This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

A Variational Bayesian Gaussian Mixture-Nonnegative Matrix Factorization Model to Extract Movement Primitives for Robust Control

Hongbo Xie, Kerrie Mengersen, Changan Di, Yongjian Zhang, Justin Lipman, *Senior Member, IEEE*, and Sabine Van Huffel, *Fellow, IEEE*

***Abstract*—Nonnegative matrix factorization (NMF) is a powerful tool for parameter estimation applied in numerous robotics applications, such as path planning, motion trajectory prediction, and motion intention detection. In particular, NMF has been successfully used to extract simplified and organized movement primitives from myoelectric signal (MES) for robust control of multi-degree of freedom humanoid robots. However, MES is typically contaminated by complex noise sources. The system performance often degrades due to the simplified Gaussian assumption of the noise distribution in existing NMF methods. Furthermore, most existing NMF models are unable to automatically determine the rank of the latent matrices. To address these issues, this article presents a hybrid variational Bayesian Gaussian mixture and NMF (GMNMF) model with a finite Gaussian mixture model adopted to fit the mixed noise density function of MES. In addition, the automatic relevant determination criterion is applied to automatically infer the number of movement primitives. The coordinate descent update rules for the proposed model are formulated by mean-field variational Bayesian inference. We assess the model performance on five synthetic noise distribution functions and an experimental MES dataset to perform six wrist movements. The results demonstrate that GMNMF yields low error and high robustness in extracting the movement primitives over four competitive methods for robust cybernetic control.**

***Index Terms*—Bayesian inference, Gaussian mixture model (GMM), humanoid robots, model learning for control, movement primitive, nonnegative matrix factorization (NMF).**

Manuscript received 11 September 2021; revised 11 February 2022 and 24 April 2022; accepted 1 July 2022. This article was recommended by Associate Editor Z. Yu. *(Corresponding author: Hongbo Xie.)*

Hongbo Xie is with the School of Information Science and Engineering, Yunnan University, Kunming 650500, China, and also with the Centre for Data Science, Queensland University of Technology, Brisbane, QLD 4001, Australia (e-mail: hongbo.xie@qut.edu.au).

Kerrie Mengersen is with the Centre for Data Science, Queensland University of Technology, Brisbane, QLD 4001, Australia (e-mail: k.mengersen@qut.edu.au).

Changan Di and Yongjian Zhang are with the School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: dichangan@njust.edu.cn; yumill629@163.com).

Justin Lipman is with the Faculty of Engineering and IT, University of Technology Sydney, Ultimo, NSW 2007, Australia (e-mail: justin.lipman@uts.edu.au).

Sabine Van Huffel is with the Department of Electrical Engineering, ESAT Stadius Division, KU Leuven, 3001 Leuven, Belgium (e-mail: sabine.vanhuffel@esat.kuleuven.be).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/THMS.2022.3194593.

Digital Object Identifier 10.1109/THMS.2022.3194593

I. Introduction

THE myoelectric control mechanism, which utilizes the myoelectric signal (MES) as control inputs, has been broadly applied to drive humanoid robots, including exoskeleton robots, teleoperation robots, prosthetic hands, and other human–machine interfaces [1]–[6]. However, commercialization of these research outcomes is not currently feasible as none of these techniques provides a reliable simultaneous control scheme as in humans. Research has shown that the human central nervous system (CNS) activates a group of muscles to simplify the construction of motor behaviors, rather than explicitly controlling individual muscles. This is referred to as motor primitives or synergies [7]–[9]. Compared with the traditional single degree of freedom (DOF) model controlled by an ON/OFF scheme, adapting a motor primitive or synergetic control paradigm can significantly improve the robot's performance during real-time control [8], [9].

Nonnegative matrix factorization (NMF) of MES is the prime technique used to explicitly extract movement primitives for cybernetic control. For a given multichannel MES matrix $\mathbf{Y}$, the nonnegative factor matrix $\mathbf{U}$ consists of the movement primitive vectors, which act as a functional unit to indicate the proportion or importance of each muscle within a specific movement primitive. Another factor matrix $\mathbf{V}$ is the activation coefficient to reflect the recruitment strength of each primitive in real time. However, existing algorithms face at least two major challenges in practical movement primitive extraction. First, most of the existing algorithms are least-squares based, and thus lack robustness to outliers in MES measurements, such as ubiquitous motion artifacts, crosstalk, and large activation spikes [2]. Second, knowledge of the matrix rank is a prerequisite to implement most of these algorithms. Unfortunately, the nonnegative rank of a matrix is NP-hard to compute [2].

To address these challenges, this article presents a Bayesian generative model to combine NMF and a Gaussian mixture model (GMM) into a unified framework for movement primitive extraction. Since a finite GMM is a universal approximator to fit any continuous probability density [10], a GMM is employed to approximate the distribution of the mixed noise components in the MES. The Dirichlet process mixture (DPM) has been proposed as a nonparametric Bayesian technique to estimate the number of Gaussians for flexible modeling of data and clustering

2168-2291 © 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 20,2022 at 09:31:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[11]. However, DPMs favor introducing new components at a log rate as the sample size increases, and therefore tend to overestimate the number of Gaussians [12], [13]. An effective and simple merging scheme is developed given the Kullback–Leibler (KL) divergence between a pair of Gaussian distributions to determine the required number of Gaussians to fit the noise. Furthermore, the automatic relevance determination (ARD) criterion is adopted to infer the rank of the low-rank nonnegative representation while pruning the columnwise sparse latent factor matrices [14]–[16].

A. Major Contribution

In comparison with existing studies, the major contributions of this article are as follows.

1) To suppress different types of noise in MES, a hybrid Bayesian NMF and a mixture of the Gaussian model is proposed.
2) A new KL divergence based metric is proposed to quantify the similarity between two Gaussian distributions for merging the GMM to fit various types of noise.
3) The ARD criterion is adopted to determine the rank without the need to be prespecified or manually pruned.
4) A new generative Gaussian mixture-nonnegative matrix factorization (GMNMF) model is developed and the corresponding variational Bayesian (VB) update rules to estimate the parameters and latent variables are derived. The movement primitives extracted by the proposed GMNMF model significantly improve the accuracy and robustness in hand motion classification for cybernetic control.

The rest of this article is organized as follows. Section II reviews the existing NMF models applied in movement primitive extraction. Section III formulizes the problem and proposes the new model. The related update rules that extract latent factors to deal with complex noise are then provided. The synthetic data and experimental MESs are introduced in Section IV. Section V shows experimental results on both datasets to demonstrate the superiority of GMNMF over four competitive methods, i.e., MUNMF, L1NMF, L21NMF, and a VB NMF counterpart [17]. Finally, Section VI concludes this article.

II. Related Work

The original NMF algorithm is relatively simple and intuitive, as it minimizes the error between the observation $\mathbf{Y}$ and the reconstructed term $\mathbf{UV}$ by multiplicative update (MUNMF) [18]. This algorithm is widely applied to investigate the performance of movement primitive-based control strategies [19]–[21]. However, the MES recordings are always contaminated by various types of noise, whereas MUNMF assumes Gaussian-distributed noise. L1 norm penalty for sparse NMF (L1NMF) has been developed to attenuate sparse noise and promote robustness [22]. Lin et al. [23] applied the L1 regularization model to extract primitives for simultaneous and proportional control of a prosthetic hand. Although L1NMF captures outliers, it neglects other types of noise [24]. Kong et al. [24] presented the concept of L21-norm, which was later applied to NMF (L21NMF) to handle noise and outliers with more consistent and accurate factors. But once the noise scale is increased, the relative reconstruction error and normalized mutual information of L21NMF drop significantly as with L1NMF [25]. There are many NMF variants that define the modified loss or gradient descent optimization method, or place additional prior knowledge of the factors into the model to accommodate specific data. This article focuses on effective NMF models for motion primitive extraction. Readers are referred to a review paper [26] for additional NMF algorithms.

Bayesian inference is a principled theory to handle uncertainty. Under this statistical framework, NMF can be naturally solved by maximizing the posterior distribution of a generative model with a nonnegativity constraint. Cemgil [27] presented a full Bayesian inference NMF model, which employs a Poisson distribution for the observations and a gamma distribution for the two latent factor matrices. However, the Poisson distribution takes on only integer values, which results in misleading outcomes if this model is applied to real-valued MESs. Devarajan and Cheung [28] described an information-theoretic approach to NMF with noise corresponding to Gaussian, gamma, and the generalized inverse Gaussian distribution, respectively, to extract motor primitives. They found that the NMF algorithm derived from a noise distribution that matches the noise type of the data yields the most accurate identification of both $\mathbf{U}$ and $\mathbf{V}$. However, when the noise assumed by the NMF algorithm and the data noise type do not completely agree, the extracted results may lose substantial information about the underlying data structure.

Regarding determining the number of motor primitives, i.e., the rank of matrix $\mathbf{X}$, it is prespecified in most of the optimization-based models [18]–[25]. Endres et al. [29] proposed a Laplace approximation to the posterior criterion to compare the statistical model performance under different rank values. In another Gaussian-process-based movement primitive model, Clever et al. [30] adopted a similar criterion to compare the model performance with the number of movement primitives ranging from two to eight.

Given complexities associated with MES, significant intraindividual variability in detectable motor primitives, and stochastic properties of MES, existing NMF models have struggled to provide robust and reliable movement features. The first challenge is that existing NMF approaches, under either an optimization or probability framework, assume a single noise distribution in the observed signal. However, MES is a weak electrophysiological recording that may be corrupted by environmental noise, motion artifacts, crosstalk, power frequency disturbance, and large activation spikes [31]. These NMF approaches are thus inadequate to fit the MES, resulting in either overfitting or underfitting [32]–[34]. Furthermore, the exact number of primitives (the rank of the model) plays a key role in accurate and robust robot control as well as indicating the neural mechanism of muscle synergies. However, the rank in existing NMF approaches either needs to be prespecified or manually pruned over multiple attempts.

