Entropy signatures of protest phase states: A step towards predictive modeling of protest synchrony and violence

Jais Adam-Troian[1,2] , Alexander Demos [3], Vivek Kant[4]

[1]LPS (EA849), Aix-Marseille University, [2]European Institute, Istanbul Bilgi University, [3]Psychology Department, University of Illinois at Chicago, [4]Industrial Design Center , IIT-Bombay

Abstract

Background: Recent work on protest violence focuses on its antecedents, behavioral intentions or on visual scene analyses (e.g. for training machine learning algorithms). For the first time, we propose to study *in vivo* protest violence relying only on auditory scene analyses in a dynamical systems perspective. **Methods**: Multiscale entropy analyses were performed on audio bands from three 2018 protest videos in France (2 against labor law reforms, 1 Yellow Vests). **Results**: Compared to surrogate and synchronous but non-violent control events, violent events were preceded by lower entropy levels on high frequency bands, which correspond to perceived activity in noise. These results held across videos and were thus robust to different violent groups and political causes. **Implications**: The present research highlight a potentially unique property of sound entropy levels for predicting display of violent collective behavior and constitute a first step towards quantitative modelling of *in vivo* violent political action.

Introduction

From the beginning of 2017 to the end of 2018, more than 11,000 protests occurred in the United States alone, involving over 9.5 million attendees (Leung & Perkins, 2018). While most protests are peaceful and unfold without noticeable incidents, they can sometimes display violence outbursts or degenerate into riots (e.g. Yellow Vests).

Currently, behavioural science research has identified a range of social psychological predictors of protest attendance and intentions to use non normative means of protesting (e.g. Mooijman, Hoover, Lin, Ji & Dehghani, 2018). However, current research on protest violence suffers from three main limitations:

1- Most research focuses on identifying distal predictors (i.e. psychological predispositions) that might be largely moderated by immediate environmental features.

2- Addressing what gets individuals to attend the protest or focus on proxy measures (e.g. number of police arrests, number of deaths) is not be precise enough to depict what goes on *in vivo*.

3- An important part of the literature relies on self-report data, which is problematic for identifying true predictors of violence.

