

## ARTICLE OPEN



# Knee flexion of saxophone players anticipates tonal context of music

Nádia Moura<sup>1,6</sup>, Marc Vidal<sup>2,3,4,6</sup>, Ana M. Aguilera<sup>3</sup>, João Paulo Vilas-Boas<sup>5</sup>, Sofia Serra<sup>1</sup> and Marc Leman<sup>2</sup>

Music performance requires high levels of motor control. Professional musicians use body movements not only to accomplish and help technical efficiency, but to shape expressive interpretation. Here, we recorded motion and audio data of twenty participants performing four musical fragments varying in the degree of technical difficulty to analyze how knee flexion is employed by expert saxophone players. Using a computational model of the auditory periphery, we extracted emergent acoustical properties of sound to inference critical cognitive patterns of music processing and relate them to motion data. Results showed that knee flexion is causally linked to tone expectations and correlated to rhythmical density, suggesting that this gesture is associated with expressive and facilitative purposes. Furthermore, when instructed to play immobile, participants tended to microflex (>1 Hz) more frequently compared to when playing expressively, possibly indicating a natural urge to move to the music. These results underline the robustness of body movement in musical performance, providing valuable insights for the understanding of communicative processes, and development of motor learning cues.

npj Science of Learning (2023)8:22; <https://doi.org/10.1038/s41539-023-00172-z>

## INTRODUCTION

Multiple mental and sensorimotor processes operate simultaneously underneath the surface of an optimal music performance. The musician's body plays a critical mediator role between the cognitive processing of musical content and its transformation into the resulting sonic and visual outcomes<sup>1,2</sup>. Research conducted over the last decades has shown that the functions associated with performative body movements go beyond the technical demands of sound production and modification, reaching facilitative, communicative, or expressive purposes<sup>3–9</sup>. The term ancillary gestures was established to define all types of motion, which are not required for playing per se, but potentially support the performance in any other way<sup>9,10</sup>. For example, a saxophone player lifting the instrument's bell to emphasize expressive intentions, or a singer leaning forward to facilitate reaching a louder sound.

Music, similarly to language, is a complex communication system in which information is conveyed through different sensory channels and not exclusively related to the auditory stimuli<sup>11</sup>. In cognitive linguistics, the power of nonverbal behavior in communicative tasks is not limited to speech comprehension and production<sup>12–16</sup>, but expanded as a tool for enhancing learning processes<sup>14,17–20</sup>. In music research, there is strong evidence that musicians' body movements affect audience perception<sup>5,21,22</sup> and reflect features of the musical speech<sup>10,23,24</sup>. However, the knowledge about the facilitative role of gestures in music playing is scarce. Literature suggests movement may aid tempo perception<sup>10,23,25</sup> and support the execution of dynamics and higher register tones<sup>9,26</sup>. In summary, it seems clear that expert performers associate expression with movement, although they can perform equally good when it is limited<sup>27</sup>. Further

research is needed to clarify if ancillary gestures constitute, indeed, an aid to musicians' technical execution.

Ancillary movement is highly idiosyncratic, as it reflects the individual style of the performer, or speaker<sup>13,28,29</sup>. Nevertheless, gestural trends have been identified among players of the same instrument<sup>10,23,30,31</sup>, and even different instruments<sup>3,29</sup>. Two major factors support this tendency of moving expressively in similar manners: the biomechanical and the musical. Biomechanical factors comprise the technical demands needed to play accurately, in respect to the weight, shape and posture of execution of the instrument. Examples include the technical difficulty of the passages<sup>10,23,24</sup> and anthropomorphic characteristics of the subjects<sup>26,32</sup>. Musical factors refer to the characteristics of the repertoire being performed, including tempo, dynamics or phrasing structure. Analogously to what occurs with spoken speech<sup>12,13,18,33</sup>, music-related gestures are molded by musical syntax. Musicians' ancillary movements have been related to phrasing structure<sup>10,34–36</sup>, repeated rhythmic patterns<sup>10,23,37</sup>, and expressive climaxes of the piece<sup>29</sup>. This phenomenon is defined as embodiment - the process of transforming inner subjective experiences, ideas or intentions in a physical manifestation<sup>1,38</sup>.

Finally, the importance of ancillary movement is reinforced by the fact that it emerges, although in smaller scales, when performers are asked to restrict to the minimal movements required for playing<sup>6,10,23</sup>. In this cases, it is argued that movement is so deeply rooted in the performers' interpretation that it becomes an intrinsic part of the task. On the other hand, several studies support the existence of an innate urge to move to music among listeners when instructed to remain still<sup>39–43</sup>. Significant differences in quantity of motion were found when comparing listeners trying to standstill in silence to when exposed to music<sup>39,40</sup>. In line with these results, involuntary entrainment to

<sup>1</sup>Research Centre for Science and Technology of the Arts, School of Arts, Universidade Católica Portuguesa, Rua de Diogo Botelho 1327, 4169-005 Porto, Portugal. <sup>2</sup>Institute for Psychoacoustics and Electronic Music, Ghent University, Miriam Makebaplein 1, 9000 Ghent, Belgium. <sup>3</sup>Department of Statistics and Institute of Mathematics, Universidad de Granada, Campus de Fuentenueva, 18071 Granada, Spain. <sup>4</sup>Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, 04103 Leipzig, Germany. <sup>5</sup>Centre of Research, Education, Innovation and Intervention in Sport (CIFIZD), Porto Biomechanics Laboratory (LABIOMEPU), Faculty of Sport, University of Porto, 4099-002 Porto, Portugal. <sup>6</sup>These authors contributed equally: Nádia Moura, Marc Vidal. ✉email: nmoura@ucp.pt; marc.vidalbadia@ugent.be; marc.leman@ugent.be

groovy music was reported in<sup>41</sup>, and discrete auditory events had a significant effect on involuntary body sway<sup>42</sup>. Thus, movement does not have to be voluntary to enhance cognitive processes related to the perception of complex rhythmic patterns<sup>43</sup>. The fact that movement rises unconsciously only strengthens its role in enhancing cognitive and motor functions implied in quotidian activities.

In this study, we focus on the knee flexion as an ancillary gesture with potential facilitative and expressive functions during saxophone performance. We selected four musical passages of increasing technical difficulty to analyze how rhythmical and tonal components of music were embodied in the performers' behavior. Each passage was recorded in an immobile (IMO) and an expressive (EXP) condition, as a means to assess to what extent knee movement is an integral part of the performance.

