

## RESEARCH ARTICLE

# Imitation Learning With Time-Varying Synergy for Compact Representation of Spatiotemporal Structures

KYO KUTSUZAWA<sup>1</sup>, (Member, IEEE), AND MITSUHIRO HAYASHIBE<sup>1</sup>, (Senior Member, IEEE)

Department of Robotics, Graduate School of Engineering, Tohoku University, Sendai 980-8579, Japan

Corresponding author: Kyo Kutsuzawa (kutsuzawa@tohoku.ac.jp)

This work was supported in part by JST, ACT-X, Japan, under Grant JPMJAX20A8; and in part by the Innovation and Technology Commission of the HKSAR Government under the InnoHK Initiative, Hong Kong.

**ABSTRACT** Imitation learning is a promising approach for robots to learn complex motor skills. Recent techniques allow robots to learn long-term movements comprising multiple sub-behaviors. However, learning the temporal structures of movements from a demonstration is challenging, particularly when sub-behaviors overlap and are not labeled in advance. This study applied *time-varying synergies*, which are representations of spatial and temporal structures in human behavior in neuroscience, to imitation learning. The proposed method extracts time-varying synergies from human demonstrations, with neural networks that learn their activation patterns. Because time-varying synergies can decompose demonstrations into linear combinations of primitives while allowing overlapping, neural networks can learn demonstrations efficiently. This would make the model compact and improve its generalization ability. The proposed method was evaluated with the task of cursive letter writing requiring overlapping sub-behaviors. Consequently, the proposed method allows a neural network to generate new movements with a higher success rate and fewer parameters than those without the proposed method. Moreover, the neural network worked robustly against control deviations and disturbances in an actual robot.

**INDEX TERMS** Imitation learning, time-varying synergy, neural networks, neuroscience.

## I. INTRODUCTION

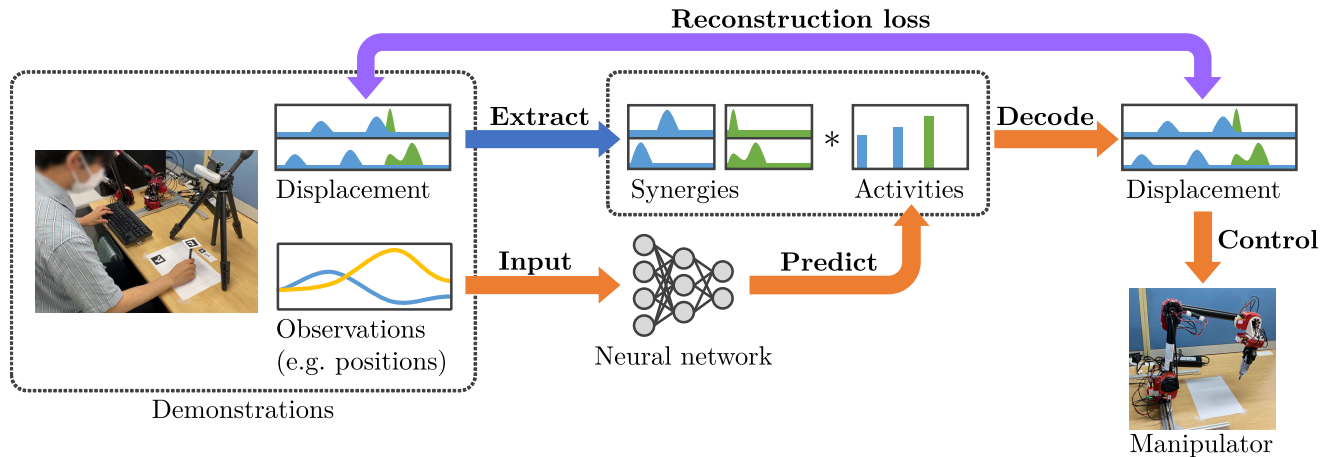
Imitation learning allows robots to learn motor skills from human demonstrations without requiring expert knowledge of robot control. The structure of demonstrated movements should be captured to learn complex motor skills through imitation. Daily human movements involve multiple short-term movements [1]. For example, a simple activity such as pick-and-place can be decomposed into sub-behaviors: reaching the hand to the object's position, grasping, moving it to the destination, and releasing the hand. Moreover, sub-behaviors often overlap and combine over time. Capturing such structures would be beneficial for learning and generalizing using a small number of demonstrations. Additionally, because collecting large numbers of demonstrations is labor-

intensive, it is necessary to learn them from small numbers of demonstrations and generalize to new behaviors.

To learn efficiently from human demonstrations, an approach based on the principle of human motor control is effective. Neuroscience has revealed that human movements can be accounted for by few primitive components, referred to as *synergies* [2], [3], [4]. Synergies efficiently represent the spatial and temporal coordinate patterns of movements using a linear combination. The central nervous system can simplify the control of complex and redundant bodies by linearly combining several synergies [5], [6]. Therefore, synergies are expected to adequately represent the structures of human movement, even in a compact model. Synergies have also been used in robotics to address joint redundancy, particularly in robotic hands [7], [8], [9], [10].

However, synergies are yet to be used to capture the temporal structures of demonstrations concerning robotic learning.

The associate editor coordinating the review of this manuscript and approving it for publication was Thomas Canhao Xu<sup>1</sup>.



**FIGURE 1. Method overview.** The Blue, orange, and purple arrows indicate preprocessing, movement generation, and learning, respectively. \* indicates convolution operation.

Although *time-varying synergies* [11], [12] indicating movement primitives activating at different onset times while overlapping exist, it has not been applied to imitation learning.

This study proposes an imitation-learning method based on time-varying synergies, that encode spatial and temporal coordination in human movements. The proposed method extracts time-varying synergies from human demonstrations for neural networks (NNs) to learn their activation patterns, as illustrated in Fig. 1. Because synergy extraction decomposes human demonstrations into short-term primitives, NNs can focus on learning the spatiotemporal structures of movements. The contributions of this study are as follows.

