Imperial College London

Final report for MSc project

Study of motor development in preterm infants using machine learning tools

María Royo

Supervisors:
Juan Álvaro Gallego
Mostafa Safaie

Submitted in partial fulfilment of the requirements for the award of MSc in Biomedical Engineering from Imperial College London

September 2021

Word count: 5987

Abstract

Preterm babies are at high risk of motor impairment, for which early diagnosis and intervention are crucial. Since the nervous system of infants is in continuous change, further understanding of early motor development will allow to define a new age-specific assessment of motor risk. The present study proposes two computational approaches to better understand early motor development in preterm infants. Both methods are computer vision-based and do not rely on the subjective assessment of experts. The first approach identifies low-dimensional movement manifolds that capture most of the infants' postural variance. The manifolds are spanned by correlated joints movement patterns that reflect motion coordination, and we assess whether these manifolds change as babies grow. The second method aims at understanding how different kinematic features that characterize motor behaviour vary with development, capturing movement variability and complexity. We applied these analyses on longitudinal recordings of preterm babies. Our results indicate that older infants move more and faster than younger ones, and their movements' variability is higher, but no pronounced coordination changes are found with development. However, our results suggest that if development was assessed within a period longer than 12 weeks, we could observe stronger coordination changes. Therefore, we show the potential of the proposed approaches to study development.

Acknowledgements

I am very grateful to my tutors Juan Álvaro Gallego and Mostafa Safaie for their support and collaboration throughout the year. I consider to have learnt a lot from them during the development of this project.

CONTENTS

LIST OF FIGURES

1. INTRODUCTION

Preterm birth, between 28 and 35 weeks of gestation, highly increases the risk of motor impairment. Many preterm children experience a delay in their motor development compared to full-term infants, or develop motor disorders like Cerebral Palsy, which is the most severe form[1]. The reason for this lies in an alteration of brain development. Sensory feedback plays an important role in the development of motor cortex, and thus, the ex-uterus experience of preterm children at an earlier stage can affect the development of their cortical topographical maps[2, 3]. Hence, understanding the development in preterm infants and monitoring them periodically to enable an early diagnosis and intervention of motor disorders will have a high impact in the preterm population[1].

Motor capacity and development of healthy preterm babies are limited compared to full-term infants during the first 18 months of adjusted age[1]. This developmental motor delay is very variable and is known to depend on factors like preterm degree, weight and environmental aspects[4]. However, there is still a lack of understanding about the developmental delay of preterm infants[4].

During the initial months of life, between 6 and 15 months of gestational age[1] approximately, infants exhibit general movements (GMs)[5, 6]. These are gross movements that involve all body parts without a particular sequence of arms, legs, neck and trunk movements, and vary in speed, amplitude and intensity. At birth, GMs shift to more forceful writhing GMs, characterized by being slow. Then, they change into fidgety GMs at around 3 months adjusted age, which consist of fluent, small movements, superimposed by large and fast ones. Finally, GMs disappear at 5 months adjusted age approximately, when isolated limbs movement emerge. Note that these ages are approximate and vary from infant to infant[5, 6].

Currently, assessing infants degree of development and motor risk is mainly done through direct observation of the spontaneous GMs, whose quality is known to reliably reflect the state of the nervous system[6, 7]. Although a definite diagnosis of motor impairment cannot be done until the age of 5 years[8], assessing babies' risk enables an early intervention, proven to be more effective the sooner it is initiated. The quality of GMs is determined by their complexity (changes in movement direction and participating body parts), and variation (set of new patterns an infant is able to generate in time). Thus, motor variability is sign of a healthy nervous system, while stereotyped movements are a sign of motor dysfunction[5–7]. In fact, cramped synchronized GMs –non-smooth, rigid movements, in which all limbs and trunk muscles contract and relax at the same time– are an indicator of motor impairment[9].

However, although current assessment for motor impairment involves direct observation of behaviour, this method is highly subjective. Therefore, we have developed two objective computer-

[1]Age minus how early the infant was

[1]Age from gestation

vision based frameworks to study the sensorimotor development in early infants, independent of an expert knowledge or experience. We tracked different body parts of four preterm babies in videos using DeepLabCut[10, 11], a deep learning software package for pose estimation. The videos were taken during the babies' initial months of life, while executing spontaneous movements. This allowed us to measure behaviour in a quantitative way and investigate how the babies' kinematics vary as they grow. Furthermore, capturing the babies poses with videos does not alter their natural behaviour as position or acceleration sensors do.

The first method focuses on infants' coordination patterns. Humans and animals perform stereotyped movements, we do not move across all the possible trajectories that our body allows, meaning that movement can be described using a reduced set of dimensions in comparison to the body's degrees of freedom (DOF)[12]. Many studies have tried to find the physiological reason behind this outcome. There is a high support to the hypothesis that in order to reduce the complexity and dimensionality of motor control, the central nervous system uses a combination of muscle synergies to construct motor behaviour, which are groups of muscle activations with similar dynamics in space or time[13–17]. Furthermore, several papers have assessed the dimensionality of the body or different body parts' kinematics[13, 14, 18–20]. This has been done through dimensionality reduction analysis that allow to describe movement as a linear combination of behavioural components or modes. Principal component analysis (PCA) finds orthogonal directions of maximum variance in a dataset, the principal components (PCs)[21]. By representing movements as linear combinations of just the PCs or behavioural modes that account for most of the motion variance, the kinematics dimensionality can be reduced to a lower-dimensional manifold within the space spanned by the DOF of the body joints or the movement space. This way, PCA can identify correlated patterns of movement that form a basis of behaviour and reflect body motion coordination.

Here, we also carry out a similar analysis in the behaviour of preterm babies', currently diagnosed with a mild motor impairment, with the aim of determining how their kinematics dimensions and modes of behaviour vary as babies grow. To our knowledge, only another study aimed at finding movement patterns in early infants, but they were only able to track hands and feet, without gathering information about single joint movement[22].

Our second approach to study early infants' development encompasses measuring different kinematic variables that summarize several aspects of their movement: velocity, acceleration, movements' variability and complexity and coordination between body parts. With the objective of understanding how the babies' kinematics vary with development, we built an age classifier based on these features, whose criteria provides insight about the differences between younger and older babies.

These approaches may contribute to the study and assessment of sensorimotor development in the preterm population, enabling an early diagnosis and treatment of developmental delays that are crucial for later motor outcomes.

1.1. Movement manifolds hypothesis

Given the inability of babies to exhibit single limbs movements, we hypothesize that their movement is lower dimensional than their bodies DOF. In particular, we expect movement to be lower dimensional than the set of joint angles we measure, which is already lower dimensional than the body DOF. The fact that as babies develop, they acquire the ability to move limbs independently and their movements become finer, may suggest that the dimensionality of their movements increases with time. This would mean that the trajectories followed in the movement space shift from a low-dimensional constricted space to a broader region encompassing the finer single limb movements. In agreement with this hypothesis, a study about hands movement dimensionality showed that the finest hand movements were built using more PCs than coarser movements, including PCs that accounted for less variance[20]. On the contrary, the subjects that we study may exhibit stereotyped or cramped synchronized GMs as they grow, since they are mildly motor impaired, which could result in a dimensionality reduction along time. Stereotyped movements follow similar paths in the movement space and can be described using a reduced set of behavioural modes. Hence, we would expect the dimensionality to increase in healthy infants but not necessarily in motor impaired ones. With these different hypotheses, the present study aims at determining how the movements dimensionality of preterm babies changes along time. In addition, this study assesses whether there are common coordination patterns for same-age infants and if they change with development.

