2021

Pursuing Collective Synchrony in Teams: A Regime-Switching Dynamic Factor Model of Speed Similarity in Soccer

Daniel M. Smith

Theodore A. Walls

University of Rhode Island, walls@uri.edu

Follow this and additional works at: https://digitalcommons.uri.edu/psy_facpubs

The University of Rhode Island Faculty have made this article openly available.
Please let us know how Open Access to this research benefits you.

This is a pre-publication author manuscript of the final, published article.

Terms of Use

This article is made available under the terms and conditions applicable towards Open Access Policy Articles, as set forth in our [Terms of Use](#).

Citation/Publisher Attribution

Smith, D.M., Walls, T.A. Pursuing Collective Synchrony in Teams: A Regime-Switching Dynamic Factor Model of Speed Similarity in Soccer. *Psychometrika* (2021). <https://doi.org/10.1007/s11336-021-09782-1>
Available at: <https://doi.org/10.1007/s11336-021-09782-1>

This Article is brought to you for free and open access by the Psychology at DigitalCommons@URI. It has been accepted for inclusion in Psychology Faculty Publications by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu.

1 Pursuing Collective Synchrony in Teams: A Regime-Switching Dynamic Factor Model of
2 Speed Similarity in Soccer

3 Daniel M. Smith

4 Springfield College

5 Theodore A. Walls

6 University of Rhode Island

7 Author Note

8 Authors' contact details are as follows: 1) Smith: dsmith19@springfield.edu,
9 +1-413-748-3177, Department of Physical Education and Health Education, 263 Alden
10 Street, Springfield, MA 01109; 2) Walls: walls@uri.edu, +1-401-874-2105, Department of
11 Psychology, 306 Chafee Hall, 142 Flagg Road, Kingston, RI 02881

Abstract

Collective synchrony refers to the simultaneous occurrence of behavior, cognition, emotion, and/or physiology within teams of three or more persons. It has been suggested that collective synchrony may emanate from the copresence of team members, from their engagement in a shared task, and from coordination enacted in pursuit of a collective goal. In this paper, a regime-switching dynamic factor analytical approach is used to examine interindividual similarities in a particular behavioral measure (i.e., speed) in a collegiate soccer team. First, the analytical approach is presented didactically, including the state space modeling framework in general, followed by the regime-switching dynamic factor model in particular. Next, an empirical application of the approach is presented. Speed similarity (covariation in speed, operationalized in two ways: running cadence and distance covered) during competitive women's soccer games is examined. A key methodological aspect of the approach is that the collective is the unit of analysis, and individuals vary about collective dynamics and their evolution. Reporting on the results of this study, we show how features of substantive interest, such as the magnitude and prevalence of behavioral similarity, can be parameterized, interpreted, and aggregated. Finally, we highlight several key findings, as well as opportunities for future research, in terms of methodological and substantive aims for advancing the study of collective synchrony.

Keywords: intensive longitudinal data, state space models, time series

Pursuing Collective Synchrony in Teams: A Regime-Switching Dynamic Factor Model of
Speed Similarity in Soccer

During team performance, teammates will at times exhibit similarities in various behavioral, cognitive, emotional, and physiological states. Such simultaneous occurrences are known as *synchrony* (Feldman, 2003; Wiltermuth & Heath, 2009). Synchrony within a group of three or more persons, which typically applies in team performance settings, is referred to as *collective synchrony* (Smith, 2018). Teams have three key characteristics that may help explain why collective synchrony can occur. That is, during competition members of a team are *copresent* in the performance setting, they are engaged in a *shared task*, and they *coordinate* with one another to achieve a common goal. Copresence means that team members are proximal to each other and are susceptible to emotional contagion, defined as “the tendency to automatically mimic and synchronize expressions, vocalizations, postures, and movements with those of another person and, consequently, to converge emotionally” (Hatfield, Cacioppo, & Rapson, 1993). While engaging in a shared task, team members may match each other’s behaviors due to specific task constraints. For example, rowing either in-phase or anti-phase (essentially, turn taking) may be a task constraint in the context of crew. Similarly, the expectation of team members to move up and down a basketball court together as a team is a structural feature of the game which will inevitably produce some degree of collective synchrony in speed, direction of movement, and by extension due to the metabolic demands of this action, physiological outcomes such as heart rate. Finally, teammates may exhibit collective synchrony in part due to their coordination, that is, their arrangement of individual actions to achieve group goals (Cannon-Bowers, Tannenbaum, Salas, & Volpe, 1995). Coordination in team sports and some performing arts (e.g., dance) stems primarily from visual perception (Araújo, Davids, & Hristovski, 2006) and both verbal and nonverbal communication such as eye contact, pointing, and body position (Duarte, Araújo, Correia, & Davids, 2012). This may also contribute to collective synchrony.

In Smith (2018), we introduced a conceptual framework proposing that the above mentioned antecedents (copresence, shared task, and coordination) each may partially explain why collective synchrony in behavior (e.g., running speed) and physiology (e.g., heart rate) materializes. In an actual performance setting, more than one of these elements is likely to be at play. For example, collective synchrony in teammates’ heart rates may reflect common emotions to some extent, but this effect is certain to be confounded by the metabolic demands of the (shared) physical activity. A long-term aim may be to parse each antecedent in order to examine the role of each in team performance. The current study is focused on developing an analytical approach to investigating the interindividual covariation of behavioral time series (i.e., *behavioral similarity*), along with substantive aims of an exploratory nature.

In this paper, we use movement data from a collegiate women’s soccer team during competition to examine the extent to which team members demonstrate speed similarity, where speed is operationalized as running cadence (a measure of steps taken per unit time) and distance covered within defined time intervals. In our approach, “similarity” is modeled as team members’ covariation in speed (i.e., not as similarity in actual speed levels). Although speed similarity is not likely to be directly predictive of goals or wins, owing to the many variables that influence these outcomes, speed similarity is useful to analyze in the context of soccer because the tendency of players to move similarly is one important part of team performance. There is empirical evidence that behavioral synchrony may be particularly valuable when soccer teams defend (López-Felip, Davis, Frank, & Dixon, 2018). Hence, speed similarity may be one indicator of the effectiveness with which a team deploys defensive tactics. Conversely, other evidence suggests that movement synchrony exists in a team independent of ball possession (Duarte et al., 2013). To meet the substantive aims of this study, a finely tuned approach to inferential modeling was required. In this study, we explored the use of regime-switching dynamic factor analysis, within a state space modeling framework, in order to quantify speed similarity

and identify a team’s discrete changes between states of high similarity and no similarity. This approach enables quantifying speed similarity’s *prevalence*, that is, the proportion of time spent in a state of high similarity, and the *magnitude*, or extent, of similarity. In this paper, we demonstrate how these features may be interpreted from the results of a regime-switching dynamic factor analysis. Useful new conceptions for considering these types of phenomena, such as dwell time and stability of the similarity, are described.

Our methodological strategy is notable for several reasons. First, although studies in various disciplines have focused on synchrony in bivariate time series (e.g., from dyads), relatively few have examined synchrony in groups of three or more persons. As a result, there are several available statistical methods for dyadic applications, but there is a shortage of multivariate options that would apply to behavioral similarity. The current study is a step toward understanding synchrony in teams of various sizes. Second, although synchrony has been acknowledged as a transient state which necessitates statistical methods accounting for changes in synchrony over time (Palumbo et al., 2017), it is most common for researchers to aggregate information about synchrony in a way that ignores its temporal dynamics. In the current research, we use an approach that accommodates temporal changes between differing states of similarity. Third, team sport scientists have called for methods that enable weighting each player’s unique influence on collective team variables, citing emerging approaches such as cluster phase, dominant region, and self-propelled particle models, which have been used to study collective behavior in schools of fish and crowds of people (Duarte et al., 2012). As we demonstrate in this paper, the use of dynamic factor analysis allows this type of weighting by producing a factor loading that quantifies the proportion of variation in each player’s signal explained by variation in a collective variable. Our strategy consolidates several techniques relevant to the study of behavioral similarity and in so doing forms an extensive case application which others may utilize in framing their analyses.

Our presentation is organized as follows. In the next section, we give an introduction

to the state space modeling framework, followed by coverage of regime-switching dynamic factor analysis in particular. Subsequently, we report on the empirical application (including Method, Results, and Discussion), followed by the Conclusion section, in which we highlight the important findings and future research directions associated with this paper. A supplementary online resource for this paper includes the R software (R Core Team, 2017) code used in the empirical application.