III. Methods

A. Problem Formulation

The muscle synergy concept provides an explanation for how the CNS deals with the complexity and high dimensionality of

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 20,2022 at 09:31:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

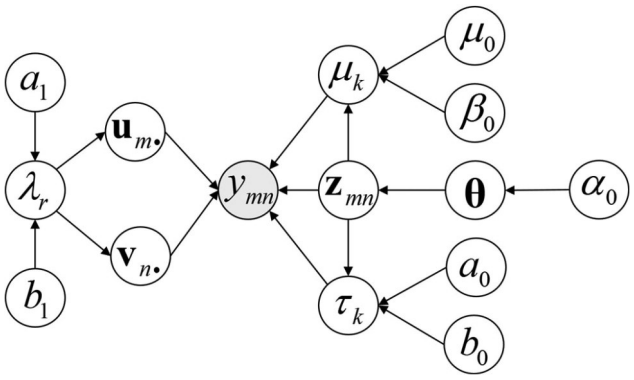


Fig. 1. Directed acyclic graph representing the GMNMF model, where $y_{mn}$is the observation, $\mathbf{z}_{mn}$ μ_k τ_k, $\boldsymbol{\theta}$, $\mathbf{u}_{m\bullet}$, $\mathbf{v}_{n\bullet}$, and λ_r are parameters and latent variables, and μ_0 β_0, a_0, b_0, a_1, b_1, and α_0 are hyperparameters.

motor control for the musculoskeletal system across multiple DoFs [2]. The concept posits that the CNS reduces the motor tasks into a lower dimensional subspace in a modular form, or simple units called movement primitives. In this framework, each module or motor primitive consists of a set of movement variables, such as joint trajectories or muscle activations acting synergistically over time. A combination of small numbers of these primitives can generate complex motor behaviors. Hence, the multichannel MES $\mathbf{Y}$ is considered as a linear mixture of movement primitives $\mathbf{U}$ with a weighting function or activation coefficients $\mathbf{V}$ across time, as formulated in the following:

$$\mathbf{Y} = \mathbf{X} + \mathbf{E} = \mathbf{U}\mathbf{V}^T + \mathbf{E} \tag{1}$$

where $\mathbf{E} \in \mathbb{R}^{m\times n}$ is the noise or residual. In practice, the MES matrix $\mathbf{Y} \in \mathbb{R}^{m\times n}$ is commonly preprocessed with nonnegative values. The problem in (1) is to analytically recover the latent factors $\mathbf{U} \in \mathbb{R}_+^{m\times r}$and $\mathbf{V} \in \mathbb{R}_+^{n\times r}$ using the NMF technique.

B. Model Specification of GMNMF

For a probabilistic interpretation, we assign priors to all latent variables and parameters, and a likelihood to the observation matrix $\mathbf{Y}$in (1). Fig. 1 shows the generative graph of the proposed GMNMF. The primary assumption of the model is that the noise is fitted by a mixture of Gaussian distributions

$$p(e_{mn}) = \sum_{k=1}^{K} \theta_k \mathcal{N}(\mu_k, \tau_k) \tag{2}$$

where θ_k represents the component weight with $\theta_k > 0$ and $\sum_{k=1}^{K} \theta_k = 1$, K is the total number of Gaussians required, and μ_k, τ_k denote the mean and precision of the kth Gaussian components, respectively. Gaussian-gamma priors are placed on μ_k and τ_k

$$p(\mu_k, \tau_k) \sim \mathcal{N}(\mu_k|\mu_0, \beta_0\tau_k)\mathcal{G}(\tau_k|a_0, b_0). \tag{3}$$

Let $\mathbf{z}_{mn}$ be the vector specifying the index of the parameter corresponding to the element e_{mn}. It is reasonable to assign a multinomial distribution with parameter $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)$ as a mixing proportion or weight $\mathbf{z}_{mn}$. To construct a conjugate Bayesian model, we assign a Dirichlet distribution to $\boldsymbol{\theta}$

$$p(\boldsymbol{\theta}) = \mathrm{Dir}(\boldsymbol{\theta}|\boldsymbol{\alpha}_0) \tag{4}$$

with hyperparameter $\boldsymbol{\alpha}_0 = (\alpha_{01}, \ldots, \alpha_{0K})$.

We introduce the truncated normal distribution

$$\mathcal{TN}(x|\mu, \lambda) = \begin{cases} \frac{\sqrt{\lambda/2\pi}\exp\{-\lambda(x-\mu)^2/2\}}{1-\Phi(-\mu\sqrt{\lambda})} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \tag{5}$$

as the prior over $\mathbf{U}$ and $\mathbf{V}$ for nonnegativity, that is

$$p(u_{mr}) \sim \mathcal{TN}(u_{mr}|\mu_{mr}^U, \lambda_r) \tag{6}$$

$$p(v_{nr}) \sim \mathcal{TN}(v_{nr}|\mu_{nr}^V, \lambda_r) \tag{7}$$

where $\mathcal{TN}(x)$is a normal distribution with zero density below $x = 0$ and renormalized to integrate to one. More details of this distribution are given in the Appendix. The precision λ_r of the rth column of $\mathbf{U}$ and $\mathbf{V}$ obeys a gamma distribution defined by shape a_1 and rate b_1

$$p(\lambda_r) \sim \mathcal{G}(a_1, b_1). \tag{8}$$

With the priors defined in (6) to (8), some of hyperparameters λ_rs will be iteratively updated to very large values. The corresponding columns of $\mathbf{U}$ and $\mathbf{V}$ are removed since they make little contribution to the approximation $\mathbf{X}$. Hence, the columnwise sparsity of the latent factors $\mathbf{U}$ and $\mathbf{V}$ and the low rank of $\mathbf{X}$ are jointly satisfied.

Considering a mixture Gaussian-distributed likelihood as in existing NMF models, we denote $\mathbf{Y}$ as

$$\begin{aligned} p(\mathbf{Y}) &= p(\mathbf{Y}|\mathbf{U}, \mathbf{V}, \mathbf{z}, \boldsymbol{\tau}) \\ &= \prod_{mn}\prod_{k=1}^{K} \mathcal{N}(y_{mn}|\mathbf{u}_{m\bullet}\mathbf{v}_{n\bullet}^T + \mu_k, \tau_k)^{z_{mnk}} \end{aligned} \tag{9}$$

where $z_{mnk}$is the indicator variable with $\sum_{k=1}^{K} z_{mnk} = 1$ and $\mathbf{z}_{mn} = (z_{mn1}, \ldots, z_{mnK}) \in \{0, 1\}^K$. The joint probability distribution of the observations, the latent factors, and the parameters can be represented as

$$\begin{aligned} &p(\mathbf{Y}, \mathbf{U}, \mathbf{V}, \boldsymbol{\mathcal{Z}}, \boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\theta}, \boldsymbol{\lambda}) \\ &\quad = p(\mathbf{Y}|\mathbf{U}, \mathbf{V}, \boldsymbol{\mathcal{Z}}, \boldsymbol{\mu}, \boldsymbol{\tau})p(\mathbf{U}|\boldsymbol{\lambda})p(\mathbf{V}|\boldsymbol{\lambda})p(\boldsymbol{\lambda}|a_1, b_1) \\ &\qquad \cdot p(\boldsymbol{\mathcal{Z}}|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{\alpha}_0)p(\boldsymbol{\mu}|\boldsymbol{\tau})p(\boldsymbol{\tau}|a_0, b_0) \end{aligned} \tag{10}$$

where $\boldsymbol{\mathcal{Z}} = \{\mathbf{z}_{mn}\}$, $\boldsymbol{\mu} = \{\mu_1, \ldots, \mu_1\}$, $\boldsymbol{\tau} = \{\tau_1, \ldots, \tau_k\}$, and $\boldsymbol{\lambda} = (\lambda_1, \cdots \lambda_r)$with r denoting the pruned rank.

C. VB Inference of GMNMF

We aim to estimate all the posterior conditional distributions for the latent variables and hyperparameters from the joint probability distribution. However, it is challenging to obtain an analytical solution of (10) containing intractable integrals. Markov chain Monte Carlo (MCMC) is a feasible technique to approximate intractable integrals through sampling. Unfortunately, the heavy computational cost limits the application of MCMC to small-scale problems. VB is a more efficient alternative to MCMC as well as Laplace approximation to compute the optimal posterior estimation, which is adopted in this article [16]. We

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 20,2022 at 09:31:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

collect latent variables into a set $\boldsymbol{\mathcal{D}} = \{\mathbf{U}, \mathbf{V}, \boldsymbol{\mathcal{Z}}, \boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\theta}, \boldsymbol{\lambda}\}$. Based on VB theory, a variational distribution $q(\boldsymbol{\mathcal{D}})$is used to approximate the true posterior distribution $p(\boldsymbol{\mathcal{D}}|\mathbf{Y})$. We use the KL divergence between $q(\boldsymbol{\mathcal{D}})$and $p(\boldsymbol{\mathcal{D}}|\mathbf{Y})$to quantify how well the approximation $q(\boldsymbol{\mathcal{D}})$fits the true posterior $p(\boldsymbol{\mathcal{D}}|\mathbf{Y})$ at each iteration, denoted by

$$\min_{q(\boldsymbol{\mathcal{D}})} \mathrm{KL}(q(\boldsymbol{\mathcal{D}}) \parallel p(\boldsymbol{\mathcal{D}}|\mathbf{Y})) = \int q(\boldsymbol{\mathcal{D}}) \ln \frac{q(\boldsymbol{\mathcal{D}})}{p(\boldsymbol{\mathcal{D}}|Y)} d\boldsymbol{\mathcal{D}}. \quad (11)$$

This is equivalent to estimating the marginal likelihood $p(\mathbf{Y})$ with a maximal lower bound. Based on the principle of mean-field approximation, we partition the latent variables and hyperparameters and assume that each partition is independent. The posterior approximation $q(\boldsymbol{\mathcal{D}})$ is therefore expressed as the product of individual $q(\boldsymbol{\mathcal{D}}_i)$ [16]

$$q(\boldsymbol{\mathcal{D}}) = \prod_i q(\boldsymbol{\mathcal{D}}_i). \quad (12)$$

Under these assumptions, the closed-form solution of $q(\boldsymbol{\mathcal{D}}_i)$, with other fixed, can be attained by

$$\ln q(\boldsymbol{\mathcal{D}}_i) = \langle \ln p(\mathbf{Y}, \boldsymbol{\mathcal{D}}) \rangle_{\boldsymbol{\mathcal{D}} \backslash \boldsymbol{\mathcal{D}}_i} + c \quad (13)$$

where $\langle \bullet \rangle$ is the expectation, c represents a constant independent of $\boldsymbol{\mathcal{D}}_i$, and $\boldsymbol{\mathcal{D}} \backslash \boldsymbol{\mathcal{D}}_i$ represents all variables except for the current one $\boldsymbol{\mathcal{D}}_i$.

