



*Citation for published version:*

Stock, HA, Van Emmerik, REA, Wilson, C & Preatoni, E 2018, 'Applying circular statistics can cause artefacts in the calculation of vector coding variability: A bivariate solution', *Gait & Posture*, vol. 65, pp. 51-56.  
<https://doi.org/10.1016/j.gaitpost.2018.06.169>

*DOI:*

[10.1016/j.gaitpost.2018.06.169](https://doi.org/10.1016/j.gaitpost.2018.06.169)

*Publication date:*

2018

*Document Version*

Peer reviewed version

[Link to publication](#)

*Publisher Rights*

CC BY-NC-ND

## University of Bath

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

## Accepted Manuscript

Title: Applying circular statistics can cause artefacts in the calculation of vector coding variability: A bivariate solution

Authors: Holly Stock, Richard van Emmerik, Cassie Wilson, Ezio Preatoni



PII: S0966-6362(18)30936-6  
DOI: <https://doi.org/10.1016/j.gaitpost.2018.06.169>  
Reference: GAIPOS 6286

To appear in: *Gait & Posture*

Received date: 20-12-2017  
Revised date: 26-6-2018  
Accepted date: 26-6-2018

Please cite this article as: Stock H, van Emmerik R, Wilson C, Preatoni E, Applying circular statistics can cause artefacts in the calculation of vector coding variability: A bivariate solution, *Gait and Posture* (2018), <https://doi.org/10.1016/j.gaitpost.2018.06.169>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Applying circular statistics can cause artefacts in the calculation of vector coding variability: A bivariate solution

Holly Stock<sup>a</sup>, Prof. Richard van Emmerik<sup>b</sup>, Dr Cassie Wilson<sup>a</sup> and Dr Ezio Preatoni<sup>a</sup>

<sup>a</sup> Department for Health, University of Bath, UK

<sup>b</sup> Department of Kinesiology, University of Massachusetts, Amherst

Corresponding Author: [holly.a.stock@gmail.com](mailto:holly.a.stock@gmail.com)

Manuscript submitted as an original research article

Word Count: 3572

### Highlights

- Vector coding variability can be artificially inflated by a data processing artefact
- The artefact is caused by the application of circular statistics in the analysis
- The artefact is greater when points on the angle-angle plot are closer together
- Running data may be contaminated by the artefact
- An approach based on the calculation of an ellipse area overcomes the problem

### **Abstract**

**Background:** Coordination variability is thought to provide meaningful insights into motor learning, skill level and injury prevention. Current analytical techniques, based on vector coding (VC) methods, use calculations from circular statistics. However a statistical artefact associated with the application of circular statistics may artificially increase the estimated coordination variability, especially when VC vectors are short.

**Research Question:** Are two popular methods for calculating vector coding coordination variability susceptible to contamination by statistical artefacts and if so, how can coordination variability be calculated without statistical artefact?

**Methods:** A combination of simulated and experimental data was used to prove the existence of the statistical artefact and to understand the extent to which it may affect experimental running gait data, respectively. An alternative approach that uses ellipse area as a bivariate measure of variability was proposed, applied to the same dataset, and compared to two popular methods of coordination variability analysis.

**Results:** The simulated data showed the existence of a statistical artefact, which was greater for shorter VC vector lengths in coordination variability measures that used circular statistics. The statistical artefact typically manifests itself as inflated peaks in the coordination variability trace. The experimental data also indicated that short vector lengths are prevalent in running gait. The Ellipse Area Method of coordination variability was not affected by the VC vector length.

**Significance:** Researchers using current VC variability measures should be particularly aware of the possible effect of the statistical artefact on their data, which is most likely to occur when vector lengths are short. The novel approach we have suggested for calculating VC coordination variability may provide the foundation for future research into vector coding coordination variability.

Keywords: Movement coordination, Coordination variability, Circular statistics, Bivariate statistics, Gait

## Introduction

Variability in human movement when repeating the same task is a certainty and can be attributed to different sources. Changes in the environment, physiological variation in sensory information and motor commands, and error in our ability to measure movement are all possible sources of variability (1). Physiological variation can occur as a result of errors (2) but some variation can have a positive effect on movement performance and can therefore be deemed functional. The functional component of movement variability might be beneficial to expert performance, motor learning, and injury prevention. For example, it can correct for errors which have already occurred (3), could allow an individual to select the best movement from exploring a range solutions (4), and may more evenly distribute loads between different tissues to prevent the accumulation of micro-trauma in one area (5).

