to unseen data. This hypothesis addresses a separate question from Part 1, which sought primarily to describe the acoustical crowd sound dynamics observed in these two basketball games. The question now for Part 2 is whether the correct crowd sound category can be predicted from RQA features alone. That is, if RQA features showing higher recurrence rate, determinism, entropy, and laminarity (reflecting higher percentages of recurring behavior; more stable sequences of behavior; more variation in possible sequences of behavior; and increased clustering of behavioral states, respectively) correlate to instances of strong coordination and synchronization within the crowd, then the classifier should be able to accurately label new, unseen instances with high RQA values as chant behavior. If intermediate values of these RQA features correlate with instances of intermittent coordination, then the classifier should be able to accurately label new, unseen instances as applause or cheering. Finally, if the lowest RQA values correlate with instances of inhibited synchronization, then the classifier should be able to accurately label new, unseen instances with the lowest RQA values as noise.

3.1. Training, validation, and testing samples

The preprocessing steps for Part 2 were identical to those in Part 1. That is, the audio signal was resampled from 50 to 5 kHz. However, as we will describe below, we used sliding windows to extract samples from this audio signal across the entire length of each crowd sound event (in contrast to nonoverlapping windows in Part 1). This allowed us to create a larger data set for training the classifier.

For each crowd sound event, 5-s samples were extracted sliding by 1-s windows at a time. This means a 7-s crowd event can generate three samples instead of a singular sample while discarding the remaining 2 s. In order to prevent over fitting, we used disjoint events for each class in the training, validation, and test sets. That is, we ensured that if any samples from a unique crowd sound event were contained in the training set, no other samples from this event were contained in either the validation or testing sets. This assures that any success in classification performance cannot be due to the similarity between two consecutive samples within the same crowd sound event. As is standard in machine learning practice, we bootstrapped samples in the training set (oversampling the minority classes) until each class had 281 samples in Game 1, and 171 samples in Game 2. This method of oversampling

allows us to address bias in parameter estimates for minority classes and minimizes any error in classification due to unbalanced classes (Mohammed, Rawashdeh, & Abdullah, 2020). Testing and validation samples were not bootstrapped.

3.2. Recurrence quantification analysis

RQA was run on the time series data extracted from the resampled audio independently for each of these 5-s samples using identical parameters as listed in Part 1, Section 2.1. However, while statistical analysis in Part 1 focused on four RQA metrics commonly studied in behavioral experiments of human interaction, the training and classification for Part 2 was computed on all 19 RQA metrics reported by PyRQA and listed below (Rawald et al., 2017):

- Minimum diagonal line length
- Maximum diagonal line length
- Minimum white vertical line length
- Recurrence rate
- Determinism
- Average diagonal line length
- Longest diagonal line length
- Divergence
- Entropy diagonal lines
- Laminarity
- Trapping time
- Longest vertical line length
- Entropy vertical lines
- Average white vertical line length
- Longest white vertical line length
- Longest white vertical line length divergence
- Entropy white vertical lines
- Ratio determinism/recurrence rate
- Ratio laminarity/determinism

3.3. SVM classification

To perform classification, we utilized an SVM classifier with an radial basis function (RBF) kernel. The SVM classifier partitions the n-dimensional feature space (in our case $n = 19$) using hyperplanes to best distinguish the data based on class. When trained on the training set, the SVM classifier learns the optimal hyperplanes (or decision boundaries) to separate different classes of data based on their features. In this case, separating samples of crowd sound based on their RQA features and associated category labels. The SVM separates the classes by maximizing the distance between the hyperplanes and opposing classes. That is, by creating the largest possible gap between the categories of crowd sound. Once the classifier has learned these categories from the training set, we can introduce new, unseen test data. When shown a sample from the new test data, the SVM classifies the new sample based on the partition it belongs to.