## RESULTS

### Average causal effect between knee flexion and short-term memory tonal expectations

We characterized pitch expectation profiles using a computational model of the auditory periphery (see Materials and Methods for technical details). These profiles represent probabilities on incoming pitch under a short-term memory model based on an autocorrelation mechanism that has been described to be coherent with Krumhansl and Kessler experiments on tonal induction<sup>44–46</sup>. Here, we use Granger causality<sup>47,48</sup> to see if knee flexion has an effect on participant's internal beliefs of prospective pitch, using a simulated echoic memory buffer generated from their own audio recordings. We performed the analyses on the average of the Hadamard product (element-wise product) between the right and left knee curves and their respective pitch expectation (PE) profiles (both onset aligned, see Materials and Methods) across the four passages (P1–P4). We observed that knee curves could predict pitch expectation as measured by our auditory model (Fig. 1b). The following interactions were found when testing for Granger causality in the EXP condition: P1 [ $p = 0.006, F = 2.344$ ], P2 [ $p = 0.063, F = 1.702$ ], P3 [ $p = 0.009, F = 2.307$ ] and P4 [ $p = 0.002, F = 3.30$ ], for a lag calculated as  $\lceil (\lambda_{\text{global}} * \lambda_{\text{local}}) f_{\text{PE}} \rceil$  ( $\approx [9, \dots, 12]$ ), where  $f_{\text{PE}}$  is the sampling rate of the pitch expectation profile,  $\lambda_{\text{global}} = 1.5$  and  $\lambda_{\text{local}} = 0.1$  (in seconds) are model parameters related to the echoic memory (see Materials and Methods). Evidence of this causal relationships was also found in the IMO condition as involuntary flexion was observed: P1 [ $p = 0.024, F = 2.025$ ], P2 [ $p = 0.009, F = 2.296$ ], P3 [ $p = 0.078, F = 1.663$ ] and P4 [ $p < 0.0001, F = 4.897$ ]; outliers were removed to perform the tests (Fig. 2a). Further analyses including the left and right knee are provided in the Supplementary Material along with an extensive discussion on the selection of the lag length.

### Saxophone players flex their knees according to rhythmical density

We generated rhythmical density profiles (see Materials and Methods) for each subject and passage (P1–P4) for the EXP condition. The knee curves (Hadamard product) were smoothed using functional multivariate principal component analysis (B-spline dimension = 19, Hubber loss = 48.43, cumulative variance = 99.52%)<sup>49,50</sup>; this procedure acts as a low-pass filter that takes into account the second order cross-dependencies across conditions. We performed this step to get a neat curvature and to reduce noise that could overlap and bias the estimation of the phase, which was calculated via discrete Hilbert transform. We then estimated the  $L^2$  norms and phase changes per second of the smoothed curves and weighted them by the inverse of the duration of each participant's performance, to correct for effects produced by the different timings (see Fig. S1a). Pearson correlation coefficient is  $r = -0.7, p < 0.0001$  for the  $L^2$  norm

scores and  $r = 0.63, p < 0.0001$  for the phase changes (Fig. 1c) with regard to the rhythmical density scores ( $L^2$  norms of these curves). We also found significant differences in comparing the scores obtained for P4 with regard to the rest of the passages (Fig. 1d).

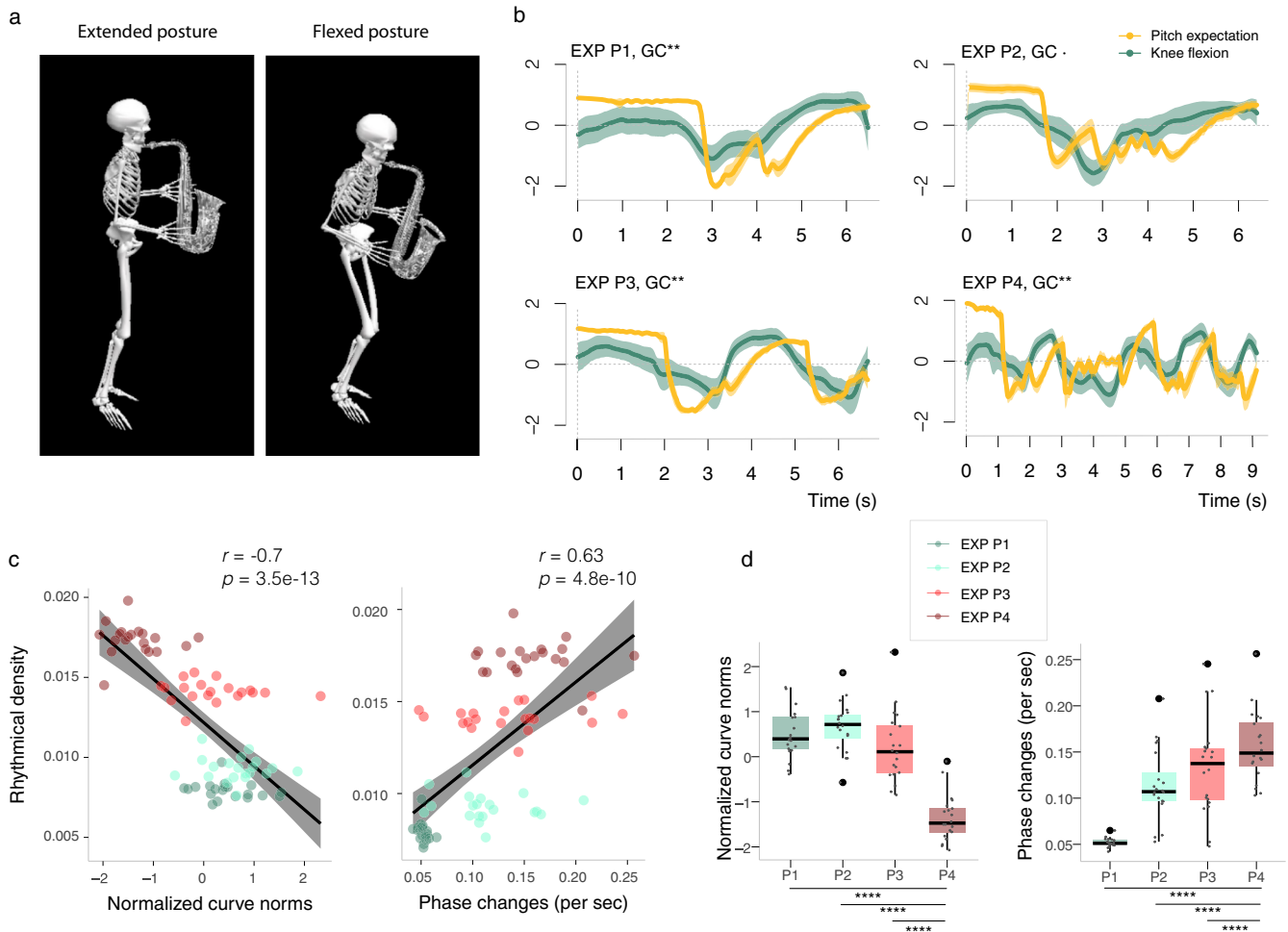
### Outlier and microflexion analyses suggest natural incline to move in immobile condition

