

Optimal Reconstruction of Human Motion From Scarce Multimodal Data

Original

Optimal Reconstruction of Human Motion From Scarce Multimodal Data / Averta, GIUSEPPE BRUNO; Iuculano, Matilde; Salaris, Paolo; Bianchi, Matteo. - In: IEEE TRANSACTIONS ON HUMAN-MACHINE SYSTEMS. - ISSN 2168-2291. - 52:5(2022), pp. 833-842. [10.1109/THMS.2022.3163184]

Availability:

This version is available at: 11583/2978584 since: 2023-05-17T14:32:03Z

Publisher:

IEEE

Published

DOI:10.1109/THMS.2022.3163184

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

IEEE postprint/Author's Accepted Manuscript

©2022 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

Optimal reconstruction of human motion from scarce multi-modal data

Giuseppe Averta, *Member, IEEE*, Matilde Iuculano, Paolo Salaris, *Member, IEEE*, and Matteo Bianchi, *Member, IEEE*

Abstract—Wearable sensing has emerged as a promising solution for enabling unobtrusive and ergonomic measurements of the human motion. However, the reconstruction performance of these devices strongly depends on the quality and the number of sensors, which are typically limited by wearability and economic constraints. A promising approach to minimize the number of sensors is to exploit dimensionality reduction approaches that fuse *prior* information with insufficient sensing signals, through Minimum Variance Estimation. These methods were successfully used for static hand pose reconstruction, but their translation to motion reconstruction has not been attempted yet. In this work, we propose the usage of functional PCA to decompose multi-modal, time-varying motion profiles in terms of linear combinations of basis functions. Functional decomposition enables the estimation of the *a priori* covariance matrix, and hence the fusion of scarce and noisy measured data with *a priori* information. We also consider the problem of identifying which elemental variables to measure as the most informative for a given class of tasks. We applied our method to two different datasets of upper limb motion D1 (joint trajectories) and D2 (joint trajectories + EMG data) considering an optimal set of measures (four joints for D1 out of seven, three joints and eight EMGs for D2 out of seven and twelve respectively). We found that our approach enables the reconstruction of upper limb motion with a median error of 0.013 ± 0.006 rad for D1 (relative median error 0.9%), and 0.038 ± 0.023 rad and 0.003 ± 0.002 mV for D2 (relative median error 2.9% and 5.1% respectively).

I. INTRODUCTION

A correct sensing of human bio-mechanics, and more specifically human body motion, is of paramount importance in many trans-disciplinary fields, which include rehabilitation and assistive applications [1], and human-robot interaction [2]. With the aim of providing unobtrusive and comfortable sensing solutions that can be suitably employed in everyday-life as well as in unstructured environments (like human-robot workspace), the research and technological effort have been focused on the development of wearable systems for evaluating human ergonomics [3]. These systems usually gather kinematic information (i.e. joint angle information) [4] and/or record human muscular activity, e.g. through surface electromyography (sEMG) electrodes [5]. This multi-modal information can be also used to devise informed models of the human neuromusculo-skeletal system [6], e.g. for a correct biomechanical risk assessment.

Giuseppe Averta is with the Department of Control and Computer Engineering of the Politecnico di Torino, Italy, e-mail: giuseppe.averta@polito.it.

Matilde Iuculano, Paolo Salaris and Matteo Bianchi are with the Research Center E. Piaggio and the Department of Information Engineering of the University of Pisa, Italy.

This project has received funding from the EU Horizon 2020 research and innovation programme under Grants No. 871237 (Sophia) and No. 101017274 (Darko), by the European Research Council (ERC) under Grant No. 810346 (Natural BionicS), and by the Italian Ministry of Education and Research (MIUR) in the framework of the CrossLab project (Departments of Excellence).

However, correctly retrieving human body motion is a challenging task because of its complex architecture. Indeed, while it is ideally feasible under a technological point of view, there are several implementation issues that are difficult to overcome, such as the difficulty of tracking, with common sensors, the movement of certain Degrees of Freedom (DoFs) such as the rotation and the elevation of the scapula, and the uncertainties due to relative motion between bones and sensors, see e.g. [7], [8].

A possible solution to deal with such abundancy - and hence to reduce the dimensionality of the problem - can come from neuroscience. Indeed, there is a vast neuroscientific literature that has highlighted the dependencies between human body DoFs [9]. These dependencies have been studied and formalized within the general concept of motor *synergies* [10], which, in a broad sense, provides a theoretical framework to analyze motion coordination in terms of goal-oriented functional couplings of elemental variables (e.g. joints, muscles) [10], [11]. This framework has found fertile ground in the study of human hands, in grasping, manipulation and haptic exploration tasks (see, among the others, [12]), producing interesting transdisciplinary outcomes in robotics [13]. Indeed, the dimensionality reduction that is intrinsic to the concept of synergies has been successfully exploited for the design [14], the control [15] and the planning [16] of artificial robotic systems [13].

Under a pure *observability* point of view, i.e. for the simplification of the problem of retrieving information on human body posture and motion, the synergy-inspired dimensionality reduction has also produced interesting results [17]. In [18], the authors presented CODE, a COordination-based action DEscriptor, which considers minimum and maximum joint velocities and the correlations between the most informative joints (i.e. the joints that are mostly involved in the execution of a certain action), for whole-body action classification. In [19], a compact 6D view-invariant skeletal feature (skeletal quad) based on a local skeleton descriptor for encoding the relative position of joint quadruples was proposed for action recognition from single depth images.

In the field of computer vision and computer graphics, data-driven approaches have been also proposed, e.g. to address the problem of *motion infilling* for 3D human motion data. In [20] the authors presented a convolutional autoencoder that, given a start and end sequence, can complete the missing gaps in between. In [21], a recurrent neural network based approach was presented, for the prediction of future motion when a seed sequence is given. Hierarchical Non Linear Principal Component Analysis NLPCA neural networks were successfully employed to reduce the dimension of the configuration space, for the representation of human/humanoid

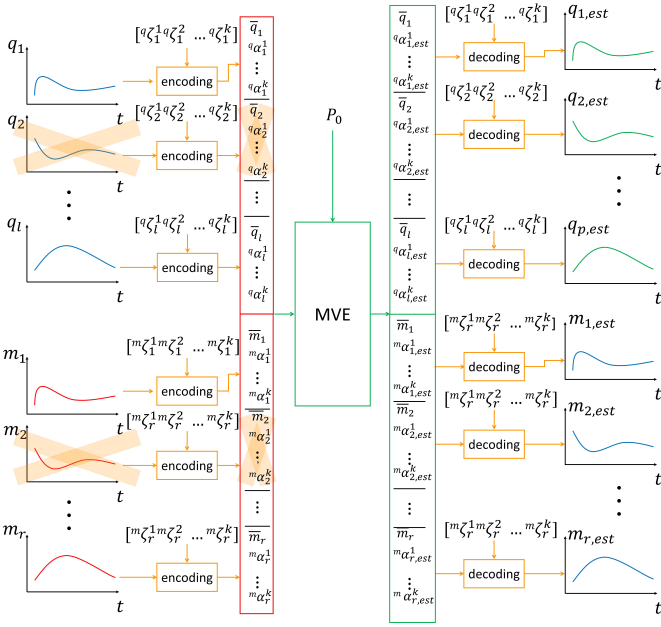


Fig. 1. Schematics of the estimation method proposed in this paper. Temporal measures of a limited number of joints and muscles are mapped (encoding) on the extended state space of weights and average trajectories/muscles envelopes, via a basis of functional Principal Components (fPCs) extracted in advance on an *a priori* dataset. Then, the missing portion of the state is estimated via MVE. Finally, the whole muscle-skeletal system temporal evolution is reconstructed by properly combining the fPCs with the estimated extended state (decoding).

motion patterns [22].

Low-dimensional feature extraction methods such as principal component analysis (PCA) and Locally Smooth Manifold Learning (LSML) were also proposed to extract linear and non linear manifolds of human motion data, e.g. for hand avatar animation [23] or for humanoid robot interface design [24].