- 1) This study applied time-varying synergies to imitation learning using NNs. Thus, NNs can focus on learning how sub-behaviors are combined; this is possible even when sub-behaviors overlap in time.
- 2) This study demonstrates that time-varying synergies can make NNs robust against fluctuations in observations caused by robot control errors. Even when the contact states of the robots with the environment fluctuate, they can continue to move without confusion.
- 3) Finally, this study demonstrates that the proposed method allows NNs to generate new movements not included in human demonstrations. Because sub-behaviors are encoded as synergies, NNs can generate new movement patterns by varying their activation timing and amplitude.

## II. RELATED WORKS

### A. IMITATION LEARNING

Recently, various studies have focused on learning movements comprising multiple short-term movements [1]. A common method involves using hierarchical architectures [13], [14]. These typically consist of low-level policies that learn primitive behaviors and high-level policies that plan sequences of low-level policies. Konidaris et al. used symbolic representations to plan upper-level policies [15]. Choi

and Kim used a Bayesian model to learn experts' behaviors [16]. Kroemer et al. used probabilistic models to decompose demonstrations into sequences of phases [17]. Lynch et al. proposed *learning from play*, which learns long-term movements [18]. In most hierarchical approaches, sub-movements are completely separated (no overlap) in time. Additionally, approaches to learning high- and low-level policies simultaneously tend to be unstable, as observed in hierarchical reinforcement learning. In contrast, our method can automatically extract primitive movements in advance and combine them with nearly arbitrary timing (allowing overlap).

Embedding short-term sequences into single latent variables [19], [20] can be useful when providing low-level behaviors explicitly or in advance. Long-term movements can be generated by combining single latent variables with model predictive control [21]. Furthermore, switching can be made between multiple skills based on observations [22] or external signals [23]. Our method also acquires low-level behaviors in advance, but does not require them to be labeled or separated manually; instead, they are extracted directly from demonstrations. Methods that generate movements by dividing them by a fixed time [24], [25] do not capture the spatiotemporal structures of the task.

### B. SYNERGIES

Synergies extracted from muscle activities or kinematic observations [26], [27] are task-specific movement primitives that account for human movements. They provide biological and neuroscientific support, suggesting the benefits of representing human demonstrations in imitation learning. Moreover, they exhibit linearity, which is advantageous for engineering applications. Additionally, synergies emerge during motor learning tasks of gait [28] and reaching [29], [30], [31].

A feature of synergies is that they can be generalized to other movements. The same synergies are often observed in different movements [32], [33], suggesting that synergies can

be reused in new situations. Furthermore, studies reported that even a few synergies can represent diverse movements [30], [33], [34]. Therefore, synergies are expected to improve the generalization ability of NNs for generating new movements.

**C. SYNERGIES IN IMITATION LEARNING**

Synergies have also been applied to imitation learning. A common application is the control of robotic hands, with many degrees of freedom (DoF) [7], [8], [9], [10]. Chai et al. used synergies in locomotion tasks to reduce the movement dimensionality [35]. Reducing the dimensionality of the control inputs with synergies is effective for learning.

Although approaches using spatial coordination patterns are common, applications of temporal coordination patterns are rare. In particular, its application to online movement generation is yet to be attempted. For example, Rückert and d’Avella used time-varying synergies to represent spatial and temporal patterns, whereas activation patterns were determined before the task [36]. Chen and Qiao developed a synergy-based musculoskeletal control system, where the temporal parameters were determined at the beginning of the task [37]. In contrast, this study aimed to determine how synergies are used online.

Synergies are similar to dynamic movement primitives (DMPs) [38], [39], which encode movements with small numbers of primitives. DMPs are often used in imitation learning because of their linearity and ability to reduce dimensionality [40], [41]. Synergies possess these features along with other unique features such as task-specificity and representing shifts in time. These features allow encoding temporal structures of human demonstrations with a smaller number of primitives compared to DMPs, making them advantageous for this study.

**III. METHOD**

**A. OVERVIEW**

The proposed method extracts time-varying synergies and performs supervised learning of synergy activities. An overview of this process is shown in Fig. 1.

We considered the displacement of the positions for each time step, as follows, to extract synergies:

$$\mathbf{x}[t] \triangleq \mathbf{p}[t + 1] - \mathbf{p}[t], \tag{1}$$

where  $\mathbf{x}[t]$  and  $\mathbf{p}[t]$  indicate the displacement and position at time  $t$ .

**B. TIME-VARYING SYNERGIES**

Time-varying synergies represent movements by superimposing primitives while varying the onset times and amplitude, as follows:

$$\tilde{\mathbf{x}}[t] \approx \sum_i c_i \mathbf{w}_{J(i)}[t - t_i], \tag{2}$$

where  $\tilde{\mathbf{x}}[t]$  indicates a non-negative state of a movement at discrete time  $t$ .  $\mathbf{w}_{J(i)}$  indicates  $J(i)$ -th synergy with length  $K$ ,

where  $J(i) \in \{1, \dots, N\}$ ;  $N$  is the number of synergies. Using  $J(i)$  instead of  $i$  allows multiple uses of the same synergy in a movement.  $c_i$  and  $t_i$  indicate the activation amplitude and onset time of  $\mathbf{w}_{J(i)}$ , respectively. An overview of the time-varying synergies is shown in Fig. 2. Time-varying synergies (hereinafter, synergies) are common in a set of movements; however, the activation amplitude and onset times are different for each movement. Therefore, diverse movements can be represented by the same repertory of synergies by varying the activation amplitude and onset times.