We examined the IMO condition to assess the capacity of participants to restrict their movement. A functional outlier analysis on the knee curves in the IMO condition across passages showed that some curves substantially deviated from a constant angular change (Fig. 2a). To demonstrate the unintentional urge of moving when performing, we conducted further analysis by high-pass filtering the knee curves (three-order Butterworth filter with cutoff at 1 Hz, see Supplementary Material), assuming participants would microflex more often while inhibiting movement than when allowed to move. The analysis of the phase counts revealed that for the IMO condition countings were significantly higher than for the EXP condition ( $\chi^2 = 31.49, p < 0.001$ , using ANOVA via Poisson regression with phase counts weighted according to performance duration). This is also confirmed by the analysis of phase changes per second ( $p < 0.001$ , Wilcoxon signed-rank test, see Fig. 2b).

## DISCUSSION

Our results explicitly demonstrate a predictive relationship between the averaged knee curves and their respective pitch expectation curves, suggesting movement not only accompanies, but anticipates pitch perception. In predictive coding frameworks<sup>2,51–55</sup>, the principle of active inference implies that the mind is constantly generating models about the environment based on the sensory inputs it receives. By encoding the discrepancy between the expected and the real outcomes of a task, the brain creates a prediction error that potentiates behavioral changes and, therefore, contributes to the refinement of its predictive model. For example, in tapping studies, it is common that participants tap ahead of the metronome<sup>56,57</sup>. Accordingly, the modeled acoustic expectations of our participants enabled them to project their movement in anticipation, finding which is particularly consistent if we take into account that they had previous contact with the piece. We identified several studies reporting correlations between the musical context and musicians' movements<sup>8,25,35,36</sup> and, specifically, demonstrating an association between the flexion-extension dynamics and the melodic contour in woodwind instrumentalists<sup>9,29</sup>, which is corroborated by our study. Nevertheless, to our knowledge, the anticipatory movement is a novel result among musicians, which may help understand how gestures can be employed to aid performance. In this sense, we propose that the motor prediction of the expected auditory outcome may, in fact, facilitate its technical execution. Similarly, research suggests that the coupling of motor and neural processes in multisensory integration contexts may potentiate learning and memory<sup>27,58</sup>. Further research is needed to explore the beneficial effect of embodiment in technical and interpretative aspects of music performing and learning.

Passages with higher rhythmical density (P3–P4) led participants to flex more often, although with smaller motion amplitude, whereas passages with lower rhythmical density (P1–P2) presented larger and less recurrent knee flexions (Fig. 1c, d). Considering that P3–P4 represent higher technical difficulty (due to the fast finger coordination required to play sequences of short-duration notes), the reduction of motion amplitude in these scenarios may be interpreted as a postural strategy with ultimate energy focusing purposes. In technically demanding sections, performers usually reduce their amplitude of motion and limit to focal movements required for accurate execution<sup>10,23,24</sup>. Since



**Fig. 1 Tonal and rhythmical relations to knee flexion.** **a** 3D model of one participant in extended (left) and flexed (right) posture. **b** Comparison of the average Hadamard product curves between right and left knees to pitch expectation profiles (curves were normalized for a better visualization). Granger causality (GC) interactions are noted for a lag equal to 0.15 s onset equivalence; error bands were calculated at a 95% bootstrap confidence interval. **c** The two plots show rhythmical density (vertical axis) compared to: weighted and normalized  $L^2$  norms (area below the curve) and phase changes per second of all curves in the expressive condition (EXP). Confidence intervals for the regression line are at 95%. **d** Boxplots showing differences across groups of passages for curve norms and phase changes. Statistical comparisons were made using the Wilcoxon signed-rank test (with Bonferroni-Holm correction for multiple comparisons). The centre line of each boxplot represents the data median and the bounds of the box show the interquartile range. The whiskers represent the bottom 25% and top 25% of the data range—excluding outliers, which are represented by a rounded point. —Significance levels are noted as follows:  $p < 0.1$ ,  $*p < 0.05$ ,  $**p < 0.01$ ,  $***p < 0.001$ ,  $****p < 0.0001$ , n.s. (not significant).

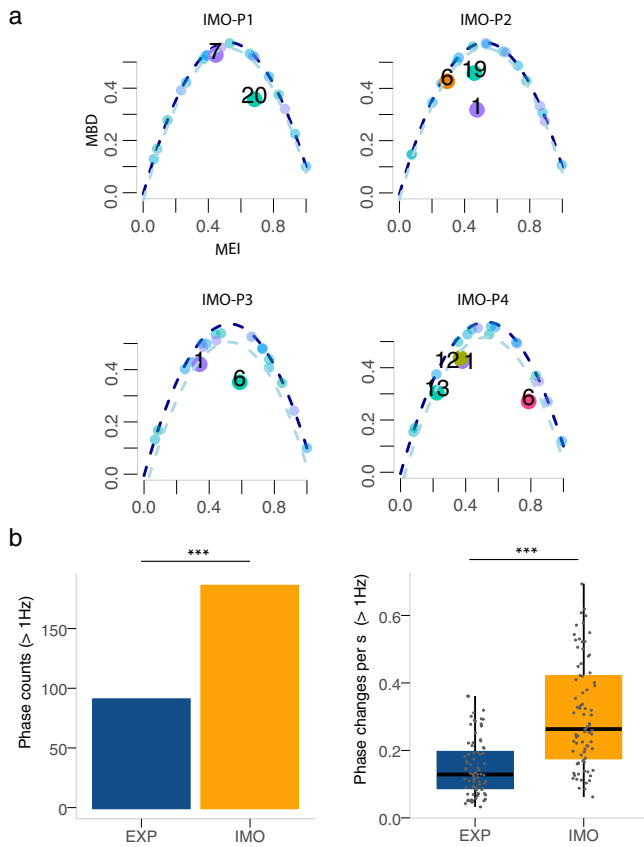
knee flexion is not directly related to sound emission in saxophone playing, we propose that it has an active role in the maintenance of balance and postural control when playing in a standing position.