2. METHODS

2.1. Participants and recordings

The dataset consists of longitudinal videos of four preterm babies, currently diagnosed with a mild motor impairment. The videos were taken while the babies were lying in a hospital crib exhibiting spontaneous movements. The parents of the babies gave consent to use the babies' videos for this project by signing an informed consent agreement.

We have two 7 minutes videos per baby, when the babies were all the same age. The first video was taken when the infants were less than 3 months old birth age, and the other one 12 weeks later. Each baby is assigned an identifier (1-4), and we refer to the videos as 'early', when the infant was younger, and 'late', when it was older.

The sampling rate was 50Hz. The videos were taken using a fixed camera that recorded the babies from above, although not from a completely top projection. The camera was closer to the infants' feet and, therefore, its orientation was not completely perpendicular to the surface where the babies were lying.

We expected to study more babies and more longitudinal recordings per baby but due to the COVID-19 pandemic, no more videos could be taken.

2.2. Pose tracking

In order to track the babies' body parts from the videos we used DeepLabCut, a deep learning software package for animal and human pose estimation[10, 11]. DeepLabCut is based on transfer learning, it allows you to further train an object recognition neural network, ImageNet, which has been trained on a massive dataset[10, 11]. This permits you to track user defined features in videos by relying only on a small set of labelled images, and without using tracking markers that could alter the babies' natural behaviour. However, video tracking has the drawback of capturing only a 2D projection of 3D motion, which results in an important information loss.

Nose, shoulders, elbows, wrists, hips, knees and ankles of the babies were tracked. DeepLabCut provides prediction likelihood estimates for each feature on each data frame. Hence, only data frames with high prediction likelihood for all the tracked features were used for the analysis.

2.3. Movement manifolds

2.3.1. Data preprocessing

The time series of different joint angles and angles created between different body parts were computed for both the left and right sides of the body (see figure 2.1): the elbows, shoulders, knees and hips joint angles, the angles created between nose and shoulders to reflect head movement, and the angles created at the trunk edges to account for trunk motion. The analysis was carried out in angular instead of cartesian coordinates to compensate for different babies' sizes and to minimize distance or projection effects from the camera to the baby. The low projection of the camera with respect to the babies made the legs relative size slightly bigger, and if the analyses were carried upon cartesian coordinates, this would have altered the results.

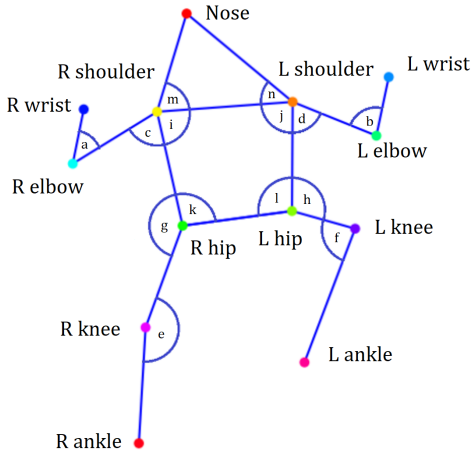


Fig. 2.1. **Tracked pose and computed angles between body parts.** DeepLabCut was used to track nose and left (L) and right (R) shoulders, hips, elbows, wrists, knees and ankles (depicted as coloured labelled points above). Then, the elbows, shoulders, knees and hips joints angles (a-h) were computed, together with the angles created at the upper and lower trunk (i-l), and the angles created between shoulders and nose (m, n), to account for trunk and head movement.

Furthermore, the computed angles time series were also normalized. While some of the computed angles correspond to the true joint range of movement, this does not hold for all. For instance, the angle created between a wrist, elbow and shoulder corresponds to the elbow joint range of motion in the flexion-extension DOF (from 0 to 180 degrees). This is true for arms and legs joints angles, which were therefore normalized between 0 and 180 degrees. On the contrary, the angles measured to reflect head movement do not correspond to true neck lateral bending, or flexion-extension DOF. Similarly, the measured trunk angles do not correspond to the spine range of motion. However, since the measured angles are proportional to the head and trunk motion in their DOF, we normalized these angles with respect to the maximum and minimum values found for them, which map to the maximum and minimum values in the head and trunk true ranges of motion.

2.3.2. Finding movement manifolds

Our dataset is composed of 14-dimensional samples that belong to the 14D movement space, the space spanned by the measured body joints angles. A point in that space corresponds to a baby posture, and its coordinate along each axis corresponds to a joint angle. If a set of baby posture samples are constrained to a low-dimensional region in the whole space, we say that the baby movement belongs to a manifold or low-dimensional subspace within the movement space.

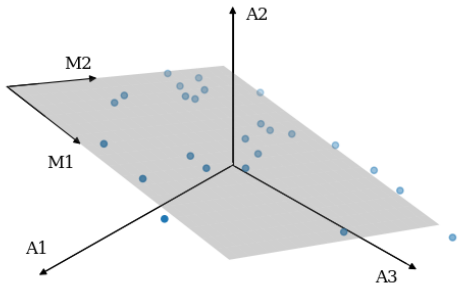


Fig. 2.2. **Movement manifold example.** Group of poses in the movement space spanned by three joint angles (A1, A2, A3). The angular postures are mostly constrained to a low-dimensional manifold, the plane, which is spanned by two behavioural modes M1 and M2.

In order to test whether the infants' behaviour dimensionality is lower than the measured joint angles, for each video, we tried to find the manifold where most of the babies' postures lie. To find the manifolds, we applied principal component analysis (PCA) to the set of angular postures captured in each video. PCA allows to project a high dimensional dataset onto a new coordinate system of orthogonal axes or principal components (PCs) such that the greatest variance of the data lies on each axis, that is, maximizing the amount of covariance[21]. These PCs correspond to the eigenvectors of the data covariance matrix, whose eigenvalues indicate the amount of variance they account for[21]. By representing the data as linear combinations of just the principal components (PCs) that account for most of the variance, the dataset dimensionality can be reduced to a lower-dimensional manifold[21].

In the movement space we refer to these PCs as behavioural or movement modes. We can describe most of the body movement in terms of these correlated joints movement patterns, that represent a low dimensional manifold within the space spanned by the measured joint angles. See an example of 2D manifold in a 3D space in figure 2.2.

2.3.3. PCA cross-projection similarity

We compared the manifolds across babies of same and different ages to test whether there are common coordination patterns for same-age infants, and to assess if they change with development.

The manifolds were compared measuring their cross-projection similarity (CPS). Given two manifolds A and B of equal dimensions, underlying datasets A and B respectively, the CPS

allows to measure the degree of alignment between both manifolds. It measures the variance in a dataset that is explained when projecting that data onto a different dataset manifold. By definition, the PCs in a dataset are their directions of maximum variance[21]. Therefore, the variance accounted for (VAF) when projecting dataset A onto its lower-dimensional manifold A (formed by its initial PCs), represents an upper bound for the VAF when projecting that dataset A onto a different manifold B. The CPS is then defined as the ratio between VAF when projecting a dataset onto the manifold of a different dataset, by the VAF when projecting it to its corresponding manifold. Note that the CPS is not symmetric, CPS(A,B) is not equal to CPS(B,A).

