Supporting Text

Synchronous and Time-Varying Synergy Models. We modeled the generation of a muscle pattern by the recruitment of N synchronous synergies as the linear combination of N nonnegative vectors $\{\boldsymbol{w_i}\}_{\boldsymbol{i}\,=\,1,\ldots,\,N}$

$$\mathbf{m}(t) = \sum_{i=1}^{N} c_i(t)\,\mathbf{w}_i \qquad \textbf{[1]}$$

where $\boldsymbol{m}(t)$ is a P-dimensional vector representing the activations of P muscles at time t; $c_i(t)$ is a nonnegative coefficient scaling the amplitude of the ith synergy at time t, and $\boldsymbol{w_i}$ is the ith muscle synergy. If we sample the muscle patterns at discrete time intervals, we can rewrite Eq. **1** as

$$\mathbf{M} = \mathbf{W}\,\mathbf{C}, \qquad \textbf{[2]}$$

where $\mathbf{M}$ has P rows and K columns (K total number of samples), $\mathbf{W}$ has P rows and N columns (N number of synergies), and $\mathbf{C}$ has N columns and K rows.

We modeled the generation of a muscle pattern by the recruitment of n instances of N time-varying synergies as the linear combination of nonnegative vectors $\mathbf{w}_i(\tau)$

$$\mathbf{m}(t) = \sum_{j=1}^{n} c_j\;\mathbf{w}_{I(j)}(t - t_j) \qquad \textbf{[3]}$$

where $\{\boldsymbol{w_i}(\tau)\}_{i\,=\,1,\ldots,\,N}$ is the ith synergy, i.e., a sequence of P-dimensional vectors representing the activations of P muscles at the time τ after the synergy onset, t_j is the synergy onset time, and c_j is a nonnegative coefficient scaling the synergy amplitude. Differently from our previous model (1), a given muscle pattern $\mathbf{m}(t)$ can be reconstructed by multiple instances of the same N synergies. The jth instance is simply a shifted version of the ith synergy, and we use an index function $i = I(j) \in [1, \ldots, N]$ to map instances into synergies. In discrete time, the ith time-varying synergy of duration T can be expressed as a matrix $\mathbf{W^i}$ with P rows and Q columns, each column representing the synergy activation vector at time t_q, $0 \le t_q < T$ ($q = 1, \ldots, Q$). We use the matrix $\mathbf{W} = [\mathbf{W}^1\ \mathbf{W}^2\ \ldots\ \mathbf{W}^N]$ with P rows and $Q \times N$ columns to compactly represent a set of N synergies. If we introduce a time-shifting matrix $\Theta_\mathbf{i}[k, K]$ to align, by matrix multiplication of $\mathbf{W}$ with $\Theta_\mathbf{i}$, the first sample of the ith synergy with the kth sample ($2 - Q \le k \le K$) of a muscle pattern (K samples long) and to truncate the synergy's samples shifted before the beginning or after the end of the pattern (1), we can rewrite Eq. **3** as

$$\mathbf{M} = \mathbf{W}\,\left(\sum_{j=1}^{n} c_j\;\boldsymbol{\Theta}_{I(j)}[k_j, K]\right) = \mathbf{W}\,\mathbf{H} \qquad \textbf{[4]}$$

Model complexity comparison. An important distinction between the two models is the number of free parameters in each one of them. Given a set of N synergies, the

reconstruction of the data with synchronous synergies requires as many combination coefficients as the number of samples (K_{tot}) times the number of synergies (N). With time-varying synergies, there is one amplitude coefficient and one timing coefficient for each instance of a synergy, thus the number of parameters depends on the number of instances. However, because the temporal overlap of different instances of the same synergy is constrained by a refractory period (see below), in general the number of instances of each synergy is less then the number of data samples divided by the number of samples for each synergy (K_{tot}/Q). Thus the number of parameters is less than K_{tot}/Q times the number of synergies times 2 (the number of parameter per instance, one amplitude and one timing coefficient), and the ratio of the number of parameters in the synchronous model over the number of parameters in the time-varying model is in general larger than the number of samples in each time-varying synergies divided by 2 ($Q/2$). Therefore, the time-varying model provides a more parsimonious description of the muscle patterns by expressing the muscle activation waveforms as a combination of the activation time courses of the synergies. In contrast, in the synchronous model the muscle activation waveforms are determined by the combination coefficients.