Using the aforementioned general results, the closed-form variational inference scheme for GMNMF model can be derived from the following factorization form"

$$\begin{aligned} & q(\mathbf{U}, \mathbf{V}, \boldsymbol{\mathcal{Z}}, \boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\theta}, \boldsymbol{\lambda}) \\ & = \prod_{m=1}^{M} \prod_{r=1}^{R} q_{u_{mr}}(\mu_{mr}^U, \tau_{mr}^U) \prod_{n=1}^{N} \prod_{r=1}^{R} q_{v_{nr}}(\mu_{nr}^V, \tau_{nr}^V) \\ & \quad \times \prod_{r=1}^{R} q_{\lambda_r}(a_r^*, b_r^*) \\ & \cdot \prod_{m=1}^{M} \prod_{n=1}^{N} q_{\mathbf{z}_{mn}}(\phi_{mnk}) \prod_{k=1}^{K} q_{(\mu,\tau)_k}(m_k, \beta_k, a_k, b_k) q_{\theta_k}(\alpha_k). \end{aligned} \quad (14)$$

The optimal posterior approximation of each factorized distribution in (14) can be analytically attained by applying the general expression in (13) with each variable iteratively updated. We present the major update rules here and list the inference details in the Appendix.

1) Estimation of Factor Matrices: The factor matrices $\mathbf{U}$ and $\mathbf{V}$ correspond to primitives and activation, respectively. The approximate posterior distribution for each entry u_{mr} can be inferred as a truncated Gaussian distribution

$$q(u_{mr}) = \mathcal{TN}(u_{mr}|\mu_{mr}^U, \tau_{mr}^U) \quad (15)$$

with mean μ_{mr}^U and precision τ_{mr}^Ugiven by

$$\mu_{mr}^U = \frac{1}{\tau_{mr}^U} \sum_n \sum_{k=1}^{K} \phi_{mnk} \langle \tau_k \rangle \langle v_{nr} \rangle (y_{mn} - \sum_{r' \neq r} \langle u_{mr'} \rangle \langle v_{nr'} \rangle - \langle \mu_k \rangle) \quad (16)$$

$$\tau_{mr}^U = \sum_n \sum_{k=1}^{K} \phi_{mnk} \langle \tau_k \rangle \langle v_{nr}^2 \rangle + \langle \lambda_r \rangle \quad (17)$$

where ϕ_{mnk} is the posterior probability of z_{mnk}, which is given later. Similarly, v_{nr}obeys the truncated Gaussian distribution

$$q(v_{nr}) = \mathcal{TN}(v_{nr}|\mu_{nr}^V, \tau_{nr}^V) \quad (18)$$

with parameters

$$\mu_{nr}^V = \frac{1}{\tau_{nr}^V} \sum_m \sum_{k=1}^{K} \phi_{mnk} \langle \tau_k \rangle \langle u_{mr} \rangle \times \left(y_{mn} - \sum_{r' \neq r} \langle u_{mr'} \rangle \langle v_{nr'} \rangle - \langle \mu_k \rangle \right) \quad (19)$$

$$\tau_{nr}^V = \sum_m \sum_{k=1}^{K} \phi_{mnk} \langle \tau_k \rangle \langle u_{mr}^2 \rangle + \langle \lambda_r \rangle . \quad (20)$$

The latent parameter λ_r controls the double columnwise sparsity, complying with a gamma distribution as its posterior

$$q(\lambda_r) = \mathcal{G}(\lambda_r|a^*, b^*) \quad (21)$$

where

$$a^* = a_1 + \frac{M+N}{2} \quad (22)$$

$$b^* = b_1 + \frac{1}{2} \sum_{m=1}^{M} u_{mr}^2 + \frac{1}{2} \sum_{n=1}^{N} v_{nr}^2. \quad (23)$$

Those columns with smaller $\sum_{m=1}^{M} u_{mr}^2 + \sum_{n=1}^{N} v_{nr}^2$ in (23) lead to larger $\mathbb{E}_q(\lambda_r)$ and, hence, larger parameters of τ_{mr}^U and τ_{nr}^V, which in turn force μ_{mr}^U ($m = 1$: M) in (16) and μ_{nr}^V ($n = 1$: N) in (19) to tend to zero. The corresponding columns can thus be removed [35].

2) Estimation of Gaussian Mixture Noise: All the remaining parameters $\boldsymbol{\mathcal{Z}}, \boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\theta}$ are directly related to the mixture Gaussian noise model. The posterior for $\mu_k$follows a Gaussian distribution

$$q(\mu_k) = \mathcal{N}(\mu_k|m_k, \beta_k) \quad (24)$$

with mean and precision

$$m_k = \frac{1}{\beta_k} \left(\beta_0 \mu_0 + \sum_{m,n} \phi_{mnk} (y_{mn} - \langle \mathbf{u}_m \rangle \langle \mathbf{v}_n^T \rangle \right) \quad (25)$$

$$\beta_k = \beta_0 + \sum_{m,n} \phi_{mnk}. \quad (26)$$

Following the relationship $\ln q(\tau_k) = \ln q(\mu_k, \tau_k) - \ln q(\mu_k|\tau_k)$, it is straightforward to obtain the gamma distribution of $q(\tau_k)$with shape and rate parameters

$$a_k = a_0 + \frac{1}{2} \sum_{m,n} \phi_{mnk} \quad (27)$$

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 20,2022 at 09:31:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

$$b_k = b_0 + \frac{1}{2}\Bigg[\sum_{m,n}\phi_{mnk}(y_{mn} - \mathbf{u}_m\mathbf{v}_n^T)^2 + \beta_0\mu_0^2 - \frac{1}{\beta_k}\Bigg(\beta_0\mu_0 + \sum_{m,n}\phi_{mnk}\left(y_{mn} - \langle\mathbf{u}_m\rangle\langle\mathbf{v}_n^T\rangle\right)^2\Bigg]. \tag{28}$$

The posterior of the indicator still follows a multinominal distribution

$$q(\mathbf{z}_{mn}) = \prod_{k=1}^{K}\phi_{mnk}^{z_{mnk}} \tag{29}$$

with the parameter

$$\phi_{mnk} = \frac{\varphi_{mnk}}{\sum_k \varphi_{mnk}} \tag{30}$$

$$\varphi_{mnk} = \langle ln\theta_k\rangle + \frac{1}{2}\langle\ln\tau_k\rangle - \frac{1}{2}\ln 2\pi\left\langle(y_{mn} - \mathbf{u}_m\mathbf{v}_n^T - \mu_k)^2\right\rangle - \frac{1}{2}\langle\tau_k\rangle. \tag{31}$$

The mixing proportion $\boldsymbol{\theta}$ obeys a Dirichlet posterior distribution

$$q(\boldsymbol{\theta}) = \mathrm{Dir}(\boldsymbol{\theta}|\boldsymbol{\alpha}) \tag{32}$$

with $\boldsymbol{\alpha} = (\alpha_1, \cdots \alpha_K)$ updated by

$$\alpha_k = \alpha_0 + \sum_{m,n}\langle z_{mnk}\rangle. \tag{33}$$

Each parameter in (15)–(33) is updated in turn while others remain fixed. As a deterministic posterior approximation method, the evidence lower bound of the model converging to a local minimum can be guaranteed following standard mean-field variational inference [16].

D. Merging Gaussian Components

To estimate the Gaussian mixture density for the noise, we first set *K* to a relatively large value and prune the number of required Gaussians during each iteration. We define a similarity measure as the arithmetic average of KL divergence between a pair of Gaussians. Other measurements, for example, resistor average or geometrical average, can also be considered. The resulting similarity is

$$S(p_{\theta_i}||p_{\theta_j}) = (\mathrm{KL}(p_{\theta_i}||p_{\theta_j}) + \mathrm{KL}(p_{\theta_j}||p_{\theta_i}))/2 \tag{34}$$

where $i, j \in [1, K]$. The KL divergence between two Gaussians is estimated from their means and precisions [16]

$$\mathrm{KL}(p_{\theta_i}||p_{\theta_j}) = \frac{1}{2}\left(\ln\frac{\tau_i}{\tau_j} + \frac{\tau_j}{\tau_i} + \tau_j(\mu_j - \mu_i)^2 - 1\right). \tag{35}$$

If the similarity $S(p_{\theta_i}||p_{\theta_j})$of the *i*th and *j*th Gaussian components is smaller than a preset threshold, we merge these two components into a new Gaussian with mean $\mu_{ij}$and precision τ_{ij} given by

$$\mu_{ij} = \frac{\theta_i\mu_i + \theta_j\mu_j}{\theta_i + \theta_j} \tag{36}$$

TABLE I
NOISE DISTRIBUTIONS AND PARAMETERS CONSIDERED IN THE SYNTHETIC DATA

Distribution	Parameters
Gaussian	$\mathcal{N}(0,0.01^2)$
Laplace	$\mathcal{L}(0,0.01)$
Gamma	$\mathcal{G}(20,\mathbf{UV}/20)$
Zero-mean mixture	20% ~ $\mathcal{U}(-0.01,0.01)$, 20% ~ $\mathcal{N}(0,0.01^2)$ 60% ~ $\mathcal{N}(0,0.015^2)$
Nonzero-mean mixture	20% ~ $\mathcal{U}(-0.01,0.01)$, 20% ~ $\mathcal{N}(0.01,0.01^2)$ 60% ~ $\mathcal{N}(-0.005,0.005^2)$

$$\tau_{ij} = \frac{\theta_i + \theta_j}{\theta_i/\tau_i + \theta_j/\tau_j}. \tag{37}$$

This merged Gaussian and other current parameters, including other Gaussians, are fed to the model as the initialization to update the parameters in the next round. The iteration is terminated when convergence criteria are met or the predefined number of iterations is reached.

IV. EXPERIMENT AND EVALUATION INDEX

A. Synthetic Data

We first carry out a comparative experiment on synthetic datasets with known statistical properties. The entries of **U** and **V** are drawn from a truncated Gaussian distribution with unit variance, which is roughly similar to the distributions observed in a previous analysis of experimental MES [20]. We set the rank $r = 4$ for both **U** and **V** since this number of primitives is frequently seen in some common human hand motions and animal locomotion models [2], [20]. The product **UV** is assumed to be the ground-truth **X**. Table I lists the typical noise distributions encountered in MES as **E**, i.e., Gaussian, Laplace, gamma, and mixed noise with zero-mean and nonzero-mean. For the two types of mixed noise, the sparse noise matrix is first simulated such that 20% of its entries are located randomly and the amplitudes are uniformly sampled with a range of [–0.01, 0.01] [36], [37]. 20% and 60% of the remaining entries are drawn from two different Gaussian distributions, respectively (see Table I).