Although promising, all the aforementioned methods lack of a theoretical understanding of the statistical properties of human muscular-skeletal data, which prevents their use to solve the twofold problem of the estimation of the temporal evolution of multi-modal data, in spite of scarce and noisy sensory information, and the definition of optimal design strategies for engineering wearable sensing devices. Furthermore, data-driven and learning approaches are usually strongly dependent on the size and quality of the training data, and have important requirements in terms of computational power, which are hardly achieved with on-board electronics used for data processing in wearable applications (whose memory resources are usually in the order of some megabytes or lower).

In [25], we proposed a Minimum Variance Estimation (MVE) technique to fuse the *a priori* information, organized in terms of inter-joint covariation patterns of the most frequent human grasping postures [12], with a reduced number of noisy measures, to reconstruct the most probable joint angles in a Bayesian sense. In [26] we pushed further this investigation to identify the optimal hand joints that should be measured to maximize the reconstruction accuracy (i.e. minimizing in average the reconstruction error), as the result of an optimization problem based on the minimization of the *a posteriori* covariance matrix. These outcomes led to the development and the design of wearable sensing gloves for hand pose reconstruction, using a reduced number of sensing elements (affected by noise) [27].

However, the methods presented in [25], [26] relied on *a priori* information of *static* postures, and then they cannot be applied in a straightforward manner for estimating temporal trajectories. Furthermore, they cannot be applied to the simultaneous reconstruction of multi-modal motion-related quantities (e.g. joint angles and EMG signals), due to the intrinsic difficulties of providing a reliable estimation of the covariance matrix from heterogeneous data [28]. Finally, an application of these techniques to human upper limb data has not been performed yet.

In this paper, we propose a method to generalize the estimation based on minimum variance to the whole upper limb temporal evolution of joint and EMG measurements, capitalizing on two pillars: (1) the existence of covariation patterns in human upper limb motions, which we demonstrated in [29]; (2) the application of functional analysis for enabling the reconstruction of the whole trajectory over time and the estimation of the covariance matrix from multi-modal data (see Fig. 1). In brief, our solution relies on three phases: *encoding phase*, *estimation phase* and *decoding phase*. The idea is to use functional Principal Component Analysis (fPCA) to identify a basis of functions, whose combination can approximate any generic joint and EMG signal temporal evolution [30]. The coefficients of this decomposition (*encoding phase*) or weights, which are extracted from an *a priori* dataset, can be organized together with the average trajectories to form an extended state space and represented by a static vector. As soon as a novel motion is performed, it can be encoded, through the basis of functional Principal Components, in terms of the extended state. We then apply MVE to estimate all the components of the state corresponding to all the joints and EMG recorded signals of the kinematic chain, even when some of them are not directly determined by the measurements or are corrupted by noise (*estimation phase*). Finally, a *decoding procedure* is applied to reconstruct the whole motion of the upper limb joints, by properly combining the functional Principal Components with the estimated extended state.

We validated our framework with two different datasets which are freely available online [31]. The first is a dataset of 30 gestures performed by a cohort of 33 healthy subjects and collected at the University of Pisa (hereinafter referred to as the dataset D1). Data consisted of the 3D positions of the markers fastened on the upper-limb links over time (see [31] for details). From the Cartesian position of the optical markers, we extracted joint angular values by solving a mapping problem on a 7 DoFs kinematic model (three for the shoulder, two for the elbow and two for the wrist, see [29]). To prove that our work can be generalized to multi-modal data, we also included a second dataset, which was recorded at the Hannover Medical School (MHH) (hereinafter referred to as the dataset D2) and also available in [31]. Also in this case, movements were recorded with a motion capture system, but also complemented with surface electromyographic recordings (sEMG) from 12 muscles. The motion capture recordings and hence the 3D optical marker positions allow the definition of a 7 DoFs kinematic model, which is however different from the one that characterizes the first dataset since it enable to retrieve two additional DoFs for the shoulder and no DoFs for the wrist. In both cases, we splitted the dataset (where the same upper limb actions were considered) in two parts, one (containing the 70% of all the entries) was used as *a priori* information, and the remaining

TABLE I
LIST OF NOTATION

Sym	Quantity	Sym	Quantity
ξ_j^s	Single fPC	P_0	A-priori covariance matrix
l	No. of DoFs	r	No. of muscles
N	No. of dataset elements	$q(t)$	Joint angular values
$m(t)$	Muscles activation envelopes	t	Normalized time
T	No. of time frames	s_{max}	Max no. of fPCs
α_i	Weight of fPCS	θ_j	weights coefficients
Q_j	Motions of joint j	Θ_j	Matrix of weights
Φ	Matrix (m.) of fPCs	k	No. of fPCs used
\bar{q}_j	Average joint trajectory	y	Sensor measures
d	No. of sensors	R	Noise covariance matrix
v	Measures noise	H	Joints-measures mapping
P_p	A posteriori covariance m.	p	p-value
x_e	Extended state space		

part for testing. We quantified the norm of the error between the real and the estimated signal on the testing data, without considering a number of independent joints/EMG signals and adding different magnitudes of noise to the measured ones. Our results demonstrate the capabilities of accurately reconstructing the trajectories of the whole kinematic chain, with respect to two alternative methods that we used for comparison, one based on the usage of the pseudoinverse (as done in [25]) and a second one which recursively estimates the current joint static posture through an iterative Minimum Variance Estimation based on a static covariance matrix [25] (the latter cannot be applied in a direct manner to multi-modal data).

Motivated by these outcomes, we designed an optimization problem to identify the optimal selection of independent joints/EMG that one should measure to maximize the estimation accuracy, by leveraging on a metric of uncertainty estimation based on the *a-posteriori* covariance matrix as in [26]. We believe that our findings can open interesting perspectives for the design of an optimized under-sensorized wearable system for multi-modal upper limb tracking, and potentially for the whole body, which could be successfully exploited for every-day life human body monitoring in different application fields, e.g. rehabilitation, assistance and advanced human-robot interaction and cobotics.

II. METHODS AND THEORETICAL FOUNDATIONS

A. Motion decomposition via fPCA: encoding phase

As anticipated before, our reconstruction of human multi-modal motion data leverages on the decomposition of signal provided by functional Principal Component Analysis (fPCA), a statistical method to identify functional primitives from time-varying data. In the following, we will briefly introduce the theory behind fPCA, while we refer the reader to [32] for more details.

Let us consider, without any loss of generality, a l -DoFs kinematic model moved by r muscles, N independent observations of joints temporal evolution $q_1(t), \dots, q_N(t)$ and (eventually) of muscles activation envelopes $m_1(t), \dots, m_N(t)$, where $q(t) : \mathbb{R} \rightarrow \mathbb{R}^l$, $m(t) : \mathbb{R} \rightarrow \mathbb{R}^r$. Let us also assume that the time-series are defined over a normalized time $t \in [0, 1]$ sampled in T evenly spaced points. The latter can be achieved for example through the solution of a Dynamic Time Warping problem, i.e. a pre-processing technique that, given a dataset of time series of arbitrary length, produces a time-normalization to generate entries with the same number of time-points. In

our implementation, given two time series, v_1 and v_2 , the affinity between the two signals is maximized by the following optimization problem:

$$(S, T) = \underset{S > 0, T}{\operatorname{argmax}} \quad \rho(v_1(t), v_2(St - T)) \quad (1)$$

where the operator ρ is the cross correlation between two vectors. Note that this implementation also imposes time-monotonicity, to preserve the coherence in time of the samples, and linear distortion of time with a constant scaling factor to the time domain. All the elements of the dataset are time-warped w.r.t. a common reference signal identified as the dataset entry of average duration. Of note, the optimal parameters S^*, T^* for each dataset entry are shared across all the DoFs to preserve the coordination of movements.