Synergy Extraction Algorithms. *Synchronous synergies.* The algorithm (2) is initialized with random nonnegative synergies (**W**) and coefficients (**C**) and proceeds to minimize the total reconstruction error by iterating two steps:

(*i*) given the synergies **W**, the coefficients **C** are updated according to the rule

$$C_{ij} = C_{ij} \frac{(\mathbf{W}^{\mathrm{T}}\mathbf{M})_{ij}}{(\mathbf{W}^{\mathrm{T}}\mathbf{W}\mathbf{C})_{ij}}; \qquad \textbf{[5]}$$

(*ii*) given the coefficients **C**, the synergies **W** are updated according to the rule

$$W_{ij} = W_{ij} \frac{(\mathbf{M}\mathbf{C}^{\mathrm{T}})_{ij}}{(\mathbf{W}\mathbf{C}\mathbf{C}^{\mathrm{T}})_{ij}}. \qquad \textbf{[6]}$$

Time-varying synergies. The algorithm starts by initializing N synergies with random nonnegative values, and it proceeds to minimize the total reconstruction error by iterating three steps.

(i) *Selection of synergy instances by matching pursuits.* For each EMG segment and given the set of N synergies, a procedure based on the matching pursuits algorithm (3) is used to select which synergy instances to use in the reconstruction. A *dictionary* for that episode is constructed with all of the possible synergy instances obtained by shifting each normalized synergy from the beginning to the end of the segment one sample at a time. For an EMG segment made of K samples and for N synergies Q samples long, allowing for partial overlap of the synergies at the segment edges, a dictionary of $N \times (K + Q - 1)$ elements is constituted in this way. The matching pursuits procedure then selects those elements in the dictionary whose combinations best match the pattern, using the following steps. First, the scalar products of the muscle pattern with all of the elements of

the dictionary are computed. Second, the dictionary element with the largest scalar product is selected. Third, the selected element is multiplied by its scalar product and subtracted from the muscle pattern. Fourth, the scalar product of the residual muscle pattern with all of the remaining dictionary elements is recomputed. At each iteration one instance is selected and these four steps are repeated until the highest scalar product is below a given threshold (set equal to 0.1). Through these iterations a variable number of instances are selected to best reconstruct the muscle pattern. To avoid selecting overlapping instances, we modified the original matching pursuits algorithm by introducing a *refractory period*. Once a synergy instance has been selected, the scalar products of the other instances of the same synergies with onset times within a fixed time interval (refractory period) around the onset time of the selected synergies are reduced. This reduction depends on the onset time difference according to a Gaussian function with a width equal to half the synergy duration (i.e., $Q/2$).

(ii) *Determination of the scaling coefficients.* For each EMG segment and given the set of N synergies, once n instances have been selected the scaling coefficients ($\{c_i\}_{i=1,\ldots,n}$) that best reconstruct that episode are determined by back-projection (3).

(iii) *Updating of the synergies.* Once the n instances and their scaling coefficients have been determined for all of the episodes, the synergies are updated by the same multiplicative update rule used in the nonnegative matrix factorization algorithm (Eq. **6**) with the matrix **H** (Eq. **4**) in place of the matrix **C**. This step corresponds to adaptively changing the dictionary used in the reconstruction.

1. d'Avella, A., Saltiel, P. & Bizzi, E. (2003) *Nat. Neurosci.* **6,** 300-308.

2. Lee, D. D. & Seung, H. S. (2001) in *Advances in Neural Information Processing Systems*, eds. Leen, T. K., Dietterich, T. G. & Tresp, V. (MIT Press, Cambridge, MA), Vol 13, pp. 556-562.

3. Mallat, S. G. & Zhang, Z. (1993) *IEEE Trans. Sign. Proc.* **41,** 3397-3415.