The body consists of multiple segments, which interact and coordinate with one another, therefore research on execution variability (i.e. the variability in how a movement is performed) has typically focused on variability in the coordination of movement rather than the variability of isolated joint or segment angle measurements. Coordination is an important aspect of human movement but it is challenging to measure and quantify because it must reflect the movements of multiple individual components. Since the 1950s, angle-angle plots (also known as relative motion plots or cyclograms) have been utilised to graphically represent coordination between two segments or joints. The creation of coupling vectors between adjacent points on the angle-angle plot became termed 'Vector Coding' (6). Two methods of analysing the variability of the coupling vectors (and therefore the coordination between the angle-angle plot variables) were put forward and have appeared frequently in the literature

since. The Tepavac Coordination Variability Method (TCVM) takes account of variation in both the length and direction of the coupling vector (7). The Heiderscheit Coordination Variability Method (HCVM) focuses solely on the variability of the coupling vector direction (8). Details of each method are provided in Appendix 1.

When performing the TCVM and HCVM on gait data, coordination variability is mostly low with sudden peaks (as shown in Figure 3B for the HCVM). This pattern was first observed by Heiderscheit et al. (8), where peaks in coordination variability in specific phases of the movement cycle were primarily explained as functional increases in variability that destabilised the dynamics of the system, thus facilitating a transition between coordination patterns. However, Heiderscheit et al. (8) also suggested the possibility of a statistical artefact by noting that the proximity of adjacent data points on the angle-angle plot (Figure 1A) may artificially affect the calculation of variability in the coupling vectors. The same observation has been reiterated by Mullineaux (9) and the mathematical principle behind this statistical artefact has also been used to justify normalisation in the calculation of continuous relative phase (3). However, the extent to which the statistical artefact affects the HCVM and TCVM data has not been investigated further and has largely been overlooked by those publications which have used these analysis techniques. Therefore, the first purpose of this work was to assess whether the HCVM and TCVM were contaminated by statistical artefact. The second was to propose a novel data analysis technique for the calculation of coordination variability, which is not susceptible to the statistical artefact. Treadmill running was used as a paradigmatic movement for the analysis.

\*\*\*\* Figure 1 near here \*\*\*\*

## Methods

Experimental and simulated approaches were utilised to investigate whether the HCVM and TCVM are subject to a statistical artefact. The experimental approach was used to inform a realistic range of coupling vector length inputs to the simulation and to indicate the potential effect of the statistical artefact on real data. The simulated approach identified the possible effect of the statistical artefact when both the signal and its possible variation were known *a priori*. We propose a new bivariate method of calculating coordination variability based upon Mullineaux's bivariate approach (9) that defines an ellipse from the coupling vector end points and calculates its area. The Ellipse Area Method was then applied to the experimental and simulated data and compared to the traditional coordination variability methods.

## Experimental data

Twenty participants (10 male, 10 female) were recorded running at 12 km/h on a treadmill (Powerjog JX100, Expert Fitness, UK) using a marker-based motion capture system operating at 200 Hz (Qualisys, AB SWEDEN). The study received ethical approval from the University of Bath, Research Ethics Approval Committee for Health, and all participants provided informed consent. An expert tester placed retro-reflective spherical markers (16 mm diameter) on the lower limbs and pelvis. The marker data were labelled and tracked in QTM and exported to Visual3D (V5, C-Motion, USA) where they were low-pass filtered with a cut-off of 8 Hz. Filtered trajectories were used to calculate joint angles for the hip, knee and ankle for all 3 joint rotations of the right leg. These data were exported to MATLAB (Mathworks, Natwick, MA) where a validated kinematic ground-contact algorithm to identify gait events (i.e. Foot Contact Algorithm (10)) was employed to identify 21 consecutive foot-strikes from

the right leg, from which 20 strides of joint angle data were created (average stride duration: 142 (SD 7) frames). Each stride was temporally registered to 101 data points. Coupling vector lengths were calculated for all 36 possible combinations of joint and rotation couplings.