In this analysis, we look at two different classification problems. In the first problem, we look at 19 RQA metrics over the six and five crowd sound classes defined in the analysis: angry noise, distraction noise, positive chant, negative chant (for Game 1 only), cheer, and applause. We also looked at a second classification problem to further probe our original hypothesis predicting three distinct coordination modes. If there are three distinct

Table 7
Overlapping 5-s samples used for RQA and SVM classification

	# of samples	
Crowd Sound Category	Game 1	Game 2
Angry Noise	37	31 (smallest class)
Applause	19 (smallest class)	36
Cheer	97	155
Distraction Noise	281 (largest class)	171 (largest class)
Negative Chant	26	0
Positive Chant	142	144

Table 8
Overlapping 5-s samples used for SVM classification after combining crowd sound classes

	# of samples	
Conglomerated Category	Game 1	Game 2
Applause/Cheer	116	191
Distraction Noise	281	171
Chant	168	144

coordination modes—strong coordination, intermittent coordination, and inhibited coordination—rather than six and five distinct coordination modes corresponding to the six and five manually labeled categories of both games, then we should observe improved classification (minimal confusion) between these three coordination categories. We combined crowd sound classes that showed a similar performance in the SVM classifier (trained on all 19 RQA metrics), and that showed an overlap in the original means comparisons of the subset of four theoretically motivated RQA metrics described in Section 2. Thus, we combined positive/negative chant into a singular chant class, cheer/applause into a singular cheer class, and kept distraction noise in a class by itself. We excluded angry noise from the conglomerated classes due to comparatively low prevalence in the data.

Table 7 lists the distribution of samples in the training set in each of the six original classes. Training samples were bootstrapped (oversampling the minority classes—a commonly accepted practice in machine learning to minimize error in classification due to unbalanced classes (Mohammed et al., 2020)), so each class has 281 samples for Game 1 and 171 samples for Game 2. Validation and testing samples were not bootstrapped.

Table 8 lists the distribution of samples in the training set by conglomerated class. As above, training samples were bootstrapped (oversampling the minority classes), so each class has 281 samples for Game 1 and 191 samples for Game 2. Testing samples were not bootstrapped.

3.4. Classification results

Results of the SVM classifier performance on the test set data are displayed in the confusion matrices in Fig. 6. Here, we have the true label on the y axis and the predicted label on the