We assess the performance of the model to extract the original primitives and activation coefficients. The similarity between extracted and ground-truth primitives or activation coefficients is defined as the average of the correlation coefficients between each pair of the matched primitives or activation coefficients [2], [20]. The baseline similarity of the extracted primitives and activation coefficients to each of 50 randomly generated counterparts is adopted to normalize the initial similarity [2], [20]

$$d_{ns} = (d_s - d_b)/(1 - d_b) \tag{38}$$

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 20,2022 at 09:31:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

where d_{ns} is the normalized similarity, and $d_s$and d_b are the nonnormalized and baseline similarity, respectively. The performance of each algorithm varies from 0 for a chance level of similarity to 1 for maximum similarity.

B. Experimental Myoelectric Recording

We use the publicly available Ninapro DB1 database, which contains the MESs of multiple upper limb movements [34]. Each motion consists of 10 repetitions performed by each of 27 healthy subjects. The muscular activity is recorded using 10 active double-differential OttoBock *MyoBock* 13E200 electrodes at 100 Hz. More details on acquisition setup and protocol can be found in [38] and [39].

We use the classification accuracy of primitives to recognize six wrist motions and the robustness of primitives across repetitions of MES recordings as the performance metrics due to the lack of ground-truth primitives for the experimental data. Here, the robustness is quantified by the cross-variance accounted for (CVAF) defined by [39]

$$\text{CVAF}^{i,j} = 1 - \frac{\sum_{c=1}^{m} \left(\mathbf{Y}_{cn}^{i} - \mathbf{X}_{cn}^{j}\right)^2}{\sum_{c=1}^{m} \left(\mathbf{Y}_{cn}^{i}\right)^2} \tag{39}$$

where $\mathbf{Y}_{cn}^{i}$ and $\mathbf{X}_{cn}^{j}$ are the original and reconstructed MES of the channel c for the ith and jth repetition, respectively. CVAF quantifies the degree of the primitives identified from the jth repetition that can reconstruct the MES that fall into the ith repetition [39]. It is expected that high CVAF can result in high classification accuracy. The sparseness measure based on the L1 and L2 norm is employed to quantify the sparsity of the movement primitives and activation coefficients [21]

$$\text{sparseness}(\mathbf{d}) = \frac{\sqrt{l} - \left(\sum |d_i|\right) \Big/ \sqrt{\sum d_i^2}}{\sqrt{l} - 1} \tag{40}$$

where l is the dimensionality of vector $\boldsymbol{d}$. This function equals unity if and only if d contains only a single nonzero component. It takes a value of zero if and only if all components are equal [21].

V. Results

We carry out comprehensive numerical experiments on both the simulated data and real MESs to assess the effectiveness and robustness of GMNMF in extraction of latent factors. We compare GMNMF with four representative baseline methods widely used in movement primitive extraction, i.e., MUNMF [18]–[21], L1NMF [22], [23], L21NMF [24], [25], and VBNMF [17], [40]. MUNMF is the vanilla NMF implemented by minimizing the mean-squared error while VBNMF implements the same model through VB update rules. L1NMF and L21NMF substitute the L0 norm for the L1 and L21 norms into the objective function, respectively, for a more robust fit of outliers and noise. The initial rank is set to be the minimum of m and n. We initialize $\mu_0$to 0, K to 30, and all other hyperparameters to 10^{-6}.

A. Results on Synthetic Data

The rank r in the three multiplicative updates based algorithms and the other VB model need to be preset to the ground-truth value since these algorithms are not able to iteratively prune the rank. The resulting mean and standard deviation of d_{ns} over 500 realizations for the five candidate algorithms based on the five different noise types is shown in Fig. 2. The left bar on each subpanel corresponds to the normalized similarity of primitives while the right bar corresponds to that of activation coefficients. For each subbar chart, the bars from left to right correspond to MUNMF, L1NMF, L2NMF, VBNMF, and GMNMF, respectively. It is apparent that MUNMF and VBNMF yield higher normalized similarity than L1NMF and L21NMF under Gaussian noise contamination. GMNMF achieves comparable performance to MUNMF and VBNMF even without the given ground-truth rank. For the two heavy-tailed noise examples, i.e., Laplace and gamma distribution, it is apparent that the normalized similarity for MUNMF and VBNMF degrades while the normalized similarity for L1NMF and L21NMF rises. In both cases, GMNMF still behaves well with d_{ns} as high as in the case of Gaussian noise. This is reasonable since only GMNMF can fit different noise types. As expected, we observe that GMNMF yields the best performance with significantly higher similarity for the two mixed noise settings. L21NMF is relatively better than the other three models, which is due to the tradeoff of both L1 and L2 norms used to define the L21NMF's cost function.

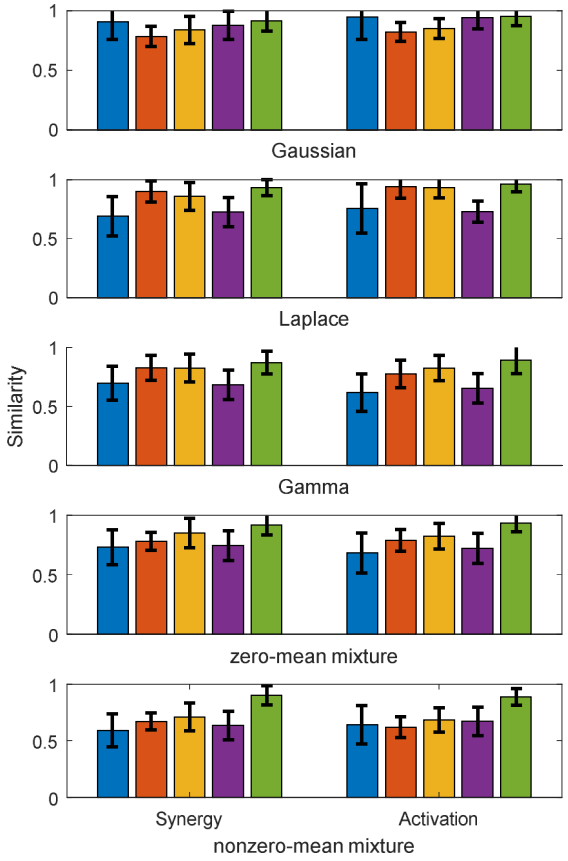


Fig. 2. Mean/standard deviation of normalized similarity for five algorithms on five different noise types is shown. For each subbar chart, bars from left to right correspond to MUNMF, L1NMF, L2NMF, VBNMF, and GMNMF, respectively.

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 20,2022 at 09:31:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

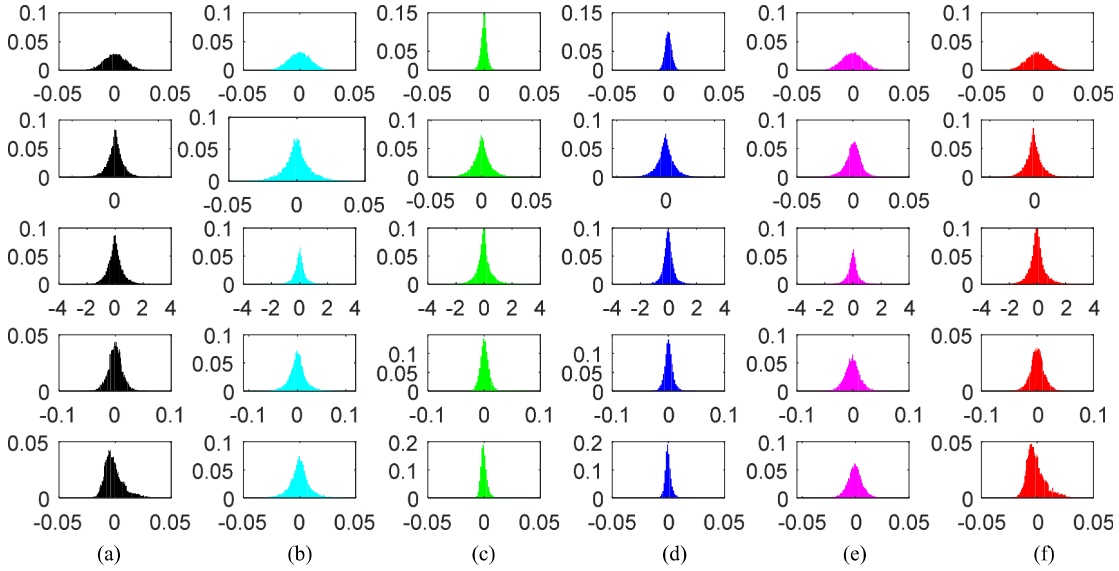


Fig. 3. Normalized histograms of noise **E** fitted by the five NMF models. From the top to bottom, each row corresponds to 1) Gaussian noise, 2) Laplace noise, 3) gamma noise, 4) mixed noise with zero mean, and 5) mixed noise with nonzero mean. (a) Ground truth. (b) MUNMF. (c) L1NMF. (d) L21NMF. (e) VBNMF. (f) GMNMF.

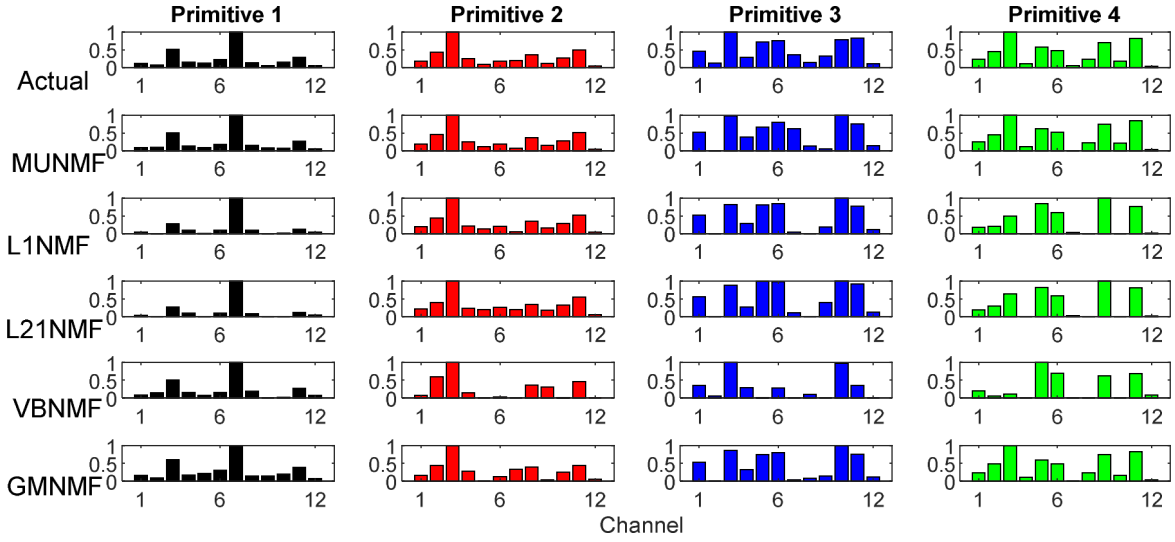


Fig. 4. Actual movement primitives (first row) and primitives extracted by GMNM and four baseline models.