### Simulation data

The movement of two pendula were modelled ((11), Appendix 2). Their reciprocal angular movements relative to the horizontal are described by an angle-angle plot with known features (Figure 1A) containing a range of coupling vector lengths and directions (Figure 1A and 1B) representative of the range of coupling vector lengths measured from experimental data (Figure 2). Twenty exact repetitions of the pendula signals were made to represent multiple cycles of the simulated signal ( $\theta_{p1,i}(t)$  and  $\theta_{p2,i}(t)$ ), where  $t = 1, \dots, 1000$  and  $i = 1, \dots, 20$ . One thousand time points were simulated in order to obtain enough data to show a clear relationship between coupling vector length and coordination variability across the range of coupling vector lengths observed in the experimental running data.

\*\*\*\* Figure 2 near hear \*\*\*\*

In experimental data, repeated cycles are not exact replicas due to variation in both the movement coordination and variation caused by measurement error. Here we have simulated the possible effect of small amounts of random variation added to the signal. In order to do this, random numbers generated from a normal distribution of mean 0 and standard deviation 0.25 degrees ( $\varepsilon_{1,i}(t)$  and  $\varepsilon_{2,i}(t)$ ) were added to the original time-series:

$$\begin{aligned}\hat{\theta}_{p1,i}(t) &= \theta_{p1,i}(t) + \varepsilon_{1,i}(t) \\ \hat{\theta}_{p2,i}(t) &= \theta_{p2,i}(t) + \varepsilon_{2,i}(t)\end{aligned}\quad (1) \quad \begin{array}{l} \text{Addition of white noise,} \\ \varepsilon \sim \mathcal{N}(0, (0.25)^2), \text{ to simulated pendulum} \\ \text{data} \end{array}$$

A standard deviation of 0.25 degrees was selected to represent a theoretical variation that caused a spread in data of no more than 1 degree at each time point.

### Traditional Coordination Variability Methods

The HCVM (8) and TCVM (7) were used to calculate coordination variability of the 20 repetitions of simulated pendulum data (Appendix 1). The TCVM value was subtracted from 1 so that higher values represented greater coordination variability (12) as is the case for the HCVM. Both the HCVM and TCVM have a minimum value of 0 but are capped at  $\sim 81^\circ$  and 1 respectively. Coupling vector length and average coupling vector length were also computed across the 20 cycles at each point in time.

### Bivariate Ellipse Area Method

By calculating the difference between adjacent data points, the vector coding approach effectively normalises coupling vectors to originate from the origin. In the HCVM and TCVM, circular statistics are then used to compare the directional spread of their end coordinates. We propose that forming an ellipse from the coupling vector end point coordinates at each time point, which have been normalised to the origin, would provide a bivariate measure of variability in the direction and length of the coupling vectors. We defined the ellipse using the same calculations as reported by Duarte et al. (13), differing only in the determination of the size of the ellipse where the chi-squared scaling factor (k, Equation 4) suggested by Mullineaux (9) was employed. These have been repeated for convenience (Equation 2-6, reported in MATLAB code format).

|   |  |
|---|--|
| $C(t) = \text{cov}(\Delta\theta_1(t), \Delta\theta_2(t))$ (2) | The covariance matrix at each time point (t) for n stride cycles   |
| $[\sim, \lambda(t)] = \text{eig}(C(t))$ (3)                   | Eigenvalues of the covariance matrix   |
| $k = \sqrt{-2 \cdot \log_e(1 - p)}$ (4)                       | The Chi-squared scaling factor where the probability (p) that a given point will lie within the defined ellipse, was set to 0.95 |
| $X(t) = k \cdot \text{svd}(\lambda(t))$ (5)                   | The magnitudes of the two ellipse axes are altered according to the constant k   |
| $A(t) = \Pi \cdot \text{prod}(X(t))$ (6)                      | The area of the ellipse, representing a bivariate measure of coordination variability  |

## Results

In the experimental running data, more than 45% of all possible coupling vector lengths had a magnitude smaller than  $1^\circ$  which is reflected in the right-skewed distribution of the histograms (Figure 2A). The minimum length observed in the experimental data was 0.0, and the maximum coupling vector length observed was  $6.4^\circ$ . In the simulated data, the minimum coupling vector length was 0.0 and maximum was  $6.1^\circ$  (Figure 1B). Depending on the particular joint angle coupling selected, coupling vector length distributions varied (Figure 2B).