The videos were divided in segments of 20 seconds, and when comparing two videos, all the possible segments combinations for both videos were compared. We checked whether the manifold of each baby changed with time. To do that, we computed the CPS between the early and late videos of the same baby. Then, we compared the CPSs distribution between all same-age babies to the CPSs between different-age babies. Furthermore, CPS was also used to test the manifolds robustness for each baby at a given age, by measuring the orientation similarity of the manifolds underlying segments of the same video. In every case, to assess the significance degree of the obtained similarities, we compared them to the values that could have been obtained by chance. The CPS by chance was computed by cross-projecting the all videos segments against random manifolds.

2.4. Development classification based on kinematic variables

A classifier that could distinguish between younger infants (early) and those same infants 12 weeks older (late) was built based on a set of kinematic features that characterize the babies' movement. Understanding the classifier criteria provides insight about the changes that occur in babies' kinematics as they develop. The computed kinematic variables were then compared among early and late babies.

2.4.1. Data preprocessing

We computed a set of kinematic features that characterize the movement in each video in terms of posture variability and complexity, kinematics, and body parts movement correlation. Firstly, from each joint angle time series in each video (obtained as described in section 2.3), the angular velocity and acceleration time series were estimated. Then, the kinematic features were computed. Most of the chosen features were computed for elbows, shoulders, knees and hips joints: angles standard deviation (measuring posture variability), angles and angular velocity and acceleration entropies (complexity), and angular velocities and accelerations interquartile range (IQR) (kinematic variability) and medians. The medians and IQR were selected to prevent outliers' effects. In addition, joints angles time series correlations were also included. Both intralimb and interlimb (ipsilateral and contralateral) correlations between shoulders, elbows, hips and knees joins movements were estimated, together with head-limb correlations. Similarly

to the movement manifolds, these correlation variables reflect coordination movement patterns.

2.4.2. Linear Discriminant Analysis and feature selection

Linear Discriminant Analysis (LDA) is a supervised method for classification which separates samples in a dataset among two possible classes based on a set of features (see example in figure 2.3). The samples are characterized by n features and hence belong to a n-dimensional space. However, LDA projects the dataset onto a hyperplane, lower dimensional than the original feature dimensionality n, where the classes become maximally separable. The hyperplane and decision boundary are chosen such that the distance between the means of the projected data belonging to each class is maximized, while the within-class variance is minimized.[23]

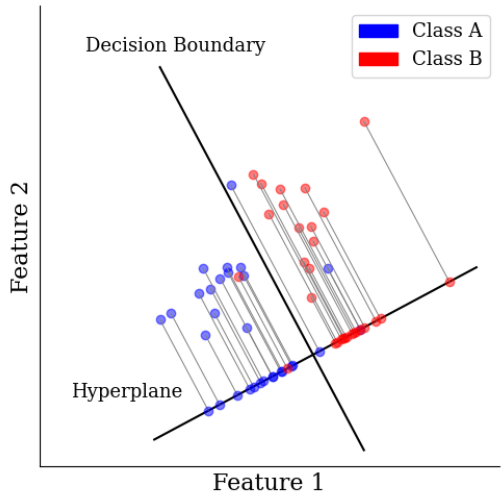


Fig. 2.3. **Linear Discriminant Analysis schematic example.** Samples belonging to two different classes A and B are characterized by features 1 and 2. LDA projects the samples onto a hyperplane (a line), lower dimensional than the original feature space (a plane), where the classes are mostly separated. The hyperplane and decision boundary are chosen such that the distance between classes A and B means is maximized while the within-class variance minimized.

In our case, each of the babies' videos represents a sample, which is characterized by 52 movement kinematic features, and the baby in that video can be classified as early or late. The LDA model projects the 52D dataset onto a 1D line, which represents the optimal linear combination of kinematic features to differentiate early and late infants. The decision boundary corresponds to a point in such line.

Firstly, we evaluated the classifier based on the 52 kinematic features to assess whether there was information about development in them. The performance of the classifier was evaluated through 4-fold cross-validation. For several iterations, three early and three late subjects were chosen at random to train the model, while the other two subjects (one early and one late) were left for testing. The overall classifier accuracy corresponded to the mean testing error obtained

for all the iterations. We also compared the classifier performance to the one that could have been obtained by chance. The accuracy by chance is estimated by shuffling the training labels (early and late) several times and then computing the performance with 4-fold cross-validation.

Then, we used a feature selection algorithm to identify the most relevant features in the classification task, and hence the most different between early and late babies. The algorithm evaluated different models adding features sequentially in a greedy fashion, selecting those that result in best classification accuracy (see [24] for details). It was run for several iterations choosing different numbers of features every time, and we recorded the number of times each feature was selected. The features selected most often were the most different between early and late babies, and therefore, the ones that changed the most with development.

3. RESULTS

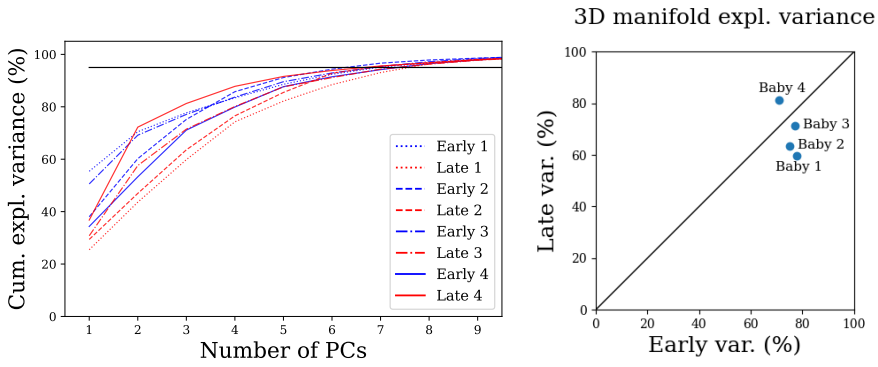
3.1. Pose tracking

A single DeepLabCut network to track all the babies' postures was trained[1], but its performance was worse than a single network per video. Each video was taken under slightly different lightning conditions, with different camera distances to be infants, and the infants looked different (some had darker skin that others). Therefore, a single network had to be trained on more babies in different conditions to further generalize, while a network per video was able to successfully learn the baby poses in uniform conditions. Hence, the data used for the analysis was obtained by training a network per video, whose tracking performance was very high (around 65% of the frames in all videos had a very high prediction likelihood for all tracked body parts). Each network was trained with around 200 labelled images per video.

3.2. Movement manifolds

As shown in figure 3.1.a, the dimensionality of babies behaviour is lower than the measured angles, since the 6 leading behavioural modes in all videos explain around 95% of the total postural variance, while 14 variables were measured. Furthermore, we observed that the early videos of babies 1-3 explain more variance than their late videos, in manifolds where less than 95% of the total variance is explained (see figures 3.1.b and A.1). Since more modes are needed in their late videos to explain the same amount of variance as in their early videos, their behaviour dimensionality seems to slightly increase with development, while the opposite happens in baby 4.

[1]The network was trained together with Michael Lawrence, another student who also works with babies' videos. The videos of Michael Lawrence and María Royo were used to train the network and labelling was shared.



(a) (b)

Fig. 3.1. **Manifolds variance explained.**

a. Cumulative variance explained as a function of the PCs or behavioural modes for each baby. Each line corresponds to a video. For all babies, 6 PCs account for around 95% of the variance.

b. Cumulative variance explained by a 3D manifold in the early vs. late videos for each baby. The early videos of babies 1-3 explain more variance than their late videos. This indicates that movement in the early videos is lower dimensional in the late videos, and hence, their dimensionality tends to slowly increase with development. The opposite is observed in baby 4. See figure A.1 for replication with more dimensions.