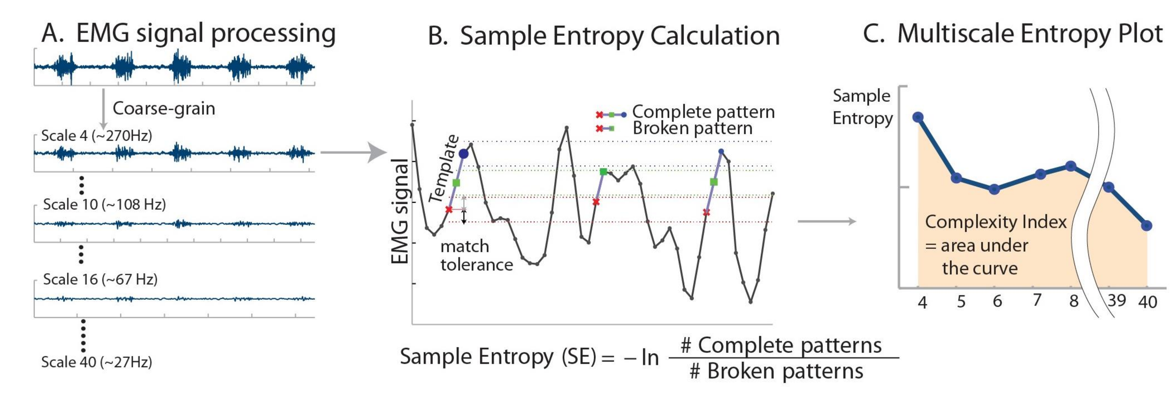
BUT

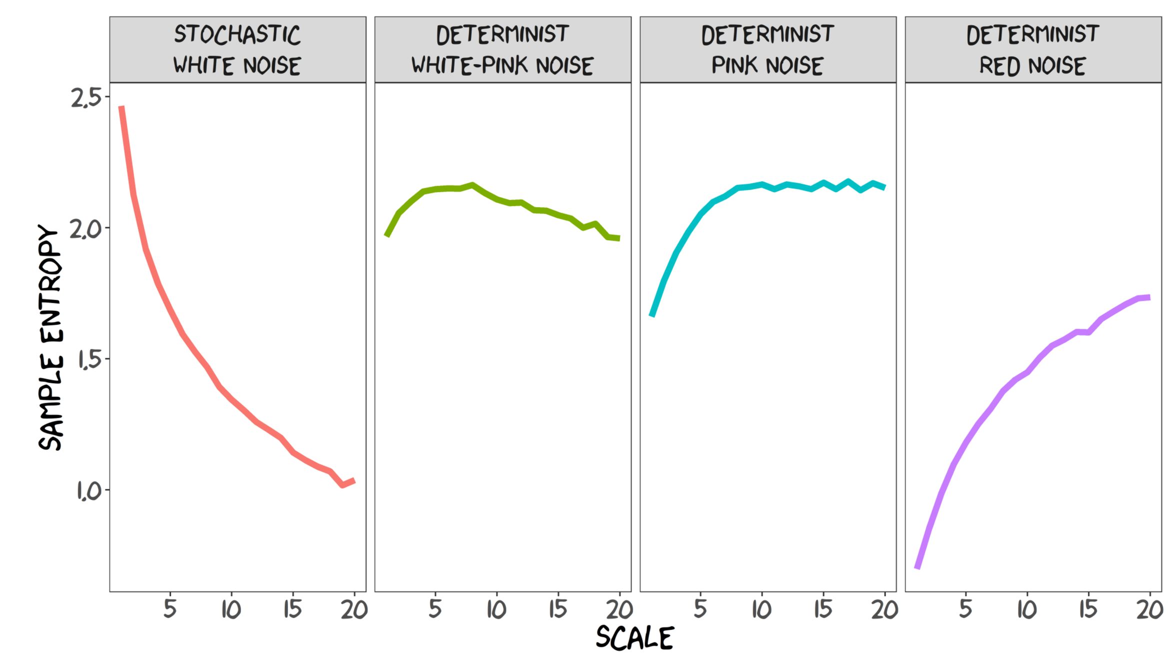
One can turn to statistical and theoretical approaches that integrate the dynamical properties of social systems to properly model protest violence, i.e. Dynamical Systems Theory (DST, see Guastello, Koopmans, & Pincus, 2009):

'DST treat interactions between components [of a system] *as time-evolving processes constrained by the context – physical, social, and mental. The focus is on change over time, requiring measures and statistical methods that are very different from those traditionally used in psychology*.' (Demos & Chaffin, 2017; p. 2).

A key property of dynamical systems is their sudden transitions from one stable state (phase) to another (called bifurcations). Thus, assuming that human coordination follows the laws of thermodynamics (dynamical systems), as we see in social insects and other natural system (Strogatz, 1993; Glass & Mackey, 1988), it becomes possible to conceptualize protests as dynamical systems that display bifurcations, transiting between disordered states to more stables ones (from marching in synchrony, to chanting and displaying violence).

In a DST framework, an important systemic indicator is *entropy* (i.e. the degree of complexity in a given signal). Low entropy in biological systems is associated with pathological states (Costa & Goldberger, 2015).





Methods

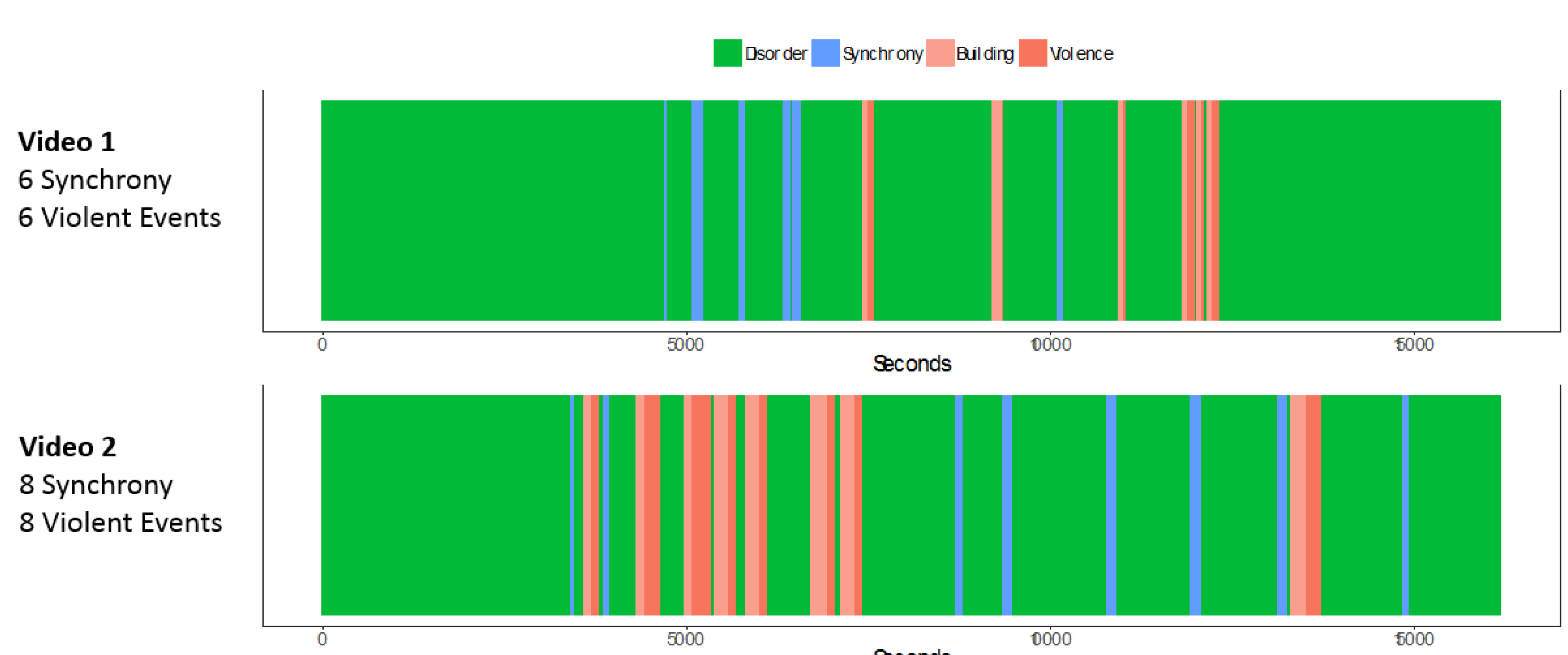
Example of violent (above) vs. synchronous (below, chanting) events from video 3.