**C. SYNERGY EXTRACTION**

The synergies were extracted using the algorithm in [32]. This algorithm can only be applied to non-negative signals; therefore, we first convert the original movements  $\mathbf{x}[t]$  into non-negative variables  $\tilde{\mathbf{x}}[t]$  as follows:

$$\tilde{\mathbf{x}}[t] = \phi(\mathbf{x}[t]) \triangleq \begin{bmatrix} \mathbf{x}_+[t] \\ \mathbf{x}_-[t] \end{bmatrix}, \tag{3}$$

where  $\phi$  is a conversion function, and

$$\mathbf{x}_+[t] \triangleq \max(\mathbf{x}[t], 0), \tag{4}$$

$$\mathbf{x}_-[t] \triangleq -\min(\mathbf{x}[t], 0). \tag{5}$$

Thus, the  $M$ -dimensional variable  $\mathbf{x}[t]$  was converted into a non-negative  $2M$ -dimensional variable  $\tilde{\mathbf{x}}[t]$ . Inverse conversion can be performed as follows:

$$\mathbf{x}[t] = \phi^{-1}(\tilde{\mathbf{x}}[t]) = \mathbf{x}_+[t] - \mathbf{x}_-[t]. \tag{6}$$

We then initialize  $N$  synergies with a length of  $K$  time steps to random positive values, where  $N$  and  $K$  are determined by the users. The synergies were optimized using the following algorithm [32]:

- 1) Determine a set of onset times  $t_i$  and synergy indices  $J(i)$  using the matching pursuit procedure [42]. This procedure selects the synergy having the highest cross-correlation with signals  $\tilde{\mathbf{x}}$  until the cross-correlation is less than the threshold value.
- 2) Calculate the activation amplitude  $c_i$  based on the cross-correlation.
- 3) Update the synergies with the gradient descent to reduce the reconstruction error as follows:

$$\mathbf{w}[t] \leftarrow \max\left(\mathbf{0}; \mathbf{w}[t] - \eta \frac{\partial}{\partial \mathbf{w}[t]} \sum_{\tau} \|E[\tau]\|^2\right), \tag{7}$$

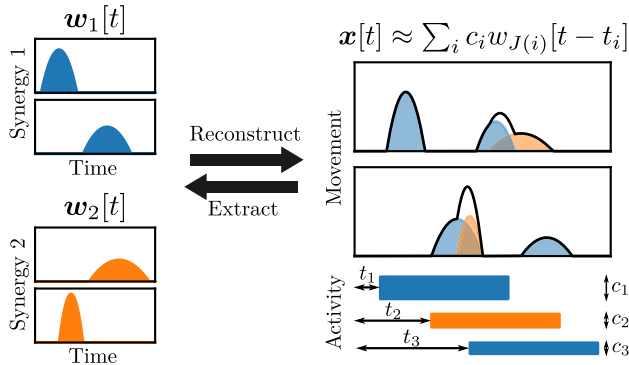
$$E[\tau] = \tilde{\mathbf{x}}[\tau] - \sum_i c_i \mathbf{w}_{J(i)}[\tau - t_i], \tag{8}$$

where the maximum operation ensures that the elements of the synergies are zero or more.

- 4) Go back to 1) and repeat the steps until convergence is achieved.

**D. LEARNING ACTIVITIES OF SYNERGIES**

After decomposing human demonstrations into synergies and their activation patterns, an NN was trained to learn the activation patterns. Here, we consider an NN that receives the



**FIGURE 2.** Overview of time-varying synergies. A movement is decomposed into two time-varying synergies with different onset times  $t_i$  and amplitudes  $c_i$ . In this case,  $J(1) = 1$ ,  $J(2) = 2$ ,  $J(3) = 1$ .

current observation and generates current synergy activities, as illustrated in Fig. 1. Training was performed using supervised learning.

To handle synergies with NNs online, we define the variables  $\boldsymbol{\gamma}[t]$  that represent synergy activity, instead of  $c_i$  and  $t_i$ .  $\boldsymbol{\gamma}[t]$  are  $N$ -dimensional vectors, where the  $n$ -th element  $\gamma_n[t]$  represents the activity of  $n$ -th synergy  $w_n$  as follows:

$$\gamma_n[t] = \begin{cases} c_i & \text{if } J(i) = n \text{ and } t = t_i, \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

The NNs receive the current observation  $\boldsymbol{o}[t]$  and predict the synergy activity  $\boldsymbol{\gamma}[t]$ . Therefore, synergy activities are decoded as follows:

$$\begin{aligned} \hat{\boldsymbol{x}}[t] &= \sum_{n=1}^N (\boldsymbol{w}_n * \gamma_n)[t] \\ &= \sum_{n=1}^N \sum_{\tau=0}^t \boldsymbol{w}_n[\tau] \gamma_n[t - \tau], \end{aligned} \quad (10)$$

where  $*$  denotes convolution operation. Finally, the non-negative  $\hat{\boldsymbol{x}}$  is converted to real-valued  $\hat{\boldsymbol{x}}$  using (6). Furthermore, the positions were reconstructed by integrating the displacements as follows:

$$\hat{\boldsymbol{p}}[t + 1] = \sum_{\tau=0}^t \phi^{-1}(\hat{\boldsymbol{x}}[\tau]). \quad (11)$$

By substituting  $t - \tau = t_i$  into (10) and (9), (10) corresponds to (2).

NNs were trained using the following loss functions:

$$L = \sum_t \left( \|\boldsymbol{p}[t] - \hat{\boldsymbol{p}}[t]\|^2 + \alpha \|\boldsymbol{\gamma}[t]\| \right), \quad (12)$$

where  $\alpha$  is the hyper-parameter. The first term allows the NNs to reproduce the demonstrated movements. The second term is a regularization term that encourages the synergy activities to be sparse (the smallest amplitude possible). This regularization term is expected to drive the NN to reproduce the ground-truth synergy activities because the synergies are mostly inactive and activate only at a single time step at any instant.

**TABLE 1.** Learned and unlearned patterns.

Learned patterns	$j-j-j$	$j-j-k$	$j-q-j$	$j-q-q$	$j-k-q$	$j-k-k$
	$q-j-q$	$q-j-k$	$q-q-q$	$q-q-k$	$q-k-j$	$k-j-j$
	$k-j-q$	$k-q-k$	$k-k-j$	$k-k-q$		
Unlearned patterns	$j-j-q$	$j-q-k$	$j-k-j$	$q-j-j$	$q-q-j$	$q-k-q$
	$q-k-k$	$k-j-k$	$k-q-j$	$k-q-q$	$k-k-k$	

## IV. EXPERIMENTS

To evaluate the proposed method, we applied it to writing cursive-script letters. This specific task has several features that are common to other tasks.