To further compare the performance of the five models, we compare the normalized histograms of the ground truth noise and the residual $\boldsymbol{E}$ obtained from each model (see Fig. 3). Visual inspection reveals that GMNMF fits the noise well regardless of the type of noise. However, the profile of the other four algorithms matches that of the ground truth only when the underlying and actual noise distributions are related. This is consistent with the results of normalized similarity in Fig. 2.

Fig. 4 shows an actual primitive sample and the primitives (P1–P4) extracted by the five algorithms under nonzero-mean mixture noise contamination. Table II shows the normalized similarity between the extracted primitives and the actual values in Fig. 4. MUNMF and VBNMF recover P1 well with normalized similarity 0.9645 and 0.9322, respectively. Errors are relatively large for the other three primitives with low d_{ns} values. However, L1NMF and L21NMF reconstruct P2 well,

TABLE II
NORMALIZED SIMILARITY d_{ns} OF FOUR PRIMITIVES IN FIG. 4 AND THEIR AVERAGE VALUE FOR FIVE ALGORITHMS

d_{ns}	P1	P2	P3	P4	Average
MUNMF	0.9322	0.8571	**0.8838**	0.9113	0.8961
L1NMF	0.7525	**0.8848**	0.8667	0.8811	0.8462
L21NMF	0.7386	0.8786	0.8719	0.8431	0.8277
VBNMF	0.8864	0.8283	0.8236	0.7925	0.8327
GMNMF	**0.9645**	0.8619	0.8797	**0.9599**	**0.9165**

The bold entity in each column indicates the maximal value.

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 20,2022 at 09:31:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

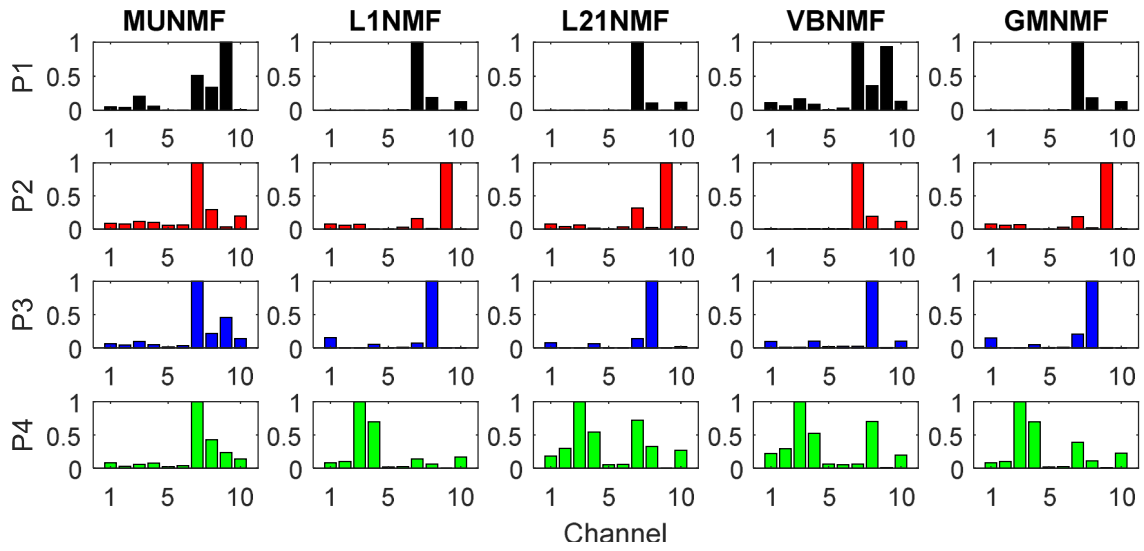


Fig. 5. Four movement primitives (P1–P4) identified by MUNMF, L1NMF, L21NMF, VBNMF, and GMNMF for a typical ten-channel MES to perform wrist flexion.

which possess the top two d_{ns} values for P2. As indicated in Table II, GMNMF has the highest average normalized similarity value 0.9165 to match all four actual primitives. In particular, P1 and P4 extracted by GMNMF possess the two largest d_{ns} values over all algorithms. This confirms the superiority of the GMNMF model over the four baseline models for extracting the primitives from the observed data contaminated by complex noise.

B. Results on Experimental Data

GMNMF and four competitive models are further evaluated to identify hand movements from the extracted primitives on experimental MES recordings, intended for cybernetic control in robotics. The Ninapro dataset is split into a training set with six repetitions of each motion and *a* testing set with four repetitions *for* each subject. We apply the five algorithms to extract primitive matrices from the training repetitions of each task independently. A support vector machine is then trained with the primitive feature to identify six wrist actions: wrist supination and pronation, wrist ulnar and radial deviation, and wrist extension and flexion. As noted earlier, GMNMF prunes the rank iteratively to reach the optimal number of primitives. However, the required number of primitives had to be manually set to four for the remaining methods [20]. The remaining four repetitions of each task are used to validate the performance of the classifiers. The training and test sets are randomly constructed ten times to attain an average of the classification performance.

The KL divergence similarity threshold plays an important role in GMNMF model fitting. If the threshold is too large, all initial Gaussians are merged into one. It is easy to see that this model is equivalent to using L2 loss, which is obviously less flexible with large error. On the other hand, if the threshold is too small and the number of initial Gaussians is large, in the extreme each entry of matrix $\mathbf{Y}$ will have its own initial Gaussian to represent the noise without any merging. This offers too much flexibility for noise modeling since it is rarely the case for any noise to affect only a single data entry. We have compared the performance of GMNMF with the threshold varying from 0.005 to 0.1. There is no substantial difference for the classification accuracy and robustness with the threshold lower than 0.05. Since the model is relatively slow to converge with a small threshold, we thus chose the threshold value 0.05 for all reported results.

Figs. 5 and 6 show the primitives and activation coefficients of a typical wrist flexion movement obtained by the five NMF algorithms. Tables III and IV provide summaries of the sparseness index for the primitives in Fig. 5 and activation coefficients in Fig. 6, respectively. More zero values in each column of Fig. 5

TABLE III
SPARSENESS VALUE FOR EACH PRIMITIVE EXTRACTED BY FIVE ALGORITHMS IN FIG. 5

Method	P1	P2	P3	P4	Average
MUNMF	0.53537	0.6287	0.6678	0.6150	0.6117
L1NMF	0.4607	**0.3631**	0.4692	0.3150	0.4020
L21NMF	0.4616	0.4675	0.4274	0.4293	0.4464
VBNMF	0.5014	0.4293	0.4686	0.4568	0.4640
GMNMF	**0.3289**	0.4145	**0.4191**	**0.2282**	**0.3476**

The bold entity in each column indicates the maximal value.

TABLE IV
SPARSENESS VALUE FOR EACH ACTIVATION EXTRACTED BY FIVE ALGORITHMS IN FIG. 6

Method	A1	A2	A3	A4	Average
MUNMF	0.6536	0.1838	0.5524	0.4358	0.4564
L1NMF	0.0932	0.4750	0.3089	0.4665	0.3359
L21NMF	0.0684	0.4781	0.2429	0.5707	0.3400
VBNMF	0.4748	**0.0849**	0.3017	0.4667	0.3320
GMNMF	**0.0383**	0.4357	**0.2400**	**0.3410**	**0.2637**

The bold entity in each column indicates the maximal value.

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 20,2022 at 09:31:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

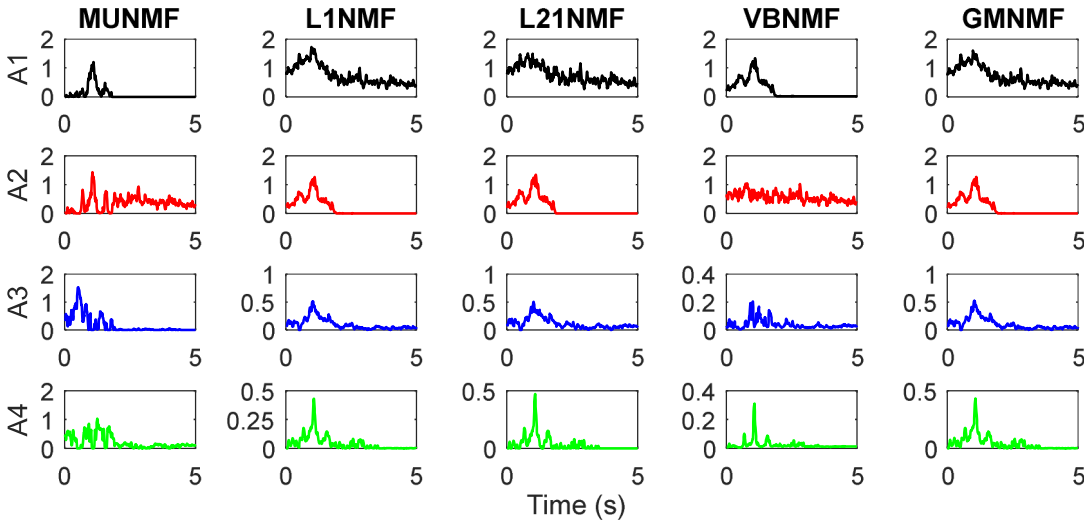


Fig. 6. Four neural activation coefficients (A1–A4) corresponding to the movement primitives in Fig. 5 identified by MUNMF, L1NMF, L21NMF, VBNMF, and GMNMF for a typical ten-channel MES to perform wrist flexion.