In the simulated data, steep rises in coordination variability (CV) were observed for both the HCVM (Figure 1D) and TCVM (Figure 1F) when the average length of the coupling vectors was low (Figure 1B). The four highest peaks in  $CV_{\text{HCVM}}$  and  $CV_{\text{TCVM}}$  coincided temporally with the lowest troughs in average length of the coupling vectors (at 4, 36, 67 and 98% time points). The standard deviation of the  $CV_{\text{TCVM}}$  signal expressed as a percentage of the possible range in that measure (0 - 1) was 5% compared to 10% for the  $CV_{\text{HCVM}}$  (possible range of 0 -  $81^\circ$ ), indicating that the HCVM was more sensitive to the presence of shorter coupling vector lengths than the TCVM.

The statistical artefact in the estimation of coordination variability for coupling vectors of smaller magnitude was noticeable for both the HCVM and the TCVM (Figure 1C and 1E) but not for the Ellipse Area Method (Figure 1G and 1H). The TCVM curve stabilised at shorter vector lengths (Figure 1E) compared to the HCVM (Figure 1C).

In the experimental data, steep rises in HCVM coordination variability (Figure 3B) coincided temporally with periods where the coupling vector length was shorter (Figure 3A). Some of the same peaks were observed in the Ellipse Area Method but they were less prominent in comparison to the rest of the signal (Figure 3B).

\*\*\*\* Figure 3 near hear \*\*\*\*

## Discussion

The aim of this paper was to demonstrate whether a statistical artefact may affect the calculation of coordination variability using conventional approaches such as Heiderscheit's (HCVM) and Tepavac's

(TCVM) vector coding methods. We have demonstrated that when a pre-set amount of variation was added to repetitions of the same simulated signal, steep peaks in coordination variability calculated using the HCVM and the TCVM are observed where a steady variability output would be expected. These peaks (i.e. inflated variability) occurred when coupling vector lengths between data points on the angle-angle plot were shorter, which we propose was the effect of the statistical artefact associated with the use of circular statistics. We have also shown that the shorter coupling vector lengths, which we have associated with the statistical artefact, are prevalent in experimental data, upon which these analyses are commonly performed. As a possible solution to the reported limitations of existing methods, we have suggested the Ellipse Area Method, which is a bivariate measure of spread, and have demonstrated that it is unaffected by coupling vector length.

In the simulated data where the variation in the signal was pre-defined, steep rises were seen in CV calculated using the HCVM and TCVM (Figure 1D and 1F). These coincided temporally with the shortest coupling vectors (Figure 1B). The relationship between CV and coupling vector length when the coupling vectors were shorter was clearly non-linear (Figure 1C and 1E) indicating the presence of the statistical artefact for the HCVM and TCVM. No relationship was apparent between the Ellipse Area Method and coupling vector length (Figure 1G). Comparing the HCVM and TCVM, coordination variability stabilised at shorter coupling vector lengths in the TCVM (Figure 1E) than in the HCVM (Figure 1C). In addition, the fluctuations in the simulated coordination variability calculated using the TCVM (Figure 1F) were less pronounced in relation to the possible range of the signal when compared to the HCVM (Figure 1D). Thus, although the TCVM was affected by the length of the coupling vector, it appeared to be more robust to the effect of the statistical artefact than the HCVM. This observation was not surprising as the HCVM solely measures the variability of the coupling vector angles and is therefore directly affected by the statistical artefact associated with the use of circular statistics. In comparison, the TCVM calculation has a component that is related to the variability in the lengths of the coupling vectors. In order for TCVM to approach its maximum value of one, the variability of the coupling vector length component would also need to be close to one (Appendix 1, equations E-I). Therefore, the variability of the vector length component is likely to have damped the increases in variability caused by the artefact.