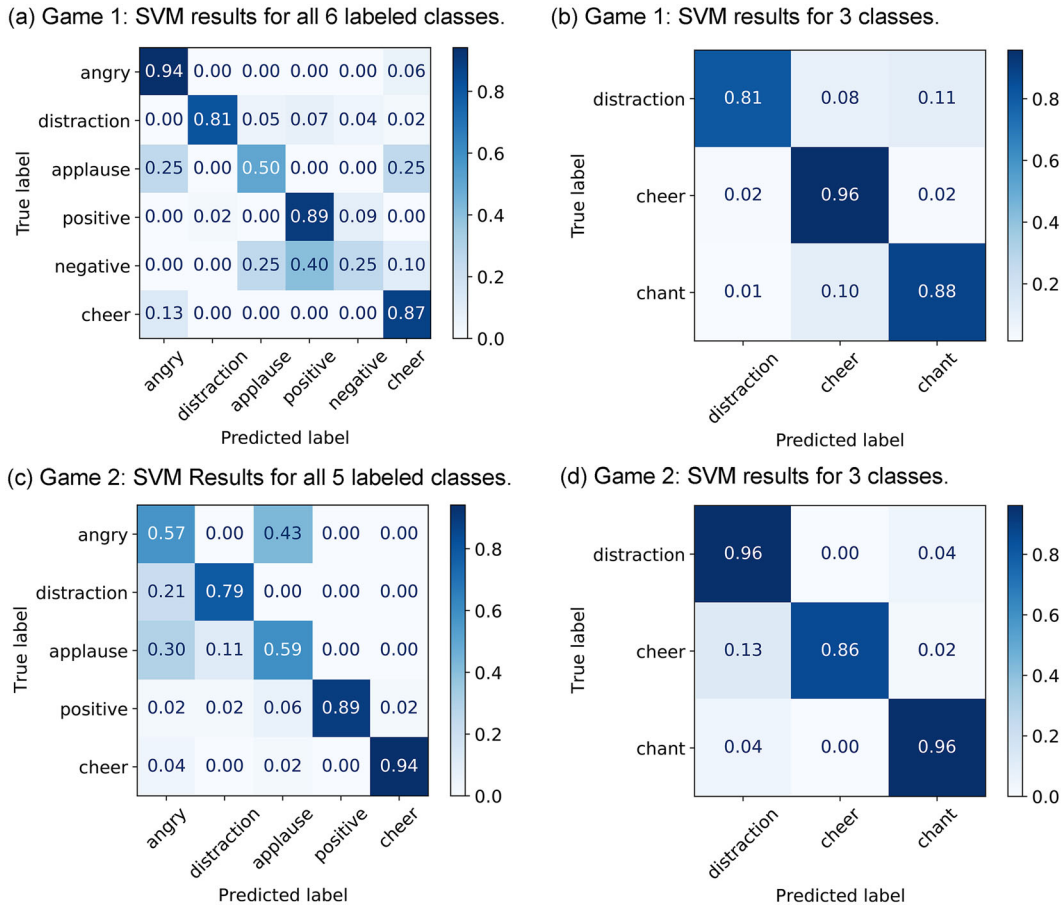

Fig. 6. SVM test set results for Game 1 (top) and Game 2 (bottom). Angry Noise was excluded from the three-class model. Applause/Cheer were combined into a single Cheer category. Positive and Negative Chant were combined into a single Chant category for Game 1. There were no instances of Negative Chant in Game 2. Three-class SVM results correspond to three proposed coordination modes: least coordinated (distraction noise), somewhat coordinated (cheer/applause), and most coordinated (positive and negative chant).

x axis. The values have been normalized across the true label, since each class contains a different number of samples because we did not bootstrap the test set. When analyzing the results of the SVM trained on six classes of data from Game 1 (Fig. 6A), we see strong distinction (above 0.8) for angry noise, distraction noise, positive chant, and cheer. We see confusions for applause with mis-classifications split between cheer and angry noise. We also see significant confusion for negative chant with a sizeable portion being classified as positive chant. When analyzing the results of the SVM trained on six classes of data from Game 2 (Fig. 6C), we see strong distinction (0.8 or above) for positive chant and cheer, with distraction noise close behind (0.79). We see confusions for angry noise and applause with mis-classification for each-other, and slight confusion for distraction noise with mis-

classification as angry noise. This indicates that crowd sound behavior in Game 2, compared to Game 1, may have been less distinct across categories.

Given our theoretical hypothesis of three distinct behavioral modes (most coordinated, somewhat coordinated, and least coordinated) and the overlap in expected marginal means from statistical analysis performed earlier, we joined classes into three subgroups: chant (most coordinated), cheer/applause (somewhat coordinated), and distraction noise (least coordinated). When we train an SVM classifier on these joined classes, we see performance above 0.8 for all classes in both Game 1 (Fig. 6B) and Game 2 (Fig. 6D), with stronger performance for classifying cheer and chant compared to distraction noise in Game 1, and stronger performance for classifying chant and distraction noise than cheer in Game 2. This indicates that the features generated by RQA are useful in predicting among broad crowd sound categories that correspond roughly to three coordination modes in classification of unseen crowd sound data.

4. Discussion

In this project, we sought to expand the application of nonlinear analysis techniques to naturalistic collective human interaction. Much work has been done to study the dynamics of human interaction in a variety of modalities, usually with the ability to record multiple signals that are generated by and recorded from each individual within a collective group (c.f. Baranowski-Pinto et al., 2022; Swarbrick et al., 2019; Gordon et al., 2020; Høffding et al., 2023). However, the ability to study human interaction in a greater variety of real-world, ecological settings can be hindered by this need to record individual-level signals. Proksch et al. (2022) demonstrated the feasibility of applying nonlinear analysis techniques to real-world social interactions that were recorded in the form of a single global-audio signal. This work showed that two coordination regimes (uncoordinated and coordinated) can be reliably differentiated using RQA. Here, we extend that previous work to a less scripted social interaction demonstrating a wider variety of interpersonal coordination regimes—the acoustical interaction of a crowd at a basketball game. To that end, we applied RQA to crowd sounds recorded from the student section of two basketball games and labeled according to a set of crowd sound categories.