Analytical Approach

State Space Modeling Framework

State space modeling is a framework for analyzing intensive longitudinal data (ILD), which arise when measurement occasions number in the tens, hundreds, thousands, etc. (Walls & Schafer, 2006). To define state space modeling, three key characteristics are notable. First, state space models are useful for modeling system dynamics, including the relationships among variables and their changes across time. Second, state space models are primarily useful when there are a large number of repeated observations across time, that is, ILD or time series data. Third, state space models refer not to a particular statistical model but rather to a framework that can be applied flexibly to deploy many types of statistical models. State space models have been described as a unified methodology for a wide range of problems in time series analysis (Durbin & Koopman, 2012) and a general model that encompasses many special cases of interest (Shumway & Stoffer, 2010). This is not unlike structural equation modeling, a framework with which state space modeling has been compared (Chow, Ho, Hamaker, & Dolan, 2010). In fact, recent software advances have enabled the implementation of state space models within a structural equation modeling environment (Hunter, 2018).

State space modeling offers several advantages over competing approaches. First, state space models can handle multiple observed and unobserved variables, much like structural equation modeling, but the former is better suited to modeling intraindividual

dynamics, especially when measurement occasions (T) outnumber subjects (N ; Chow et al., 2010). Second, many ILD models, such as time-varying regression, ARMA models, linear mixed models, and dynamic factor models, can be deployed in state space form (Ho, Shumway, & Ombao, 2006). Third, the methods enabled by state space models are able to address a wider range of problems than more traditional approaches to time series analysis (e.g., the Box-Jenkins ARIMA method; Durbin & Koopman, 2012). Fourth, state space models require a lower computational burden than Laird-Ware linear mixed models, efficiency which is valuable when analyzing ILD with large T (Ho et al., 2006; Jones, 1993). Fifth, state space modeling handles three issues common in ILD applications, that is, unevenly spaced data, missing data, and forecasting.

In the name “state space modeling”, “state” refers to an unobserved, or latent, variable that characterizes a dynamic system, and “space” refers to a vector space, or a collection of vectors. Put together, “state space” reflects the central component of the model, that is, a set of state vectors, one per time point, containing the latent state variables. In state space modeling it is assumed that a dynamic system’s evolution over time is characterized by these state vectors, denoted in this paper as $\boldsymbol{\eta}_t$. These are associated with vectors of the observed variables ($\mathbf{Y}_t$), and the nature of this relationship is defined by a loading matrix ($\mathbf{\Lambda}$). The relationship among state variables, from one time point to the next, is defined by an autoregression matrix ($\mathbf{\Phi}$). Paramount to the framework, the state vector ($\boldsymbol{\eta}_t$) appears in both equations used to specify a state space model, namely the observation equation and the state equation, which are, respectively

$$\mathbf{Y}_t = \mathbf{\Lambda}\boldsymbol{\eta}_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim N(0, \boldsymbol{\Theta}) \quad (1)$$

$$\boldsymbol{\eta}_t = \mathbf{\Phi}\boldsymbol{\eta}_{t-1} + \boldsymbol{\zeta}_t, \quad \boldsymbol{\zeta}_t \sim N(0, \boldsymbol{\Psi}) \quad (2)$$

where

- $\mathbf{Y}_t$ is a $p \times 1$ vector of observations at the current time t

- $\boldsymbol{\eta}_t$ and $\boldsymbol{\eta}_{t-1}$ are $k \times 1$ state vectors at the current and previous times, respectively
- $\boldsymbol{\Lambda}$ is a $p \times k$ loading matrix relating the state variables to the observed values
- $\boldsymbol{\epsilon}_t$ is a $p \times 1$ vector of measurement errors with zero mean and covariance matrix $\boldsymbol{\Theta}$
- $\boldsymbol{\Phi}$ is a $k \times k$ autoregression matrix reflecting the dependence of the current state vector on the previous one
- $\boldsymbol{\zeta}_t$ is a $k \times 1$ vector of innovation errors with zero mean and covariance matrix $\boldsymbol{\Psi}$.

Owing to the flexibility of state space modeling mentioned above, Equations 1 and 2 can be customized by defining the contents of the model matrices, in particular $\boldsymbol{\Lambda}$ and $\boldsymbol{\Phi}$. These equations can also be extended in order to estimate intercepts and regression coefficients relating a vector of covariates to either the observation vector or state vector.

The unknown parameters of a state space model are estimated using a recursive procedure called the Kalman filter (Kalman, 1960). In this algorithm, information about the system state vector ($\boldsymbol{\eta}_t$) is updated at each time step, as a vector of observed data ($\mathbf{Y}_t$) is introduced. The Kalman filter is initiated with user-defined starting values for the state vector and its covariance matrix at $t = 0$ (i.e., $\boldsymbol{\eta}_{0|0}$ and $\mathbf{P}_{0|0}$, respectively). The Kalman filter's steps are as follows:

$$\boldsymbol{\eta}_{t|t-1} = \boldsymbol{\Phi}\boldsymbol{\eta}_{t-1|t-1} \quad (3)$$

$$\mathbf{P}_{t|t-1} = \boldsymbol{\Phi}\mathbf{P}_{t-1|t-1}\boldsymbol{\Phi}' + \boldsymbol{\Psi} \quad (4)$$

$$\mathbf{e}_t = \mathbf{Y}_t - \mathbf{Y}_{t|t-1} = \mathbf{Y}_t - \boldsymbol{\Lambda}\boldsymbol{\eta}_{t|t-1} \quad (5)$$

$$\mathbf{D}_t = \boldsymbol{\Lambda}\mathbf{P}_{t|t-1}\boldsymbol{\Lambda}' + \boldsymbol{\Theta} \quad (6)$$

$$\mathbf{K}_t = \mathbf{P}_{t|t-1}\boldsymbol{\Lambda}'\mathbf{D}_t^{-1} \quad (7)$$

$$\boldsymbol{\eta}_{t|t} = \boldsymbol{\eta}_{t|t-1} + \mathbf{K}_t\mathbf{e}_t \quad (8)$$

183

$$\mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t\mathbf{\Lambda})\mathbf{P}_{t|t-1} \quad (9)$$

184

185

186

187

188

189

190

191

192

193

194

The Kalman filter proceeds at each time step by computing one-step-ahead predictions of the state vector and its covariance matrix (Equations 3 and 4), a vector of one-step-ahead prediction errors and its covariance matrix (Equations 5 and 6), and a matrix called the Kalman gain (Equation 7). Those five equations make up the “prediction step” of the Kalman filter. In Equations 8 and 9, the state vector and its covariance matrix are updated based on the one-step-ahead prediction errors and the Kalman gain matrix (note: $\mathbf{I}$ is an identity matrix). These two equations make up the “update step” of the Kalman filter. The byproducts $\mathbf{e}_t$ and $\mathbf{D}_t$, which are the one-step-ahead prediction error vector and its covariance matrix, respectively, are passed to a likelihood function known as the prediction error decomposition function (Equation 10; Harvey, 1989). Optimizing this function returns the model parameter estimates:

$$\frac{1}{2} \sum_{t=1}^T [-p \log(2\pi) - \log |\mathbf{D}_t| - \mathbf{e}_t' \mathbf{D}_t^{-1} \mathbf{e}_t] \quad (10)$$

195

196

Next, we demonstrate how a regime-switching dynamic factor model (RSDFM) can be represented as a state space model.

197

Regime-Switching Dynamic Factor Analysis

198

199

200

201

202

203

204

Regime-switching state space models (Kim & Nelson, 1999) are useful for applications in which a dynamic system transitions (i.e., switches) between two or more discrete stages (i.e., regimes). For example, this approach has been used to detect switches between regimes of high and low pain, abrupt mood changes during a major depressive episode, and changes between high and low performance in basketball field goal attempts (Hamaker & Grasman, 2012); as well as between regimes of facial electromyography activation and nonactivation (Yang & Chow, 2010). In the current study, we use a

regime-switching state space modeling approach to analyze transitions between high and zero speed similarity during soccer competition.