- 1) The task consisted of sub-behaviors, that is, movements of writing individual letters or basic curves.
- 2) The order of the sub-behaviors (i.e., the order of letters) varies.
- 3) The boundaries of each sub-behavior is uncertain and often overlap.

Therefore, evaluating this task would demonstrate the potential for various imitation-learning applications.

### A. TASK SETUP

The robot grasps the pen and moves it to write letters  $j$ ,  $q$ , and  $k$ . We chose these letters because their shapes were significantly different, making it easy to determine successful completion of the task. The robot writes the three letters in one stroke based on a task signal, as explained in Section IV-C.

### B. DATA COLLECTION AND SYNERGY EXTRACTION

A motion capture system collects the demonstrations, as shown in Fig. 3. A marker attached to the pen records human movements. Another marker on the paper identifies the paper's position and pose. The Intel RealSense D455 captured the movements at 30 Hz. A third marker specifies the pattern currently being measured.

We collected all combinations of three letters (27 patterns in total) with three trials for each, that is, 81 demonstrations in total. Fig. 4 shows the collected demonstrations. The time series length is 9.57 s on average, with a standard deviation of 1.69 s. Sixteen out of 27 patterns used for synergy extraction and NN training are given in Table 1.

We extracted synergies from the displacements of positions  $\Delta \boldsymbol{p}$  to capture writing movements regardless of their positions as follows:

$$\boldsymbol{x}[t] = \Delta \boldsymbol{p}[t] \triangleq \boldsymbol{p}[t + 1] - \boldsymbol{p}[t]. \quad (13)$$

The trajectories of  $\Delta \boldsymbol{p}$  were low-pass filtered at 1.1 Hz and down-sampled to 20 Hz. Subsequently, synergies were extracted using the method described in Section III-C.

### C. NEURAL-NETWORK SETUP

The single-layer NN architecture is illustrated in Fig. 5. The model consists of 256-long short-term memory (LSTM) [43], [44] and a linear output layer with softplus activation. For

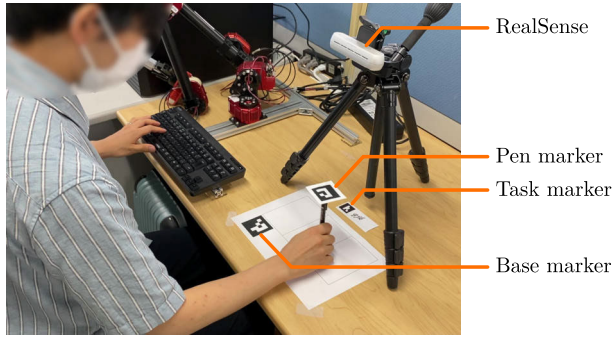


FIGURE 3. Demonstration setup.



FIGURE 4. Collected demonstrations. Three demonstrations are collected for each pattern (overlaid). The orange patterns are excluded from synergy extraction or training.

each time step, this model receives a five-dimensional observation consisting of a two-dimensional hand position  $p[t]$  and a three-dimensional task signal  $z$ . This model then generates synergy activity  $\gamma$  for each time step. The synergy activities are then decoded into positions as follows:

$$\Delta \hat{p}[t] = \phi^{-1} \left( \sum_{n=1}^N \sum_{\tau=0}^t w_n[\tau] \gamma_n[t - \tau] \right), \quad (14)$$

$$\hat{p}[t + 1] = \sum_{\tau=0}^t \Delta \hat{p}[\tau], \quad (15)$$

where  $\phi^{-1}$  is defined in (6). During the training, we used the loss function described in (12).

Task signal  $z$  represents the pattern of letters to be written. It is a three-dimensional variable; the  $n$ -th component is  $-1$ ,  $0$ , or  $+1$  if the  $n$ -th letter is  $j$ ,  $q$ , or  $k$ , respectively. For example, pattern  $j-q-k$  is expressed by  $[-1, 0, +1]$ , and pattern  $q-k-q$

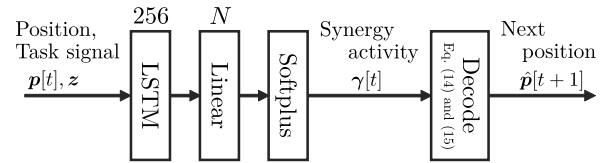


FIGURE 5. Neural-network architecture.

TABLE 2. Hyperparameters for training.

Item	Proposed	Baseline-A	Baseline-B
# Epochs	100000	100000	100000
Mini-batch size	4	4	4
Optimizer	Adam	Adam	Adam
Learning rate	$5 \times 10^{-5}$	$10^{-4}$	$10^{-3}$
Synergy penalty $\alpha$	0.1	-	-

is expressed by  $[0, +1, 0]$ . It is expected that NNs can learn the correspondence between task signals and demonstrated trajectories.

We additionally trained two baseline models, *baseline-A* and *baseline-B*, for comparison with the proposed model. The *baseline-A* model consisted of a single layer with 256 LSTMs, followed by a linear output layer. This architecture is nearly identical to that of the proposed model except for the synergy decoder. The *baseline-B* model consisted of three layers with 256 LSTMs, followed by a linear output layer. This model has more parameters than the proposed and *baseline-A* models. We used this model because the *baseline-A* model resulted in unsatisfactory performance when used in the robot experiments, as described in Section IV-E. Both baseline models directly generated two-dimensional displacement commands of the hand,  $\Delta p[t]$ , instead of synergy activities. The generated commands are then converted to position commands as described in (15), and passed to the robot controller. The hyperparameters used for training are listed in Table 2. The codes are publicly available.<sup>1</sup>

The main objective of the experiments was to compare the performance with and without synergies. Accordingly, we employed LSTM as a general method for time-series generation, which is widely used in various fields, including robotics.

#### D. ROBOT SETUP

We used a five-DoF robot with a gripper, as shown in Fig. 6. This robotic arm has series elastic actuators provided by HEBI Robotics, Inc. This robot was connected to a desktop computer with Intel Core i7-10700 CPU using Ethernet. The computer ran the control system and neural networks.