Previous studies on quiet human stance have demonstrated that the knee joint takes part in the stabilization of the centre of mass<sup>59,60</sup> and in balance control, in synchronization with the ankle and hip joints<sup>61</sup>. It is now acknowledged that postural control implies multi-joint activation involving the knee, even though the ankle and hip joints are the most activated in this context, and widely studied for decades<sup>62,63</sup>. Postural joint synergy is reinforced by the fact that when the knee joint is constrained, the reliance on ankle strategy increases<sup>64</sup>, and, inversely, when the ankle joint is constrained, adaptation shifts to knee-engaging dynamics<sup>65,66</sup>. Additionally, in the specific case of focal arm movement tasks, which somehow share similarities to the playing task, the activation of leg musculature acts as a compensation of external destabilizing forces<sup>67</sup>.

Microflexions occurred more often during the IMO condition when compared to the EXP, suggesting that saxophone players

presented an unintentional inner urge to move during performance, although they were instructed to limit to the minimal movements required for playing. A corpus of research validates music induces involuntary movement, even when subjects try to standstill<sup>39–43</sup>. It is possible that movement is so integrated in the participants' performative habits that it becomes difficult to erase in a preliminary approach. In order to automatize the playing process, musicians develop motor programs where executive and interpretative functions coexist<sup>5</sup>. After repeated practice of the same piece, these motor programs become ingrained and hard to remove<sup>6,10</sup>. Small movements have emerged before in performances with movement restriction<sup>6,10,23,68</sup>. Nonetheless, in order to further strengthen this finding, it would be pertinent to compare the immobile condition with several kinds of resting state behavior (e.g., holding or not the instrument, etc.) or behavioral data assessment.

Preserving the natural expressive behavior of the participants was a major consideration in our design. For that, we used musical excerpts from standard saxophone repertoire, assuring high degrees of familiarity with the interpreted music. The passages



**Fig. 2 Analyses of participants in immobile condition (IMO) during performance.** **a** Functional outlier analysis representation for each passage P1–P4. The vertical axis is the modified band depth (MBD) and the horizontal the modified epigraph index (MEI); larger points and numbers represent outliers related to participants. **b** Boxplots of weighted phase counts and phase changes per second comparing expressive (EXP) and immobile (IMO) conditions. Statistical comparisons were made using the one-sided Wilcoxon signed-rank test (with Bonferroni-Holm correction). The centre line of each boxplot represents the data median and the bounds of the box show the interquartile range. The whiskers represent the bottom 25% and top 25% of the data range. –Statistical significance is measured as detailed in Fig. 1.

were selected considering a criteria of redundancy, which allowed for fair in-between comparisons. In this sense, the musical contexts we studied were characterized by starting with descendant pitch lines and presenting low rhythmical density associated with higher pitches. Follow-up studies including other variations in pitch, rhythm, but also structure or articulation, are necessary to clarify the role of knee flexion in diverse musical contexts. Similarly, other musical styles and genres should be considered.

In conclusion, our study provides novel evidence about the predictive role of the sensorimotor system in music performance. We propose that anticipatory knee movement facilitates the cognitive processing involved in motor adaptation, and consequently supports technical execution. This, in addition to the correlations found between rhythmical density profiles of the music and phase and amplitude of knee movement, and the underlying urge to move in immobile conditions, reinforces the central role of the body in musical expression. We suggest these results constitute a valuable contribution for other fields involving communication, such as the performing arts, linguistics, or human-computer interaction. Furthermore, this study provides insights for the development of motor learning programs focused on proprioceptive and kinaesthetic awareness.

## METHODS

### Participants

The recordings used for this study were retrieved from the database reported in ref. <sup>69</sup>, including a sample of 20 participants (9 female, 11 male; mean age: 26.3 years  $\pm$  5.4 years; mean duration of saxophone practice: 16.9 years  $\pm$  5.6 years). This study was approved by the Ethics Committee for Health of Universidade Católica Portuguesa (protocol number 137/2021), and the procedures were carried out according to the Declaration of Helsinki. Prior to the recording session, all participants were informed about the data collection and protection protocols, signed a written informed consent form and were sent the repertoire for preparation. Also, at this point, we confirmed that they had already studied the piece before.

### Musical repertoire

The musical passages used for this study were extracted from the piece *Concertino da Camera* for solo saxophone and orchestra, by Ibert<sup>70</sup>. They were selected as representative of the two musical contexts planned for our study: melodic lines comprising an interval jump, and technical passages with rhythmical density. Two passages per context were chosen for redundancy purposes, resulting in a total of four passages for analysis. Music scores and analyses are presented in the Supplementary Material.

### Data collection

Motion data was recorded with an optical 3D motion capture system with 9 infrared cameras, at a sampling frequency of 240 Hz. The marker setup adopted included 58 retroreflective markers placed in the body of the subject and 9 markers placed in the saxophone (for more details, see ref. <sup>69</sup>). Motion data was then processed using Qualisys Track Manager 2021 (QTM). Knee angular values were extracted from QTM at a sampling rate of 240 Hz, by calculating the angle between the lateral markers of the right and left ankle, knee and hip. Audio data was recorded with a Zoom H4N recorder, at a frequency rate of 44.1 Hz. Additionally, to align all data channels, we recorded video using a Canon EOS 100D with a 18-55mm lens, introducing an audiovisual signal with a clapperboard in the beginning and end of each take (2 markers attached). We then synchronized timestamps for the multiple devices retrospectively in respect to this signal.