Then, a generic signal (representing either a joint temporal profile or a muscle activation envelope) can be decomposed as a weighted sum of basis elements $\xi_j^s(t)$, aka functional Principal Components (fPCs):

$$\begin{aligned} q(t) &\simeq \bar{q}(t) + \sum_{i=1}^{s_{max}} [q \alpha^i \circ^q \xi^i(t)] \\ m(t) &\simeq \bar{m}(t) + \sum_{i=1}^{s_{max}} [m \alpha^i \circ^m \xi^i(t)] \end{aligned} \quad (2)$$

where $q/m \alpha^i \in \mathbb{R}^{l/r}$ is a vector of weights, $q/m \xi^i(t) \in \mathbb{R}^{l/r}$ is the i^{th} basis element or fPC and s_{max} is the number of basis elements considered. The operator \circ is the element-wise product (Hadamard product), $\bar{q} \in \mathbb{R}^l$ and $\bar{m} \in \mathbb{R}^r$ is the average of $q(t)$ and $m(t)$ respectively. The output of fPCA, which is calculated independently for each joint or muscle j , is a basis of functions $\{\xi_j^1, \dots, \xi_j^{s_{max}}\}$ that maximizes the explained variance of upper limb motions collected in the dataset. Therefore, the first fPC $\xi_j^1(t)$ is the function that solves the following problem:

$$\begin{aligned} \max_{\xi_j^1} \quad & \sum_{i=1}^N \left(\int \xi_j^1(t) q_j^i(t) dt \right)^2 \\ \text{s. t.} \quad & \|\xi_j^1(t)\|_2^2 = \int_0^1 [\xi_j^1(t)]^2 dt = 1. \end{aligned} \quad (3)$$

Subsequent fPCs $\xi_j^s(t)$ are the functions that solve the same problem of Eq. 3, with an additional constraint of orthogonality, i.e.:

$$\int_0^1 \xi_j^s(t) \xi_j^i(t) dt = 0, \quad \forall i \in \{1, \dots, k-1\}. \quad (4)$$

Note that Eqs. 3 and 4 refer to joint angular values, but the extension to muscles is straightforward and here omitted to improve readability. For practical implementations of fPCA that bypass the solution of the optimization problem the interested reader can refer to [32].

Given the optimal selection of basis elements, the movements of the upper limb are now described by an average configuration plus a reduced number of weighting coefficients, i.e. a generic trajectory $q_j^i(t)$ is now represented by the average \bar{q}_j^i and by the vector ${}^q \theta_j^i = [{}^q \alpha_j^1, \dots, {}^q \alpha_j^k] \in \mathbb{R}^k$, where k is the cardinality of the fPCs basis and $(\bullet)'$ is the transpose operation (an analogous nomenclature holds for the characterization of muscles activation envelopes). Note that k can be intended as a free integer parameter, arbitrarily selected in the interval

$[1, s_{\max}]$. It is important to note the significant reduction of dimensionality, given that in practical cases k (the number of fPCs, in our case equal to 5, see Sec. III-B) is usually at least two order of magnitude smaller than T (the number of time frames, in our case 100 samples for each second of recordings, i.e. 9.7s for the signal used for the Dynamic Time Warping procedure and ≈ 4.7 s for the shortest signal in our dataset, see Sec. III-A). More importantly, this change of domain comes with the additional benefit that a movement can be now described by a single static vector of coefficients or weights. Clearly, higher values of k enable a more precise approximation of the real motion, yet augmenting the dimensionality of the state in the weights domain. For each joint, the relationship in matrix form between the joint and the weights domains is then:

$$Q_j(t) = \bar{q}_j + {}^q\Phi_j^q \Theta_j \quad (5)$$

where

$$Q_j = \begin{bmatrix} q_j^1(t) & \dots & q_j^N(t) \\ \downarrow & \dots & \downarrow \end{bmatrix} \in \mathbb{R}^{T,N}, \quad (6)$$

$${}^q\Phi_j = \begin{bmatrix} {}^q\xi_j^1(t) & \dots & {}^q\xi_j^k(t) \\ \downarrow & \dots & \downarrow \end{bmatrix} \in \mathbb{R}^{T,k}$$

and ${}^q\Theta_j \in \mathbb{R}^{k,N}$ is composed by columns collecting the weights θ_j^i associated to the dataset elements. An equivalent matrix form exists for muscles activation envelopes:

$$M_j(t) = \bar{m}_j + {}^m\Phi_j^m \Theta_j. \quad (7)$$

B. Estimating non-measured motion-related quantities via MVE (estimation and decoding phase)

As briefly introduced in the previous sections, in [25] we proposed a method to estimate the posture of human hands leveraging on a set of observations previously collected and used as *a priori* information. This information was organized and described in terms of joint mean μ_0 and covariance matrix P_0 . When some of the joints are not directly measured or the measurements are corrupted by noise, we can fuse such insufficient sensory information with the prior data to estimate the whole posture through Minimum Variance Estimation MVE. In the following, we will provide only a basic description of how to use MVE to improve the posture estimation, while we invite the reader to refer to [25] for more theoretical details. Let us consider a set of measures $y \in \mathbb{R}^d$ provided by a selection of d sensors, and let us also assume a linear relationship between the joint variables $q \in \mathbb{R}^l$ and the measures y :

$$y = Hq + v, \quad (8)$$

where $H \in \mathbb{R}^{d,l}$ is a full row rank measurement matrix (hereinafter we consider, without loss of generality, H as a selection matrix or *discrete measurement matrix* [26]) and v is the measurement noise. The goal is to estimate q given y when $d < l$. Because the number of unknown is higher than the number of measures, the problem admits an infinite number of solutions:

$$\hat{q} = H^\dagger y + N_H \psi, \quad (9)$$

where H^\dagger stands for the pseudo-inverse of H , N_H is the null space of H and $\psi \in \mathbb{R}^{l-d}$ a free vector of parameters. The least-squared solution is:

$$\hat{q} = H^\dagger y. \quad (10)$$

Yet, the output of Eq. 10, although it represents the vector of minimum Euclidean norm, can be very far from the real one. However, from neuroscientific studies it is known that human movements are characterized by stable and coherent covariation patterns between joints [12]. This is indeed a source of information that can be used to improve the estimation in Eq. 9. Given a (large enough - [33]) number of realizations of q^i collected in a matrix of *a priori* $X \in \mathbb{R}^{l,N}$, we can compute the covariance matrix

$$P_0 = \frac{(X - \bar{q})(X - \bar{q})'}{N-1}, \quad (11)$$

where \bar{q} is a matrix whose columns contain the average μ_0 of X . Given P_0 , the best estimate \hat{q} of q is the vector that solves the following optimization problem:

$$\hat{q} = \operatorname{argmin}_q \frac{1}{2} (q - \mu_0)' P_0^{-1} (q - \mu_0) \quad \text{s.t.} \quad y = Hq. \quad (12)$$

Under the hypothesis of v with zero mean and Gaussian distribution with covariance matrix R , the solution of Eq. 12 can be found in closed form and is:

$$\hat{q} = (P_0^{-1} + H'R^{-1}H)^{-1} (H'R^{-1}y + P_0^{-1}\mu_0) \quad (13)$$

and the *a posteriori* covariance matrix (which brings information about the uncertainty of the corresponding state estimation) is

$$P_P = (H'R^{-1}H + P_0^{-1})^{-1}. \quad (14)$$

This work generalizes the MVE-based posture estimation method to heterogenous data and motion profiles by applying Eq. 13 to the extended state of the weights ${}^q\Theta$ for joint angular values, eventually complemented with ${}^m\Theta$ for the estimation of muscles activation. Indeed, considering the kinematic domain only, we can leverage on the relationship introduced in Eq. 5 to build an extended state vector, where a general motion trajectory for the whole upper limb can be represented as in the following:

$$x_e = [\bar{q}_1^q \alpha_1^1 \dots {}^q\alpha_1^k | \bar{q}_2^q \alpha_2^1 \dots {}^q\alpha_2^k | \dots | \bar{q}_p^q \alpha_p^1 \dots {}^q\alpha_p^k]', \quad (15)$$

where the subscript $(\bullet)_j$ stands for the j -th joint, \bar{q}_j is the mean value of the j -th joint and ${}^q\alpha_j^k$ are the coefficients in weight space identified in Sec. II-A. Obviously, the extension to the muscle-skeletal case is straightforward and can be obtained by expanding x_e with the terms associated to muscles activation envelopes (Eq. 7). x_e can be considered as an extended space, represented by a vector $\in \mathbb{R}^{(k+1)*l}$ for the kinematics-only case and $\in \mathbb{R}^{(k+1)*(l+r)}$ for the muscle-skeletal case (i.e. joint and muscle activation signals). If some of the ${}^q\alpha_j^k$ -related components cannot be directly computed (because the related measurements are missing) or are affected by noise, we can perform an estimation of the whole extended state by applying the static formalization of [25]. The estimated coefficients can then be used to reconstruct the motion of all the elements of the kinematic chain according to Eq. 5.