Equations 1 and 2 can be modified as follows to reflect regime dependency, where the subscript R_t indicates matrices that may contain regime-varying parameters:

$$\mathbf{Y}_t = \mathbf{\Lambda}_{R_t} \boldsymbol{\eta}_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim N(0, \boldsymbol{\Theta}_{R_t}) \quad (11)$$

$$\boldsymbol{\eta}_t = \boldsymbol{\Phi}_{R_t} \boldsymbol{\eta}_{t-1} + \boldsymbol{\zeta}_t, \quad \boldsymbol{\zeta}_t \sim N(0, \boldsymbol{\Psi}_{R_t}) \quad (12)$$

Within a regime-switching paradigm, the state space framework remains flexible to fit many special cases of statistical models. Dynamic factor analysis (Ho et al., 2006; Engle & Watson, 1981; Molenaar, 1985; Nesselroade, McArdle, Aggen, & Meyers, 2002) generalizes conventional multi-subject cross-sectional factor analysis to ILD in order to capture common dependence among multiple time series. One novel aspect of our approach is that the collective is treated as the unit of analysis, with speed similarity operationalized as the latent structure that drives multiple time series (i.e., one time series per individual). Another novel aspect is the use of a regime-switching model to account for transitions between a regime of “high” similarity, that is, one in which the observed time series are driven by a common latent factor; and “zero” similarity, that is, a regime in which there is assumed to be no correlation among the multiple time series.

The RSDFM used in the current investigations is written

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{pt} \end{bmatrix} = \begin{cases} \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} \begin{bmatrix} C_t \\ C_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \vdots \\ \epsilon_{pt} \end{bmatrix}, & \Theta = \begin{bmatrix} \theta_{11} & & & \\ & \theta_{21} & & \\ & & \ddots & \\ & & & \theta_{p1} \end{bmatrix}; & \text{if } R_t = 1 \\ \begin{bmatrix} 1 & 0 \\ \lambda_2 & 0 \\ \vdots & \vdots \\ \lambda_p & 0 \end{bmatrix} \begin{bmatrix} C_t \\ C_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \vdots \\ \epsilon_{pt} \end{bmatrix}, & \Theta = \begin{bmatrix} \theta_{12} & & & \\ & \theta_{22} & & \\ & & \ddots & \\ & & & \theta_{p2} \end{bmatrix}; & \text{if } R_t = 2 \end{cases} \quad (13)$$

$$\begin{bmatrix} C_t \\ C_{t-1} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} C_{t-1} \\ C_{t-2} \end{bmatrix} + \begin{bmatrix} \zeta_t \\ 0 \end{bmatrix}, \quad \Psi = \begin{bmatrix} \psi & 0 \\ 0 & 0 \end{bmatrix} \quad (14)$$

where the observation vector $[y_{1t} \ y_{2t} \ \dots \ y_{pt}]'$ is a multivariate time series consisting of data from p persons¹. Regime 1 ($R_t = 1$) is defined as the “zero” similarity regime, and Regime 2 ($R_t = 2$) is defined as the “high” similarity regime. This is apparent in the disparate loading matrices (Λ_{R_t}) in Equation 13. Whereas in Regime 1 the loadings are set to zero signifying that the observations are not driven by a latent collective process, in Regime 2 the loadings are estimated parameters, with the exception of the first one being set to 1 for the purpose of scaling. Additionally, Equation 13 reflects that the measurement error variance matrix (Θ_{R_t}) is estimated separately for each regime.

In Equations 13 and 14, for illustration, a second-order autoregressive process, or AR(2), is specified for the collective state variable (C_t). However, these equations can be modified to specify any chosen order of process. In the case application reported in this

¹ Although in multivariate applications the symbols p and n conventionally refer to the number of variables and persons, respectively, in this model formulation the number of time series (p) is equal to the number of persons in the collective (i.e., one time series per person). We have decided to keep with the convention of the state space modeling framework, in which the observation vector is $p \times 1$, and hence throughout this paper, p refers to the number of persons in the collective.

paper, AR(1) and AR(2) models are used, as we explain under the heading “Data Processing and Analysis”. To re-formulate the model depicted in Equations 13 and 14 as an AR(1) process, the state vector would become simply $[C_t]$; Φ and Ψ would become 1×1 matrices (i.e., $[\phi_1]$ and $[\psi]$, respectively) in Equation 14; and the second column (of zeros) in each regime-dependent Λ matrix in Equation 13 would be omitted.

In order to infer the regime in which a system resides at each time point (i.e., R_t), it is necessary to estimate a transition probability matrix (Π), which contains values indicating the probability that the system (e.g., collective) is in a particular regime conditional upon the regime at the previous time point. This is a square matrix whose dimensions equal the number of regimes. For a two-regime model, to which the scope of this paper is limited, the transition probability matrix can be written as

$$\begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix} \quad (15)$$

where each π_{ij} is the probability of Regime j at time t , given Regime i at time $t - 1$, or expressed in notation, $\pi_{ij} = \Pr[R_t = j | R_{t-1} = i]$. For example, π_{11} is the probability of staying in Regime 1, while π_{12} is the probability of switching from Regime 1 to Regime 2. Hence, these values must sum to 1, and more generally, all row sums of Π must equal 1. Therefore, it is only necessary to estimate one probability per row. In our application, the natural log odds of π_{11} and π_{21} were estimated.

Estimation of the state vector and regime at each time step, as well as the model parameters, is performed using the Kim filter (Kim & Nelson, 1999) and maximum likelihood estimation. The Kim filter is a combination of the Kalman filter (Kalman, 1960) and the Hamilton filter (Hamilton, 1989). In a regime-switching model, the Hamilton filter enables the probabilistic inference of the regimes, which are also unobserved, based on the behavior of the observed time series. The Kim filter deploys these algorithms in three steps. First, the Kalman filter is used to generate an estimate of the state vector and its

covariance matrix. Second, the Hamilton filter is used to obtain the joint probability of Regime i at time $t - 1$ and Regime j at time t (i.e., $\Pr[R_{t-1} = i, R_t = j | \mathbf{Y}_t]$), as well as the probability of Regime j at time t (i.e., $\Pr[R_t = j | \mathbf{Y}_t]$). Third, a so-called collapsing process combines the estimates from the first two steps. Prediction errors are obtained as byproducts of the Kim filter and passed to the prediction error decomposition function (Equation 10), which is entered into an optimization step to obtain the parameter estimates. For each iteration of the optimization routine, the Kim filter is carried out recursively for all t ($1, 2, \dots, T$) so that the state vector and regime probabilities have been estimated at each time step. Although a full and detailed coverage of these algorithms is beyond the scope of this paper, details can be obtained from other sources (Kim & Nelson, 1999; Yang & Chow, 2010). The estimated parameters of the RSDFM include the factor loadings for Regime 2 ($\lambda_2, \dots, \lambda_p$), measurement error variances for each regime ($\theta_{11}, \dots, \theta_{p1}, \theta_{12}, \dots, \theta_{p2}$), one or two autoregression coefficients for the latent collective variable (ϕ_1, ϕ_2), the innovation error variance (ψ), and the natural log odds of the regime transition probabilities ($\ln(\frac{\pi_{ij}}{1-\pi_{ij}})$).

Interpreting the Parameters of the RSDFM

It may be of substantive value to researchers to quantify the *magnitude* and *prevalence* of behavioral similarity. Magnitude is the extent to which the individuals in a collective are synchronized in terms of the variable of interest. Within the RSDFM approach, magnitude may be interpreted from the effect sizes attributed to the individuals in a collective. Effect size refers to the proportion of variance in each observed time series explained by the collective state variable, and as such, its value may range from 0 to 1. In the current formulation of the RSDFM, effect size for each individual is equal to 1 minus the unexplained variance (i.e., measurement error variance) in Regime 2, which is also equal to the Regime 2 standardized factor loading squared. Regime 1 is formulated with zero factor loadings (i.e., no collective process driving the observed time series), and hence,

should have 100% unexplained variance. That is, the confidence intervals of the Regime 1 measurement error variances should all include 1. In sum, an effect size between 0 and 1 will be estimated for each individual for Regime 2, and this will quantify the magnitude of speed similarity, or the extent to which each individual's time series reflects the collective, within the high similarity regime. Individual effect sizes can be averaged to obtain an aggregate measure of magnitude for the collective. As an alternative approximation of magnitude, and/or for comparison, the researcher may assess the correlation coefficients for all pairs of collective members (e.g., teammates) separately for each of the predicted regimes.

It may also be useful to quantify speed similarity's prevalence, which we define as the proportion of time in which a collective resides in Regime 2 (i.e., the high similarity regime). The RSDFM approach yields a prediction of Regime 1 or Regime 2 at every time point, making it straightforward to assess the prevalence of high similarity. This can be easily computed as the number of time points at which Regime 2 was predicted, divided by the total number of time points. This proportion (or percentage, if reported as such) is often referred to as the *dwelt time* of a system within a particular state, in this case the high similarity regime. In the next sections, dwell time will be reported as the main metric of prevalence. The estimated regime transition probabilities (π_{ij}) can also indicate whether Regimes 1 and 2 are well balanced or one regime is relatively dominant over an analyzed time interval. For example, if π_{11} is estimated to be .75 and π_{22} is estimated to be .99, this suggests that Regime 2 is dominant. That is, Regime 2 is so prevalent that when the collective resides in this high similarity state, there is only a .01 probability of switching to Regime 1 at the next time point. In contrast, when the collective is classified as residing in Regime 1, there is a .25 probability of switching to Regime 2. These approaches to examining and reporting the magnitude and prevalence of behavioral similarity are demonstrated in the empirical study reported in the next section.