The principle which causes the statistical artefact in the HCVM and TCVM lies in the treatment of these data as circular. If the variability of coupling vectors on the angle-angle plot are analysed with circular statistics, it must be accepted that if the variation of the data causes the spread of coupling vector end points to be centred round the origin, then the angular deviation will approach its maximum value. The same linear variation in abscissa and ordinate coordinates of the coupling vector end points not centred about 0 will have a lower angular deviation (Figure 4A). This is regulated by a non-linear relationship, similar to that displayed in Figure 1C. Regardless of whether the variation in the data is only small, as was simulated in this paper, or larger, as may be the result of actual variation in the movement performance, the use of circular statistics elicits a non-linear relationship between coupling vector length and angular deviation. This characteristic of circular statistics has a direct effect on the calculation of coordination variability using the HCVM and TCVM and is most likely to occur when coupling vector lengths are shorter.

\*\*\*\* Figure 4 near here \*\*\*\*



Possible solutions to this problem might have been to correct for the relationship or to discount data that was under a certain coupling vector length threshold (e.g. starting from the relationships in Figure 1C and 1E). The limiting factor with such approaches is that the relationship between coupling vector length and coordination variability is moderated by the magnitude of the variation in the data at each time point, which is difficult to model and likely impossible to predict *a priori*. In this paper the simulated variation had a SD of ~0.25 but variation in real data may increase or decrease over the course of the movement. The statistical artefact was still apparent when the simulation was run with different magnitudes of variation added to the pendulum signal. However, the coupling vector length threshold under which the statistical artefact dominated, increased when the variation was greater (Figure 5). Thus, with different variations in the data at each part of the signal, it is challenging to define a consistently valid threshold for a coupling vector length under which we are confident the result is dominated by the statistical artefact. In addition to this, the simulation we have presented models a situation where the variance in the abscissa and ordinate components of the angle-angle plot are equal. In reality, unequal variances can and will occur in ways that are currently impossible to predict, which adds another layer of complexity to the problem. If the main axis of variation at one time point across multiple cycles is perpendicular to the average vector, the effect of the statistical artefact is high. This decreases as the axis of variation becomes parallel to the mean vector (Figure 4B-C). This consideration makes both understanding whether data are affected by the statistical artefact and the possibility of applying a correction to the HCVM and TCVM yet more complicated and reinforces the need of a more robust estimator of coordination variability. Artificially increasing vector length would also be ineffective, as it would proportionally increase the variation in the data leaving the statistical artefact unchanged.

\*\*\*\* Figure 5 near here \*\*\*\*

The combination of simulated and experimental data presented in this article suggested that experimental data has the strong potential to be affected by statistical artefact, but it is difficult to quantify its effect. Looking at the running data we collected in more detail, the steep rises in coordination variability coincided temporally with periods where the coupling vector length was shorter (Figure 3), which could be indicative of contamination of the coordination variability signal due to the statistical artefact. We believe the statistical artefact is problematic in the data we collected, but the severity of its effect is ultimately dependent on the coupling vector length and the characteristics of the variation in the data. Therefore the phase of the movement studied, the coupling that was analysed, the number of data points signals were time registered to, and the participant and movement in question can all affect the magnitude of the statistical artefact. In our data, the shortest coupling vectors often coincided with times when one or both of the component variables of the coordination coupling changed direction (e.g. when a joint moves from flexion to extension). Thus, if an analysis of coordination variability were to focus on a period of the movement where a change in joint motion occurred, then the statistical artefact would likely have a greater influence on the outcome than if the average of the entire gait cycle were taken. Shorter vectors were also more likely throughout the signal when one or both of the component variables of the coordination coupling had a small range of motion. Considering the many factors that can impact upon the coupling vector length, we recommend that researchers using the HCVM and TCVM for the assessment of coordination variability should investigate their own data to understand the possible effect of the artefact or consider other methods of analysing coordination variability.