We assessed whether the manifold found for each dataset was consistent along the whole video and not a product of undersampling, since a 7-minute video may not be long enough to represent the actual babies' motor behaviour. As described in section 2.3.3, the similarity between manifolds underlying segments of the same video –the within-video similarity– was computed. It was then compared to the similarity obtained when cross-projecting against random manifolds. As the distributions did not conform to normality, we used the two-tailed Wilcoxon rank sum test ($p<0.001$). Since the within-video similarity was higher than by chance, the robustness of the manifold underlying every video could be confirmed. Figure 3.2 shows the pooled within-video CPSs (same baby, same age) and CPSs obtained by chance (random) for all videos, where it can be clearly observed that the highest similarity is achieved between within-video manifolds and the lowest one by chance. Since the movement manifold underlying different video segments is preserved, we may argue that the manifold underlying the whole video represents an estimate of the true movement manifold underlying each baby behaviour.

We also found the manifolds underlying limbs and head movement, excluding the trunk angles. For those manifolds, between 5 and 6 behavioural modes account for around 95% of the variance (see figure A.2) out of the 10 measured joint angles. In comparison to the results shown in figure 3.1.a, this indicates that the trunk explains a very small amount of variance. In this case, the dimensionality of babies 1, 2 and 3 also slightly increases, and decreases in baby 4.

3.3. Manifolds comparison

In order to assess whether there are common coordination patterns underlying equally developed babies, and if they change with development, we compared the manifolds underlying different babies' videos. The resulting CPS was contrasted to the similarity that could have been obtained by chance and to the similarity for a single baby at a particular age (the within-video similarity). As shown in figure 3.2, the similarity obtained by comparing the manifolds to random ones represents a lower bound (grey distribution), while the within-video similarity (green distribution), an upper bound.

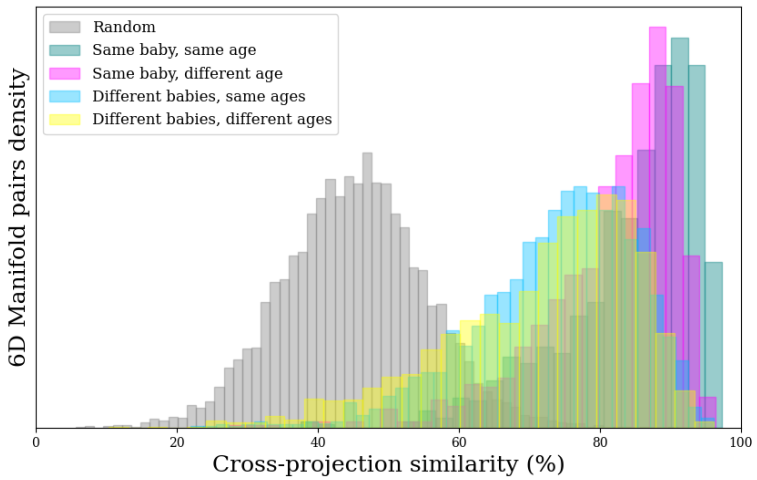


Fig. 3.2. **6D manifolds comparison.** All the videos were divided in segments of 20 seconds. When comparing two videos, all the possible combinations of segments within each video were compared through CPS.
Same baby, same age: Pooled within-video similarities. Manifolds comparisons were performed within the segments of a single video.
Same baby, different age: For each baby, the early and late videos were compared.
Different babies, same ages: All the videos of babies of a given age were compared.
Different babies, different ages: All possible comparisons between early and late videos that did not belong to the same baby were performed.
Random: Distribution resulting of cross-projecting the all videos segments against random manifolds.
All the similarities are higher than random. Therefore, the random CPS distribution represents a lower bound. Although there is overlap among the distributions, the within-video similarity (same baby, same age) is the highest, representing an upper bound; followed by the within-baby similarity (same baby, different age); and with lowest similarity, comparisons among different babies. There is no difference between different babies' comparisons both for same and different ages.

The early and late videos of each of the babies were compared (pink distribution). Their similarity was slightly less than the within-video similarity (rank sum test $p<0.001$), which could

suggest the babies' manifolds vary as they grow, but higher than comparing different babies' manifolds (blue and yellow distributions) (rank sum test $p<0.001$). This indicates that although the manifold underlying each baby's behaviour may vary along time, is more similar at two different time points for the same baby, than for two different babies of same age.

Furthermore, the similarity between manifolds of same-age babies (blue distribution) was compared to the one between manifolds of different-age infants (yellow distribution), but no difference was found. This might happen just because 12 weeks is not an age difference long enough to observe coordination differences between babies of different ages. Another explanation could be that the infants are developing different degrees of motor disorders or developmental delays, and thus, not all of them are exhibiting the same coordination patterns at similar ages. The fact that different dimensionality trends were found among the four babies points to the last reason. Therefore, since the dimensionality of babies 1, 2 and 3 seemed to increase, we performed the same manifolds comparisons just between them, and the similarity between same-age babies became slightly higher than between different-age babies (see supplementary figure A.4). Still, there was a very high overlap among all the similarity distributions, meaning that no strong coordination changes were found with development within a 12-week period.

All the similarities were higher than by chance. The distributions were compared to the random CPS distribution using a two-tailed Wilcoxon rank sum test ($p<0.001$).

Figure 3.2 shows 6D manifolds comparisons, but these results were consistent for different manifolds dimensions (see supplementary figures A.3 and A.4).

3.4. Kinematic features of development

Firstly, we evaluated a classifier that could distinguish between younger (early) and older (late) babies, built upon all the kinematic features, to ensure they contained information about development. The classifier had an accuracy of 94%, much higher than the one obtained by chance with an equivalent model, 50%. Then, the feature selection algorithm was run to identify the features with most classification power and, hence, more different among early and late babies.

As shown in figure 3.3, the features selected more often after running the feature selection algorithm several times are mostly related to velocity and acceleration (median, IQR and entropy), meaning that early and late babies differ in terms of these variables. Some of the body correlation features were selected often as well, mostly intralimb and head-limb movements correlations. Angles entropies were selected less often, suggesting there is not much difference in terms of postural complexity between early and late babies. Elbows and knees joints angles standard deviation (SD) were selected often as well, but not shoulders and hips joint angles SD.