Application: Speed Similarity in Women's Soccer Players

Method

Participants, Procedures, and Materials. Varsity women's soccer players were recruited from a National Collegiate Athletic Association (NCAA) Division I team in the United States. The university's Institutional Review Board approved the study protocol detailing the recruitment of participants and data collection procedures. The number of players who gave informed consent to participate in the study was 25. Data were collected during the team's competitive 2017 season, including all 18 regular season home and away games. Only outfield players were included in the study; goalkeepers were excluded. For each game, participants who started the game and played without substitution until halftime were included. Therefore, out of the ten outfield players starting each game, some were excluded due to first half substitutions and the fact that some starters may not have been consenting participants. The actual sample size for each of the 18 games ranged from 3 to 9 participants (median = 6). Data from the second half of games were not used due to practical issues such as the halftime break and the prevalence of second half substitutions.

In this study, the collective is the unit of analysis, and data were collected in the team's natural competitive setting without any researcher interference. That is, in each game it was solely the team's coaching staff who determined which individuals played, so the study participants vary from game to game. A unique identification code was randomly assigned to each participant for the purpose of recording which individuals started each game. However, for the purposes of the analyses performed, the identities of the individuals participating in each game and their playing positions (e.g., defender, midfielder, forward) are not accounted for. In terms of how the analyses were carried out, the individual participants can be assumed to be interchangeable. For example, the symbol used to represent player 4 (i.e., y_4) may, in different games, refer to different individuals. Likewise, data from the same individual may, for example, be denoted as y_1 in one game and y_3 in another game.

Data were collected using the Polar® Team Pro system (Polar Electro, Inc., Kempele, Finland). This system consists of a chest strap monitor worn by each participant and a tablet computer application with an interface that enables real-time performance tracking. The wearable devices include GPS tracking, accelerometers, and heart rate monitoring, and the data are delivered to the online application using Bluetooth technology. The system was owned and used regularly by the team during training and competitive games. Team members each had their own numbered device, and at the outset of the study all participants already had training and experience wearing the monitors properly. Data streams including acceleration, running cadence, cumulative distance, and heart rate were available for download after each game. In this study the variables of interest are cadence and distance. Data sets were downloaded following each game, then processed and analyzed, as detailed next.

Data Processing and Analysis. Cadence data streams were recorded at a rate of 1 Hz in units of revolutions per minute (rpm), where one revolution equals two steps (e.g., 80 rpm = 160 steps per minute). Cumulative distance was recorded at 10 Hz in units of yards. The cumulative distance time series were converted to distances covered within defined time intervals, or bins, by differencing the cumulative values. Similarly, cadence time series were aggregated by taking the mean cadence within each bin. Time series were examined for order of ARMA process using plots of the autocorrelation function (ACF) and partial ACF (PACF) and by running univariate ARMA models on individual time series. The R (R Core Team, 2017) functions `acf`, `pacf`, and `arma` were used to perform these diagnostics. If the ACF has a significant autocorrelation persisting over many lags (i.e., decays gradually), and the PACF becomes non-significant abruptly after a smaller number of lags, then this is indicative of an AR process (Bowerman, O’Connell, & Koehler, 2005). ACF and PACF plots for this study are included in Figure 1. Inspecting these plots to guide our assessment of ARMA order, we treated the significance thresholds (dashed lines) as approximate but not strict cutoffs. Most of the individual time series, both for

cadence and distance, exhibited ACFs and PACFs reflecting an AR(1) process (Panels D & F of Figure 1) or an AR(2) process (Panels B, H, L). Much less common were AR(3) processes (Panel J), so only AR(1) and AR(2) models were used in this study.

(FIGURE 1 NEAR HERE)

Determining an appropriate sampling rate (i.e., bin size) for the data requires balancing a tradeoff between scientific and practical considerations. In terms of scientific considerations, it is desirable to sample data frequently enough to reflect the time scale of interest to examine changes in the observed variables (Collins & Graham, 2002; Smith & Walls, 2016; Walls, Barta, Stawski, Collyer, & Hofer, 2012). In terms of practical considerations, data sampled very close together may have features such as repetition of the same or very similar values (i.e., high autocorrelation) and may therefore exhibit nonstationarity. Ultimately, both cadence and distance time series were aggregated in bins of 3 seconds due to issues with nonstationarity that became apparent when using bins of 1 or 2 seconds. This was evident in part by the large number of models that failed to converge. Of the models that did successfully converge, the estimated AR coefficients were very close to, and their confidence intervals covered, the boundaries of stationarity conditions. That is, for AR(1) models, the parameter ϕ_1 estimates were close to 1, and for AR(2) models, the sums of the parameter ϕ_1 and ϕ_2 estimates were close to 1. These problems were no longer apparent after reducing the sampling rate by increasing the bin size to 3 seconds. Given that observations were taken from the first half (45 minutes) of each soccer game, using a bin size of 3 seconds yielded time series each with 900 observations. Finally, before analysis each time series was intraindividually standardized, that is, converted to z-scores. Doing so enables straightforward interpretation of each measurement error variance estimate as the proportion of variance in the observations not explained by the latent collective variable and is consistent with previous regime-switching applications (e.g., Chow, Witkiewitz, Grasman, Hutton, & Maisto, 2014; Chow, Witkiewitz, Grasman, & Maisto, 2015). Moreover, speed similarity should not be affected

by intraindividual standardization because similarity is modeled as covariation, not as sameness of players' actual levels of cadence or distance. However, it is unknown what effect standardization might have on the standard error estimates, so this is an important area of future research (see Conclusion).

In this study, in addition to a two-regime RSDFM characterizing the presence of zero and high similarity regimes, it is also considered that a one-regime model, that is, a non-switching (high similarity only) dynamic factor model (DFM), may provide better fit than the RSDFM. For all analyses, we used the **dynr** R package (Ou, Hunter, & Chow, 2018); code included in supplementary online resource. The variables cadence and distance were analyzed separately for each of the 18 games (i.e., 36 unique data sets). With a typical laptop, analyzing each variable (fitting 4 models for each of 18 games) took approximately 1.5 hours. For each data set, the best fitting model was selected by comparing Akaike Information Criterion (AIC; Akaike, 1998) and Bayesian Information Criterion (BIC; Schwarz, 1978) fit indices. When comparing AIC or BIC values for models fit to a given data set, smaller values indicate better model fit. In sum, one best fitting model (AR(1) DFM, AR(2) DFM, AR(1) RSDFM, or AR(2) RSDFM) was selected for each game/variable combination based on the lowest AIC/BIC value (see Tables 1 and 2). There were some models that converged but had non-positive definite Hessian matrices, which meant that the standard errors were computed using a nearest positive definite approximation to the Hessian matrix, and hence were not trustworthy. These models were discarded and not considered for selection.

Results

In general, AIC and BIC were in agreement of the best fitting model for each data set, so only AIC values are displayed in Tables 1 and 2. One exception was the analysis of cadence data from Game 17, where the AR(2) RSDFM produced the smallest AIC, while the AR(2) one-regime model produced the smallest BIC. In that case, the RSDFM was

Game	N	AR(1) DFM	AR(2) DFM	AR(1) RSDFM	AR(2) RSDFM	Selected Model
1	8	17105	16969	-	-	AR(2)DFM
2	6	12702	12586	12634	12530	AR(2)RSDFM
3	9	-	19116	19168	-	AR(2)DFM
4	5	10956	-	-	-	AR(1)DFM
5	5	10286	10204	10211	10125	AR(2)RSDFM
6	6	11808	11674	11761	-	AR(2)DFM
7	7	15258	15162	-	15098	AR(2)RSDFM
8	9	19396	19262	-	-	AR(2)DFM
9	7	-	15144	-	-	AR(2)DFM
10	6	12995	12872	12956	14580	AR(2)DFM
11	8	-	-	-	17591	AR(2)RSDFM
12	7	14240	14116	14104	13984	AR(2)RSDFM
13	4	7798	7727	7791	7667	AR(2)RSDFM
14	6	12934	12816	12891	-	AR(2)DFM
15	7	-	15179	-	-	AR(2)DFM
16	5	11400	11284	11342	11238	AR(2)RSDFM
17	4	8566	8483	8556	8468	AR(2)RSDFM
18	3	-	6666	6674	6596	AR(2)RSDFM

Note: The selected model for each data set (table row) is that with the smallest AIC value; missing AIC indicates that the model failed to converge during estimation or was discarded

Table 1

Cadence models: sample sizes (N), AIC, selected model.

selected. AR(2) models were more commonly selected than AR(1) models for both variables analyzed. One striking difference between the analyses for cadence and distance is the number of regime-switching models (RSDFMs) selected over one-regime models (DFMs). Out of the 18 selected models used to analyze cadence, there were 9 RSDFMs and 9 DFMs. Out of the 17 selected models for distance (for Game 15, no model was selected), 14 were RSDFMs and 3 were DFMs.