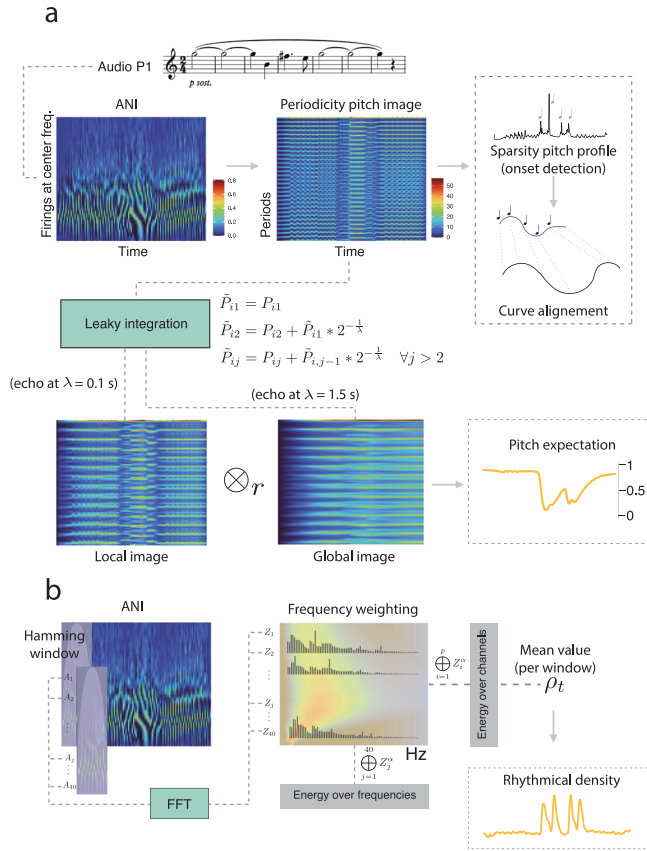
In the recordings, participants were asked to play in two conditions: immobile (restricting their movement to the minimum required for effective playing) and expressive (moving as idealized in a concert situation). Conditions and passages were randomized. Players had the freedom to perform in the tempo they found appropriate to their interpretation.

### Curve registration using onset positions in the audio signal

Here, we describe the procedure to align the knee curves to a set of audio features by transforming their abscissa variables. Registration was conducted by retrieving strategic note onsets (see Supplementary Fig. 2 in the Supplementary Material) from the audio using computational modeling of the human auditory periphery (Fig. 3a). The challenge here is the homogeneity of tone changes. The merging of fundamental and harmonic frequencies in the spectral envelope, along with small oscillations in the air column that are characteristic of the saxophone acoustic signal<sup>71</sup> made it difficult to detect note onsets with some existent methods.

To simulate the auditory brainstem response to sound, the acoustic signal is firstly converted into a multivariate pattern of neural impulses<sup>72,73</sup>. The algorithm generates, from a mono audio signal  $s_t$  ( $t = 1, \dots, N$ ), a matrix (or auditory nerve image, ANI)  $A \in \mathbb{R}^{q \times n}$  with  $q = 40$  channels representing the instantaneous amplitude across frequency sub-bands half a critical band apart





**Fig. 3 Schematic of the auditory modeling framework. a** The auditory nerve image (ANI) is estimated from a mono audio signal and then converted into periodicity pitch image from which we extract different pitch profiles for further analysis. Leaky integration is performed for echoes of 0.1, 1.5 s on the periodicity pitch image to obtain two images of different echo whose Pearson's correlation index between running columns gives the pitch expectation profile. **b** Calculation of the rhythmical density from ANI. The operator  $\oplus$  is used to denote the column/row sums of the matrix  $Z^a$ .

from each other, with center frequencies on the range 140 Hz to 8877 Hz. Thus,  $A$  stands for a time-frequency representation with frequencies spanning the audible range, capturing the information in  $s_t$  as an auditory percept that simulates the peripheral auditory system function. The ANI sampling rate was determined as  $f_{\text{ANI}} = 11025/4$ , so that  $n = (N * f_{\text{ANI}})f_s^{-1}$ , where  $f_s$  denotes the sampling rate of  $s_t$ .

A second order Butterworth band-pass filter (80–1250 Hz) is applied to each row of the ANI matrix  $A_i \rightarrow \tilde{A}_i$  ( $i = 1, \dots, q$ ) to retrieve pitch periodicity patterns. Then, a frame-based convolution analysis and a coincidence mechanism are conducted by the summation of the convolution results over all channels. For a fixed frame width  $\delta$  (usually of 60 ms<sup>44</sup>) and frame lag  $w$  (10 ms), each column of the periodicity pitch image is obtained from a filtered ANI image  $\tilde{A}$  as

$$\begin{aligned} P_1 &= \sum_{i=1}^q \tilde{A}_{i,1:\delta} \tilde{A}_{i,(\delta+1):2\delta} \\ P_2 &= \sum_{i=1}^q \tilde{A}_{i,(w+1):(w+\delta)} \tilde{A}_{i,(w+\delta+1):(w+2\delta)} \\ &\vdots \\ P_j &= \sum_{i=1}^q \tilde{A}_{i,((j-1)w+1):((j-1)w+\delta)} \tilde{A}_{i,((j-1)w+\delta+1):((j-1)w+2\delta)}, \end{aligned} \quad (1)$$

for all  $j \leq w^{-1}(n - 2\delta) + 1$ , ( $j = 1, \dots, \lceil w^{-1}(n - 2\delta) + 1 \rceil$ ). Note that the periodicity pitch image represents estimated periodicities in the firing patterns of the ANI. Then, to estimate pitch onsets, a sparsity profile is defined as

$$\kappa_j = \text{kurt}(P_j) = \frac{\sum_{i=1}^{\delta} (P_{ij} - \bar{P}_j)^4}{\left(\sum_{i=1}^{\delta} (P_{ij} - \bar{P}_j)^2\right)^2}. \quad (2)$$

The resulting time series peaks at image times of maximum kurtosis, meaning that it captures transitions from tone to tone based on the assumption that periodicities decrease at high velocity in such locations. For this particular case, we constructed ANI using first differences in  $s_t$ , as derivatives are natural tools for pre-whitening, which is commonly used to enhance the estimation of fourth-order moments<sup>74</sup>. Then, to detect the note onsets, we select a number of expected outcomes and find iteratively maximas in  $\kappa_j$  by defining the threshold  $\bar{\kappa}_j * (1.5 - 0.001) * \text{SD}(\kappa_j)$  with a decreasing factor of 0.001 until the expected number of peaks is achieved.