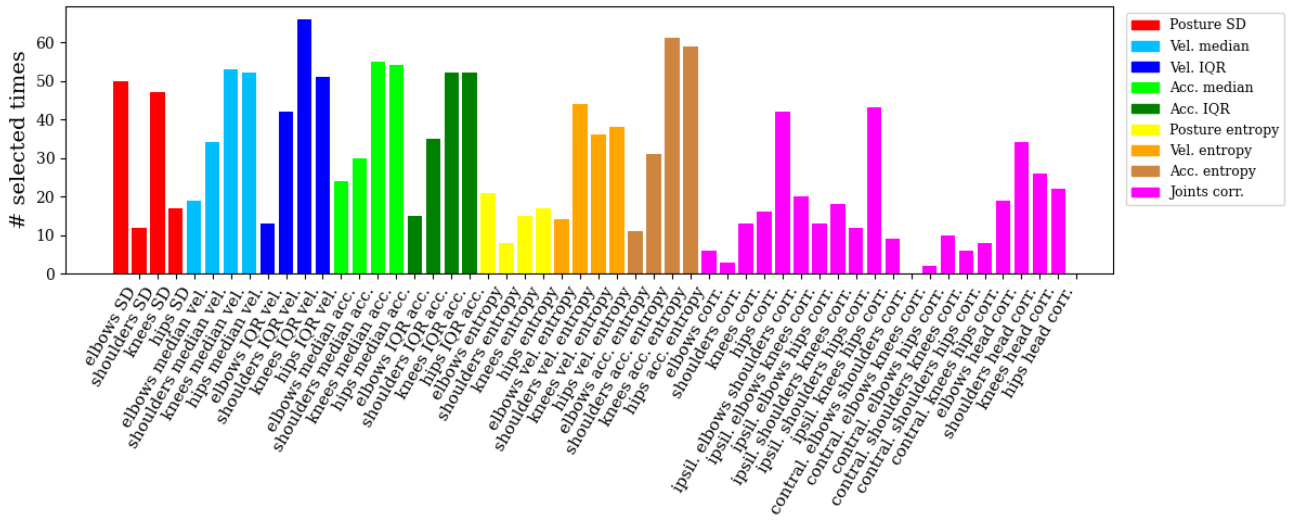
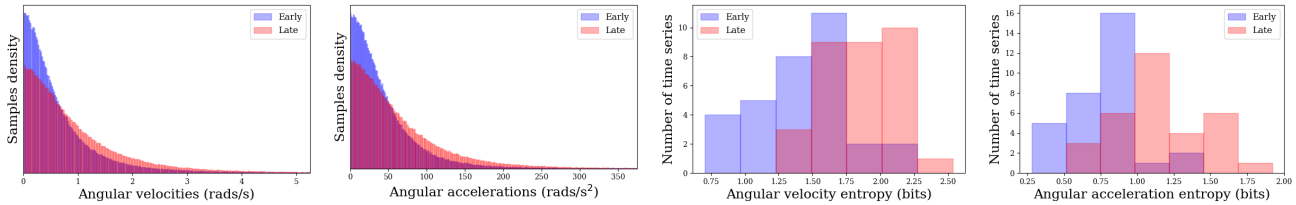


Fig. 3.3. **Kinematic features relevance in age classification.** Number of times each feature was selected during the feature selection process, indicating the importance of such feature in the classification task. Velocities and acceleration-related variables were the ones selected more often. Intralimb and head-limb correlations were selected often but not the rest of correlations. Joints angles entropies were selected the least. While elbows and knees angles SD were highly selected, shoulder and hips angles SD were not selected often.

We assessed potential differences between early and late babies in the computed kinematic features and contrasted them to the insight provided by the classifier. By comparing the pooled angular velocities and accelerations from the elbows, arms, knees and legs joints time series (figures 3.4.a, 3.4.b), it can be seen that older babies move more and faster than younger babies (both distributions were compared using a two-sided Wilcoxon rank sum test $p<0.001$). Older babies velocity and acceleration have also a higher entropy (figures 3.4.c, 3.4.d), meaning that their movements repertoire is higher in terms of velocity and acceleration (rank sum test $p<0.001$). These results agree with the classifier findings since velocities and acceleration-related variables were selected most often. Furthermore, the pooled interlimb, intralimb and head-limbs movement correlations were also compared between early and late babies (see supplementary figures A.5, A.6). Babies 1, 2 and 3 showed higher correlations in their early state than in their late state (rank sum tests $p<0.005$) and the contrary was observed in baby 4 (rank sum test $p<0.001$). Since not all the babies showed the same coordination trend with development, coordination features were not selected that often for classification. Besides, the difference between early and late babies was accentuated in the correlation between head and limbs movement (rank sum test $p<0.001$). However, although elbows and knees joints angles SD were selected often for classification, no significant differences in posture variability or complexity, represented by joints angles SD and entropy, were found (see figure A.7).



(a) (b) (c) (d)

Fig. 3.4. **Joints angular velocities, accelerations and their entropies, compared among early and late infants.**

a,b. Pooled angular velocities and accelerations of all limbs joint angles time series (elbows, shoulders, knees and legs) in all early and late infants. Distributions in late infants are wider. Velocities and accelerations in early infants are lower.

c,d. Pooled entropies of each of the limbs' angular velocities and accelerations time series. Late babies' entropies are shown to be higher.

4. DISCUSSION

The present study has proposed two lines of computational methods to study motor development in babies based on computer vision. Various analyses were carried out with four preterm babies mildly motor disabled, but could be further extended to the full-term, typical or severely affected populations.

The first approach focused on finding movement manifolds that capture most of the babies' postural variance, reflecting movement coordination. This allowed to represent movement as linear combinations of behavioural modes, lower dimensional than the measured joints angles. For all babies of all ages, 6 modes were enough to explain around 95% of the postural variance, out of 14 measured variables. This result is consistent with the fact that babies exhibit GMs in which all body parts participate, instead of single limbs[5, 6].

In three of the four babies (1-3), we observed that with development, their movement dimensionality tends to slowly increase. In accordance with this result, the correlation between the movement of the different body parts in babies 1-3 decreased as they grown, and in baby 4 increased. The higher the correlations, the less behavioural modes are needed to explain behaviour. These findings are also consistent with the fact that as the infants approach an adjusted age of 5 months, their GMs disappear, and isolated limbs movement emerge[5, 6]. The marginal dimensionality decrease of subject 4 could be an indicator of stereotypy and higher motor risk[5–7, 9], but this hypothesis should be contrasted with the future diagnosis of infant 4. The dimensionality increase hypothesis is then held for three out of the four subjects, but the analysis should be run in a larger number of participants and with more longitudinal recordings per participant for further support. In fact, the analysis should be run with typical full-term infants to observe the standard dimensionality trend in healthy development. Furthermore, the dimensionality of behaviour in preterm and full-term infants should be compared, where we would expect the dimensionality of full-term babies to be higher, since most preterm babies suffer a developmental delay[4].

Apart from assessing how movement dimensionality changes with development, we also compared the manifolds across ages to find possible common coordination patterns for same-age babies and assess whether they change with development. Although the manifold underlying each baby behaviour seemed to slightly change as the infants grow, we expected to find stronger coordination changes with development. When comparing all the babies, the similarity between same-age babies' manifolds was not higher than between different-age babies. However, when comparing only babies 1, 2 and 3, which showed the same dimensionality trend, the similarity between same-age babies' manifolds became slightly higher than different-age ones. Although the similarity difference in this case is not totally obvious, these results suggest we might be able to find more pronounced coordination differences if we compared infants with an age difference larger than 12 weeks. For instance, if the analysis compared very early infants with infants older than 6 months, when their GMs are supposed to disappear[5, 6], we might have observed stronger

motor differences.

In the context of muscles synergies theory, since the mapping between the movement kinematics and the muscular activity should be stable, a change in coordination patterns may be explained by a change in muscle synergies[25]. However, finding behavioural manifolds underlying spontaneous movement does not imply the existence of muscle synergies, the fact that body movement mostly occurs in groups of limbs or body parts does not necessarily mean that muscles activate in groups as well[15, 16]. Therefore, this hypothesis should be assessed, for instance, trying to relate muscle synergies found in EMG recordings[15–17, 25] with behavioural manifolds.