As an alternative to considering angle-angle plot data as circular, Mullineaux (9) recently suggested a bivariate approach to analyse differences between angle-angle plots. In a similar vein, we have suggested that bivariate statistics may be appropriate for the analysis of coordination variability. The ellipse method accounts for variability in both the length and direction of the coupling vectors, comparable to the TCVM. We have demonstrated that the Ellipse Area Method is unaffected by coupling vector length (Figures 1G and 1H). We believe this will make this method popular for calculating coordination variability in the future as it negates the need to understand and account for the complex nature of the statistical artefact associated with using circular statistics.

To demonstrate the Ellipse Area Method further we applied it to experimental data from a single participant and compared the results to those from the HCVM (Figure 3B). Some of the signal's pattern was maintained but the effect was not consistent across the signal. It is possible that the peaks caused by the statistical artefact in the HCVM may have detracted attention from other features in the data and so these features became more apparent when the Ellipse Area Method was used. Further work will be required to confirm the differences between methods on a larger sample of data than is presented here for the individual and to determine the best method of defining the ellipse. In the examples provided, ellipse areas were calculated from covariance matrices, but this method has been suggested as sensitive to outliers (9), therefore it may be important to investigate alternative techniques to define the ellipse or to measure bivariate spread.

### **Conclusion**

This is the first study which has investigated the statistical artefact affecting popular biomechanical techniques that quantify coordination variability. We have demonstrated via simulated data that when circular statistics are employed to calculate coordination variability, a statistical artefact can substantially inflate coordination variability values at shorter coupling vector lengths. We have also shown that the coupling vector lengths, where the statistical artefact is high, can be common in experimental gait data. Thus, we recommend future studies should not use methods based on circular statistics without investigating the possible impact of the statistical artefact on their results and conclusions. As a solution to the statistical artefact, we have advocated an alternative method for calculating coordination variability by showing that a bivariate approach is robust to the presence of different coupling vector lengths. An initial example of this method has been presented that may form a new basis for vector coding coordination variability research. Though the bivariate method appears to be a viable replacement for measuring coordination variability, it is important to note that some previous findings, which suggest the importance of coordination variability, are based on results from HCVM and TCVM methods. It may therefore be prudent to confirm these findings remain unchanged in light of this new evidence regarding the methods used to calculate them. In conclusion, this paper has highlighted a possible longstanding issue with current vector coding measures of coordination variability (specifically the HCVM and TCVM) and has also provided additional evidence for an alternative method for continuing research on this topic.

### **Conflicts of interest**

The authors certify they have no involvement with any organisation that has financial interest in the subject matter of this manuscript.

### **Acknowledgements**

We would like to acknowledge the English Institute of Sport for match funding this research with the University of Bath. Thanks also to Nicos Haralabidis and Ben Thomas for proof reading the manuscript.

## References

1. Preatoni E, Hamill J, Harrison AJ, Hayes K, Van Emmerik RE, Wilson C, et al. Movement variability and skills monitoring in sports. *Sports Biomechanics*. 2013;12(2):69-92.
2. Harris CM, Wolpert DM. Signal-dependent noise determines motor planning. *Nature*. 1998;394(6695):780.
3. Button C, Davids K, Schollhorn W. Coordination profiling of movement systems. *Movement system variability*. 2006:133-52.
4. Newell KM, Kugler P, Van Emmerik RE, McDonald P. Search strategies and the acquisition of coordination. *Advances in psychology*. 1989;61:85-122.
5. Hamill J, Van Emmerik R, Heiderscheit BC, Li L. A dynamical systems approach to lower extremity running injuries. *Clin Biomech (Bristol, Avon)*. 1999;14:297-308.
6. Sparrow W, Donovan E, Van Emmerik R, Barry E. Using relative motion plots to measure changes in intra-limb and inter-limb coordination. *Journal of Motor Behavior*. 1987;19(1):115-29.
7. Tepavac D, Field-Fote EC. Vector coding: a technique for quantification of intersegmental coupling in multicyclic behaviors. *Journal of Applied Biomechanics*. 2001;17(3):259-70.
8. Heiderscheit BC, Hamill J, Van Emmerik R. Variability of Stride Characteristics and joint Coordination Among Individuals. *Journal of Applied Biomechanics*. 2002;18:110-21.
9. Mullineaux DR. CI2 for creating and comparing confidence-intervals for time-series bivariate plots. *Gait & posture*. 2017;52:367-73.
10. Maiwald C, Sterzing T, Mayer T, Milani T. Detecting foot-to-ground contact from kinematic data in running. *Footwear Science*. 2009;1(2):111-8.
11. Kolukula SS. Simple Pendulum 2011. Available from: <https://uk.mathworks.com/matlabcentral/fileexchange/33082-simplependulum>.
12. Mullineaux DR, Uhl TL. Coordination-variability and kinematics of misses versus swishes of basketball free throws. *Journal of Sports Sciences*. 2010;28(9):1017-24.
13. Duarte M, Zatsiorsky VM. Effects of body lean and visual information on the equilibrium maintenance during stance. *Experimental brain research*. 2002;146(1):60-9.
14. Samaan MA, Teng H-L, Kumar D, Lee S, Link TM, Majumdar S, et al. Acetabular cartilage defects cause altered hip and knee joint coordination variability during gait. *Clinical Biomechanics*. 2015.
15. Brown AM, Zifchock RA, Hillstrom HJ, Song J, Tucker CA. The effects of fatigue on lower extremity kinematics, kinetics and joint coupling in symptomatic female runners with iliotibial band syndrome. *Clinical Biomechanics*. 2016;39:84-90.