After the pitch onsets were detected, auditory checking and manual correction was performed in SonicVisualizer<sup>75</sup>. Performance metrics were derived: about 92.40%, 71.66%, 76.80%, and 76.875% of correct onsets were achieved for P1–P4, respectively.

Curve registration was conducted by aligning the abscissa variables of P1–P4 group of curves to their respective vector of mean onsets by resampling each snippet to  $(240 * [\text{onset}])\ell^{-1}$  where  $\ell$  is the length for each snippet and  $[\text{onset}]$  is a vector of rounded mean onsets of size the number of expected onsets per passage. All curves were finally sampled to a fine grid of equal time points.

### Pitch expectation index

This index measures the pitch commonality using two running leaky integrated periodicity pitch images, each one having a possible different echo (Fig. 3a). Local images (short echo) are compared with the global images (long echo) at run time. This comparison generates tone expectations based on constant updates according to the spectral properties related to changes in pitch. The formal expression for this index is

$$\tilde{P}_{\lambda_{\text{local}}} \otimes_r \tilde{P}_{\lambda_{\text{global}}} \rightarrow [-1, \dots, 1], \quad (3)$$

where  $\tilde{P}$  is a leaky integrated pitch image,  $\lambda$  stand for the respective echoes in seconds ( $\lambda_{\text{local}} = 0.1$  for local,  $\lambda_{\text{global}} = 1.5$  for global) and  $\otimes_r$  is an operator that calculates the Pearson's correlation coefficient between running columns of both matrices. Echoes durations' are determined according to the short-time memory model formulated in ref. <sup>44</sup>. Leaky integration is exemplified for a null factor of enlargement as detailed in Fig. 3a.

In summary, given a sequence of pitches, the pitch expectation index shows a curve of correlations going up and down, depending on how well the subsequent individual pitches and the previous pitches (in a given time window) match.

### Rhythmical density

Here we define rhythmical density ( $\rho$ ) by the energy of relevant beating frequencies over each ANI row (channels). As shown in Fig. 3b, a set of ( $k$ ) ANI overlapping frames mapped by a Hamming window are transformed to the frequency domain per channel using fast discrete Fourier transform (FFT). We obtain a matrix of size  $q \times p$ , where  $p$  is the length of the spectrum, that is weighted more prominently at channel center frequencies below 800 Hz and attenuated for high spectral frequencies as well, to reproduce psychoacoustic sensation of roughness or dissonance, a critical dimension of timbre<sup>76</sup>. Without loss of generality, the weighed matrix  $Z \in \mathfrak{R}^{q \times p}$  is usually exponentiated at  $\alpha = 1.6$  by the power law. The sums of the weighted bins across each channel produces

a vector  $1 \times q$  whose mean value is the rhythmical density  $\rho_t$  for each ANI frame ( $t=1, \dots, k$ ). Rhythm is therefore represented continuously according to the roughness properties of sound which makes this index particularly sensitive to pitch changes when there is homogeneity in timbre. Note that this index increases and keeps sustained according to pitch appearance and decays regardless of their duration and volume.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### DATA AVAILABILITY

Restrictions apply to the availability of these data, which were used under licence for the current study and so are not publicly available. However, anonymized motion and audio data that support the findings of this study are available from the corresponding author N.M. upon reasonable request for research purposes.

### CODE AVAILABILITY

The routines for the computational modeling of the auditory periphery are freely available as an R-package at a Github repository<sup>74</sup>. ANIs are computed with the function `CalcANI`, periodicity pitch images with `PeriodicityPitch`, pitch expectation profiles with `ContextualityIndex` (Comparison) and rhythmical density with `RoughnessFFT`.

Received: 3 November 2022; Accepted: 14 June 2023;