Our second analysis aimed at finding differences among early and late babies using a set of kinematic features that characterize their motor behaviour. The kinematic features were shown to contain enough information to reliably classify early and late subjects, and therefore provided development insight. This analysis revealed that late babies move more and faster than early ones, and that they exhibit a larger repertoire of movements in terms of velocity and acceleration. These results are also consistent with the fact that infants' movements are shifting from slow writhing GMs to fidgety GMs, which include both small, fluent movements and large, fast ones[5, 6]. In addition, the analysis showed that the movement of different body parts becomes less correlated with development, in accordance with the fact that individual limbs movement appears as infants grow[5, 6]. However, no difference between early and late babies was found in terms of postural complexity and variation. These features might be useful in risk assessment, but did not show to vary with a 12-week development in these four preterm infants.

The present work has its advantages and limitations. If more subjects had participated with larger longitudinal recordings, together with uniform lightning conditions and a constant position of babies with respect to the camera, our analyses would have had a higher statistical significance, and our models would have generalized better across development groups. Nonetheless, we were able to observe motor changes with development, which show the potential of the proposed approaches to study development. Furthermore, although we only captured a 2D projection of the babies' 3D pose, tracking body parts from videos requires a simple experimental setup, as only a normal camera is needed, and does not disrupt the babies' natural behaviour. In addition, computer-vision based methods for neuro-motor assessment can be easily automated, which increases their viability as future clinical tools.

To summarize, this work contributes to finding a cost-effective and objective assessment method to determine development degree in a quantitative way, independent of a human knowledge or experience. Such method would highly benefit preterm babies, at high risk of suffering motor impairment, since a better understanding of early motor development will enable to define new age-specific assessments of neuromotor risk.

REFERENCES

[1] J. Williams, K. J. Lee, and P. J. Anderson, "Prevalence of motor-skill impairment in preterm children who do not develop cerebral palsy: A systematic review," *Developmental medicine and child neurology*, vol. 52, no. 3, pp. 232–237, 2010. DOI: 10.1111/j.1469-8749.2009.03544.x. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-8749.2009.03544.x.

[2] G. Ball *et al.*, "The effect of preterm birth on thalamic and cortical development," *Cerebral cortex (New York, N.Y. 1991)*, vol. 22, no. 5, pp. 1016–1024, 2012. DOI: 10.1093/cercor/bhr176. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/21772018.

[3] S. Dall'Orso *et al.*, "Somatotopic mapping of the developing sensorimotor cortex in the preterm human brain," *Cerebral cortex (New York, N.Y.: 1991)*, vol. 28, no. 7, pp. 2507–2515, 2018. DOI: 10.1093/cercor/bhy050[doi].

[4] R. D. N. Fuentefria, R. C. Silveira, and R. S. Procianoy, "Motor development of preterm infants assessed by the alberta infant motor scale: Systematic review article," *Jornal de pediatria*, vol. 93, no. 4, pp. 328–342, 2017. DOI: S0021-7557(16)30203-0[pii].

[5] M. Hadders-Algra, "General movements: A window for early identification of children at high risk for developmental disorders," *The Journal of pediatrics*, vol. 145, no. 2 Suppl, s12–s18, 2004. DOI: 10.1016/j.jpeds.2004.05.017.

[6] C. Einspieler, A. F. Bos, M. E. Libertus, and P. B. Marschik, "The general movement assessment helps us to identify preterm infants at risk for cognitive dysfunction," *Frontiers in psychology*, vol. 7, p. 406, 2016. DOI: 10.3389/fpsyg.2016.00406.

[7] H. F. Prechtl, "Qualitative changes of spontaneous movements in fetus and preterm infant are a marker of neurological dysfunction," eng, *Early human development*, vol. 23, no. 3, pp. 151–158, 1990. DOI: 10.1016/0378-3782(90)90011-7.

[8] Can Child, *Developmental coordination disorder assesment and diagnosis*, [Online; accessed 26-August-2021], 2021. [Online]. Available: https://www.canchild.ca/en/diagnoses/developmental-coordination-disorder/assessment-diagnosis.

[9] F. Ferrari *et al.*, "Cramped synchronized general movements in preterm infants as an early marker for cerebral palsy," eng, *Archives of Pediatrics Adolescent Medicine*, vol. 156, no. 5, pp. 460–467, 2002. DOI: 10.1001/archpedi.156.5.460.

[10] A. Mathis *et al.*, "Deeplabcut: Markerless pose estimation of user-defined body parts with deep learning," *Nature neuroscience*, vol. 21, no. 9, pp. 1281–1289, 2018. DOI: 10.1038/s41593-018-0209-y.

[11] T. Nath *et al.*, "Using deeplabcut for 3d markerless pose estimation across species and behaviors," *Nature Protocols*, vol. 14, no. 7, pp. 2152–2176, 2019. DOI: 10.1038/s41596-019-0176-0. [Online]. Available: https://doi.org/10.1038/s41596-019-0176-0.

[12] G. J. Berman, "Measuring behavior across scales," *BMC Biology*, vol. 16, no. 1, p. 23, 2018. DOI: 10.1186/s12915-018-0494-7. [Online]. Available: https://doi.org/10.1186/s12915-018-0494-7.

[13] T. D. Sanger, "Human arm movements described by a low-dimensional superposition of principal components," *The Journal of neuroscience*, vol. 20, no. 3, pp. 1066–1072, 2000. DOI: 10.1523/JNEUROSCI.20-03-01066.2000. [Online]. Available: http://www.jneurosci.org/cgi/content/abstract/20/3/1066.

[14] R. Vinjamuri *et al.*, "Dimensionality reduction in control and coordination of the human hand," *IEEE transactions on biomedical engineering*, vol. 57, no. 2, pp. 284–295, 2010. DOI: 10.1109/TBME.2009.2032532. [Online]. Available: https://ieeexplore.ieee.org/document/5272370.

[15] E. Bizzi and V. Cheung, "The neural origin of muscle synergies," *Frontiers in computational neuroscience*, vol. 7, p. 51, 2013. DOI: 10.3389/fncom.2013.00051. [Online]. Available: https://www.frontiersin.org/article/10.3389/fncom.2013.00051.

[16] P. Saltiel, A. d'Avella, and E. Bizzi, "Combinations of muscle synergies in the construction of a natural motor behavior," *Nature neuroscience*, vol. 6, no. 3, pp. 300–308, 2003. DOI: 10.1038/nn1010. [Online]. Available: http://dx.doi.org/10.1038/nn1010.

[17] M. C. Tresch and A. Jarc, "The case for and against muscle synergies," *Current opinion in neurobiology*, vol. 19, no. 6, pp. 601–607, 2009. DOI: https://doi.org/10.1016/j.conb.2009.09.002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S095943880900124X.

[18] V. Patel and M. Burns, "Linear and nonlinear kinematic synergies in the grasping hand," *Journal of Bioengineering Biomedical Science*, vol. 05, no. 3, 2015. DOI: 10.4172/2155-9538.1000163.

[19] M. Santello, M. Flanders, and J. F. Soechting, "Postural hand synergies for tool use," *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 18, no. 23, pp. 10 105–10 115, 1998. DOI: 10.1523/JNEUROSCI.18-23-10105.1998.

[20] Y. Yan, J. M. Goodman, D. D. Moore, S. A. Solla, and S. J. Bensmaia, "Unexpected complexity of everyday manual behaviors," *Nature Communications*, vol. 11, no. 1, p. 3564, 2020. DOI: 10.1038/s41467-020-17404-0. [Online]. Available: https://doi.org/10.1038/s41467-020-17404-0.