C. Optimal sensors placement

As anticipated in the previous section, MVE relies on a dataset of complete measures available *a priori* to provide, given partial measurements of the current state, a complete estimation of the whole muscle-skeletal configuration. Such an estimation is affected by a certain degree of uncertainty, which

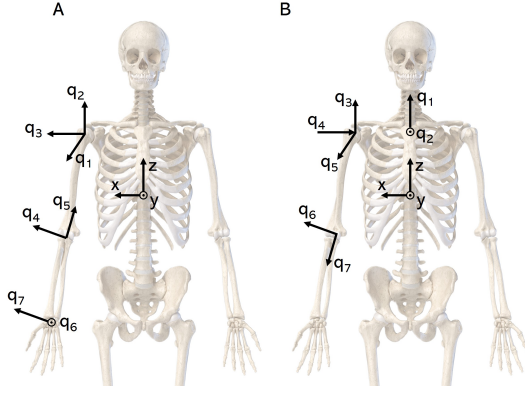


Fig. 2. A) Kinematic model used to model the dataset D1. Three DoFs are used to describe the rotation of the shoulder (in order abduction–adduction, external–internal rotation, flexion–extension), two for the elbow (in order flexion–extension and pronation–supination) and two for the wrist (in order abduction–adduction and flexion–extension); B) Kinematic model used to model the dataset D2. Two DoFs are used to describe the rotation of the scapula (in order abduction–adduction and elevation–depression), three for the shoulder (in order external–internal rotation, flexion–extension and abduction–adduction) and two for the elbow (in order flexion–extension and pronation–supination).

is directly related to the amount of noise (with covariance matrix R) that characterizes the measurements y , and, most importantly, to the choice of matrix H , i.e. the choice of which degrees of freedom to measure (see eq. 14). For this reason, a proper selection of the DoFs to be measured can improve the accuracy of the estimation. This problem, which is dual to the estimation problem described in the previous section, targets the optimal placement of a given number of sensors, with the goal of minimizing the estimation uncertainty, and therefore maximizing the accuracy. To do this, one could rely on the *a posteriori* covariance matrix P_P and minimize a cost function associated to a given norm of P_P , where the argument of the minima is the matrix H itself [26]. Among the possible choices of norm definitions available, it is worth mentioning the Frobenius Norm [26], which leads to the definition of a matrix H that - used with the estimation approach described in Eq. 12 leads to the minimization of the average uncertainty across joints. However, with this choice it could happen that some DoFs could be affected by large uncertainties, compensated by other DoFs where the estimation is more accurate. To avoid this problem, in this work we rely on a differentiable approximation of the maximum eigenvalue of P_P , which we can get considering the Schatten norm

$$\|P_P\|_p = \left(\sum_i \sigma_i^p(P_P) \right)^{\frac{1}{p}} \quad (16)$$

when $p \gg 1$ (ideally $p = \infty$) [34]. A minimization of the index defined in Eq. 16 with H as the argument of the minima, then, will converge toward a matrix H , i.e. a sensor placement distribution, which minimizes the uncertainty in the worst case. Exploiting the differentiability properties of the Schatten norm, it is possible to calculate a gradient of the cost function by differentiating Eq. 16 after substituting Eq. 14. However, because of the particular structure of H for our extended state, which is composed of blocks of squared k -dimensional diagonal matrices, it is not possible to introduce a constraint in the gradient-based optimization procedure, which can also guarantee that H preserves the block structure. For this reason,

TABLE II
LIST OF MUSCLES CONSIDERED IN THIS STUDY

#	Muscles
1	M. Deltoideus pars clavicularis (DC)
2	M. Biceps brachii (BB)
3	M. Triceps brachii (TB)
4	M. Flexor digitorum superficialis (FDS)
5	M. Extensor digitorum (DE)
6	M. Brachioradialis (BR)
7	M. Flexor carpi ulnaris (FCU)
8	M. Extensor carpi ulnaris (ECU)
9	M. Pronator teres (PT)
10	M. Flexor carpi radialis (FCR)
11	M. Abductor pollicis brevis (APB)
12	M. Abductor digiti minimi (ADM)

gradient-free methods should be applied to minimize Eq. 16 in our case.

III. IMPLEMENTATION

A. A dataset of upper limb actions

To test our framework over a wide range of functional tasks, we considered a dataset of upper limb movements collected during daily living activities [31]. More specifically, we picked from [31] two blocks of data. The first, which hereinafter we refer to as the dataset D1, consists of motions performed by 33 healthy subjects (17 Female, age 26.56 ± 2.77 , all right-handed) who were asked to execute three times with their right arm a selection of 30 daily living tasks (enumerated in [31]), while the motion of the upper limb was recorded via optical motion tracking (PhaseSpace, sampling frequency 100Hz). The 3D position of the 20 optical markers used to track the movement were given as input to a two-step procedure: i) an optimization routine to estimate the subject-specific physical parameters of a pre-defined 7 DoFs kinematic model (i.e. length of the links, distances between joints and markers etc. – see Fig. 2-A); ii) an Extended Kalman filter to estimate the time evolution of joint profiles that generate the measured markers' motion [16].

To further validate our method with multi-modal data, we complemented our analysis with a second dataset, the dataset D2, also available in [31]. This is a collection of signals acquired by the Hannover Medical School and consists of motions performed by 20 healthy subjects (12 Female, age 46.8 ± 15.3 , of which eight right-handed) who were asked to perform the same set of tasks of the first dataset, but in this case the placements of optical markers was different. In this manner, the estimation of the relative motion of the shoulder with respect to (w.r.t.) the torsum, but not of the hand w.r.t. the forearm can be obtained. This resulted in a kinematic model that is different from the one used for the dataset D1 (see Fig. 2-B). The Cartesian position of markers was processed using the same procedure introduced above for the dataset D1 to extract joint angular values. In addition, the activation of 12 different muscles was also collected during the execution of the tasks (see list of involved muscles in Tab. II). Muscular data were pre-processed using standard techniques to extract muscle activation envelopes, namely: rectification, filtering (low-pass, cut-off frequency equal to 20Hz), Resampling. These data - and the related information - are also available at [31].

To enable a meaningful comparison between different time-series (different joint trajectories and/or EMG signals) for each dataset we applied Dynamic Time Warping (DTW)

independently. The output of this procedure is a dataset of signals with the same number of time frames T , which for the dataset D1 is equal to 970 - duration of the reference signal used for DTW [30].

Finally, for each dataset independently, the total number recordings was splitted in two blocks, one (containing a random selection of 70% of all samples) that was then used to build the *a priori* dataset (see Sec. II-B) and the second (containing the remaining entries) that served for testing.

Once data were properly warped in time w.r.t. a common reference, we used functional Principal Component Analysis on the *a priori* dataset to identify a low-dimensional set of primitives/basis functions that can be combined with the measurements of few joint/EMG trajectories to optimally reconstruct the joint evolutions and/or the muscles activation envelopes in time. This set of basis elements extracted from data can then be used to map any new movement from the testing dataset to the weight space by simply inverting Eq. 5 and 7 - see Section II-A.