The robot is velocity controlled, using the control system illustrated in Fig. 7, and the joint angle  $\theta$  is observed. The controller calculates the angular velocity command to the

<sup>1</sup><https://github.com/kyo-kutsuzawa/synergy-based-imitation-learning>

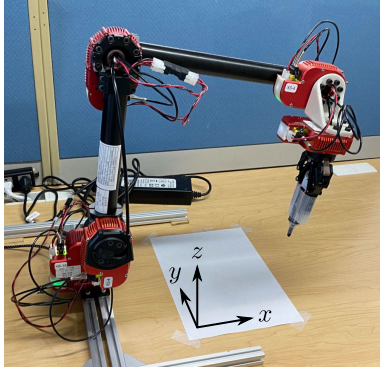


FIGURE 6. Robot appearance.

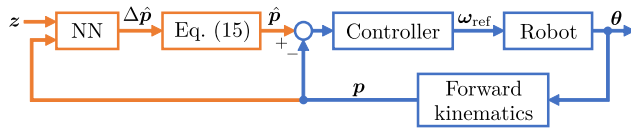


FIGURE 7. Control system. Blue blocks run at 1 kHz, while orange blocks run at 10 Hz.

actuators,  $\omega_{\text{ref}}$ , as follows:

$$\omega_{\text{ref}} = \mathbf{J}^{-1} \mathbf{K}_p (\mathbf{p}_{\text{cmd}} - \mathbf{p}) \quad (16)$$

where  $\mathbf{J}$  is the Jacobian matrix of the robot,  $\mathbf{K}_p = \text{diag}[40, 50, 40, 20, 20]$ , and  $\mathbf{p}_{\text{cmd}}$  indicates the positional command. The  $x$ - $y$  components of the position were computed based on  $\hat{\mathbf{p}}$  and the  $z$ -axis component was set to a constant value to maintain contact with the paper. The pitch and yaw components of the posture were set to ensure that the pen was perpendicular to the paper. The NNs were run at 20 Hz and the controller was run at 1 kHz; first-order interpolation was applied to  $\hat{\mathbf{p}}$  before passing to the controller to account for the difference in computation intervals and smoothen the position command values.

## E. RESULTS

### 1) EXTRACTED SYNERGIES AND THEIR ACTIVITIES

First, we show how synergies encode demonstrations. We also demonstrated that synergies could represent new movement patterns. Figure 8 shows the extracted synergies and examples of their activities in the demonstrations. The convergence criterion was made using the variance accounted for (VAF), denoted by  $R^2$ :

$$R^2 = 1 - \frac{\sum_m \sum_t \|\tilde{\mathbf{x}}_m[t] - \hat{\tilde{\mathbf{x}}}_m[t]\|^2}{\sum_m \sum_t \|\tilde{\mathbf{x}}_m[t] - \tilde{\boldsymbol{\mu}}_m\|^2}, \quad (17)$$

where  $m$  is the index of the demonstrations and  $\tilde{\boldsymbol{\mu}}_m$  is the mean vector of  $\tilde{\mathbf{x}}_m$  over the demonstrations. The value ranges from 0 to 1; the higher this value, the more accurate the reconstruction with the synergies. The synergy extraction is completed when the changes in  $R^2$  is less than  $10^{-5}$ .

Table 3 lists the reconstruction performance using the VAF defined in (17) and validates how the synergies could account

TABLE 3. Reconstruction performance of demonstrations by synergies.

Parameters	$R^2$	
	Learned	Unlearned
$N = 4, K = 25$	$0.865 \pm 0.0225$	$0.865 \pm 0.0222$
$N = 6, K = 10$	$0.936 \pm 0.0080$	$0.934 \pm 0.0078$
$N = 3, K = 80$	$0.713 \pm 0.0912$	$0.746 \pm 0.0685$

for the demonstrations. In  $N = 4, K = 25$  and  $N = 6, K = 10$ , the VAFs were almost identical in the learned and unlearned patterns, with more than 85% on average. For  $N = 3, K = 80$ , the VAFs were below 80%. Although we cannot immediately state whether these values are sufficient, in most cases, synergies could encode sub-behaviors accurately and even represent new movements.