[21] J. Lever, M. Krzywinski, and N. Altman, "Principal component analysis," *Nature Methods*, vol. 14, no. 7, pp. 641–642, 2017. DOI: 10.1038/nmeth.4346. [Online]. Available: https://doi.org/10.1038/nmeth.4346.

[22] M. Kato, M. Hirashima, H. Oohashi, H. Watanabe, and G. Taga, "Decomposition of spontaneous movements of infants as combinations of limb synergies," *Experimental brain research*, vol. 232, no. 9, pp. 2919–2930, 2014. DOI: 10.1007/s00221-014-3972-3[doi].

[23] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification (2nd Edition)*, 2nd ed. Wiley-Interscience, 2000.

[24] F. J. Ferri, P. Pudil, and M. Hatef, "Comparative study of techniques for large-scale feature selection," *Pattern Recognition in Practice, IV: Multiple Paradigms, Comparative Studies and Hybrid Systems*, vol. 16, pp. 403–413, 1994. DOI: 10.1016/B978-0-444-81892-8.50040-7. [Online]. Available: https://doi.org/10.1016/B978-0-444-81892-8.50040-7.

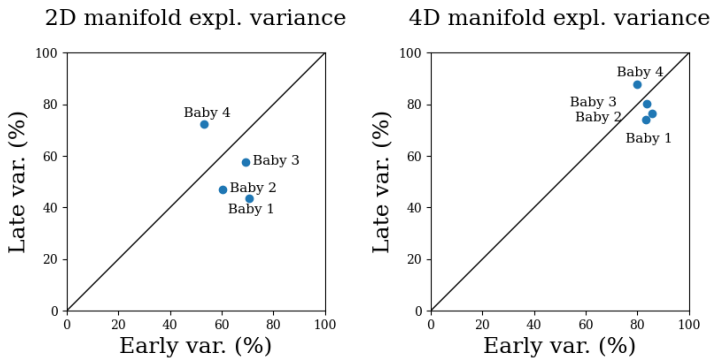
[25] V. C. K. Cheung *et al.*, "Plasticity of muscle synergies through fractionation and merging during development and training of human runners," *Nature Communications*, vol. 11, no. 1, p. 4356, 2020, ID: Cheung2020. DOI: 10.1038/s41467-020-18210-4. [Online]. Available: https://doi.org/10.1038/s41467-020-18210-4.

APPENDIX

Code availability

The analysis was implemented in python and is available in Github at https://github.com/mariaroyo/babies-motor-development.git and https://github.com/BeNeuroLab/neonate-motor.git.

Supplementary figures



(a) (b)

Fig. A.1. Cumulative variance explained by 2D (a) and 4D (b) manifolds in the early vs. late videos of each baby. Replication of figure 3.1.b for 2D and 4D manifolds.
The early videos of babies 1, 2 and 3 explain more variance than their late videos. This indicates that movement is lower dimensional in their early videos, and hence, their dimensionality tends to slowly increase with development. The opposite is observed in baby 4.

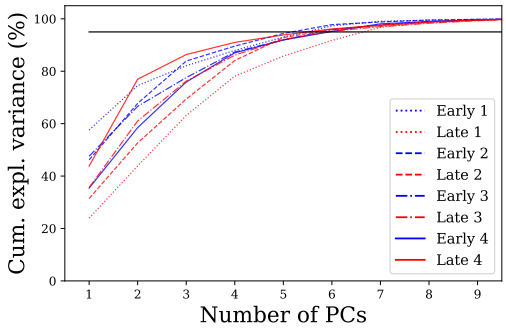
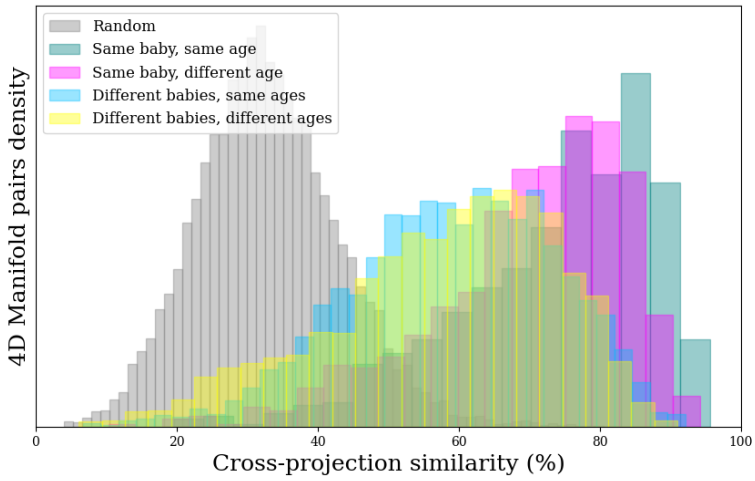
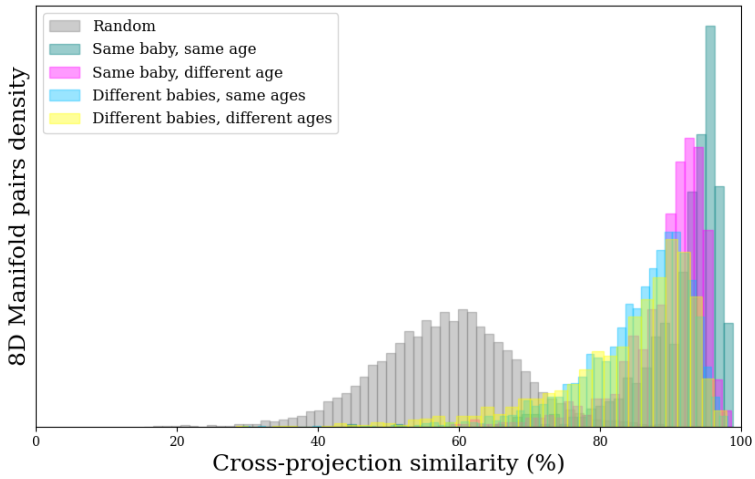


Fig. A.2. Manifolds' variance explained (excluding trunk angles). Cumulative variance explained as a function of behavioural modes for each baby. Each line corresponds a video. For all babies, 5-6 PCs account for more than 95% of the variance.

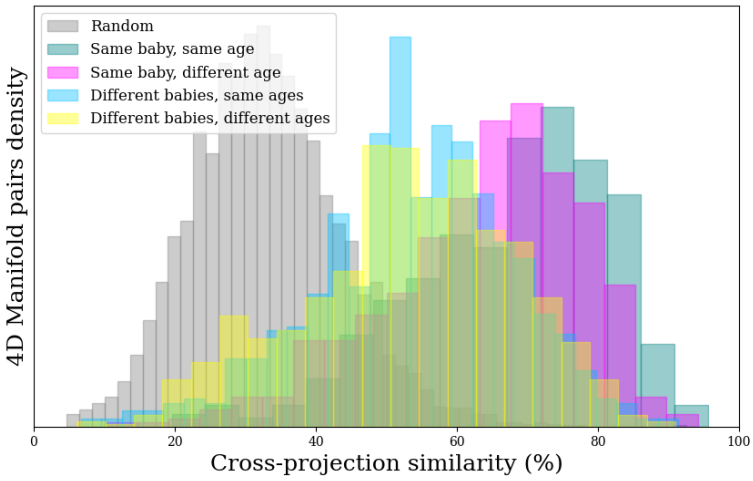


(a)