Two videos of recent protest against labor reforms in France (approx. 8 hours) were selected on specific criteria (continuous filming, no intervention of film maker, no interviews of protesters).

Events were then identified and coded within each video (surrogates vs. synchronous non-violent vs. synchronous violent; N = 66; Median length: 142.5). Audio signals were then extracted from these events using MATLAB's MIR toolbox 1.7 and decomposed into 10 bands (Gammatone filterbank decomposition; Auditory Toolbox built in MIR), which simulate the response of the basilar membrane.

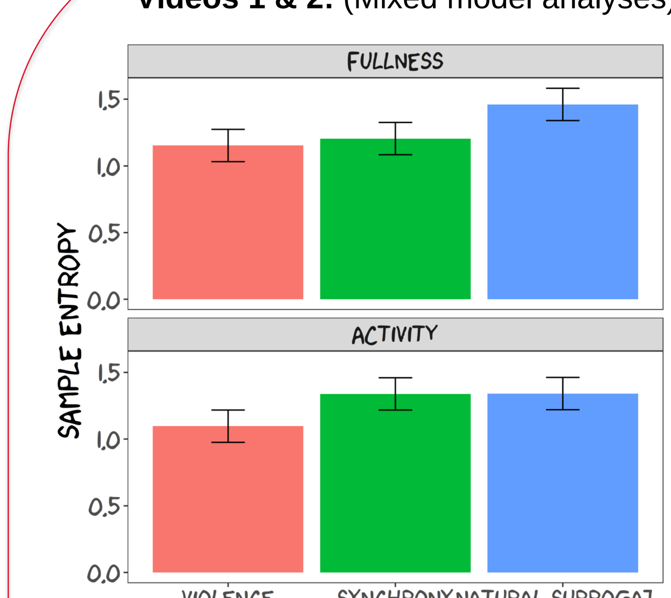
Bands corresponding to two psychoacoustic and perceptual properties were extracted: low frequency bands 2-3: 50-200 Hz (perceived as "Fullness" in sound), and high frequency ones 7-8: 1600-6400 Hz (perceived as "Activity" in sound). Root-Mean Squared of the signal (loudness) at 1/100 second with 50% overlap was then computed.

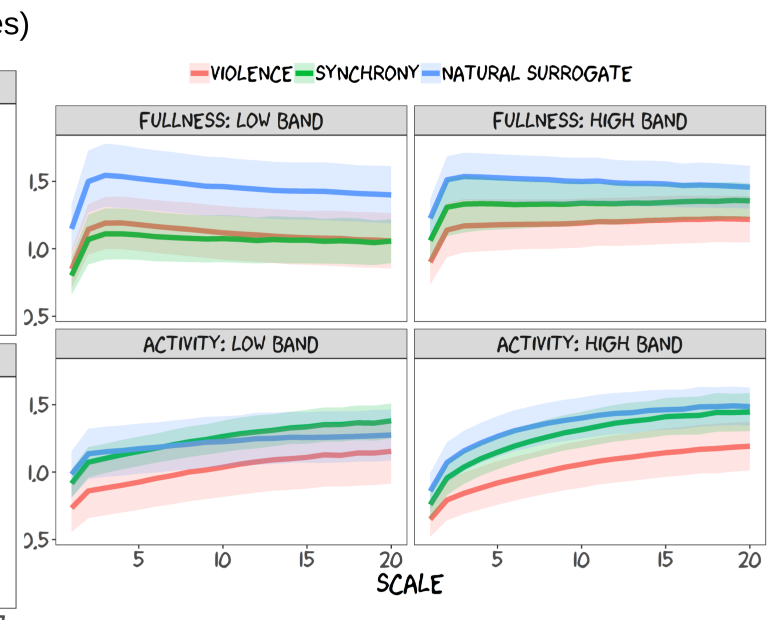
Multiscale Entropy of that signal was calculated using the PhysioNet toolbox (Costa, Goldberger, & Peng, 2005) on each frequency bands of interest (2,3,7,8). These indices were then computed for a third video filming a Yellow Vests protest in Paris for replication purposes (events N = 24; length approx. 8 hours).