TABLE V
CLASSIFICATION ACCURACY AND ROBUSTNESS OF FIVE MODELS TO CLASSIFY SIX WRIST MOVEMENTS ACROSS 27 SUBJECTS

Method	Accuracy (%)	Robustness (%)
MUNMF	76.68 ± 2.56	74.86 ± 6.86
L1NMF	81.49 ± 3.48	77.82 ± 8.21
L21NMF	83.15 ± 3.05	78.19 ± 5.06
VBNMF	79.88 ± 2.72	75.75 ± 4.55
GMNMF	88.25 ± 2.55	83.44 ± 3.57

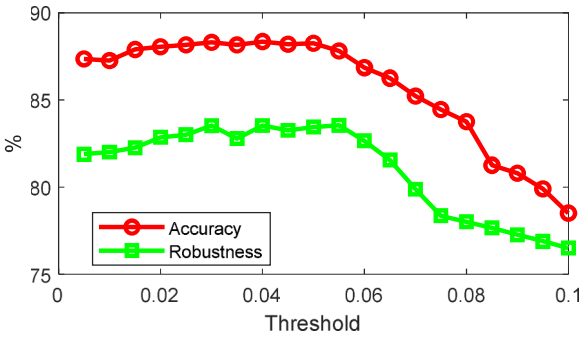


Fig. 7. Effect of KL divergence threshold on classification accuracy and robustness.

or low sparseness values in Table III indicate that fewer active channels are needed in each primitive, highlighting the sparseness of primitives. From Fig. 5, primitives extracted by GMNMF as well as L1NMF and L21NMF indicate this property, i.e., each action requiring a limited number of DOFs [41], whereas MUNMF and VBNMF appear to be inferior to indicate this neural coding mechanism with relatively high average sparsity values. Furthermore, the sparseness of the neural activation scheme, i.e., a sparse combination of primitives, can be observed in most of all methods in Fig. 6 while GMNMF possesses the lowest average sparsity value 0.2637. This result demonstrates that GMNMF is better able to fit the experimental MES with low sparsity values, as shown in Tables III and IV. The L1 norm or the corresponding Laplace distribution can be expressed as an integrated Gaussian mixture with mixing distribution equal to the exponential distribution. This means that each entry in the matrix has its own GMM to represent the noise. On the other hand, L2 loss is equivalent to using a single Gaussian to fit the noise. The GMNMF model can be viewed as a tradeoff between L1 and L2 loss, which shares the L1 property to promote sparseness. The second column of Table V is a summary of the classification results for five models with GMNMF achieving the highest average accuracy of 88.25%. The third column shows the robustness index CVAF for the five models. It is not surprising that CVAF is correlated to classification accuracy with GMNMF achieving the highest average CVAF. Fig. 7 shows the effect of the KL divergence similarity $S(p_{\theta_i}||p_{\theta_j})$ threshold varying from 0.005 to 0.1 with steps of 0.005 on the classification accuracy and robustness. For values from 0.005 to 0.05, the performance for both measures is relatively stable. However, when the threshold exceeds 0.05, both performance metrics degrade slightly, but still comparable to the four baseline algorithms when the threshold is less than 0.1.

In order to further investigate the model performance and clarify the superiority of GMNMF over the competitive methods, we restore the dynamic muscle activities using the generated primitives and activation coefficients. Fig. 8 shows the short-time Fourier transform spectrograms of the reconstructed MES corresponding to channel 3 of Figs. 5 and 6. There is no muscle activity at left and right side of each subplot, where noise is the major power source. In particular, Gaussian white noise spans from normalized frequency 0 to 1 while sparse noise and movement artifacts are mainly located in low frequency. Compared with the same area of other subplots, more yellow dots within the two red frames on GMNMF subplot indicate more Gaussian white noise attenuated by GMNMF.

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 20,2022 at 09:31:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

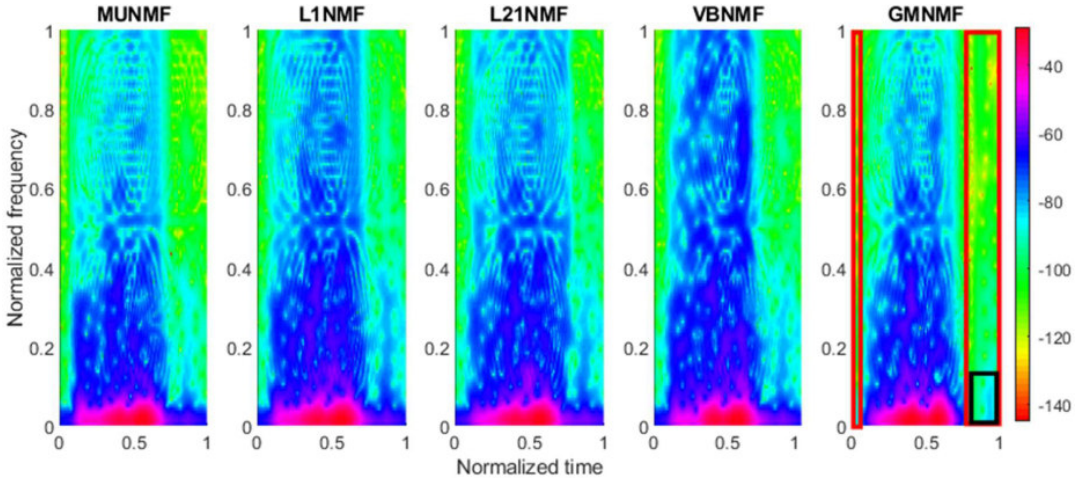


Fig. 8. Time-frequency distribution of restored MES of channel 3 in Figs. 5 and 6 by five algorithms.

Furthermore, we find that the complex noise components, including the movement artifacts and sparse noise, highlighted in the black frame, are remarkably compressed by GMNMF with light blue color within the black frame. MUNMF and VBNMF achieve a similar Gaussian noise reduction effect compared to GMNMF, which is evidenced by the green block and yellow dots before and after the major muscle activation. However, their effect to remove movement artifacts and sparse noise is suspicious since these noise components still exist at left and right bottom of both spectrograms in deep blue color. Regarding L1NMF and L21NMF, noise components still occupy the right part of two spectrograms shown in blue and green colors. The subtle details of the time-frequency distribution in these spectrograms further illustrate the superiority of GMNMF to fit the experimental MESs containing multiple noise ingredients. It should be noted that the rank r for this dataset varies from 3 to 7 with most movements requiring four primitives in the GMNMF model. The required number of Gaussians needed to fit the number varies from 4 to 10. In comparison with the Gaussian or sparse noise assumption in other models, we conclude that the improved classification accuracy and robustness come from the improved noise and rank fitting of GMNMF.

VI. Conclusion

The brain coordinates a group of muscles through muscle synergies or motor primitives to generate the desired movement. This has been broadly applied in robotics, brain science, clinical diagnosis, human–machine interface, and rehabilitation engineering. However, there is a significant gap between such important applications and the availability of sophisticated model to extract movement primitives. Here, we develop a generative NMF model by employing a finite GMM to fit complex noise in the MESs within a VB framework. To determine the required number of Gaussians to fit the observed data, we define a similarity measure based on KL divergence to iteratively merge mixture components. In comparison with DPMs, our new KL divergence based criterion avoids overestimation that can degrade the performance, and the model is relatively simpler with fewer parameters. We have demonstrated the high fidelity of GMNMF to identify latent variables from the simulated data with five different noise distributions. We also evaluated the ability of GMNMF model to classify six wrist movements and showed that it outperforms five baseline models in terms of accuracy and robustness. In future work, we will embed GMNMF into a robot system or other multiple DOFs myoelectric interfaces to assess its actual performance for robust simultaneous control. Another extension is to modify the current model to an online version to improve motor learning from streaming MESs. Although we limit the application of GMNMF to the specific robot movement primitive problem in this study, the proposed VB framework of GMNMF can be applied to a broader range of problems in parameter estimation, system identification and prediction, and intelligent control.

Appendix

Here, we provide the details of VB inference for the GMNMF model.

Infer the entry u_{mr} and v_{nr}

$$\begin{aligned}
\ln q(u_{mr}) &\propto \mathbb{E}_q[\ln p(\mathbf{Y}|\mathbf{U},\mathbf{V},\boldsymbol{\mathcal{Z}},\boldsymbol{\mu},\boldsymbol{\tau}) + \ln p(u_{mr}|\lambda_r)] \\
&\propto \mathbb{E}_q\Bigg[\sum_n\sum_{k=1}^{K} p(z_{mn}=k) \\
&\ln\left[\sqrt{\frac{\tau_k}{2\pi}}\exp\left\{-\frac{\tau_k}{2}(y_{mn}-\mathbf{u}_m\mathbf{v}_n^T-\mu_k)^2\right\}\right] \\
&\quad + \ln\left[\sqrt{\frac{\lambda_r}{2\pi}}\exp\left\{-\frac{\lambda_r}{2}u_{mr}^2\right\}\right] \times h(u) \\
&\propto \mathbb{E}_q\left[\sum_n\sum_{k=1}^{K} p(z_{mn}=k)\left\{-\frac{\tau_k}{2}(y_{mn}-\mathbf{u}_m\mathbf{v}_n^T-\mu_k)^2\right\}\right] \\
&\quad - \frac{\lambda_r}{2}u_{mr}^2\Bigg\}\Bigg] \times h(u) \\
&\propto \left\{-\frac{u_{mr}^2}{2}\left(\sum_n\sum_{k=1}^{K}\phi_{mnk}\langle\tau_k\rangle\langle v_{nr}^2\rangle + \langle\lambda_r\rangle\right)\right.
\end{aligned}$$

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 20,2022 at 09:31:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

$$+ u_{mr} \sum_{n} \sum_{k=1}^{K} \phi_{mnk} \langle \tau_k \rangle \langle v_{nr} \rangle$$

$$\bullet \left(y_{mn} - \sum_{r' \neq r} \langle u_{mr'} \rangle \langle v_{nr'} \rangle - \langle \mu_k \rangle \right) \Bigg\} \times h(u)$$

$$\propto \left\{ -\frac{\tau_{mr}^U}{2} (u_{mr} - \mu_{mr}^U)^2 \right\} \times h(u).$$

One can find $q(u_{mr})$obeys a truncated normal distribution

$$q(u_{mr}) \propto \mathcal{TN}\left(u_{mr} | \mu_{mr}^U, \tau_{mr}^U\right)$$

with parameters

$$\tau_{mr}^U = \sum_{n} \sum_{k=1}^{K} \phi_{mnk} \langle \tau_k \rangle \langle v_{nr}^2 \rangle + \langle \lambda_r \rangle$$

$$\mu_{mr}^U = \frac{1}{\tau_{mr}^U} \sum_{n} \sum_{k=1}^{K} \phi_{mnk} \langle \tau_k \rangle \langle v_{nr} \rangle \times \left(y_{mn} - \sum_{r' \neq r} \langle u_{mr'} \rangle \langle v_{nr'} \rangle - \langle \mu_k \rangle \right)$$

and the required expectation and variance are

$$\langle u_{mr} \rangle = \mu_{mr}^U + \frac{1}{\tau_{mr}^U} f\left(-\mu_{mr}^U \sqrt{\tau_{mr}^U} \right)$$

$$\mathrm{var}[u_{mr}] = \frac{1}{\tau_{mr}^U} \left(1 - g\left(-\mu_{mr}^U \sqrt{\tau_{mr}^U} \right) \right)$$

where $f(u_{mr}) = \psi(u_{mr})/(1 - \Psi(u_{mr}))$, and $g(u_{mr}) = f(u_{mr})(f(u_{mr}) - u_{mr}))$. $\Psi(\bullet)$and $\psi(\bullet)$ are the cumulative distribution function and probability density function of a standard $\mathcal{N}(0, 1)$, respectively.