Next, we present detailed results of one exemplar analysis each for cadence and distance. These include the Game 7 cadence data and Game 12 distance data. In both of the examples, the AR(2) RSDFM was the selected model. Parameter estimates from the AR(2) RSDFM fit to the Game 7 cadence data can be found in Table 3. Some individual differences are apparent in the magnitudes (effect sizes) associated with individual players.

Game	N	AR(1) DFM	AR(2) DFM	AR(1) RSDFM	AR(2) RSDFM	Selected Model
1	8	17026	16934	16651	16565	AR(2)RSDFM
2	6	12656	12590	12517	12440	AR(2)RSDFM
3	9	18934	18837	18745	-	AR(1)RSDFM
4	5	-	10822	10753	-	AR(1)RSDFM
5	5	10115	10066	10004	9958	AR(2)RSDFM
6	6	11935	11874	11778	11687	AR(2)RSDFM
7	7	14935	14877	-	14755	AR(2)RSDFM
8	9	18809	18687	18699	-	AR(2)DFM
9	7	-	14918	14892	14804	AR(2)RSDFM
10	6	13159	13044	12948	12828	AR(2)RSDFM
11	8	17263	17141	17147	-	AR(2)DFM
12	7	14319	14193	14087	13965	AR(2)RSDFM
13	4	7742	7695	7621	7567	AR(2)RSDFM
14	6	-	13302	13228	-	AR(1)RSDFM
15	7	-	-	-	-	None
16	5	11399	11265	11307	-	AR(2)DFM
17	4	-	8634	8235	8200	AR(2)RSDFM
18	3	6703	6628	6598	6532	AR(2)RSDFM

Note: The selected model for each data set (table row) is that with the smallest AIC value; missing AIC indicates that the model failed to converge during estimation or was discarded

Table 2

Distance models: sample sizes (N), AIC, selected model.

That is, the collective state variable explains a higher proportion of variance in some individuals' cadence compared to others. This is reflected by the standardized loadings, which can be squared to obtain effect sizes. For example, the standardized loading associated with y_2 is .20 (i.e., effect size of .04, unexplained variance of .96 in Regime 2), which may point to this individual lacking in similarity, in terms of cadence, with the rest of her teammates. This is also evident in Table 4, which displays the empirical bivariate correlations in cadence time series during predicted Regime 1 ("Zero"; Panel a) and predicted Regime 2 ("High"; Panel b) intervals. In Panel b, it is clear that correlations involving y_2 tend to be smaller than others. As expected, the correlations displayed in Table 4 tend to be higher in predicted Regime 2 intervals, compared to their Regime 1 counterparts in Panel a, which tended to be smaller and, in some cases, negative. The

Parameter	Symbol	Est.	SE	t	95% CI	p
L1 fixed (std.)	λ_1	1.00 (.81)	-	-	-	-
L2 unstd. (std.)	λ_2	.30 (.20)	.05	6.6	(.21 , .39)	<.001
L3 unstd. (std.)	λ_3	.75 (.62)	.04	16.7	(.66 , .83)	<.001
L4 unstd. (std.)	λ_4	.90 (.70)	.04	21.1	(.82 , .98)	<.001
L5 unstd. (std.)	λ_5	1.10 (.90)	.04	29.4	(1.03 , 1.18)	<.001
L6 unstd. (std.)	λ_6	.60 (.54)	.04	13.9	(.52 , .68)	<.001
L7 unstd. (std.)	λ_7	.92 (.77)	.04	22.5	(.84 , 1.00)	<.001
MEV1 reg. 1	θ_{11}	.87	.14	6.4	(.61 , 1.14)	<.001
MEV2 reg. 1	θ_{21}	.86	.15	5.8	(.57 , 1.15)	<.001
MEV3 reg. 1	θ_{31}	1.09	.18	6.1	(.74 , 1.44)	<.001
MEV4 reg. 1	θ_{41}	.65	.11	5.8	(.43 , .87)	<.001
MEV5 reg. 1	θ_{51}	.91	.14	6.7	(.64 , 1.17)	<.001
MEV6 reg. 1	θ_{61}	1.28	.22	5.7	(.84 , 1.72)	<.001
MEV7 reg. 1	θ_{71}	1.24	.19	6.6	(.87 , 1.61)	<.001
MEV1 reg. 2	θ_{12}	.34	.02	14.0	(.29 , .39)	<.001
MEV2 reg. 2	θ_{22}	.96	.05	18.9	(.86 , 1.06)	<.001
MEV3 reg. 2	θ_{32}	.61	.04	15.8	(.54 , .69)	<.001
MEV4 reg. 2	θ_{42}	.51	.03	16.4	(.45 , .57)	<.001
MEV5 reg. 2	θ_{52}	.19	.02	11.5	(.16 , .23)	<.001
MEV6 reg. 2	θ_{62}	.71	.04	16.2	(.62 , .80)	<.001
MEV7 reg. 2	θ_{72}	.40	.03	14.5	(.35 , .46)	<.001
AR1 coefficient	ϕ_1	1.30	.05	27.8	(1.21 , 1.39)	<.001
AR2 coefficient	ϕ_2	-.49	.04	-12.5	(-.57 , -.41)	<.001
IE var.	ψ	.12	.01	8.4	(.09 , .15)	<.001
Log odds 1→1	$\ln(\frac{\pi_{11}}{1-\pi_{11}})$	1.23	.28	4.5	(.69 , 1.78)	<.001
Log odds 2→1	$\ln(\frac{\pi_{21}}{1-\pi_{21}})$	-3.38	.28	-11.9	(-3.94 , -2.82)	<.001

Note: CI = confidence interval; Est. = estimate; IE = innovation error; L = loading; MEV = measurement error variance; p = p-value; reg. = regime; SE = standard error; std. = standardized; t = Student's t test statistic; unstd. = unstandardized; var. = variance

Table 3

Parameter estimates from AR(2) RSDFM fit to Game 7 cadence data.

	y_1	y_2	y_3	y_4	y_5	y_6
y_2	.19					
y_3	-.20	.10				
y_4	-.06	-.11	.35			
y_5	.45	.08	-.38	-.10		
y_6	.16	.19	-.06	-.11	.00	
y_7	-.27	-.25	.34	.22	-.13	-.39

(a) “Zero” similarity intervals (predicted Regime 1)

	y_1	y_2	y_3	y_4	y_5	y_6
y_2	.19					
y_3	.39	.18				
y_4	.51	.20	.51			
y_5	.74	.17	.48	.63		
y_6	.42	.11	.34	.36	.43	
y_7	.56	.16	.54	.49	.67	.33

(b) “High” similarity intervals (predicted Regime 2)

Table 4

Empirical correlations in cadence in Game 7, by predicted regime.

empirical correlations in Panel a demonstrate that a zero similarity regime may not fit the data well in terms of covariation in pairs of teammates (e.g., a correlation between y_1 and y_5 of .45 and between y_6 and y_7 of -.39).

The log odds of the regime transition probabilities suggest that Regime 2 was dominant, according to the AR(2) RSDFM. The estimated value of -3.38 for $\ln(\frac{\pi_{21}}{1-\pi_{21}})$ equates to a probability of switching from Regime 2 to Regime 1 of only .034, and therefore, the probability of staying in the high similarity Regime 2 is estimated to be .966. On the other hand, the estimated value of 1.23 for $\ln(\frac{\pi_{11}}{1-\pi_{11}})$ in Table 3 would convert to $\pi_{11} = .774$, which is not a very high probability of staying in the zero similarity Regime 1. The prevalence of Regime 2 is quite apparent in the top panel of Figure 2, which shows the Game 7 cadence time series superimposed on the predicted regimes. In this analysis, the team was predicted to reside in the high similarity regime with a dwell time of .91 (i.e., 820 out of the 900 time points). In the bottom panel of Figure 2, the plot is zoomed in on 30

time points to show time series data in each predicted regime. The individual time series appear to vary independently of one another when Regime 1 is predicted (white region) and show more similar patterns of variation when Regime 2 is predicted (shaded region).