We hypothesized that we would observe three overall coordination modes in the acoustical behavior of crowds attending a basketball game. Chanting behavior requires whole-crowd coordination to produce specific words for semi-rehearsed chants, and would exhibit the highest levels of coordination. Cheering and applause involves locally coordinated patterns of clapping or vocalizing within an individual's earshot and eyeshot, but this behavior will vary across the crowd as a whole, resulting in some coordination, but less than chanting. Finally, noise-making behaviors generated by a crowd in an effort to distract players on the court will result in little structured variability in the acoustic signal, and will exhibit the lowest level of coordination.

We hypothesized that these coordination modes would be reflected in nonlinear RQA metrics which measure the recurrence of behavior over time (Recurrence Rate), sequences of behavior over time (Determinism), variation in these sequences (Entropy), and revisiting the

same behavioral states in time (Laminarity). Initial statistical analysis of RQA metrics previously used to study acoustical coordination of a large interacting group (Recurrence Rate, Determinism, Entropy, and Laminarity, c.f. Proksch et al., 2022) indicate that chant categories do indeed exhibit higher RQA metrics than other categories of acoustical crowd behavior. However, it is less clear from these analyses whether there are distinct semi-coordinated and low-coordinated categories of acoustical crowd behavior, or merely two overarching categories for coordinated and relatively uncoordinated behavior.

In addition to the descriptive question asked in Part 1, we asked a practical question of whether unseen samples of crowd sound could be reliably predicted as belonging to distinct crowd sound categories in Part 2. Across the 6 and 5 labeled categories, we again observed chant categories being the most accurately predicted. However, dividing classes into three categories representing the most (chant), some (applause/cheer), and least (distraction noise) coordinated acoustical behavior did result in a high classification accuracy (above 0.8) across the three coordination modes.

The results here and from Proksch et al. (2022) indicate that RQA (and phase space reconstruction) can be meaningfully applied to global acoustic recordings when individual recordings are not available. Resulting RQA metrics reflect specific coordination modes hypothesized to exist in the crowd, from the most coordination/synchronization in instances of chanting to the least synchronization in instances of noise. The collegiate basketball crowd sound data sets analyzed here are exemplars of naturalistic crowd interactions that are not dictated by a set of predetermined instructions, like in a musical score. While crowd behavior is influenced by the events of the basketball game over time, the specific acoustic output of the crowd is neither rehearsed nor explicitly dictated via shared access to a "behavioral score." Instead, individuals of the crowd "softly assemble" into certain functional patterns acoustic of behavior, which emerge from local interactions among fans and are influenced by external events of the environment. Further, we have demonstrated that a combination of RQA and SVM classifier can effectively differentiate between at least a subset of acoustic crowd responses. Future research involving classification of acoustic crowd behavior can test whether classification accuracy of SVMs, Naive Bayes Classifiers, or even Convolutional Neural Networks—which are typically trained on more standard measures of acoustic analysis (such as spectral features and mel cepstral coefficients)—may be improved by incorporating numerical RQA metrics or even Recurrence Plot images into the classification.