**Figure 1.** Pendulum simulation graphs for 1000 time points (A) Angle-angle plot demonstrating variation in the coupling vector lengths between adjacent data points for the single pendulum simulation upon which the analysis is based. The longest coupling vectors are found on the straights and correspond to the peaks in Figure 1B. The shortest vectors occur around the turning points and correspond to the troughs in Figure 1B (B) Average coupling vector lengths from the 20 repetitions of

the pendulum signal (C) Relationship between the average coupling vector length and HCVM coordination variability (D) HCVM coordination variability (E) Relationship between the average coupling vector length and TCVM coordination variability (F) TCVM coordination variability (G) Relationship between the average coupling vector length and the ellipse method (H) Coordination variability from the newly proposed Ellipse Area Method.

**Figure 2.** Normalised distributions of coupling vector lengths from experimental running data of all time points in the 20 stride cycles of the 20 participants (A) The distribution when all 36 joint couplings from the possible combination of tri-planar hip, knee and ankle rotations are considered together, and (B) Examples of individual coupling distributions for hip ab/adduction – knee rotation (dark grey) and hip flexion/extension – knee flexion/extension (light grey) couplings. These couplings were chosen to demonstrate the diversity of distributions associated with different couplings. They have also been used in other coordination variability literature, e.g. (14) and (15).

**Figure 3.** Example data from an individual participant of hip flexion/extension – knee flexion/extension coupling during treadmill running at 12 km/h calculated from 20 stride cycles (A) Average coupling vector length (B) HCVM coordination variability (solid line) and Ellipse Area Method coordination variability method (dashed line).

**Figure 4** The effect of the average coupling vector length (ACVL) on angular deviation (AD) when variance in the  $\Delta\theta_1$  (abscissa) and  $\Delta\theta_2$  (ordinate) coordinates are equal (A), and when the main axis of variation in the bivariate data is perpendicular (B) and parallel (C) to the average coupling vector. On each plot, circles/ellipses represent three separate average length scenarios that are demonstrated independently of each other (i.e. they are not consecutive time points). Each dot represents the end point of a coupling vector that has been normalised to have its origin at (0,0). Shaded areas are visual representations of the angular variation, which is maximal ( $\sim 81^\circ$ ) when the coupling vectors are spread about the origin (dark grey), but is smaller if vector end points are offset from the origin (mid grey) and smaller again when vector end points are far away from the origin (light grey). Within each figure the AD is very different in the three ACVL scenarios but the circle/ellipse areas around the coupling vector end points are constant.

**Figure 5** The relationship between HCVM coordination variability and average coupling vector length when the standard deviation of the normal distribution from which the random error was calculated and added to both of the simulated pendulum angle signals (Equation 1) were:  $0.10^\circ$  (black),  $0.25^\circ$  (dark grey) and  $0.40^\circ$  (light grey).