Similarly, $q(v_{nr}) \propto \mathcal{TN}(v_{nr} | \mu_{nr}^V, \tau_{nr}^V)$can be reached.

Infer λ_r

$$\ln q(\lambda_r) \propto \mathbb{E}_q \left(\sum_{m=1}^{M} \log p(u_{mr} | \lambda_r) \right) + \mathbb{E}_q \left(\sum_{n=1}^{N} \log p(v_{nr} | \lambda_r) \right) + \mathbb{E}_q (\log p(\lambda_r | a_1, b_1)$$

$$\propto \exp \left\{ \left(a_1 + \frac{M+N}{2} - 1 \right) \log \lambda_r - \left(b_1 + \frac{1}{2} \sum_{m=1}^{M} u_{mr}^2 + \frac{1}{2} \sum_{n=1}^{N} v_{nr}^2 \right) \lambda_r \right\}$$

which means

$$q(\lambda_r) = \mathcal{G}(\lambda_r | a^*, b^*)$$

with $a^* = a_1 + \frac{M+N}{2}$, and $b^* = b_1 + \frac{1}{2} \sum_{m=1}^{M} u_{mr}^2 + \frac{1}{2} \sum_{n=1}^{N} v_{nr}^2$.

Infer μ_k, τ_k

$$\ln q(\mu_k) \propto \sum\nolimits_{m,n} q(z_{mn} = k) \mathbb{E}_q[\ln p(y_{mn} | \mathbf{U}, \mathbf{V}, \mathbf{z}, \boldsymbol{\mu}, \boldsymbol{\tau})] + \mathbb{E}_q[\ln p(\mu_k | \mu_0, \beta_0 \tau_k)]$$

$$\propto \left\{ -\frac{\mu_k^2}{2} \tau_k \left(\beta_0 + \sum_{m,n} q(z_{mn} = k) \right) + \mu_k \tau_k \left(\beta_0 \mu_0 + \sum_{m,n} q(z_{mn} = k) \left(y_{mn} - \langle \mathbf{u}_m \rangle \left\langle \mathbf{v}_n^T \right\rangle \right) \right) \right\}.$$

This is a Gaussian distribution with

$$\beta_k = \beta_0 + \sum_{m,n} \phi_{mnk}$$

$$m_k = \frac{1}{\beta_k} \left(\beta_0 \mu_0 + \sum_{m,n} \phi_{mnk} \left(y_{mn} - \langle \mathbf{u}_m \rangle \left\langle \mathbf{v}_n^T \right\rangle \right) \right.$$

Since $\ln q(\tau_k) = \ln q(\mu_k, \tau_k) - \ln q(\mu_k | \tau_k)$, we have

$$\ln q(\tau_k) \propto \left\{ \sum_{m,n} q(z_{mn} = k) \mathbb{E}_q[\ln p(y_{mn} | \mathbf{U}, \mathbf{V}, \mathbf{z}, \boldsymbol{\mu}, \boldsymbol{\tau})] + \mathbb{E}_q[\ln p(\mu_k | \mu_0, \beta_0 \tau_k)] \right\} + \mathbb{E}_q[\ln p(\tau_k | a_0, b_0)]$$

$$\propto \left[\frac{1}{2} \beta_k \tau_k (\mu_k - m_k)^2 + \ln \tau_k - \ln q(\mu_k | \tau_k) \right] + \frac{1}{2} \tau_k \left(2b_0 + \sum_{m,n} \phi_{mnk} \left(y_{mn} - \mathbf{u}_m \mathbf{v}_n^T \right)^2 + \beta_0 \mu_0^2 - \beta_k m_k^2 \right) + \left(a_0 + \frac{1}{2} \sum_{m,n} \phi_{mnk} - 1 \right) ln \tau_k$$

τ_k obeys a gamma distribution with parameters

$$a_k = a_0 + \frac{1}{2} \sum_{m,n} \phi_{mnk}$$

$$b_k = b_0 + \frac{1}{2} \left[\sum_{m,n} \phi_{mnk} (y_{mn} - \mathbf{u}_m \mathbf{v}_n^T)^2 + \beta_0 \mu_0^2 - \frac{1}{\beta_k} \left(\beta_0 \mu_0 + \sum_{m,n} \phi_{mnk} \left(y_{mn} - \langle \mathbf{u}_m \rangle \left\langle \mathbf{v}_n^T \right\rangle \right)^2 \right] \right.$$

The expectations are estimated by

$$\langle \mu_k \rangle = m_k$$

$$\left\langle \mu_k^2 \right\rangle = (\beta_k \tau_k)^{-1} + m_k^2$$

$$\langle \tau_k \rangle = a_k / b_k$$

$$\langle \ln \tau_k \rangle = \gamma(a_k) - \ln b_k$$

where $\gamma(\bullet)$is the digamma function.

Infer $\mathbf{z}$

$$\ln q(\mathbf{z}) \propto \mathbb{E}_q[\ln p(\mathbf{z} | \boldsymbol{\theta}) + \ln p(\mathbf{Y} | \mathbf{U}, \mathbf{V}, \boldsymbol{\mathcal{Z}}, \boldsymbol{\mu}, \boldsymbol{\tau})]$$

$$\propto \sum_{m,n} \sum_{k=1}^{K} z_{mnk} [\mathbb{E}[ln \theta_k] + \frac{1}{2} \mathbb{E}[\ln \tau_k] - \frac{1}{2} \ln 2\pi - \frac{1}{2} \mathbb{E} \left[(y_{mn} - \mathbf{u}_m \mathbf{v}_n^T - \mu_k)^2 \tau_k \right]$$

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 20,2022 at 09:31:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

$$q(\mathbf{z}_{mn}) = \prod_{k=1}^{K} \phi_{mnk}^{z_{mnk}}$$

with $\phi_{mnk} = \frac{\varphi_{mnk}}{\sum_k \varphi_{mnk}}$

$$\varphi_{mnk} = \langle ln\theta_k \rangle + \frac{1}{2}\langle \ln \tau_k \rangle - \frac{1}{2}\ln 2\pi \left\langle (y_{mn} - \mathbf{u}_m \mathbf{v}_n^T - \mu_k)^2 \right\rangle - \frac{1}{2}\langle \tau_k \rangle$$

and the expectation $\langle z_{mnk} \rangle = \phi_{mnk}$.

Infer $\boldsymbol{\theta}$

$$p(\mathbf{z}_{mn}) = Multi(z_{mn}|\boldsymbol{\theta}) = \prod_{k=1}^{K} \theta_k^{z_{mnk}}$$

$$p(\boldsymbol{\theta}) = \mathcal{D}ir(\boldsymbol{\theta}|\boldsymbol{\alpha}_0) = \frac{\prod_{k=1}^{K}\Gamma(\alpha_{0k})}{\Gamma\left(\sum_{k=1}^{K}\alpha_{0k}\right)}\theta_k^{\alpha_0 - 1}$$

$$\ln q(\boldsymbol{\theta}) \propto \mathbb{E}_{\mathbf{z}}[\ln p(\mathbf{z}|\boldsymbol{\theta})] + \ln p(\boldsymbol{\theta}) \propto \sum_{k=1}^{K} \ln \theta_k^{\sum_{m,n}\langle z_{mnk}\rangle + \alpha_0 - 1} + \text{const}$$

$$q(\boldsymbol{\theta}) = \text{Dir}(\boldsymbol{\theta}|\boldsymbol{\alpha})$$

with $\boldsymbol{\alpha} = (\alpha_1, \cdots \alpha_K)$

$$\alpha_k = \alpha_0 + \sum_{m,n} \langle z_{mnk} \rangle$$

and the expectation

$$\langle \theta_k \rangle = \gamma(\alpha_k) - \gamma(\bar{\alpha}), \bar{\alpha} = \sum_{k=1}^{K} \alpha_k.$$

REFERENCES

[1] I. Riccardo, M. Francesca, S. Marianna, L. Matteo, and M. Maddalena, "Perspectives and challenges in robotic neurorehabilitation," *Appl. Sci.*, vol. 9, no. 15, pp. 3183-1–3183-29, 2019.

[2] V. C. K. Cheung and K. Seki, "Approaches to revealing the neural basis of muscle synergies: A review and a critique," *J. Neurophysiol.*, vol. 125, no. 5, pp. 1580–1597, 2021.

[3] A. Chiavenna, A. Scano, M. Malosio, L. T. Molinari, and F. Molteni, "Assessing user transparency with muscle synergies during exoskeleton-assisted movements," *Appl. Bionics Biomech.*, vol. 3, pp. 7647562-1–7647562-10, 2018.

[4] Y. Hu, K. Sirlantzis, G. Howells, N. Ragot, and P. Rodriguez, "An online background subtraction algorithm deployed on a NAO humanoid robot based monitoring system," *Robot. Auton. Syst.*, vol. 85, pp. 37–47, 2016.

[5] J. P. Eggert, S. Hellbach, A. Kolarow, E. Korner, and H. Gross, "Prediction and classification of motion trajectories using spatio-temporal NMF," in *Proc. 32nd Annu. German Conf. Artif. Intell.*, 2019, pp. 597–606.

[6] P. Kim, K. Kim, and S. Kim, "Modified nonnegative matrix factorization using the Hadamard product to estimate real-time continuous finger-motion intentions," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 6, pp. 1089–1099, Dec. 2017.

[7] M. Ison and P. Artemiadis, "The role of muscle synergies in myoelectric control: Trends and challenges for simultaneous multifunction control," *J. Neural Eng.*, vol. 11, no. 5, pp. 051001-1–051001-23, 2014.

[8] S. Muceli, N. Jiang, and D. Farina, "Extracting signals robust to electrode number and shift for online simultaneous and proportional myoelectric control by factorization algorithms," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 3, pp. 523–633, May 2014.

[9] M. Ghislieri, V. Agostini, and M. Knaflitz, "Muscle synergies extracted using principal activations: Improvement of robustness and interpretability," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 2, pp. 453–460, Feb. 2020.