### 2) AUTOREGRESSIVE TRAJECTORY GENERATION

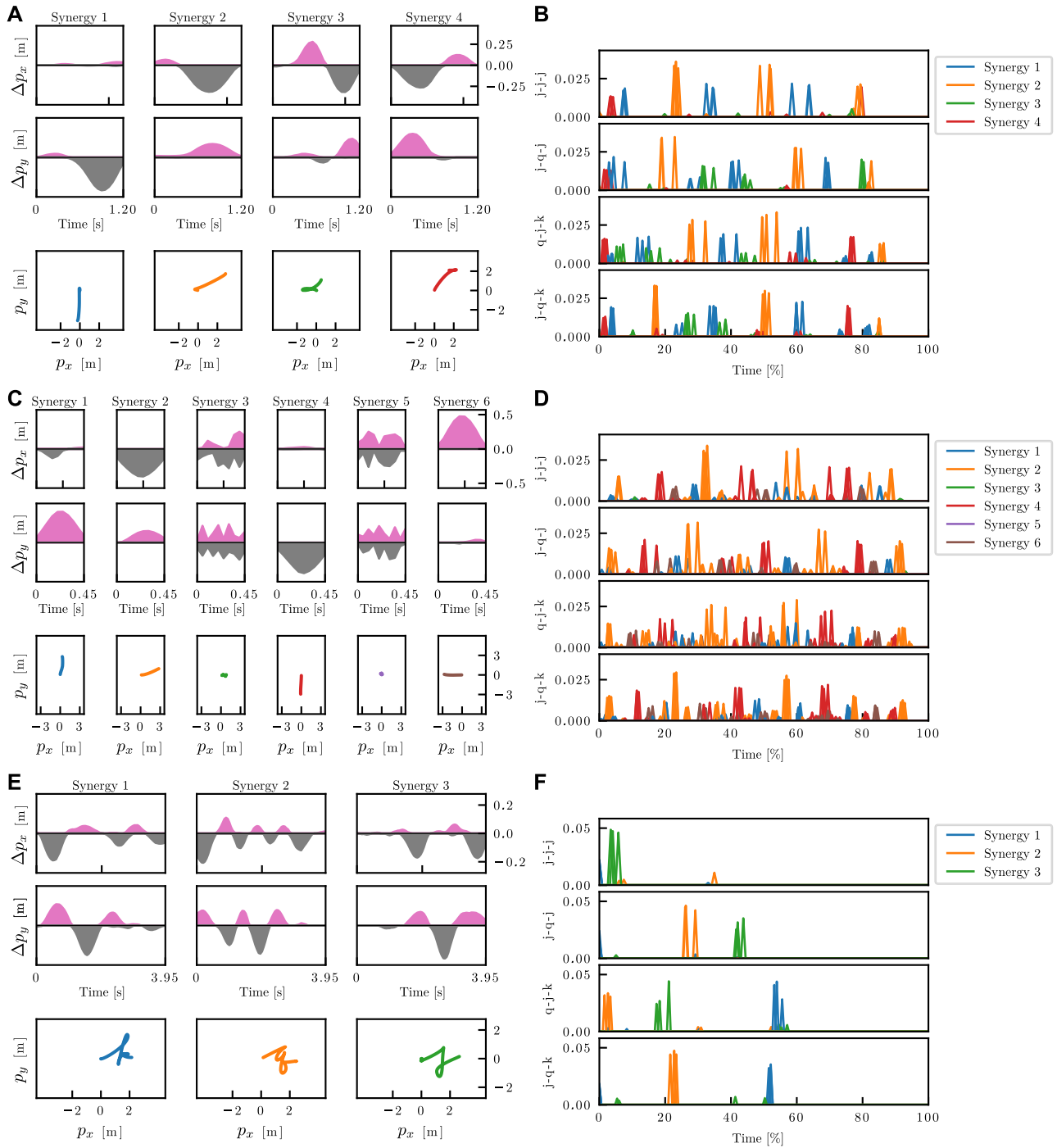
Here, we test the three NN models described in Section IV-C. To evaluate the performance of the models without robots, we let the models generate trajectories in an autoregressive manner. Instead of receiving responses  $\mathbf{p}[t]$  from robots, the NNs received their previous position outputs,  $\hat{\mathbf{p}}[t]$ . This setup corresponded to a case in which robots were ideally controlled without any control delays or disturbances, which was impossible in actual experiments. The initial position was selected based on demonstrations.

Figures 9 and 10 show the letters generated by the NNs. The success rates computed using demonstrations are summarized in Table 4. A trajectory is labeled a success if more than half of the top-two similarity demonstrations are correct. The similarity was measured using the mean squared error; when the two trajectories had different lengths, we truncated them to match the shorter ones. In the baseline-A model, the trajectories were significantly unstable in the learned and unlearned patterns. In contrast, trajectories were generated stably in the baseline-B and proposed models, despite the proposed model having a significantly smaller NN size. Moreover, the success rates of the proposed models were higher than that of the baseline-B model, except for the case of  $N = 3, K = 80$ , suggesting that synergies influence learning.

### 3) APPLICATION TO ROBOTS

Finally, we evaluate the performance of the proposed method using an actual robot. Additionally, we assessed the generalization ability by commanding the model to generate unlearned patterns (orange patterns in Fig. 4). We deployed the proposed and baseline-B models to the robot and not the baseline-A model, which could not generate stable movements, even in the autoregressive scenario, as shown in Fig. 9A. Figure 11 shows the hand-tip trajectories of the robot. Furthermore, Fig. 12 shows examples of actual letters written on paper. Figure 13 shows snapshots of the trial using the proposed method.

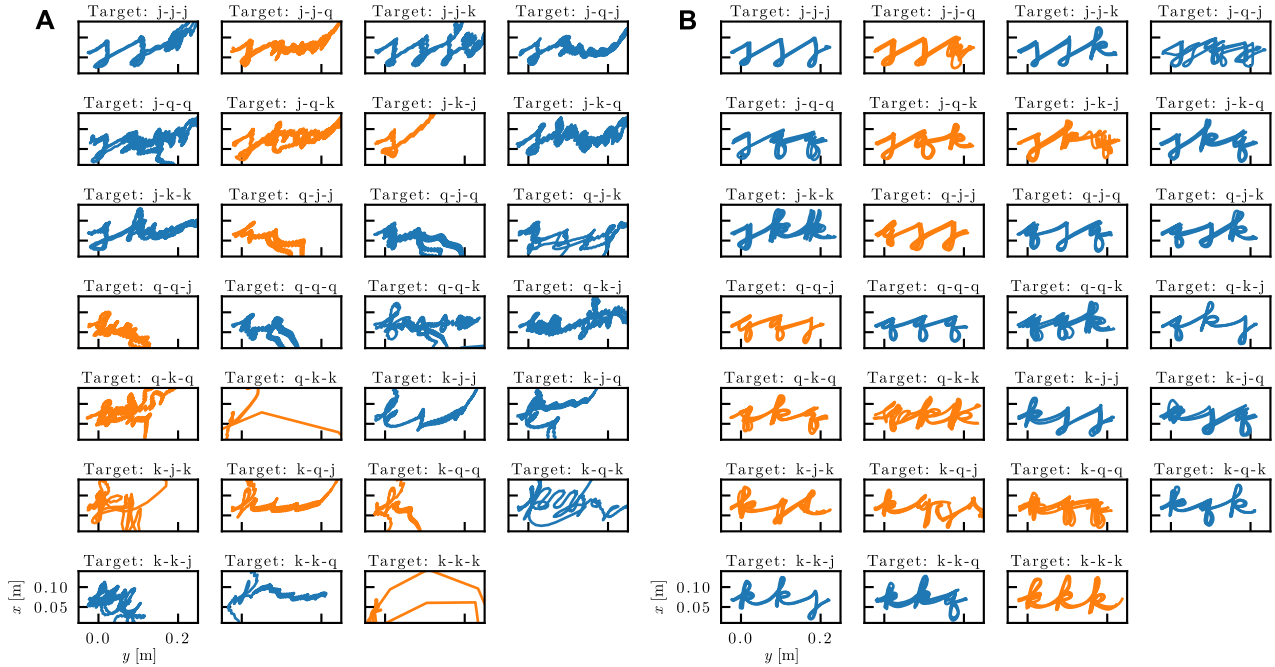
The baseline model generated oscillating behaviors and could not reproduce the letters correctly, despite generat-