B. Estimation and metrics for performance evaluation

As already mentioned, the usage of MVE as in [25] cannot be used for estimating time-varying and multi-modal signals, and the aim of this paper is to propose a solution which generalizes to the estimation of multi-modal temporal data. Our approach relies on the decomposition of signals via fPCA, which enables the definition of an extended state space x_e in place of joint trajectories. All the experimental validation performed in this paper used the first five functional PCs only (i.e. $k=5$), because these are sufficient to explain over 93% of the total variance in the *a priori* dataset for all the joints [16]. Of note, the number of fPCs used is to be considered as a trade-off between accuracy and computational burden. To quantify the effectiveness of our estimation method, we used as quality index the norm of the estimation error over all the degrees of freedom w.r.t. the real motion. For performance comparison, we considered two different alternative estimation methods, i.e. a direct application of static MVE over each time frame, and a standard pseudo-inverse (Eq. 10) as in [25]. The first is a straightforward application of [25] and consisted in the definition of a static covariance matrix, which we built starting from the matrix of average poses (i.e. for each movement, we picked the average joint configuration). Note that this method can be used for uni-modal data only, due to the problems of computing the matrix P_0 from heterogeneous data. To this aim, we implemented a recursive MVE algorithm in which each time frame is estimated through standard MVE relying on the updated measure (i.e. values of joint angular values/EMG measured at the specific time frame) and the same a-priori information. The second, instead, refers to the implementation of Eq. 10 applied to the extended state. All the methods were tested over multiple examples, where we discarded the information coming from one or more DoFs, and considering random values of noise amplitude. For each testing condition, we evaluated the norm of the reconstruction error (in rad) for all the entries of the testing dataset. Statistical significance of differences between estimations performed through pseudo-inverse, static MVE and the proposed method was evaluated using the Wilcoxon signed-rank test applied to the norm of the reconstruction error.

It is worth noticing that the whole procedure described in this paper relies on the information of the set of movements considered in the *a priori* dataset. In our case, these are complete actions, i.e. the initial and final arm kinematic postures are close to each other. Therefore, the description of generic movements in terms of functional components is more effective when the new signal considered for testing is a complete action as well. It is important to state that this is not a limitation of the work per-se, since it depends on the particular selection of tasks included in the *a priori* dataset. The choice we made in our paper makes our method particularly suitable to estimate cyclic motions (as the ones commonly performed by humans in human-robot industrial collaborative tasks [3]) and, given a novel stream of sensory data, to (likely) automatically detect the end of a cycle and the beginning of the following one, based on the estimation performance. This point will be further discussed later in the manuscript and analyzed as future work.

C. Minimization of a posteriori covariance matrix: optimal sensing design

As mentioned in Sec. II-C, we target the identification of an optimal placement of sensors to minimize the estimation uncertainty. Unfortunately, because of the block structure of H , it is not possible to employ gradient based optimization algorithms, and search methods over discrete states should be considered. We developed an implementation in *Matlab* which uses a genetic algorithm (Population size = 150, Max number of generations = 200, Elite Count = 10) to solve the optimization problem:

$$H^* = \operatorname{argmin} V_1(P_0, H, R), \quad (17)$$

where

$$V_1(P_0, H, R) = \left\| \left(H'R^{-1}H + P_0^{-1} \right)^{-1} \right\|_p = \left\| P_0 - P_0H' (HP_0H' + R)^{-1}HP_0 \right\|_p. \quad (18)$$

IV. RESULTS

To verify the accuracy in reconstruction, we applied our method and the two alternative solutions to estimate all the entries of the test dataset D1, assuming to measure only a limited number of DoFs and artificially adding a Gaussian noise on the measured ones, with standard deviation (std) equal to 0.1 rad ([25]) in the exemplary tests, and in the range $[0, 2]$ rad to test how the results are affected by the Signal-to-Noise (SN) ratio. As anticipated in the previous section, the accuracy of the estimation highly depends on the number and the selection of joints that are measured. For this reason, without any loss of generality, we will first discuss the optimization of the sensor placement, i.e. the solution of the minimization problem of Eq. 17, considering a trade off between number of sensors used and accuracy of the estimation. Then, we will use the optimal placement of sensors to evaluate the estimation accuracy of our method compared to the alternative solutions.

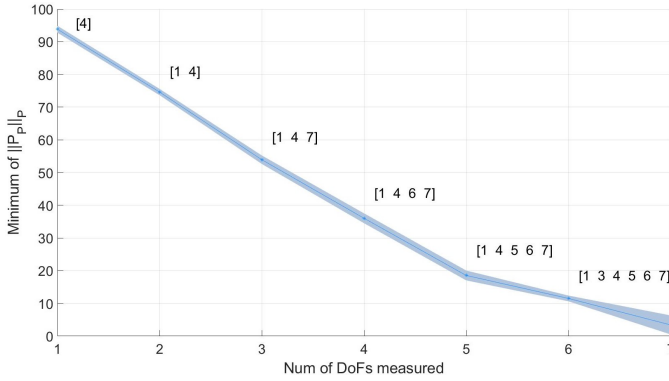


Fig. 3. Minimum value of the P Schatten norm calculated over the *a posteriori* covariance matrix (i.e. an approximation of the maximum estimation uncertainty) versus the number of different DoFs measured for the dataset D1. For each point, the corresponding label reports the n-tuple of measured joints that provides the minimum estimation uncertainty. The label associated to seven DoFs is trivial and omitted for sake of clarity. Results are drawn considering a random added noise in the range $[0, 2]$ rad. The filled line refers to the average minimum of $\|P_p\|_P$ w.r.t. different instances of noise, while the coloured area marks the standard deviation of $\|P_p\|_P$ over the mean.

A. Optimal sensing design

As mentioned above, we can identify an optimal placement of sensors by minimizing a norm of the *a posteriori* covariance matrix reported in Eq. 18. However, the number of DoFs to be measured must be known in advance, and this is a free parameter of the problem that has to be set as a trade-off between estimation uncertainty and complexity of the sensing setup. Indeed, it is reasonable to expect that an higher number of available measures contributes toward a reduction of the maximum uncertainty error (see Fig. 3). Interestingly, there is a consistency over the joints that are subsequently enrolled. In other words, if the optimal placement of three sensors is the one that measures joints number one, four and seven, this n-tuple will be part of all the optimal n-tuples with higher cardinality. Furthermore, we also studied the effect of noise on the identification of the optimal set of DoFs to measure, and verified that a reduction in the signal-to-noise ratio (coherently across joints) does not affect the resulting optimal design, as depicted with the shaded area in Fig. 3, while we anticipate that it will play a crucial role in the accuracy of the estimation itself. For this reason, and without any loss of generality, hereinafter we will consider and discuss the optimal sensing setup corresponding to four measured DoFs. In this case, we obtained that the optimal placement that minimizes the objective function of Eq. 18 is composed by the n-tuple 1,4,6,7 (see. Fig. 2 for joints definition). Of note, the most important joints in terms of reduction of the estimation accuracy is the flexion-extension of the elbow, which indeed represents one of the DoFs that is largely involved in reaching tasks. Then, the algorithm gives particular relevance to one of the joints that move the shoulder and the two DoFs of the wrist, which are highly variable across different movements. Of note, although we can explain the biomechanical reason underlying the particular selection of the most informative joints, their identification is not feasible a priori without our analysis.

Considering the multi-modal case (dataset D2), similar observations could be made, and are here omitted for sake of

clarity, leaving space to discuss the optimal setup we identified. Indeed, in this case it is important to mention that, considering the strong causality relationship between muscle activation and consequent joint temporal evolution, one could impose a number of constraints on a minimum selection of muscles/joints that needs to be recorded. For example, given the list of muscles available in the dataset D2, it is possible to observe that four of them (i.e. Flexor digitorum superficialis, Extensor carpi ulnaris, Abductor pollicis Brevis and Abductor digiti minimi) are specifically enrolled in the movement of the wrist and of the fingers and, therefore, cannot be estimated relying on the kinematic counterpart. Without any loss of generality, we report here on the optimal sensor placement that minimizes the *a-posteriori* covariance matrix when only three joints and eight muscles are directly measured. In this case, the optimal sensor placement resulted in the joint n-tuple 3,5,6 complemented with the muscles n-tuple 1,2,4,7,8,9,11,12. It is interesting to observe that our optimization converged to an optimal sensing setup in which the system exploits the readings of M1 and M2, i.e. the Deltoideus pars clavicularis and the Biceps brachii muscles, which are involved in most of the movements considered in the dataset. In addition, also M7 (Flexor carpi ulnaris) and M9 (Pronator teres), which are responsible for elbow pronation-supination, are enrolled in the optimal setup.