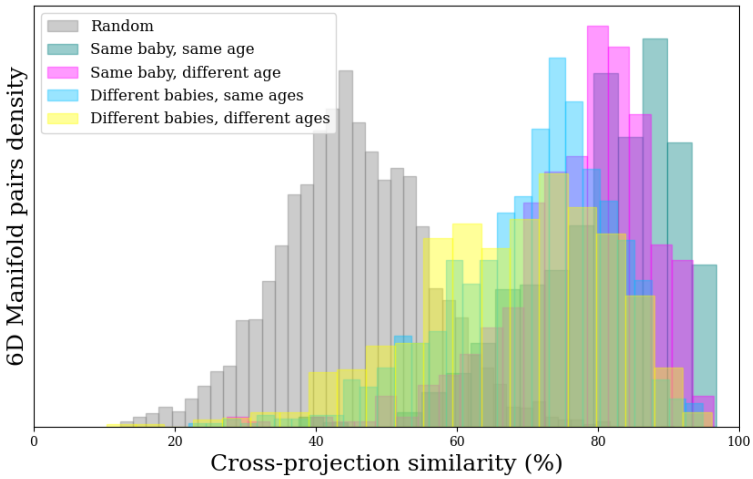


(b)

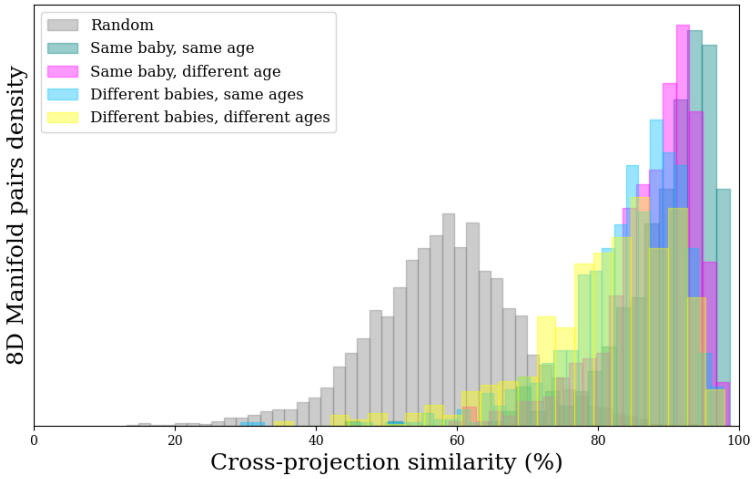
Fig. A.3. **4D, 8D manifolds comparison.** Replication of figure 3.2 for 4D and 8D manifolds. All the videos were divided in segments of 20 seconds. When comparing two videos, all the possible combinations of segments within each video were compared through CPS.
Same baby, same age: Pooled within-video similarities. Manifolds comparisons are performed within the segments of a single video.
Same baby, different age: For each baby, the early and late videos were compared.
Different babies, same ages: All the videos of babies of a given age were compared.
Different babies, different ages: All possible comparisons between early and late videos that did not belong to the same baby were performed.
Random: Distribution resulting of cross-projecting the all videos segments against random manifolds.
All the similarities are higher than random. Therefore, the random CPS distribution represents a lower bound. Although there is overlap among the distributions, the within-video similarity (same baby, same age) is the highest, representing an upper bound; followed by the within-baby similarity (same baby, different age); and with lowest similarity, comparisons among different babies. There is no difference between different babies' comparisons both for same and different ages.



(a)

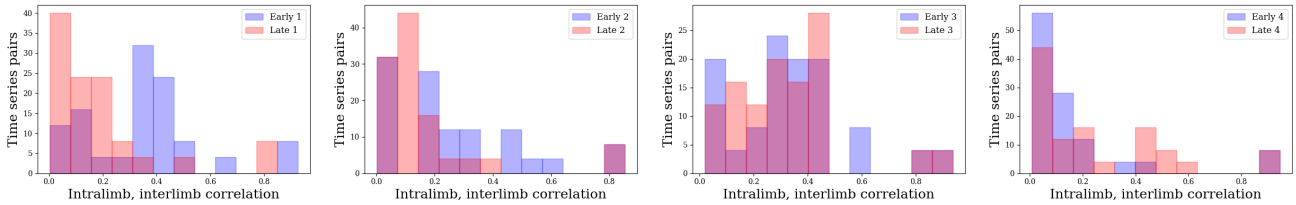


(b)



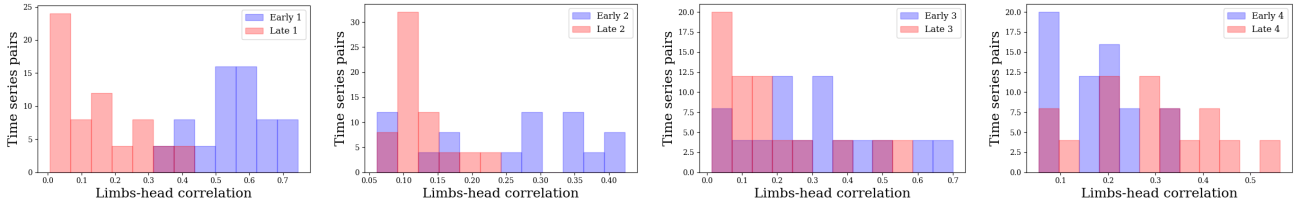
(c)

Fig. A.4. **4D, 6D and 8D manifolds comparison for babies 1, 2 and 3.** Replication of figures 3.2 and A.3 excluding baby 4. In contrast to figures 3.2 and A.3, the similarity between different babies of the same age is slightly higher than between different babies of different ages.



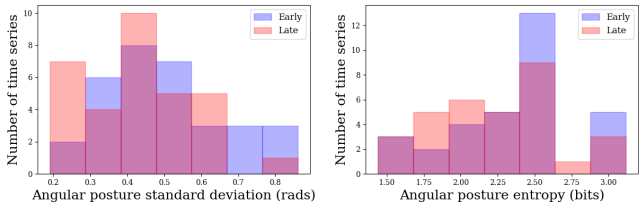
(a) (b) (c) (d)

Fig. A.5. Pooled intralimb and interlimb absolute correlation in babies 1, 2, 3 and 4 at the early and late states. Intralimb correlation includes correlation values between the joint angles time series in the same limb (i.e. left shoulder and left elbow time series correlation); interlimb correlation values consist of the correlation between all possible pair combinations between left and right elbows, shoulders, knees and hips joints angles time series. This includes ipsilateral (i.e. left knee and left elbow) and contralateral (i.e. left hip with right shoulder) comparisons. Babies 1, 2 and 3 correlation values decrease as they develop (higher correlation values in early state than in late state). The opposite is observed for baby 4.



(a) (b) (c) (d)

Fig. A.6. Pooled head-limbs correlations in babies 1, 2, 3 and 4. Head-limb correlations consist of the absolute correlations between each joint angle time series with the angles reflecting head movement time series. Babies 1, 2 and 3 correlation values decrease as they develop (higher correlation values in early state than in late state). The opposite is observed for baby 4.



(a) (b)

Fig. A.7. Pooled entropies and standard deviations of each of the limbs' angular posture time series in all early and late babies. There is no difference among early and late babies.

Note

Aspects of the literature in this thesis are reproduced verbatim from my own writing included in my planning/interim report.