Published online: 27 June 2023

### REFERENCES

- Leman, M. *Embodied Music Cognition and Mediation Technology* (MIT Press, 2008).
- Leman, M. *The Expressive Moment: How Interaction (with Music) Shapes Human Empowerment* (MIT Press, 2016).
- Chang, A., Kragness, H. E., Livingstone, S. R., Bosnyak, D. J. & Trainor, L. J. Body sway reflects joint emotional expression in music ensemble performance. *Sci. Rep.* **9**, 1–10 (2019).
- Chang, A., Livingstone, S. R., Bosnyak, D. J. & Trainor, L. J. Body sway reflects leadership in joint music performance. *Proc. Natl. Acad. Sci. USA* **114**, E4134–E4141 (2017).
- Tsay, C.-J. Sight over sound in the judgment of music performance. *Proc. Natl. Acad. Sci. USA* **110**, 14580–14585 (2013).
- Davidson, J. W. & Correia, J. S. In *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (Oxford University Press, 2002).
- Godoy, R. I. & Leman, M. (eds.) *Musical Gestures: Sound, Movement, and Meaning* (Routledge, 2009).
- Huberth, M., Davis, S. & Fujioka, T. Expressing melodic grouping discontinuities: Evidence from violinists' rubato and motion. *Mus. Sci.* **24**, 494–514 (2020).
- Nussek, M., Czedik-Eysenberg, I., Spahn, C. & Reuter, C. Associations between ancillary body movements and acoustic parameters of pitch, dynamics and timbre in clarinet playing. *Front. Psychol.* **13** <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.885970> (2022).
- Wanderley, M. M., Vines, B. W., Middleton, N., McKay, C. & Hatch, W. The musical significance of clarinetists' ancillary gestures: an exploration of the field. *J. New Music Res.* **34**, 97–113 (2005).
- Kraus, N. & Slater, J. in *Handbook of Clinical Neurology* (eds. Celesia, G. & Hicock, G.) vol. 129, 207–222 (2015).
- McNeill, D. *Gesture and Thought* (The University of Chicago Press, 2005).
- McNeill, D. *Hand and Mind: What Gestures Reveal about Thought* (The University of Chicago Press, 1992).
- Goldin-Meadow, S. & Alibali, M. W. Gesture's role in speaking, learning, and creating language. *Annu. Rev. Psychol.* **64**, 257–83 (2013).
- Hostetter, A. B. When do gestures communicate? a meta-analysis. *Psychol. Bull.* **137**, 297–315 (2011).
- Ekman, P. & Friesen, W. V. The repertoire of nonverbal behavior: categories, origins, usage, and coding. *Semiotica* **1**, 49–98 (1969).
- Li, P. & Jeong, H. The social brain of language: grounding second language learning in social interaction. *NPJ Sci. Learn.* **5**, 8 (2020).
- Goldin-Meadow, S. Using our hands to change our minds. *Wiley Interdiscip. Rev. Cogn.* **8**, 1–2 (2017).
- Goldin-Meadow, S. et al. Doing gesture promotes learning a mental transformation task better than seeing gesture. *Dev. Sci.* **15** **6**, 876–84 (2012).
- Goldin-Meadow, S., Cook, S. W. & Mitchell, Z. Gesturing gives children new ideas about math. *Psychol. Sci.* **20**, 267–272 (2009).
- Weiss, A. E., Nussek, M. & Spahn, C. Motion types of ancillary gestures in clarinet playing and their influence on the perception of musical performance. *J. New Music Res.* **47**, 129–142 (2018).
- Moura, N. et al. The impact of visual display of human motion on observers' perception of music performance. *PLoS ONE* **18**, e0281755 (2023).
- Massie-Laberge, C., Cossette, I. & Wanderley, M. M. Kinematic analysis of pianists' expressive performances of romantic excerpts: applications for enhanced pedagogical approaches. *Front. Psychol.* **9** <https://doi.org/10.3389/fpsyg.2018.02725> (2019).
- Nussek, M. & Wanderley, M. M. Music and motion—how music-related ancillary body movements contribute to the experience of music. *Mus. Percept.* **26**, 335–353 (2009).
- Thompson, M. R. & Luck, G. Exploring relationships between pianists' body movements, their expressive intentions, and structural elements of the music. *Mus. Sci.* **16**, 19–40 (2012).
- Bejjani, F. & Halpern, N. Postural kinematics of trumpet playing. *J. Biomechanics* **22**, 439–446 (1989).
- Vidal, M. et al. Cholinergic-related pupil activity reflects level of emotionality during motor performance. *Eur. J. Neurosci.* <https://doi.org/10.1111/ejn.15998> (2023).
- Broughton, M. C. & Davidson, J. W. An expressive bodily movement repertoire for marimba performance, revealed through observers' laban effort-shape analyses, and allied musical features: two case studies. *Front. Psychol.* **7** <https://doi.org/10.3389/fpsyg.2016.01211> (2016).
- Davidson, J. W. Bodily movement and facial actions in expressive musical performance by solo and duo instrumentalists: two distinctive case studies. *Psychol. Music* **40**, 595–633 (2012).
- Broughton, M. C. & Davidson, J. W. Action and familiarity effects on self and other expert musicians' laban effort-shape analyses of expressive bodily behaviors in instrumental music performance: a case study approach. *Front. Psychol.* **5** <https://doi.org/10.3389/fpsyg.2014.01201> (2014).
- Teixeira, E., Loureiro, M., Wanderley, M. & Yehia, H. C. Motion analysis of clarinet performers. *J. New Music Res.* **44**, 1–15 (2014).
- Turner, C., Visentin, P., Oye, D., Rathwell, S. & Shan, G. Pursuing artful movement science in music performance: single subject motor analysis with two elite pianists. *Percept. Mot. Skills* **128**, 1252–1274 (2021).
- Goldin-Meadow, S. The role of gesture in communication and thinking. *Trends Cogn. Sci.* **3**, 419–429 (1999).
- Buck, B., MacRitchie, J. & Bailey, N. The interpretive shaping of embodied musical structure in piano performance. *EMR* **8**, 92–119 (2013).
- MacRitchie, J., Buck, B. & Bailey, N. J. Inferring musical structure through bodily gestures. *Mus. Sci.* **17**, 86–108 (2013).
- Demos, A. P., Chaffin, R. & Logan, T. Musicians body sway embodies musical structure and expression: a recurrence-based approach. *Mus. Sci.* **22**, 244–263 (2018).
- Spiech, C., Sioros, G., Endestad, T., Danielsen, A. & Laeng, B. Pupil drift rate indexes groove ratings. *Sci. Rep.* **12**, 11620 (2021).
- Gibbs, Jr., R. W. *Embodiment and Cognitive Science* (Cambridge University Press, 2005).
- Jensenius, A. R., Zelechowska, A. & Sánchez, V. E. G. The musical influence on people's micromotion when standing still in groups. In: *Proceedings of the 14th SMC Conference* (2017).
- Gonzalez-Sanchez, V. E., Zelechowska, A. & Jensenius, A. R. Correspondences between music and involuntary human micromotion during standstill. *Front. Psychol.* **9**, 1382 (2018).
- Ross, J. M., Warlaumont, A. S., Abney, D. H., Rigoli, L. M. & Balasubramaniam, R. Influence of musical groove on postural sway. *J. Exp. Psychol.* **42** **3**, 308–19 (2016).
- Coste, A., Salesse, R. N., Gueugnon, M., Marin, L. & Bardy, B. G. Standing or swaying to the beat: discrete auditory rhythms entrain stance and promote postural coordination stability. *Gait Posture* **59**, 28–34 (2018).
- Phillips-Silver, J. & Trainor, L. J. Vestibular influence on auditory metrical interpretation. *Brain Cogn.* **67**, 94–102 (2008).
- Leman, M. An auditory model of the role of short-term memory in probe-tone ratings. *Mus. Percept.* **17**, 481–509 (2000).
- Janata, P. et al. The cortical topography of tonal structures underlying western music. *Science* **298**, 2167–2170 (2002).
- Collins, T., Tillmann, B., Barrett, F. S., Delbé, C. & Janata, P. A combined model of sensory and cognitive representations underlying tonal expectations in music: from audio signals to behavior. *Psychol. Rev.* **121** **1**, 33–65 (2014).
- Granger, C. W. J. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **37**, 424–438 (1969).