Results

Videos 1 & 2: (Mixed model analyses)





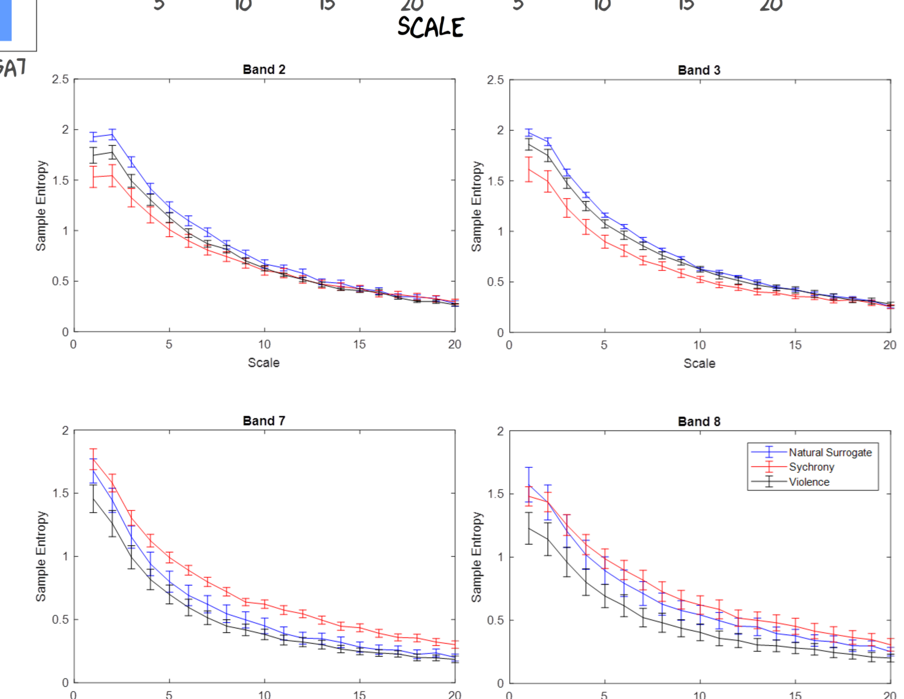
Fullness (Bands 2+3)

Surrogate > Synchrony, p =.02
Surrogate > Violence, p =.004
Synchrony = Violence, p =.83

Activity Bands (7+8)

Surrogate = Synchrony, p =.99
Surrogate > Violence, p =.02
Synchrony > Violence, p =.03

Video 3: (Replication)



Discussion

By taking a dynamical systems perspective on protest violence, we accounted for the role of emergent systemic states (index by entropy) and tested whether these can be used as reliable signatures of different protest behaviors. So far, most of the work investigating human social phenomena from a dynamical and ecological perspective (Gibson, 2014) relied on visual perception processes (e.g. Moussaïda, Helbing, & Theraulaza, 2011). Similarly, machine learning approaches to identification of violent crowd behaviour use mostly visual data (see Hassner, Itcher, & Kliper-Gross, 2012; Mohammadi, Perina, Kiani, & Murino, 2016).

Nevertheless, from a first person perspective to protest situations, one cannot help noticing that vision in protests is often obstructed and covers only a small range. Visual information transmission is thus slower and has the disadvantage of requiring directed attention. 'Eyes can be closed', ears cannot. Sound is our primary "danger detection" mechanism and stream segregation processes ensure that individuals know *where* the sound came from and *what* produced it (see Gaver, 1993; Kim, Zahorik, Carney, Bishop, & Kuwada, 2015). Moreover, sound is a key element that enables social connection (deaf individuals report more social isolation than blind individuals, see Mohr et al, 2000).

Some studies even show that only hearing a partners' *rhythmic* movements is sufficient to cause spontaneous synchronization of whole body movements (Demos et al., 2012; Demos & Chaffin, 2018). Furthermore, Auditory Scene analysis of protest contexts (Bregman, 1990; Bregman, 2007) makes it clear that sound in crowds is omnipresent and can be quickly transmitted with a large range. In sum, sound can constitute an important cue for individuals in protests. Because violent events are preceded by lower sound entropy levels, we argue systemic sound entropy may be a useful predictor of violence outbursts in protests.

.