(FIGURE 2 NEAR HERE)

Results of the analysis of Game 12 distance data, using an AR(2) RSDFM, can be found in Table 5. Unlike in the previous example, the standardized loadings are all relatively high, ranging from .57 to .87 (.32 to .75 when squared to evaluate effect size, or magnitude). This suggests a large proportion of variance in players' distance explained by the collective state variable, that is, a large magnitude of speed similarity. The smallest of the standardized loadings (.57, or 32% explained variance) is consistent with the slightly lower correlation coefficients including y_5 in Table 6(b), although this individual difference is not as striking as in the previous example. In other words, there is some indication that y_5 may not be synchronized to her teammates as well as they are to each other, in terms of distance covered in Game 12. As expected, the correlation coefficients in Table 6(a) are low and/or negative, but their counterparts in Panel b are higher and entirely positive.

As in the previous example, "staying probabilities" predicted by the AR(2) RSDFM for the Game 12 distance data ($\pi_{11} = .579$; $\pi_{22} = .946$), and the predicted regimes themselves (see Figure 3, top panel), show that the high similarity Regime 2 is far more prevalent than Regime 1. In this example, the team was predicted to reside in Regime 2 for 857 out of 900 time points (i.e., dwell time = .95). In the bottom panel of Figure 3, the plot is zoomed in on 30 time points to juxtapose the behavior of time series in predicted Regime 1 against Regime 2. Again, the time series appear to behave more similarly in the shaded region and less so in the area with a white background.

(FIGURE 3 NEAR HERE)

Having inspected the results of two models in detail, next we present aggregate information about the magnitude and prevalence of speed similarity reflected by the selected model results, across the entire season. In particular, to summarize the magnitude

Parameter	Symbol	Est.	SE	t	95% CI	p
L1 fixed (std.)	λ_1	1.00 (.71)	-	-	-	-
L2 unstd. (std.)	λ_2	.94 (.81)	.05	19.7	(.84, 1.03)	<.001
L3 unstd. (std.)	λ_3	1.15 (.84)	.05	22.6	(1.05, 1.25)	<.001
L4 unstd. (std.)	λ_4	.94 (.81)	.05	2.2	(.85, 1.04)	<.001
L5 unstd. (std.)	λ_5	.81 (.57)	.05	16.0	(.71, .91)	<.001
L6 unstd. (std.)	λ_6	1.03 (.86)	.05	21.9	(.93, 1.12)	<.001
L7 unstd. (std.)	λ_7	.99 (.87)	.05	21.8	(.90, 1.08)	<.001
MEV1 reg. 1	θ_{11}	.69	.14	5.0	(.42, .96)	<.001
MEV2 reg. 1	θ_{21}	1.97	.31	6.3	(1.36, 2.59)	<.001
MEV3 reg. 1	θ_{31}	.67	.13	5.3	(.43, .92)	<.001
MEV4 reg. 1	θ_{41}	1.85	.30	6.2	(1.26, 2.43)	<.001
MEV5 reg. 1	θ_{51}	.61	.13	4.8	(.36, .85)	<.001
MEV6 reg. 1	θ_{61}	2.14	.35	6.0	(1.44, 2.83)	<.001
MEV7 reg. 1	θ_{71}	2.26	.36	6.3	(1.56, 2.96)	<.001
MEV1 reg. 2	θ_{12}	.49	.03	17.3	(.43, .54)	<.001
MEV2 reg. 2	θ_{22}	.35	.02	16.0	(.31, .40)	<.001
MEV3 reg. 2	θ_{32}	.30	.02	15.2	(.26, .34)	<.001
MEV4 reg. 2	θ_{42}	.35	.03	13.1	(.29, .40)	<.001
MEV5 reg. 2	θ_{52}	.68	.04	18.1	(.60, .75)	<.001
MEV6 reg. 2	θ_{62}	.26	.02	15.1	(.23, .30)	<.001
MEV7 reg. 2	θ_{72}	.25	.02	14.8	(.21, .28)	<.001
AR1 coefficient	ϕ_1	1.26	.04	29.4	(1.18, 1.34)	<.001
AR2 coefficient	ϕ_2	-.50	.04	-12.4	(-.58, -.42)	<.001
IE var.	ψ	.12	.01	8.8	(.10, .15)	<.001
Log odds 1→1	$\ln(\frac{\pi_{11}}{1-\pi_{11}})$	.32	.24	1.3	(-.15, .79)	.091
Log odds 2→1	$\ln(\frac{\pi_{21}}{1-\pi_{21}})$	-2.91	.19	-15.3	(-3.28, -2.53)	<.001

Note: CI = confidence interval; Est. = estimate; IE = innovation error; L = loading; MEV = measurement error variance; p = p-value; reg. = regime; SE = standard error; std. = standardized; t = Student's t test statistic; unstd. = unstandardized; var. = variance

Table 5

Parameter estimates from AR(2) RSDFM fit to Game 12 distance data.

	y_1	y_2	y_3	y_4	y_5	y_6
y_2	.01					
y_3	.03	.14				
y_4	.22	.15	.10			
y_5	.16	-.37	-.24	.19		
y_6	.10	-.11	-.02	-.23	.00	
y_7	.18	.27	.15	.26	-.04	.07

(a) “Zero” similarity intervals (predicted Regime 1)

	y_1	y_2	y_3	y_4	y_5	y_6
y_2	.48					
y_3	.61	.64				
y_4	.52	.55	.63			
y_5	.48	.35	.48	.47		
y_6	.56	.62	.71	.57	.39	
y_7	.52	.66	.65	.73	.42	.62

(b) “High” similarity intervals (predicted Regime 2)

Table 6

Empirical correlations in distance in Game 12, by predicted regime.

of similarity, we focus on mean effect sizes (i.e., proportion of variance explained by the collective state variable) and bivariate correlations between teammate pairs (i.e., in zero vs. high predicted regimes). To summarize the prevalence of similarity, we present dwell time proportions in the high similarity regime. In Figure 4, overlapping histograms depict the distribution of bivariate correlations between teammate pairs for selected regime-switching models. The histograms represent bivariate correlations from all 9 of the selected RSDFMs analyzing cadence data and all 14 of the selected RSDFMs analyzing distance data. As expected, the distributions of correlation coefficients for Regime 1 are approximately centered on zero, and the distributions for Regime 2 are approximately centered on .5, with only positive values.

(FIGURE 4 NEAR HERE)

Figures 5 and 6 display prevalence and magnitude, respectively, over the course of the 18-game season. Although it is not possible to examine temporal trends due to the

many variables (individuals involved, opponent, location, conditions, etc.) that may differ for each game, it is clear that Regime 2 was overwhelmingly prevalent in all of the analyses (i.e., greater than .88 dwell time for all 23 of the selected RSDFMs). A possible explanation for the dominance of Regime 2 is addressed in the Discussion section. The magnitudes in Figure 6 are displayed as the mean effect sizes, that is, the means of the squared standardized loadings for all selected models, both DFMs and RSDFMs. The mean effect sizes mostly range from .4 to .6 throughout the season.

(FIGURES 5 AND 6 NEAR HERE)

Discussion

The analysis of cadence and distance data sets from college women's soccer produced several important outcomes worthy of discussion, both at the level of specific analyses and at the aggregate level. To exemplify the substantively relevant details that can be extracted from each analysis, we presented two sets of results using the RSDFM approach (Game 7 cadence and Game 12 distance). One characteristic of a team's speed similarity that is likely to have scientific and practical value is the magnitude of similarity. In particular, standardized loadings, when squared, equal the effect size, the proportion of variance in each observed time series explained by the collective state variable. This can be interpreted as the extent of each individual's similarity with the collective. In the Game 7 example, it was noted that one individual's cadence (y_2) was not as well synchronized to the collective as the cadences of the other six members. The small standardized loading (.20) and small bivariate correlations between y_2 and each other time series in predicted Regime 2 supported this finding. Hence, with this analytical approach it is possible to identify individuals who contribute more or less than others to behavioral similarity, which has practical value to those affiliated with teams such as coaches, performance analysts, sport psychologists, and the athletes themselves. Additionally, presenting specific examples made it possible to show didactically how the bivariate correlations among time series can

be inspected separately for the two predicted regimes for comparison, and how the estimated regime switching probabilities (π_{ij}) and proportion of dwell time reflect the prevalence of each regime. Plotting the predicted regimes is also useful to visualize the prevalence of each regime.