B. Estimation accuracy

To assess the estimation accuracy of our method, we used the optimal placement of sensors identified in the previous section. Considering the dataset D1 and joint trajectory estimation corrupted by Gaussian noise with standard deviation equal to 0.1 rad, numerical results demonstrate that the n-tuple that yields the minimum estimation uncertainty (1,4,6 and 7) provides a norm of estimation error equal to 0.013 ± 0.006 rad (expressed in median \pm interquartile range, values are normalised over T). To quantify the relative error, we calculated the average absolute difference between maxima and minima of each reconstruction (1.43 rad) and verified that the relative median error of our estimation for D1 is equal to 0.9%. Our approach outperforms the pseudoinverse method (0.05 ± 0.014 rad, i.e. on average 280% higher than our approach) and the static mve (0.017 ± 0.08 rad, i.e. on average 35% higher than our approach). A Wilcoxon signed-rank test applied to the norm of the reconstruction error, with Bonferroni correction to compensate for multiple tests, verified that the difference in performances is statistically significant for all the three methods ($p < 0.0001$), with higher performances provided by our method. In Fig. 4 we show a random task selected from the test dataset. It is possible to observe that the estimation performance between the pseudo-inverse, the static MVE and our approach are similar for the joints that can be directly measured, while our method strongly outperforms the alternative solutions in case of no direct measure of the joints. Both the methods based on MVE are able to provide an estimation of the joint temporal evolutions in case of missing data, however only our method based on fPCA+MVE can guarantee the likelihood of the estimated movement, since it is directly embedded in the functional components that encode the *a-priori* information [16]. Furthermore, it is worth noticing that, although on average the estimation accuracy of the static MVE is close to our approach yet statistically significantly different, it provides

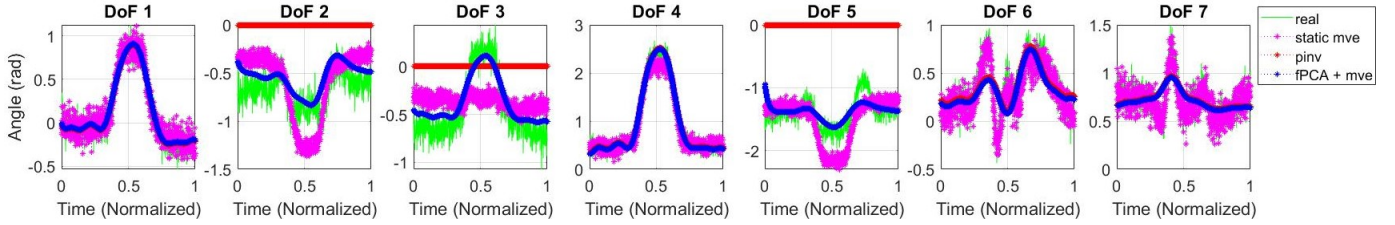


Fig. 4. Motion estimation for a random entry of the test dataset (D1) when the n -tuple of joints $[2, 3, 5]$ is assumed to be not measured and the remaining measurements are corrupted by Gaussian noise (std. 0.1 rad). In this plot we compare the estimation provided by three different methods: the pseudo-inverse (in red) a static MVE (in magenta) and our fPCA+MVE approach (in blue). Note how our method (blue lines) provides a good and clean approximation of the real signal (green line) unlike the pseudo-inverse technique, which cannot estimate joints not measured, and the static MVE, which is highly affected by noise.

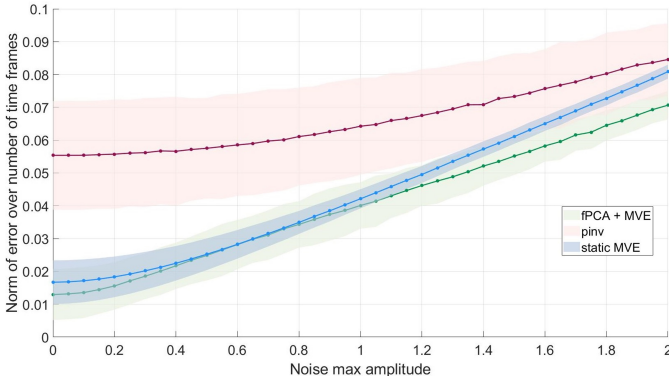


Fig. 5. Norm of the estimation error (n -tuple of joints not measured $[2, 3, 5]$) (normalized over T) vs. noise standard deviation (in rad).

a very noisy estimation (see Fig. 4), which depends on the SN ratio and is intrinsically unavoidable, because there is no direct relationship between subsequent time frames. We also verified the performances of our method for different values of standard deviation of the Gaussian noise applied to the data, testing values ranging in the interval $[0, 2]$ rad. Results are depicted in Fig. 5. It is possible to observe that, although - as expected - the error increases for all the methods at high noise level, numerical values obtained with the pseudo-inverse and with the static MVE are consistently higher than our approach, especially in the range $[0, 0.2]$ deg, which corresponds to a reasonable amount of noise observed in common sensors [35] and therefore will be the proper working range of our system. Statistical significance of the differences in estimation error was proved also in this case through Wilcoxon signed-rank test with Bonferroni correction, proving the difference between the three conditions ($p < 0.0001$).

Finally, to evaluate the effectiveness of our approach with different sensing sources, we also quantified the estimation accuracy obtained with our method using the dataset D2 and obtained a reconstruction error of 0.038 ± 0.023 rad for the kinematic joints and 0.003 ± 0.002 mV for the muscles (values expressed in median \pm interquartile range and normalised over T). In this case, considering that for D2 we observed an average absolute difference between maxima and minima of each reconstruction equal to 1.27 rad and 0.059 mV, the relative error for joints and muscles is equal to 2.9% and 5.1% respectively. An example of motion estimation carried out on a random entry of the test dataset D2 is given in Fig. 6. Using the pseudo-inverse, instead, we obtained 0.224 ± 0.046 rad for the kinematic joints and 0.005 ± 0.002 mV for the muscles. Of

note, in this second dataset the reconstruction error obtained on the kinematic temporal profiles is slightly higher than in the D1 case. This is caused by the fact that here we measure only three joints. It is also interesting to observe that our approach is able to provide a reasonable estimation of muscles envelopes, which is not feasible with the static MVE implementation.

V. DISCUSSIONS AND CONCLUSIONS

In this paper, we presented an integrated framework which extends the Minimum Variance Estimation to the more challenging case of multi-modal time-varying temporal kinematics of multi-modal upper limb muscular-skeletal data. The idea is to leverage on fPCA to establish a mapping from the domain of the time series to a static extended state of weighting coefficients. We demonstrated that, despite the measurement noise, our approach can provide an accurate estimation of all motion profiles, including those referred to the joints not directly measured, leveraging on the covariation patterns identified on *a priori* dataset of motions. However, since the estimation strongly relies on the set of movements used as *a priori* information, it is reasonable to expect that the estimation accuracy could degrade when the novel signal is significantly different with respect to the ones that are present in the *a priori* dataset. This is an intrinsic limitation of minimum variance estimation procedures, and it could be particularly relevant for the muscular domain, where the signals shape may vary significantly as a consequence of stiffness modulation during the action.