This work adds to the growing body of research on joint action and coordination among groups, and the need to combine methods from both coordination dynamics and dynamical systems (Schiavio et al., 2022). We discussed a collection of studies examining physiological and behavioral synchrony arising from group interaction through the analysis of individual signals from members of interacting groups. Baranowski-Pinto et al. (2022) found heart-rate interdependence between fans attending live (vs. televised) basketball games, but did not measure the synchrony in behavioral (i.e., movement or acoustic) dynamics. Swarbrick et al. (2019) demonstrated that movement vigor and engagement was enhanced by attending a live concert compared to a prerecorded concert, but did not analyze the coordination in these dynamics between audience members. And Gordon et al. (2020) demonstrated an independence between behavioral and physiological synchrony—such that physiological and

behavioral synchrony are not always coordinated. Going forward, we argue that incorporating analysis of coordination measured from global acoustical signals (as shown in our work) to group interaction studies such as these will help shed light on the role of acoustical behavior in joint action, and whether that role is meaningfully correlated with movement, physiological, and psychological dynamics.

To conclude, we sought to describe emergent coordination dynamics in the acoustical behavior of a crowd in a naturalistic setting. We presented a case study of acoustical behavior of a crowd at two collegiate men's basketball games. Specifically, we performed phase space reconstruction and RQA on acoustic data recorded from fans attending the two basketball games. While there was an overlap in some categories, we found reliable differences in recurrence measures after SVM classification for three conglomerated categories of crowd activity (i.e., chant [most coordinated], cheer [somewhat coordinated], and distraction noise [least coordinated]). There is likely substantial individual variability between basketball games in part dependent on variables, such as type of game (Men's vs. Women's game), attendance, team performance, and even within a specific game between halves or between specific plays. Further research should include data from a larger variety of basketball games to evaluate any trends in basketball crowd behavior overall, and to evaluate differences based on these listed variables. In the future, it would be beneficial to analyze how these recurrence measures extend to additional basketball games or to acoustical behavior of crowds at different sporting events. Analyzing these signals over the time course of a game may shed light into how joint acoustical behavior changes over time. Further, it would be insightful to relate these acoustical behavioral dynamics with coordination dynamics across modalities. The presentation of the case study and extension of dynamical systems methods presented here will enhance research into the acoustical behavior of large group interpersonal coordination—particularly when researchers lack access to record and analyze individual-level signals.

Author contributions

S.P. devised the project, the main conceptual ideas, and led writing the manuscript. M.R. performed data cleaning and preparation. S.P. and M.R. performed the nonlinear analysis, S.P performed statistical analysis. M.R. performed machine learning analysis. K.G. and M.T. collected and provided the original data set, and consulted on analysis methods and results. R.B. and C.K. verified the analytical methods and supervised the findings of this work. All authors discussed the results and contributed to the final manuscript.

Acknowledgments

We thank Sean Warnick (Computer Science, Brigham Young University) for assistance in providing the basketball crowd sound data set and for early discussions during this project.

This material is based upon work supported by the National Science Foundation under Grant No. DGE-1633722. Any opinions, findings, and conclusions or recommendations

expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Citation diversity statement

Recent work in several fields of science has identified a bias in citation practices such that papers from women and other minority scholars are under-cited relative to the number of such papers in the field (Bertolero et al., 2020; Caplar, Tacchella, & Birrer, 2017; Chatterjee & Werner, 2021; Dion, Sumner, & Mitchell, 2018; Dworkin et al., 2020; Fulvio, Akinnola, & Postle, 2021; Mitchell, Lange, & Brus, 2013; Maliniak, Powers, & Walter, 2013; Wang et al., 2021). Here, we sought to proactively consider choosing references that reflect the diversity of the field in thought, form of contribution, gender, race, ethnicity, and other factors. First, we obtained the predicted gender of the first and last author of each reference by using databases that store the probability of a first name being carried by a woman (Dworkin et al., 2020; Zhou et al., 2020). By this measure (and excluding self-citations to the first and last authors of our current paper), our references contain 4.55% woman(first)/woman(last), 11.36% man/woman, 25.87% woman/man, and 58.23% man/man. This method is limited in that (a) names, pronouns, and social media profiles used to construct the databases may not, in every case, be indicative of gender identity and (b) it cannot account for intersex, non-binary, or transgender people. Second, we obtained predicted racial/ethnic category of the first and last author of each reference by databases that store the probability of a first and last name being carried by an author of color (Ambekar, Ward, Mohammed, Male, & Skiena, 2009; Sood & Laohaprapanon, 2018). By this measure (and excluding self-citations), our references contain 14.53% author of color (first)/author of color (last), 12.31% white author/author of color, 19.09% author of color/white author, and 54.08% white author/white author. This method is limited in that (a) names and Florida Voter Data to make the predictions may not be indicative of racial/ethnic identity, and (b) it cannot account for Indigenous and mixed-race authors, or those who may face differential biases due to the ambiguous racialization or ethnicization of their names. We look forward to future work that could help us to better understand how to support equitable practices in science.