48. Granger, C. W. J. *Testing for Causality: A Personal Viewpoint*, vol. 2 of *Econometric Society Monographs*, 48–70 (Cambridge University Press, 2001).
49. Ramsay, J. & Silverman, B. W. *Functional Data Analysis* (Springer, New York, 2005).
50. Acal, C. et al. Functional ANOVA approaches for detecting changes in air pollution during the Covid-19 pandemic. *Stoch. Environ. Res. Risk. Assess.* **36**, 1083–1101 (2021).
51. Clark, A. Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behav. Brain. Sci.* **36**, 181–204 (2013).
52. Friston, K. J., Mattout, J. & Kilner, J. M. Action understanding and active inference. *Biol. Cybern.* **104**, 137–160 (2011).
53. Brown, H., Friston, K. & Bestmann, S. Active inference, attention, and motor preparation. *Front. Psychol.* **2** <https://doi.org/10.3389/fpsyg.2011.00218> (2011).
54. Friston, K. J. & Kiebel, S. J. Predictive coding under the free-energy principle. *Philos. Trans. R. Soc. Series B* **364**, 1211 – 1221 (2009).
55. Parr, T., Limanowski, J., Rawji, V. & Friston, K. J. The computational neurology of movement under active inference. *Brain* **144**, 1799–1818 (2021).
56. Parr, T., Rees, G. & Friston, K. J. Computational neuropsychology and bayesian inference. *Front. Hum. Neurosci.* **12** <https://doi.org/10.3389/fnhum.2018.00061> (2018).
57. Rosso, M., Maes, P.-J. & Leman, M. Modality-specific attractor dynamics in dyadic entrainment. *Sci. Rep.* **11**, 18355 (2021).
58. van Atteveldt, N., Murray, M. M., Thut, G. & Schroeder, C. E. Multisensory integration: flexible use of general operations. *Neuron* **81**, 1240–1253 (2014).
59. Hsu, W.-L., Scholz, J. P., Schöner, G., Jeka, J. J. & Kiemel, T. Control and estimation of posture during quiet stance depends on multijoint coordination. *J. Neurophysiol.* **97** **4**, 3024–3035 (2007).
60. Yamamoto, A., Sasagawa, S., Oba, N. & Nakazawa, K. Behavioral effect of knee joint motion on body's center of mass during human quiet standing. *Gait Posture* **41** **1**, 291–4 (2015).
61. Günther, M., Putsche, P., Leistriz, L. & Grimmer, S. Phase synchronisation of the three leg joints in quiet human stance. *Gait Posture* **33** **3**, 412–7 (2011).
62. Bronstein, A. M. & Pavlou, M. Balance. In *Neurological Rehabilitation*, vol. 110 *Handbook of Clinical Neurology* (eds. Barnes, M. P. & Good, D. C.) pp. 189–208 (Elsevier, 2013). <https://www.sciencedirect.com/science/article/pii/B9780444529015000162>.
63. Nashner, L. M. Adapting reflexes controlling the human posture. *Exp. Brain Res.* **26**, 59–72 (1976).
64. Tsai, Y.-Y., Chang, G.-C. & Hwang, I.-S. Changes in postural strategy of the lower limb under mechanical knee constraint on an unsteady stance surface. *PLoS ONE* **15**, e0242790 (2020).
65. Fasola, J., Vouga, T., Baud, R., Bleuler, H. & Bourri, M. Balance control strategies during standing in a locked-ankle passive exoskeleton. In: *IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*, 593–598 (IEEE, 2019).
66. Tsai, Y.-Y., Chang, G.-C. & Hwang, I.-S. Adaptation of kinematic synergy and postural control to mechanical ankle constraint on an unsteady stance surface. *Hum. Mov. Sci.* **60**, 10–17 (2018).
67. Cordo, P. & Nashner, L. M. Properties of postural adjustments associated with rapid arm movements. *J. Neurophysiol.* **47** **2**, 287–302 (1982).
68. Davidson, J. W. Visual perception of performance manner in the movements of solo musicians. *Psychol. Music* **21**, 103–113 (1993).
69. Moura, N., Vilas-Boas, J. P. & Serra, S. In *Advances in Design, Music and Arts II*. (eds Raposo, D., Neves, J., Silva, R., Correia Castilho, L. & Dias, R.) 655–673 (Springer International Publishing, Cham, 2023).
70. Ibert, J. *Concertino da Camera Pour Saxophone et Orchestre* (Alphonse Leduc Éditions Musicales, 1935).
71. Benade, A. H. & Lutgen, S. The saxophone spectrum. *JASA* **83**, 1900–1907 (1988).
72. Martens, J.-P. & Immerseel, L. V. An auditory model based on the analysis of envelope patterns. In *JCASSP*, vol. 1, 401–404 (IEEE, 1990).
73. Immerseel, L. V. & Martens, J.-P. Pitch and voiced/unvoiced determination with an auditory model. *JASA* **91**, 3511–3526 (1992).
74. Vidal, M. & Aguilera, A. M. Novel whitening approaches in functional settings. *Stat* **12**, e516 (2023).
75. Cannam, C., Landone, C. & Sandler, M. Sonic visualiser: an open source application for viewing, analysing, and annotating music audio files. In: *Proceedings of the ACM Multimedia 2010 International Conference*, 1467–1468 (2010).
76. Leman, M. Visualization and calculation of the roughness of acoustical musical signals using the Synchronization Index Model (SIM). In: *COST-G6 Conference on Digital Audio Effects (DAFx-00)*, 125–130 (2000).

## ACKNOWLEDGEMENTS

This research was partially supported by the individual grant 2020.05257.BD of the Portuguese Foundation for Science and Technology (also supported by the FSE programme through Programa Operacional Regional Norte), the Methusalem funding from the Flemish Government, the project FQM-307 of the Government of Andalusia (Spain) and by the project PID2020-113961GB-I00 of the Spanish Ministry of Science and Innovation (also supported by the FEDER programme). We also acknowledge the financial support of the Consejería de Conocimiento, Investigación y Universidad, Junta de Andalucía (Spain) and the FEDER programme for the project A-FQM-66-UGR20 and the IMAG-Maria de Maeztu grant CEX2020-001105-M/AEI/10.13039/501100011033.

## AUTHOR CONTRIBUTIONS

N.M., S.S., and J.P.V.B. developed the study concept and design. Testing and data collection was performed by N.M. and M.V. analyzed the data and contributed with methodological tools under the supervision of A.M.A. and M.L.. N.M. and M.V. interpreted the data and drafted the manuscript, and S.S., J.P.V.B., A.M.A., and M.L. provided critical revisions. All authors approved the final version of the manuscript for submission.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41539-023-00172-z>.

**Correspondence** and requests for materials should be addressed to Nádia Moura, Marc Vidal or Marc Leman.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023