**FIGURE 8. Results of synergy extraction varying parameters. Left column: Synergies extracted from demonstrations. The pink and gray waves correspond to the positive and negative components,  $x_+$  and  $x_-$ , respectively. The bottom row has the trajectories drawn by each synergy. Right column: Examples of synergy activities in the demonstrations. The activities for all three trials were overlaid. The pattern  $j-q-k$  was not used in synergy extraction. A-B:  $N = 4, K = 25$ . C-D:  $N = 6, K = 10$ . E-F:  $N = 3, K = 80$ .**

ing the correct trajectories in the autoregressive situation. In addition, most trials were halted during the task because of strong oscillations that could damage the surrounding environment. This was mainly due to the control deviations of the robot, which did not appear in training. Moreover, friction between the pen and paper makes it difficult to

model and compensate perfectly. In contrast, the proposed models, except for  $N = 3, K = 80$ , could write letters in the correct order in most trials, although not in perfect shapes. In addition, these models can write letters in new orders. Even when the contact states occasionally fluctuated owing to the stick-slip phenomenon, as observed in



**FIGURE 9.** Generated trajectories with the baseline models. The robot was not used in these results. The trajectories in orange were not used in training. **A:** Baseline-A. **B:** Baseline-B.

Fig. 12, the proposed models could continue to write without confusion.

The success rates are summarized in Table 4. The baseline-B model could not achieve the task in most trials owing to the oscillation. However, the proposed models could write letters, although the success rate was slightly reduced for unlearned patterns. In addition, the shapes of the written letters barely oscillate. The proposed models with  $N = 4, K = 25$  and  $N = 6, K = 10$  resulted in almost consistent success rates with those in the autoregressive scenario, whereas those with  $N = 3, K = 80$  performed poorly.

The proposed model failed in some unlearned patterns; for example, in  $N = 6, K = 10$ , the generated trajectories resembled  $q-q-k$  for the task signal  $k-j-k$ . Because the model could generate nearly correctly in an autoregressive scenario (cf. Fig. 10), the model may have gotten confused during the experiment owing to the control deviations. However, despite such failure cases, the proposed model did not behave destructively, such as vibratory behaviors; instead, the mistakes were at an abstract level (i.e., writing a different letter).

Figure 14 shows that the proposed model activated synergies similar to the demonstrations. Although the synergies are occasionally activated for several time steps, unlike in the demonstrations, the activation patterns are similar to those of the demonstrations.

## V. DISCUSSION

### A. REPRESENTATIONS OF EXTRACTED SYNERGIES

As shown in Fig. 8, the extracted synergies represent the spatial features of the task at different scales depending on

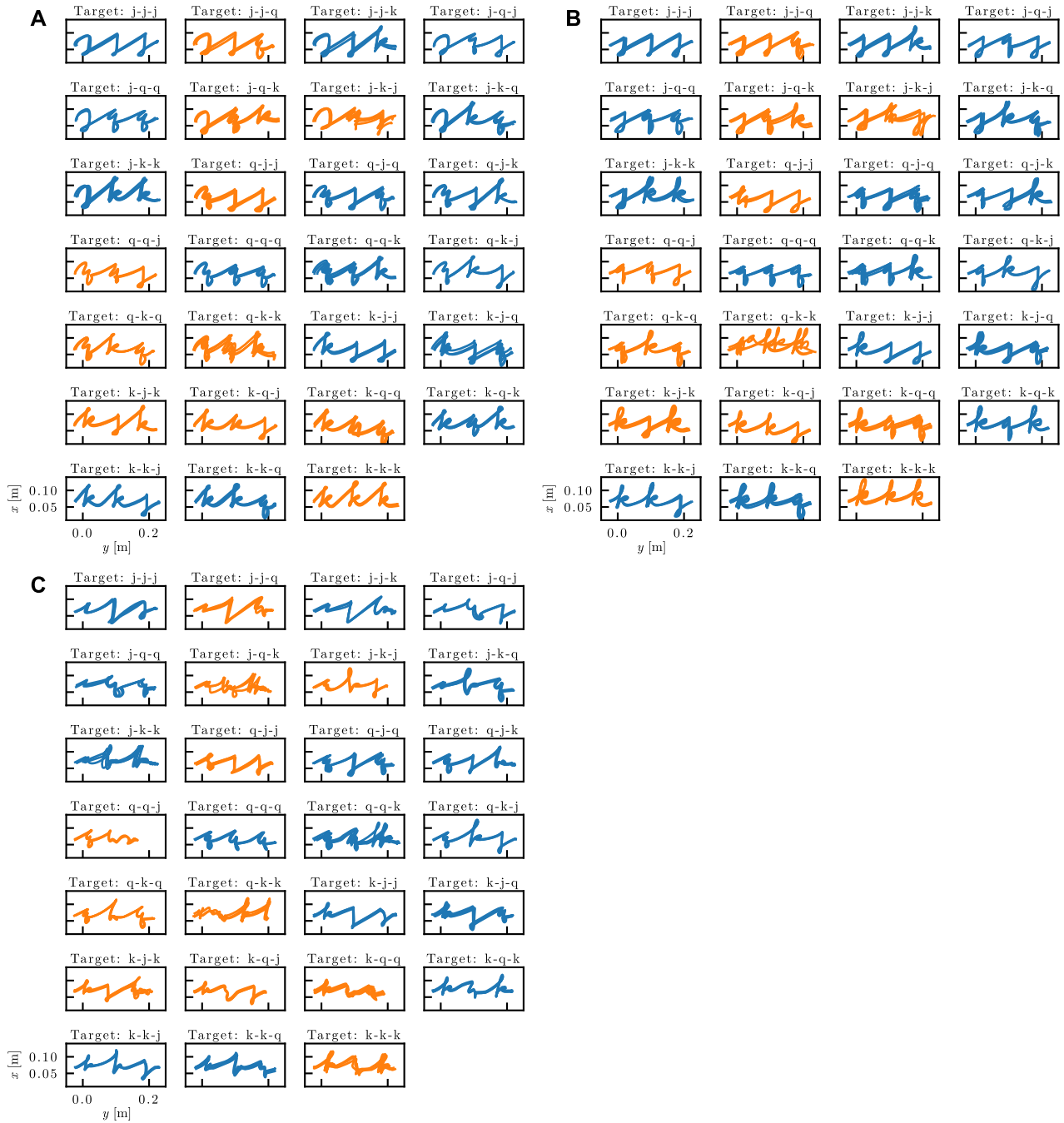
**TABLE 4.** Success rate of the models.