[10] V. Maz'ya and G. Schmidt, "On approximate approximations using Gaussian kernels," *IMA J. Numer. Anal.*, vol. 16, no. 1, pp. 13–29, 1996.

[11] A. Jara, "Theory and computations for the Dirichlet process and related models: An overview," *Int. J. Approx. Reasoning*, vol. 81, no. 1, pp. 128–146, 2017.

[12] J. W. Miller and M. T. Harrison, "A simple example of Dirichlet process mixture inconsistency for the number of components," in *Proc. Adv. Neural Inf. Process. Syst.*, 2013, pp. 199–206.

[13] C. Li et al., "Bayesian nonnegative matrix factorization with Dirichlet process mixtures," *IEEE Trans. Signal Process.*, vol. 68, pp. 3860–3870, 2020.

[14] H. B. Xie, C. Li, R. Xu, and K. Mengersen, "Robust kernelized Bayesian matrix factorization for video background/foreground separation," in *Proc. 5th Int. Conf. Mach. Learn., Optim., Data Sci.*, 2019, pp. 484–495.

[15] C. Li et al., "Image denoising based on nonlocal Bayesian singular value thresholding and Stein's unbiased risk estimator," *IEEE Trans. Image Process.*, vol. 28, no. 10, pp. 4899–4911, Oct. 2019.

[16] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.

[17] J. L. Hinrich, K. H. Medsen, and M. Mørup, "The probabilistic tensor decomposition toolbox," *Mach. Learn., Sci. Technol.*, vol. 1, 2020, Art. no. 025011.

[18] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2001, pp. 556–562.

[19] C. Camardella, M. Junata, K. C. Tse, A. Frisoli, and R. K. Y. Tong, "How many muscles? Optimal muscles set search for optimizing myocontrol performance," *Front. Comput. Neurosci.*, vol. 15, 2021, Art. no. 668579.

[20] A. Ebied, E. Kinney-Lang, L. Spyrou, and J. Escudero, "Evaluation of matrix factorisation approaches for muscle synergy extraction," *Med. Eng. Phys.*, vol. 57, pp. 51–60, 2018.

[21] P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," *J. Mach. Learn. Res.*, vol. 5, pp. 1457–1469, 2004.

[22] B. Shen, B. D. Liu, Q. Wang, and R. Ji, "Robust nonnegative matrix factorization via L1 norm regularization by multiplicative updating rules," in *Proc. IEEE Int. Conf. Image Process.*, 2014, pp. 5282–5286.

[23] C. Lin, B. Wang, N. Jiang, and D. Farina, "Robust extraction of basis functions for simultaneous and proportional myoelectric control via sparse non-negative matrix factorization," *J. Neural Eng.*, vol. 15, no. 2, 2017, Art. no. 026017.

[24] D. Kong, C. Ding, and D. H. Huang, "Robust nonnegative matrix factorization using L21-norm," in *Proc. 20th ACM Conf. Inf. Knowl. Manage.*, 2011, pp. 673–682.

[25] A. Diaz and D. Steele, "Analysis of the robustness of NMF algorithms," 2021, *arXiv:2106.02213*.

[26] G. Zhou, Q. Zhao, Y. Zhang, T. Adalı, S. Xie, and A. Cichocki, "Linked component analysis from matrices to high-order tensors: Applications to biomedical data," *Proc. IEEE*, vol. 104, no. 2, pp. 310–331, Feb. 2016.

[27] A. T. Cemgil, "Bayesian inference for nonnegative matrix factorisation models," *Comput. Intell. Neurosci.*, vol. 2009, 2009, Art. no. 785152.

[28] K. Devarajan and V. C. Cheung, "On nonnegative matrix factorization algorithms for signal-dependent noise with application to electromyography data," *Neural Comput.*, vol. 26, no. 6, pp. 1128–1168, 2014.

[29] D. M. Endres, E. Chiovetto, and M. A. Giese, "Model selection for the extraction of movement primitives," *Front. Comput. Neurosci.*, vol. 7, 2013, Art. no. 185.

[30] D. Clever, M. Harant, H. Koch, K. Mombaur, and D. Endres, "A novel approach for the generation of complex humanoid walking sequences based on a combination of optimal control and learning of movement primitives," *Robot. Auton. Syst.*, vol. 83, pp. 287–298, 2016.

[31] S. Thongpanja, A. Phinyomark, C. Limsakul, and P. Phukpattaranont, "A robust measure of probability density function of various noises in electromyography (EMG) signal acquisition," in *Proc. 7th Int. Conf. Knowl. Smart Technol.*, 2015, pp. 19–24.

[32] I. Delis, S. Panzeri, T. Pozzo, and B. Berret, "Task-discriminative space-by-time factorization of muscle activity," *Front. Hum. Neurosci.*, vol. 9, 2015, Art. no. 399.

[33] J. Roh, W. Z. Rymer, E. J. Perreault, S. B. Yoo, and R. F. Beer, "Alterations in upper limb muscle synergy structure in chronic stroke survivors," *J. Neurophysiol.*, vol. 109, no. 3, pp. 768–781, 2013.

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 20,2022 at 09:31:47 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[34] P. M. Hilt, I. Delis, T. Pozzo, and B. Berret, "Space-by-time modular decomposition effectively describes whole-body muscle activity during upright reaching in various directions," *Front. Comput. Neurosci.*, vol. 12, pp. 201–2019, 2018.
[35] Q. Zhao, L. Zhang, and A. Cichocki, "Bayesian CP factorization of incomplete tensors with automatic rank determination," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 3, no. 9, pp. 1751–1763, Sep. 2015.
[36] H. Zhang, W. He, L. Zhang, H. Shen, and Q. Yuan, "Hyperspectral image restoration using low-rank matrix recovery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4729–4743, Aug. 2014.
[37] Y. Wang, C. Xu, X. C. Xu, and D. Tao, "Beyond RPCA: Flattening complex noise in the frequency domain," in *Proc. 31st AAAI Conf. Artif. Intell.*, 2017, pp. 2761–2767.
[38] M. Atzori et al., "Characterization of a benchmark database for myoelectric movement classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 1, pp. 73–83, Jan. 2015.
[39] M. Ghislieri, V. Agostini, and M. Knaflitz, "Muscle synergies extracted using principal activations: Improvement of robustness and interpretability," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 2, pp. 453–460, Feb. 2020.
[40] T. Brouwer, J. Frellsen, and P. Lio, "Comparative study of inference methods for Bayesian nonnegative matrix factorization," in *Lecture Notes in Computer Science*, vol. 10534, M. Ceci, J. Hollmen, L. Todorovski, C. Vens, and S. Dvzeroski, Eds. Cham, Switzerland: Springer, 2017, pp. 513–529.
[41] R. Prevete, F. Donnarumma, A. d'Avella, and G. Pezzulo, "Evidence for sparse synergies in grasping actions," *Sci. Rep.*, vol. 8, no. 1, pp. 616-1–616-16, 2018.



Hongbo Xie received the Ph.D. degree in biomedical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2006.

He has authored/coauthored more than 80 peer-reviewed journal and conference papers. His research interests include machine learning, signal and image processing, nonlinear time series analysis, and their applications in biomedical engineering.



Kerrie Mengersen received the Ph.D. degree in mathematical statistics from the University of New England, Armidale, Australia, in 1989.

She is currently a Full Professor with the School of Mathematical Sciences, Queensland University of Technology, Brisbane, QLD, Australia. In 2016, QUT awarded the title of Distinguished Professor to her in recognition of her outstanding achievements, both nationally and internationally, in mathematics and statistical research. She has more than 300 refereed journal publications, more than 40 keynote and invited international conference presentations, and attraction of more than 30 large research grants.

Prof. Mengersen was elected as a Fellow of the Australian Academy of Science in 2018. From 2011 to 2012, she was the National President of the Statistical Society of Australia, and from 2015 to 2017, the President of the International Society for Bayesian Analysis.



Changan Di received the Ph.D. degree in mechanical engineering from the Nanjing University of Science and Technology (NJUST), Nanjing, China, in 2003.

He is currently a Professor with NJUST. He is also the Director of the Instrument Science and Technology Experimental Center, NJUST. His research interests include measurement technology, human–computer interaction, and sensor technology.



Yongjian Zhang is currently working toward the Doctoral degree in mechanical engineering with the School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing, China.

His research interests include measurement technology, sensor technology, and signal processing.



Justin Lipman (Senior Member, IEEE) received the Ph.D. degree in telecommunications engineering from the University of Wollongong, Wollongong, NSW, Australia, in 2004.

He is currently an Industry Associate Professor with the University of Technology Sydney (UTS), Ultimo, NSW, Australia, and a Visiting Associate Professor with Hokkaido University's Graduate School of Engineering, Sapporo, Japan. He is also the Director of Research Translation with the Faculty of Engineering and IT and the Director of the RF Communications Technologies Laboratory, where he leads industry engagement in RF technologies, cybersecurity, privacy preserving technologies, Internet of Things, and tactile Internet. He serves as a committee member in Standards Australia, Sydney, NSW, Australia, contributing to International IoT standards and Digital Twins. He was the Deputy Chief Scientist with the Food Agility Cooperative Research Center, Sydney, NSW, Australia. Prior to joining UTS, over a 12 years period, he held a number of senior management and technical leadership roles with Intel and Alcatel driving research and innovation, product development, architecture, and IP generation. His research interests include "things" adaptive, connected, distributed, ubiquitous, and secure.



Sabine Van Huffel (Fellow, IEEE) bachelor's and doctoral degrees in electronic engineering from KU Leuven, in 1981 and 1987.

She has been a Full Professor with the Department of Electrical Engineering, KU Leuven, Leuven, Belgium, since 2002. She was a Guest Professor with Stanford University, Stanford, CA, USA, and with Uppsala University, Uppsala, Sweden, and a Visiting Fellow and Scientist with the University of Minnesota, Minneapolis, MN, USA. She has the following publications record: two monographs, about 304 articles in peer-reviewed international journals, about 257 conference papers, 4 edited books, 7 edited journal special issues, and 21 book chapters. Her research topics include fundamental/theoretical as well as application oriented and are performed in the domain of (multi)linear algebra, (non)linear signal analysis, classification and system identification with special focus to the development of numerically reliable and robust algorithms for improving medical diagnostics.

Prof. Van Huffel was elected as a Fellow of the Royal Flemish Academy of Belgium for Sciences and the Arts in 2017.

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 20,2022 at 09:31:47 UTC from IEEE Xplore. Restrictions apply.