With the choice of *a priori* data reported in this paper, our method seems particularly suitable to estimate cyclic tasks, as the ones commonly performed in industrial human-robot collaboration. The development of an index based on the estimation performance could be used to identify whether the analysed motion is a complete action or not, and - on line - detect the end of a cycle. Furthermore, this could also provide important information on how a specific cycle is performed w.r.t. the previous iterations. Indeed, if significant changes in the estimated parameters can be observed while analyzing subsequent iterations of a given task, it is reasonable to expect that these changes could be correlated with modifications in the execution of the cycle, which can be caused for example by fatigue or other non-ergonomic conditions. Our future work will investigate these aspects together with the capability to generalize to other, significantly different, classes of motions and prior information, for on-line motion reconstruction and prediction. The extension to whole-body reconstruction is also envisioned. Another interesting result is that specific joints seem

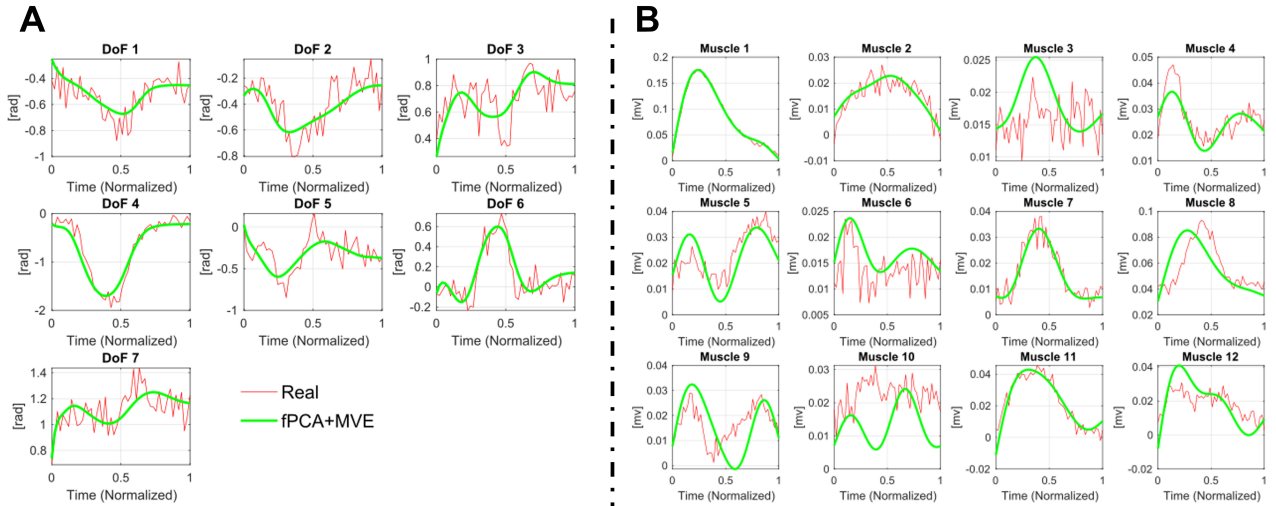


Fig. 6. Motion estimation for a random entry of the test dataset (D2) when we use a noised recording of joints [3,5,6] and muscles [1,2,4,7,8,9,11,12] (Gaussian noise with std. 0.1 rad for joints, with std. 10^{-2} mV for muscles). In this plot we compare the estimation provided our fPCA+MVE approach (in green) versus the real signal. In A the kinematic signals, in B the muscles.

to be more informative than others, and hence they should be preferably measured, if some constraints due to wearability as well as to economic and environmental conditions impose limits on the number of sensing elements and their accuracy. This opens to the perspective of devising optimal design strategies for the development of under-sensorized wearable sensing systems [26].

Finally, it is worth mentioning that there are no constraints on the kind of signals used with our method. We demonstrated this by applying our method to a second dataset that included both kinematic and muscular data. We observed that, while a static implementation of static MVE cannot provide an estimation of multi-modal data because of the difficulties in defining a reliable covariance matrix, with our approach we are able to provide a reasonable estimation of the whole muscle-skeletal temporal evolution relying on a reduced set of measures. However, we noticed that the estimation of muscles profiles is more difficult than the kinematic counterpart. This could be also partially ascribed to the fact that not all the muscular activity is devoted to generate motion. Indeed, a not negligible part of muscular activation is used to preserve muscular tone and to regulate the stiffness during the movement. Our method is not able to discriminate between the different components of the muscle activity and, therefore, this could result in reduced estimation performance for certain muscles and tasks. It is reasonable to expect that this phenomenon can be accentuated when the kinematic component of the signal cannot compensate for the missing information, as for Muscle 10 (Flexor carpi radialis FCR) in Fig. 6. Indeed, FCR muscle is mostly involved in the flexion and abduction of the wrist, which are not part of the available *a priori* kinematic data in the dataset D2. It is likely that the availability of joint angular values of wrist flexion and abduction could also improve the muscular estimation. This point is left for future developments of our work.

Interestingly, in principle, using as *a-priori* information a set of data extracted in a low-stiffness condition, one could also use our method to develop a reliable on-line estimation of arm stiffness even in functional tasks. This will be one of the focuses of future works of our group, in conjunction with

the extension of our framework to other sensing modalities, such as pressure forces at the feet level (which we can record with sensorized insoles), electroencephalographic data etc. This extension will also take into consideration that the relationship between measurements and DoFs could be, in general, non linear (i.e. H will be a block-based matrix of non-linear terms). This will require the integration in our framework of a mapping tool from measurement recordings to the DoFs, as for example done in [8] and in our previous work [30], [29]. We strongly believe that this work can provide interesting insights for the development of under-sensorized tracking solutions for multi-domain applications, opening interesting perspectives in different scenarios, such as advanced human robot interaction (including brain machine interfaces), tracking of working conditions, assessment of muscular fatigue, rehabilitation and modelling of the neuro-muscular system.

REFERENCES

- [1] L. Bai, M. G. Pepper, Y. Yan, M. Phillips, and M. Sakel, "Low cost inertial sensors for the motion tracking and orientation estimation of human upper limbs in neurological rehabilitation," *IEEE Access*, vol. 8, pp. 54 254–54 268, 2020.
- [2] A. Ranavolo, G. Chini, F. Draicchio, A. Silveti, T. Varrecchia, L. Fiori, A. Tattarelli, P. H. Rosen, S. Wischniewski, P. Albrecht, L. Vogt, M. Bianchi, G. Averta, A. Cherubini, L. Fritzsche, M. Sartori, B. Vanderborgh, R. Govaerts, and A. Ajoudani, "Human-robot collaboration (hrc) technologies for reducing work-related musculoskeletal diseases in industry 4.0," in *Proceedings of the 21st Congress of the International Ergonomics Association (IEA 2021)*, N. L. Black, W. P. Neumann, and I. Noy, Eds. Cham: Springer International Publishing, 2022.
- [3] A. Ranavolo, A. Ajoudani, A. Cherubini, M. Bianchi, L. Fritzsche, S. Iavicoli, M. Sartori, A. Silveti, B. Vanderborgh, T. Varrecchia *et al.*, "The sensor-based biomechanical risk assessment at the base of the need for revising of standards for human ergonomics," *Sensors*, vol. 20, no. 20, 2020.
- [4] M. Lorenzini, W. Kim, E. De Momi, and A. Ajoudani, "An online method to detect and locate an external load on the human body with applications in ergonomics assessment," *Sensors*, vol. 20, no. 16, p. 4471, 2020.
- [5] L. Peternel, C. Fang, N. Tsagarakis, and A. Ajoudani, "A selective muscle fatigue management approach to ergonomic human-robot co-manipulation," *Robotics and Computer-Integrated Manufacturing*, vol. 58, pp. 69–79, 2019.
- [6] C. Varotto, Z. Sawacha, L. Gizzi, D. Farina, and M. Sartori, "Closed-loop emg-informed model-based analysis of human musculoskeletal mechanics on rough terrains," in *2017 International Conference on Rehabilitation Robotics (ICORR)*, 2017, pp. 364–368.