Other important outcomes of the study can be examined at the aggregate level. This is useful, as in the case of this study, when analyzing behavioral similarity within many measurement epochs (e.g., games) spanning a longer time period (e.g., a sports season). First, to aggregate information about the magnitude of speed similarity over all games in which a regime-switching model was selected, we generated overlapping histograms to compare the distribution of bivariate correlations in the predicted Regime 1 versus Regime 2. This confirmed what one might expect, that is, smaller correlations centered on zero in Regime 1, and larger positive correlations centered on .5 in Regime 2. Second, the mean effect sizes, computed by averaging the squared standardized loadings from each selected model, can be plotted for the 18-game season. In this application, there was substantial heterogeneity in games (in terms of players, opponents, conditions, etc.) that would have hindered our ability to assess longitudinal trends in speed similarity, but under more controlled conditions, researchers may choose to investigate questions such as: Does the average magnitude of a group's behavioral similarity change over time? Third, we emphasized the proportion of time spent in each regime (i.e., dwell time) as a substantively valuable outcome to assess the prevalence of high similarity when using RSDFMs. At the aggregate level, dwell time proportions can be plotted, for example, for all games in a season to check for changes over time.

We reported a large percentage of time spent in the high similarity regime (i.e., greater than 88%). This finding is likely due to the constraints of competitive soccer. That is, particularly at the highest levels of competitive sport, the movements of teammates are often constrained to be highly similar. The observed variables cadence and distance were used in this study to operationalize speed of movement. In competitive soccer, in which

team members are typically arranged in and trained to maintain a particular formation, it would be expected, and beneficial to performance, for teammates to move in similar directions and at similar speeds in response to events in the game such as the position of the ball and the team in possession. Indeed, it is difficult to envision a team achieving success with some players standing still, others walking, others jogging, and others sprinting. Put another way, it would be detrimental to the team's overall performance if individuals' speeds were uncorrelated, that is, if they did not reside in the high similarity regime for a large proportion of time points. Given the context of NCAA Division I competition, it is appropriate that the proportions of dwell time in the high similarity regime were so high. In other contexts or at other levels of expertise, it may be of substantive interest to rigorously test whether a collective can show improvements over time, in terms of dwell time in a high similarity state.

Conclusion

This paper featured didactic presentation of a regime-switching dynamic factor analytic approach and an empirical application. This inquiry produced several important developments in the study of collective synchrony. First, unlike most other synchrony applications, which tend to focus on dyads, here we have employed a multivariate approach to enable the examination of behavioral similarity in groups of three or more. Our data included teams of various sizes ranging from three to nine. Second, as opposed to studies that have used metrics to quantify synchrony over the duration of a time interval as a single aggregate value, we used a regime-switching approach to account for temporal changes between states of high and zero similarity. Third, the dynamic factor modeling approach used in this study enabled the weighting of each individual player's unique influence on the team's behavioral similarity. These weights, or factor loadings, can be squared to obtain effect sizes (i.e., proportions of explained variance), which quantify the magnitude of similarity. These can be examined on an individual basis and summarized for

all team members by computing the mean of their values, for example. Fourth, by categorizing time intervals as residing within either a high or zero similarity regime, this modeling approach allows the researcher to extract information about the prevalence of behavioral similarity. That is, what proportion of time (i.e., dwell time) is spent in the high similarity regime over a given time interval analyzed? We have shown how these features, magnitude and prevalence, can be aggregated and depicted graphically in order to summarize multiple epochs of observation over a longer time period.

In terms of the methodological aim, this investigation has demonstrated the value of the RSDFM approach to analyzing behavioral similarity. In our application, it was apparent that the parameter estimates and predicted regimes can be useful to directly interpret about a single time interval (e.g., dwell time within one game; magnitude of similarity of one player's behavior with that of the collective), and to aggregate for a larger set of events (e.g., many games within a season). Substantively, two findings are particularly noteworthy. First, each player's unique contribution to behavioral similarity can be detected in the form of a standardized factor loading. The practical significance of this cannot be overstated. This implies that stakeholders interested in team performance (e.g., coaches, analysts, players, support staff) could use this information to identify and address possible weaknesses in terms of behavioral similarity. Second, the large dwell time proportions of the high similarity regime observed in college women's soccer players' running cadences and distances is notable. In earlier text, we suggested that this is likely due to the task constraints of high-level competitive soccer.

The above methodological and substantive findings are salient in the fledgling science of collective synchrony, and they point to multiple avenues for future work. Regarding the methodological approach, first, simulation studies are needed to systematically evaluate the RSDFM approach under varying conditions such as number of persons, number of time points, frequency of regime switching, and other model parameters. Simulations in which skewness, kurtosis, and model error are manipulated would shed light on the robustness of

estimation and model selection when the assumption of normality is violated. It will also be important to investigate whether intraindividual standardization affects the standard error estimates. This issue is well documented in structural equation modeling (Cudeck, 1989; Krane & McDonald, 1978), but it is unclear whether it applies in state space modeling. Second, more application is needed with multivariate time series exhibiting greater balance in behavioral similarity, unlike in our empirical application where the high similarity regime dominated. Third, it would be worthwhile to assess the value of approaching the regime-switching framework as a continuous time model (Voelkle & Oud, 2013). This may be particularly advantageous when observations are not equally spaced and/or when individuals within a collective are observed over the same time period but not at the exact same time points. Fourth, multifactor models may be useful to assess whether there are “sub-collectives” within a team. In other words, are there subgroups within a team that demonstrate similarity such as attackers/defenders or left/central/right positions? Fifth, beyond *magnitude* and *prevalence*, which we highlighted as two important features, it may be worthwhile to explore the *stability* of behavioral similarity. That is, for what duration does a team typically reside in one regime before switching to the other? Sixth, selecting an overall ARMA order for multi-subject time series models is a broad and difficult challenge (see also Rovine & Walls, 2006) that needs to be addressed. Finally, from a methods perspective, it is important to consider whether a RSDFM can be formulated to effectively detect more subtle regime changes (e.g., medium vs. low vs. none). It may be advantageous to establish guidelines for determining absolute cutoffs to differentiate between low similarity and no similarity. In this paper, we have used effect size (explained variance) and correlation to quantify the magnitude of similarity. Leveraging null hypothesis significance testing (i.e., p-values and confidence intervals) is a possibility for establishing a cutoff between low and no similarity. For example, one possibility is to create pseudo-collectives by randomly drawing time series from different games, then computing confidence intervals of correlations among these unrelated time

series, in order to characterize a state of no similarity.

There are numerous possible substantive directions for future research. More studies are needed to understand the relationship between team performance and collective physiological synchrony, for example. This has been the focus of a few studies (Elkins et al., 2009; Henning, Boucsein, & Gil, 2001; Strang, Funke, Russell, Dukes, & Middendorf, 2014), but this relationship is still not well understood. As it was suggested in the introduction of this paper, an important future direction will be to control for copresence, characteristics of the shared task, and/or coordination, in order to tease out the effect of each on collective behavioral and physiological synchrony. Understanding the unique role of each antecedent in the context of team performance would have tremendous scientific and practical importance. For example, to what extent is collective synchrony in physiology related to interindividual matching of emotion, and not simply a byproduct of the metabolic demands of physical exertion? It would also be useful to investigate collective synchrony in other variables not included in this paper such as players' direction of movement (i.e., change in longitudinal and lateral position). Another critical question that remains is whether behavioral similarity can be developed within a team (i.e., can a team gain expertise in it, and can it be trained?). Finally, another important remaining issue for future research is to clarify what specifically is the relationship between behavioral similarity and collective flow. If collective flow is an outcome of behavioral similarity as we have rendered previously (Smith, 2018), then it is plausible that regime changes in behavioral similarity may reflect the emergence/departure of collective flow states. Although there is still much to uncover about how and why behavioral similarity manifests during team performance, this paper has taken some very important steps to establish a framework for scholarly inquiry in the field.