Open Research Badges

This article has earned Open Data and Open Materials badges. Data and materials are available at https://osf.io/t2shj/?view_only=db68b2d885ab4fc68956340cb2ac8617

References

Albarracin, M., Constant, A., Friston, K. J., & Ramstead, M. J. D. (2021). A variational approach to scripts. *Frontiers in Psychology*, *12*, 58549.

Ambekar, A., Ward, C., Mohammed, J., Male, S., & Skiena, S. (2009). Name-ethnicity classification from open sources. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 49–58).

Baranowski-Pinto, G., Profeta, V. L., Newson, M., Whitehouse, H., & Xygalatas, D. (2022). Being in a crowd bonds people via physiological synchrony. *Scientific Reports*, *12*(1) .

Bertolero, M. A., Dworkin, J. D., David, S. U., Lloreda, C. L., Srivastava, P., Stiso, J., Zhou, D., Dzirasa, K., Fair, D. A., Kaczkurkin, A. N., Marlin, B. J., Shohamy, D., Uddin, L. Q., Zurn, P., & Bassett, D. S. (2020). Racial and ethnic imbalance in neuroscience reference lists and intersections with gender. *bioRxiv*. https://doi.org/10.1101/2020.10.12.336230

Blake, E. C., & Cross, I. (2015). The acoustic and auditory contexts of human behavior. *Current Anthropology*, *56*(1), 81–103.

Butler, B. A., Pedersen, K., Cook, M. R., Wadsworth, S. G., Todd, E., Stark, D., Gee, K. L., Transtrum, M. K., & Warnick, S. (2018). Classifying crowd behavior at collegiate basketball games using acoustic data. In *Proceedings of Meetings on Acoustics* (volume 35).

Caplar, N., Tacchella, S., & Birrer, S. (2017). Quantitative evaluation of gender bias in astronomical publications from citation counts. *Nature Astronomy*, *1*(6), 0141.

Chatterjee, P., & Werner, R. M. (2021). Gender disparity in citations in high-impact journal articles. *JAMA Network Open*, *4*(7), e2114509.

Connell, J. P., DiMercurio, A., & Corbetta, D. (2017). Dynamic systems theory. In J. Vonk& T. Shackelford (Eds.), *Encyclopedia of Animal Cognition and Behavior* (pp. 1–8). New York: Springer International Publishing

Cummins, F. (2013). Joint speech: The missing link between speech and music? *PERCEPTA-Revista de Cognição Musical*, *1*(1), 17.

Cummins, F. (2020). The territory between speech and song: A joint speech perspective. *Music Perception*, *37*(4), 347–358.

Demos, A. P., & Palmer, C. (2022). Social and nonlinear dynamics unite: Musical group synchrony. *Trends in Cognitive Sciences*.

Dion, M. L., Sumner, J. L., & Mitchell, S. M. (2018). Gendered citation patterns across political science and social science methodology fields. *Political Analysis*, *26*(3), 312–327.

dos Santos, L., Barroso, J. J., Godoy, M. F. d., Macau, E. E., & Freitas, U. S. (2014). Recurrence quantification analysis as a tool for discrimination among different dynamics classes: The heart rate variability associated to different age groups. In N. Marwan, M. Riley, A. Giuliani & C. L. Webber (Eds.), *Translational recurrences: from mathematical theory to real-world applications* (pp. 125–136). Springer.