- [7] M. Gabbicini, G. Stillfried, H. Marino, and M. Bianchi, in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE.
- [8] E. Todorov, "Probabilistic inference of multijoint movements, skeletal parameters and marker attachments from diverse motion capture data," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 11, pp. 1927–1939, 2007.
- [9] S. Ambike and J. P. Schmiedeler, "The leading joint hypothesis for spatial reaching arm motions," *Experimental brain research*, vol. 224, no. 4, pp. 591–603, 2013.
- [10] M. L. Latash, J. P. Scholz, and G. Schöner, "Toward a new theory of motor synergies," *Motor control*, vol. 11, no. 3, pp. 276–308, 2007.
- [11] A. d'Avella, P. Saltiel, and E. Bizzi, "Combinations of muscle synergies in the construction of a natural motor behavior," *Nature neuroscience*, vol. 6, no. 3, p. 300, 2003.
- [12] M. Santello, M. Flanders, and J. F. Soechting, "Postural hand synergies for tool use," *The Journal of Neuroscience*, vol. 18, no. 23, pp. 10 105–10 115, 1998.
- [13] M. Santello, M. Bianchi, M. Gabbicini, E. Ricciardi, G. Salvietti, D. Prattichizzo, M. Ernst, A. Moscatelli, H. Jörnbell, A. M. Kappers *et al.*, "Hand synergies: integration of robotics and neuroscience for understanding the control of biological and artificial hands," *Physics of life reviews*, 2016.
- [14] M. G. Catalano, G. Grioli, E. Farnioli, A. Serio, C. Piazza, and A. Bicchi, "Adaptive synergies for the design and control of the pisa/iit soffhand," *The International Journal of Robotics Research*, vol. 33, no. 5, pp. 768–782, 2014.
- [15] F. Ficuciello, G. Palli, C. Melchiorri, and B. Siciliano, "Experimental evaluation of postural synergies during reach to grasp with the ub hand iv," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2011, pp. 1775–1780.
- [16] G. Averta, C. Della Santina, G. Valenza, A. Bicchi, and M. Bianchi, "Exploiting upper-limb functional principal components for human-like motion generation of anthropomorphic robots," *Journal of NeuroEngineering and Rehabilitation*.
- [17] F. Ficuciello, P. Falco, and S. Calinon, "A brief survey on the role of dimensionality reduction in manipulation learning and control," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 2608–2615, 2018.
- [18] P. Falco, M. Saveriano, E. G. Hasany, N. H. Kirk, and D. Lee, "A human action descriptor based on motion coordination," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 811–818, 2017.
- [19] G. Evangelidis, G. Singh, and R. Horaud, "Skeletal quads: Human action recognition using joint quadruples," in *2014 22nd International Conference on Pattern Recognition*, 2014, pp. 4513–4518.
- [20] M. Kaufmann, E. Aksan, J. Song, F. Pece, R. Ziegler, and O. Hilliges, "Convolutional autoencoders for human motion infilling," in *2020 International Conference on 3D Vision (3DV)*. IEEE, 2020, pp. 918–927.
- [21] Z. Li, Y. Zhou, S. Xiao, C. He, Z. Huang, and H. Li, "Auto-conditioned recurrent networks for extended complex human motion synthesis," *arXiv preprint arXiv:1707.05363*, 2017.
- [22] K. Tatani and Y. Nakamura, "Dimensionality reduction and reproduction with hierarchical nlpc neural networks-extracting common space of multiple humanoid motion patterns," in *2003 IEEE International Conference on Robotics and Automation (Cat. No. 03CH37422)*, vol. 2. IEEE, 2003, pp. 1927–1932.
- [23] S. Mulatto, A. Formaglio, M. Malvezzi, and D. Prattichizzo, "Animating a synergy-based deformable hand avatar for haptic grasping," in *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. Springer, 2010, pp. 203–210.
- [24] Y. Arik, T. Inamura, and J. Morimoto, "Observing human movements to construct a humanoid interface," in *2014 IEEE-RAS International Conference on Humanoid Robots*. IEEE, 2014, pp. 342–347.
- [25] M. Bianchi, P. Salaris, and A. Bicchi, "Synergy-based hand pose sensing: Reconstruction enhancement," *The International Journal of Robotics Research*, vol. 32, no. 4, pp. 396–406, 2013.
- [26] —, "Synergy-based hand pose sensing: Optimal glove design," *The International Journal of Robotics Research*, vol. 32, no. 4, pp. 407–424, 2013.
- [27] S. Ciotti, E. Battaglia, N. Carbonaro, A. Bicchi, A. Tognetti, and M. Bianchi, "A synergy-based optimally designed sensing glove for functional grasp recognition," *Sensors*, vol. 16, no. 6, 2016. [Online]. Available: <https://www.mdpi.com/1424-8220/16/6/811>
- [28] M. O. Kuismin and M. J. Sillanpää, "Estimation of covariance and precision matrix, network structure, and a view toward systems biology," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 9, no. 6, p. e1415, 2017.
- [29] G. Averta, G. Valenza, V. Catrambone, F. Barontini, E. P. Scilingo, A. Bicchi, and M. Bianchi, "On the time-invariance properties of upper limb synergies," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 7, pp. 1397–1406, 2019.
- [30] G. Averta, C. Della Santina, E. Battaglia, F. Felici, M. Bianchi, and A. Bicchi, "Unveiling the principal modes of human upper limb movements through functional analysis," *Frontiers in Robotics and AI*, vol. 4, p. 37, 2017.
- [31] G. Averta, F. Barontini, V. Catrambone, S. Haddadin, G. Handjaras, J. P. Held, T. Hu, E. Jakobowitz, C. M. Kanzler, J. Kühn *et al.*, "U-limb: A multi-modal, multi-center database on arm motion control in healthy and post-stroke conditions," *GigaScience*, vol. 10, no. 6, p. giab043, 2021.
- [32] J. O. Ramsay, G. Hooker, and S. Graves, *Functional data analysis with R and MATLAB*. Springer Science & Business Media, 2009.
- [33] F. Camastra, "Data dimensionality estimation methods: a survey," *Pattern recognition*, vol. 36, no. 12, pp. 2945–2954, 2003.
- [34] R. Spica and P. R. Giordano, "Active decentralized scale estimation for bearing-based localization," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 5084–5091.
- [35] L. K. Simone, N. Sundarajan, X. Luo, Y. Jia, and D. G. Kamper, "A low cost instrumented glove for extended monitoring and functional hand assessment," *Journal of neuroscience methods*, vol. 160, no. 2, pp. 335–348, 2007.



Giuseppe Averta (male) is currently a postdoctoral researcher at the Polytechnic of Turin. He got his B.S. in Biomedical Engineering and M.S. in Robotics, in 2013 and 2016, respectively, from the University of Pisa. In 2020, he obtained a Ph.D. in Robotics cum laude from the same institution with a dissertation that has been selected as the winner of the Georges Giralt PhD Award 2021 (for the best European PhD thesis). In 2019, he was a visiting student at the Eric P. and Evelyn E. Newman Laboratory for Biomechanics and Human Rehabilitation Group at MIT. Giuseppe's research interests are in the study of human movement in healthy and pathological conditions and in the implementation of human-inspired design, planning, and control guidelines for autonomous, collaborative, assistive, and prosthetic robots.



Matilde Iuculano (female) is currently a master's student in Robotics and Automation Engineering at the University of Pisa. Her research interests are in the study and in the mathematical modeling of human movements, with specific focus on the upper limb. The application or her work is toward the development of multi-domain optimal sensing systems for motion tracking, and in the natural control of multi-dof hand-wrist prostheses exploiting shared control policies.



Paolo Salaris (male) received the M.Sc. degree in Electrical Engineering in 2007 and his PhD in Automation and Bioengineering in 2011, both from the university of Pisa. He has been Visiting Scholar at Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana-Champaign in 2009. He has been a PostDoc at the Research Center "E.Piaggio" from March 2011 to January 2014 and at LAAS-CNRS in Toulouse from February 2014 to July 2015. From October 2015 to August 2019 he has been a permanent researcher (CRCN) at Inria Sophia Antipolis Méditerranée. From September 2019 he is an assistant professor at the Department of Information Engineering and the research Center "E. Piaggio", University of Pisa. His main research interests are optimal motion planning and control, active sensing, optimal sensors placements and optimal estimation, multi-robot and nonholonomic systems. ϵ



Matteo Bianchi (male) is currently an Associate Professor at the Research Centre "E. Piaggio" and the Department of Information Engineering (DII) of the Università di Pisa. He is also clinical research affiliate at Mayo Clinic (Rochester, USA) and serves as co-Chair of the RAS Technical Committee on Robot Hands, Grasping and Manipulation and served as Vice-Chair for Information and Dissemination of the RAS Technical Committee on Haptics (2018–2021). His research interests include haptic interface and sensor design; human and robotic hands: optimal sensing and control. He acts as the Principal Investigator of national and EU grants, and research contracts with companies in the field of human-machine interaction. He is author or co-author of more than 100 peer-reviewed contributions and serves as member of the editorial/organizing board of international conferences and journals. He is the recipient of several national and international awards.