References

- Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected papers of hirotugu akaike* (pp. 199–213). Springer.
- Araújo, D., Davids, K., & Hristovski, R. (2006). The ecological dynamics of decision making in sport. *Psychology of Sport and Exercise*, 7(6), 653–676.
- Bowerman, B., O’Connell, R. T., & Koehler, A. B. (2005). *Forecasting, time series, and regression: An applied approach* (4th ed.). Cincinnati, OH: South-Western College Pub.
- Cannon-Bowers, J., Tannenbaum, S., Salas, E., & Volpe, C. (1995). Defining competencies and establishing team training requirements. In R. Guzzo & E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp. 333–380). San Francisco, CA: Jossey-Bass.
- Chow, S.-M., Ho, M.-H. R., Hamaker, E. L., & Dolan, C. V. (2010). Equivalence and differences between structural equation modeling and state-space modeling techniques. *Structural Equation Modeling*, 17, 303–332.
- Chow, S.-M., Witkiewitz, K., Grasman, R., Hutton, R. S., & Maisto, S. A. (2014). A regime-switching longitudinal model of alcohol lapse-relapse. In P. C. M. Molenaar, R. M. Lerner, & K. M. Newell (Eds.), *Handbook of developmental systems theory and methodology* (pp. 397–424). Guilford Press.
- Chow, S.-M., Witkiewitz, K., Grasman, R., & Maisto, S. A. (2015). The cusp catastrophe model as cross-sectional and longitudinal mixture structural equation models. *Psychological Methods*, 20(1), 142–164.
- Collins, L. M., & Graham, J. W. (2002). The effect of the timing and spacing of observations in longitudinal studies of tobacco and other drug use: Temporal design considerations. *Drug & Alcohol Dependence*, 68, 85–96.
- Cudeck, R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin*, 105(2), 317–327.

- Duarte, R., Araújo, D., Correia, V., & Davids, K. (2012). Sports teams as superorganisms: implications of sociobiological models of behaviour for research and practice in team sports performance analysis. *Sports Medicine*, 42(8), 633–642. doi: 10.2165/11632450-000000000-00000
- Duarte, R., Araújo, D., Correia, V., Davids, K., Marques, P., & Richardson, M. J. (2013). Competing together: Assessing the dynamics of team-team and player-team synchrony in professional association football. *Human Movement Science*, 32(4), 555–566. doi: 10.1016/j.humov.2013.01.011
- Durbin, J., & Koopman, S. J. (2012). *Time series analysis by state space methods* (Second ed.).
- Elkins, A. N., Muth, E. R., Hoover, A. W., Walker, A. D., Carpenter, T. L., & Switzer, F. S. (2009). Physiological compliance and team performance. *Applied Ergonomics*, 40(6), 997–1003. doi: 10.1016/j.apergo.2009.02.002
- Engle, R., & Watson, M. (1981). A one-factor multivariate time series model of metropolitan wage rates. *Journal of the American Statistical Association*, 76(376), 774–781. doi: 10.1080/01621459.1981.10477720
- Feldman, R. (2003). Infant–mother and infant–father synchrony: The coregulation of positive arousal. *Infant Mental Health Journal*, 24(1), 1–23.
- Hamaker, E. L., & Grasman, R. P. P. P. (2012). Regime switching state-space models applied to psychological processes: Handling missing data and making inferences. *Psychometrika*, 77(2), 400–422.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, 57, 357–384.
- Harvey, A. C. (1989). *Forecasting, structural time series models and the kalman filter*. Cambridge, UK: Cambridge University Press.
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. *Current*

703 *Directions in Psychological Science*, 2(3), 96–100.

704 Henning, R. A., Boucsein, W., & Gil, M. C. (2001). Social-physiological compliance as a
705 determinant of team performance. *International Journal of Psychophysiology*, 40(3),
706 221–232. doi: 10.1016/S0167-8760(00)00190-2

707 Ho, M.-H. R., Shumway, R., & Ombao, H. (2006). The state space approach to modeling
708 dynamic processes. In T. A. Walls & J. L. Schafer (Eds.), *Models for intensive*
709 *longitudinal data* (pp. 148–175). New York, NY: Oxford. doi:
710 10.1093/acprof:oso/9780195173444.001.0001

711 Hunter, M. D. (2018). State space modeling in an open source, modular, structural
712 equation modeling environment. *Structural Equation Modeling: A Multidisciplinary*
713 *Journal*, 25(2), 307–324.

714 Jones, R. H. (1993). *Longitudinal data with serial correlation: a state-space approach*.
715 London, UK: Chapman & Hall Press.

716 Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal*
717 *of Basic Engineering*, 82(1), 35–45.

718 Kim, C. J., & Nelson, C. R. (1999). *State-space models with regime switching: classical and*
719 *gibbs-sampling approaches with applications*. Cambridge, MA: MIT Press.

720 Krane, W. R., & McDonald, R. P. (1978). Scale invariance and the factor analysis of
721 correlation matrices. *British Journal of Mathematical and Statistical Psychology*,
722 31(2), 218–228.

723 López-Felip, M. A., Davis, T. J., Frank, T. D., & Dixon, J. A. (2018). A cluster phase
724 analysis for collective behavior in team sports. *Human Movement Science*, 59,
725 96–111. doi: 10.1016/j.humov.2018.03.013

726 Molenaar, P. C. M. (1985). A dynamic factor model for the analysis of multivariate time
727 series. *Psychometrika*, 50(2), 181–202. doi: 10.1007/BF02294246

728 Nesselroade, J. J., McArdle, J. J., Aggen, S. H., & Meyers, J. H. (2002). Dynamic factor
729 analysis models for representing process in multivariate time-series. In

D. M. Moskowitz & S. L. Hershberger (Eds.), *Modeling intraindividual variability with repeated measures data: Methods and applications* (pp. 233–266). Mahwah, NJ: Erlbaum.

Ou, L., Hunter, M. D., & Chow, S.-M. (2018). dynr: Dynamic modeling in r [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=dynr> (R package version 0.1.12-5)

Palumbo, R. V., Marraccini, M. E., Weyandt, L. L., Wilder-Smith, O., McGee, H. A., Liu, S., & Goodwin, M. S. (2017). Interpersonal autonomic physiology: A systematic review of the literature. *Personality and Social Psychology Review*, 21(2), 99–141. doi: 10.1177/1088868316628405

R Core Team. (2017). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from <http://www.R-project.org/>

Rovine, M. J., & Walls, T. A. (2006). Multilevel autoregressive modeling of interindividual differences in the stability of a process. In T. A. Walls & J. L. Schafer (Eds.), *Models for intensive longitudinal data* (pp. 127–147). New York, NY: Oxford.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461–464.

Shumway, R. H., & Stoffer, D. S. (2010). *Time series analysis and its applications: with r examples*. Springer Science & Business Media.

Smith, D. M. (2018). *Collective synchrony in team sports* (Unpublished doctoral dissertation). University of Rhode Island.

Smith, D. M., & Walls, T. A. (2016). Multiple time scale models in sport and exercise science. *Measurement in Physical Education and Exercise Science*, 20(4), 185–199.

Strang, A. J., Funke, G. J., Russell, S. M., Dukes, A. W., & Middendorf, M. S. (2014). Physio-behavioral coupling in a cooperative team task: Contributors and relations. *Journal of Experimental Psychology: Human Perception and Performance*, 40(1), 145–159.

- 757 Voelkle, M. C., & Oud, J. H. (2013). Continuous time modelling with individually varying
758 time intervals for oscillating and non-oscillating processes. *British Journal of*
759 *Mathematical and Statistical Psychology*, 66(1), 103–126.
- 760 Walls, T. A., Barta, W. D., Stawski, R. S., Collyer, C., & Hofer, S. M. (2012).
761 Time-scale-dependent longitudinal designs. In B. Laursen, T. D. Little, & N. A. Card
762 (Eds.), *Handbook of developmental research methods* (pp. 46–64). New York, NY:
763 Guilford.
- 764 Walls, T. A., & Schafer, J. L. (Eds.). (2006). *Models for intensive longitudinal data*. New
765 York, NY: Oxford.
- 766 Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science*,
767 20(1), 1–5.
- 768 Yang, M., & Chow, S.-M. (2010). Using state-space model with regime switching to
769 represent the dynamics of facial electromyography (EMG) data. *Psychometrika*,
770 75(4), 744–771.

Figure 1. Plots of ACF and PACF of cadence (upper 3 rows) and distance (lower 3 rows) time series data; each row shows ACF (left) and PACF (right) side-by-side for a randomly selected participant; dashed lines indicate $p = .05$ significance limits.

Figure 2. Game 7 cadence data superimposed on predicted regimes; full time series (top panel) and zoomed in (bottom panel).

Figure 3. Game 12 distance data superimposed on predicted regimes; full time series (top panel) and zoomed in (bottom panel).

Figure 4. Overlapping histograms of empirical bivariate correlations between all teammate pairs; zero vs. high similarity regimes for selected RSDFMs only.

Figure 5. Proportion dwell time in high similarity regime, over 18 games.

Figure 6. Mean effect sizes, over 18 games.