Dworkin, J. D., Linn, K. A., Teich, E. G., Zurn, P., Shinohara, R. T., & Bassett, D. S. (2020). The extent and drivers of gender imbalance in neuroscience reference lists. *Nature Neuroscience*, *23*, 918–926.

Eckmann, J.-P., Kamphorst, S. O., & Ruelle, D. (1987). Recurrence plots of dynamical systems. *Europhysics Letters*, *4*(9), 973–977.

Farkas, I., Helbing, D., & Vicsek, T. (2002). Mexican waves in an excitable medium. *Nature*, *419*(6903), 131–132.

Farrera, A., & Ramos-Fernández, G. (2022). Collective rhythm as an emergent property during human social coordination. *Frontiers in Psychology*, *12*, 772262.

Fulvio, J. M., Akinnola, I., & Postle, B. R. (2021). Gender (im)balance in citation practices in cognitive neuroscience. *Journal of Cognitive Neuroscience*, *33*(1), 3–7.

Fusaroli, R., & Tylén, K. (2016). Investigating conversational dynamics: Interactive alignment, interpersonal synergy, and collective task performance. *Cognitive Science*, *40*, 145–171.

Gordon, I., Gilboa, A., Cohen, S., & Kleinfeld, T. (2020). The relationship between physiological synchrony and motion energy synchrony during a joint group drumming task. *Physiology & Behavior*, *224*, 113074.

Høffding, S., Yi, W., Lippert, E., Sanchez, V. G., Bishop, L., Laeng, B., Danielsen, A., Jensenius, A. R., & Wallot, S. (2023). Into the hive-mind: Shared absorption and cardiac interrelations in expert and student string quartets. *Music & Science*, *6*, 20592043231168597.

Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition*, *27*(6), 949–960.

Kello, C. T., & Van Orden, G. C. (2009). Soft-assembly of sensorimotor function. *Nonlinear Dynamics, Psychology, and Life Sciences*, *13*, 57–78.

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, *31*(5), 354–364.

Kok, V. J., Lim, M. K., & Chan, C. S. (2016). Crowd behavior analysis: A review where physics meets biology. *Neurocomputing*, *177*, 342–362.

Lenth, R. V. (2022). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.7.5.

Maliniak, D., Powers, R., & Walter, B. F. (2013). The gender citation gap in international relations. *International Organization*, *67*(4), 889–922.

Mann, R. P., Faria, J., Sumpter, D. J., & Krause, J. (2013). The dynamics of audience applause. *Journal of the Royal Society Interface*, *10*(85), 20130466.

Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science*, *1*(2), 320–339.

Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, *438*, 237–329.

Marwan, N., Wessel, N., Meyerfeld, U., Schirdewan, A., & Kurths, J. (2002). Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. *Physical Review E*, *66*(2), 026702.

McNeill, W. H. (2022). *Keeping together in time*. Harvard University Press.

Michard, Q., & Bouchaud, J.-P. (2005). Theory of collective opinion shifts: From smooth trends to abrupt swings. *European Physical Journal B-Condensed Matter and Complex Systems*, *47*(1), 151–159.

Mitchell, S. M., Lange, S., & Brus, H. (2013). Gendered citation patterns in international relations journals. *International Studies Perspectives*, *14*(4), 485–492.

Mogan, R., Fischer, R., & Bulbulia, J. A. (2017). To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect. *Journal of Experimental Social Psychology*, *72*, 13–20.

Mohammed, R., Rawashdeh, J., & Abdullah, M. (2020). Machine learning with oversampling and undersampling techniques: Overview study and experimental results. In *2020 11th International Conference on Information and Communication Systems (ICICS)* (pp. 243–248). IEEE.

Néda, Z., Ravasz, E., Bréchet, Y., Vicsek, T., & Barabási, A.-L. (2000). The sound of many hands clapping. *Nature*, *403*, 849–850.

Paxton, A., & Dale, R. (2013). Argument disrupts interpersonal synchrony. *Quarterly Journal of Experimental Psychology*, *66*, 2092–2102.

Proksch, S., Reeves, M., Spivey, M., & Balasubramaniam, R. (2022). Coordination dynamics of multi-agent interaction in a musical ensemble. *Scientific Reports*, *12*(1), 421.

Rawald, T., Sips, M., & Marwan, N. (2017). PyRQA - Conducting recurrence quantification analysis on very long time series efficiently. *Computers and Geosciences*, *104*, 101–108.

Reddy, V. R., Sinha, A., & Seshadri, G. (2013). Fusion of spectral and time domain features for crowd noise classification system. In *2013 13th International Conference on Intelligent Systems Design and Applications* (pp. 1–6). IEEE.

Riley, M., Richardson, M., Shockley, K., & Ramenzoni, V. (2011). Interpersonal synergies. *Frontiers in Psychology*, *2*, 38.

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*, 104038.

Schiavio, A., Maes, P.-J., & van der Schyff, D. (2022). The dynamics of musical participation. *Musicae Scientiae*, *26*(3), 604–626.

Schinkel, S., Dimigen, O., & Marwan, N. (2008). Selection of recurrence threshold for signal detection. *European Physical Journal Special Topics*, *164*(1), 45–53.

Singh, P. P., & Rani, P. (2014). An approach to extract feature using MFCC. *IOSR Journal of Engineering*, *4*(8), 21–25.

Sood, G., & Laohaprapanon, S. (2018). Predicting race and ethnicity from the sequence of characters in a name. *arXiv preprint arXiv:1805.02109*.

Swarbrick, D., Bosnyak, D., Livingstone, S. R., Bansal, J., Marsh-Rollo, S., Woolhouse, M. H., & Trainor, L. J. (2019). How live music moves us: Head movement differences in audiences to live versus recorded music. *Frontiers in Psychology*, *9*, 2682.

Swathi, H., Shivakumar, G., & Mohana, H. (2017). Crowd behavior analysis: A survey. In *2017 International Conference on Recent Advances in Electronics and Communication Technology (ICRAECT)* (pp. 169–178).

Takens, F. (1981). Detecting strange attractors in turbulence. In D. Rand & L. S. Young (Eds.), *Dynamical Systems and Turbulence, Warwick 1980. Lecture Notes in Mathematics* (volume 898, p. 366). Berlin, Heidelberg: Springer.

Thiel, M., Romano, M. C., Kurths, J., Meucci, R., Allaria, E., & Arecchi, F. T. (2002). Influence of observational noise on the recurrence quantification analysis. *Physica D: Nonlinear Phenomena*, *171*(3), 138–152.

Vlachos, I., & Kugiumtzis, D. (2010). Nonuniform state-space reconstruction and coupling detection. *Physical Review E*, *82*, 016207.

Wallot, S., & Leonardi, G. (2018). Analyzing multivariate dynamics using cross-recurrence quantification analysis (CRQA), diagonal-cross-recurrence profiles (DCRP), and multidimensional recurrence quantification analysis (MdRQA)-A tutorial in R. *Frontiers in Psychology*, *9*, 2232.

Wang, X., Dworkin, J. D., Zhou, D., Stiso, J., Falk, E. B., Bassett, D. S., Zurn, P., & Lydon-Staley, D. M. (2021). Gendered citation practices in the field of communication. *Annals of the International Communication Association*.

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science*, *20*(1), 1–5.

Zhang, L., Liu, X., Zhang, X., & Bu, B. (2011). Audio classification algorithm based on nonlinear characteristics analysis. In *Proceedings of 2011 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*.

Zhou, D., Cornblath, E. J., Stiso, J., Teich, E. G., Dworkin, J. D., Blevins, A. S., & Bassett, D. S. (2020). Gender diversity statement and code notebook v1.0.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1: Pairwise comparison results for Game 1
Table S2: Pairwise comparison results for Game 2