Model	Autoregressive generation		
	Learned data	Unlearned data	Total
Proposed ( $N = 4, K = 25$ )	100.0%	57.6%	82.7%
Proposed ( $N = 6, K = 10$ )	100.0%	69.7%	87.7%
Proposed ( $N = 3, K = 80$ )	100.0%	27.3%	70.4%
Baseline-A	41.7%	12.1%	29.6%
Baseline-B	83.3%	39.4%	65.4%
Model	Robot experiment		
	Learned data	Unlearned data	Total
Proposed ( $N = 4, K = 25$ )	100.0%	54.5%	81.5%
Proposed ( $N = 6, K = 10$ )	100.0%	45.5%	77.8%
Proposed ( $N = 3, K = 80$ )	93.8%	9.1%	59.3%
Baseline-B	12.5%	9.1%	11.1%

$N$  and  $K$ . In summary, the longer the length of the synergies, the more global the features they represent. In addition, when the number of synergies increases, some may not be used.

In  $N = 4, K = 25$  (Fig. 8A), the synergies represented complex curves, as is often observed with cursive letters. Additionally, most synergies move the pen to the right ( $+x$  direction), whereas none moves the pen to the left ( $-x$  direction). This observation makes logical sense because the movement is to the right when writing Latin letters in a cursive style. The synergies indicate smooth acceleration and deceleration patterns, as observed in human writing





**FIGURE 10.** Generated trajectories with the proposed model. The robot was not used in these results. **A:**  $N = 4, K = 25$ . **B:**  $N = 6, K = 10$ . **C:**  $N = 3, K = 80$ .

movements, suggesting that they can be reused in different orders from the demonstrations.

Moreover, their activation patterns correspond to the temporal structures of the task (Fig. 8B). First, the activation patterns are nearly regular and correspond to the letters. For example, when writing the letter *j*, the first and second synergies are activated in a specific order. The third and fourth synergies were predominantly activated when writing *q* and *k*, respectively. Moreover, the activation orders within

individual letters were nearly consistent, even when varying the letter orders. However, certain modifications could be observed depending on the previous and next letters. The same tendency is observed for the pattern *j-q-k*, which was not used in synergy extraction.

Synergies represent different types of features when using different numbers and lengths. In  $N = 6, K = 10$  (Fig. 8C), the synergies represent simpler curves than those for  $N = 4, K = 25$ . The first, second, fourth, and sixth synergies

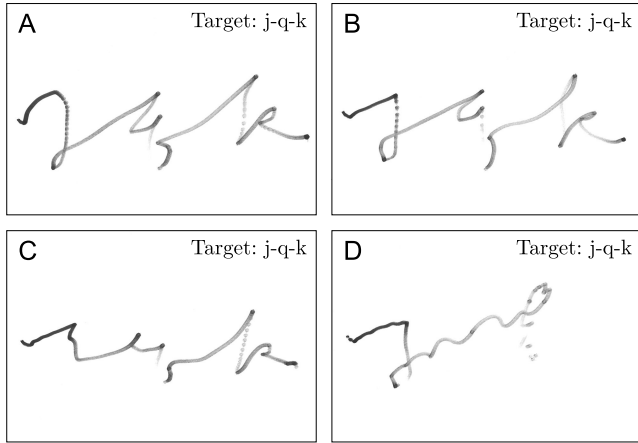


**FIGURE 11.** Generated trajectories by the robot. The trajectories in orange were not used in training. Proposed model with **A:**  $N = 4, K = 25$ , **B:**  $N = 6, K = 10$ , and **C:**  $N = 3, K = 80$ ; **D:** baseline-B model.

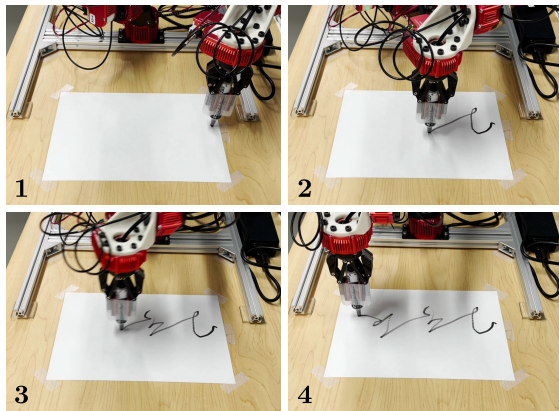
represent straight movements toward the  $-y$ ,  $+x$ ,  $+y$ , and  $-x$  directions, respectively. In contrast, because linear combinations of these synergies can express movement in any direction, the third and fifth synergies represent meaningless movements that remain at the origin. Their activation patterns correspond to the temporal structures of the letter orders, although they look complicated (Fig. 8D).

Each synergy clearly represents an individual letter in  $N = 3, K = 80$  (Fig. 8E). This captures the structure

of the demonstrations. However, the activation patterns in Fig. 8F did not adequately correspond to the temporal structures. For example, in  $j-j-j$ , the third synergy is expected to be used thrice; however, it was used only once at the beginning of the task. Moreover, the first synergy, which represents the shape of  $k$ , was used initially, although the commanded pattern did not include the letter  $k$ . Similarly, the activation patterns did not correspond to the letter orders in  $j-q-j$  and  $j-q-k$ . A possible reason for these failures is



**FIGURE 12.** Examples of actual letters written by the robot. Proposed model with **A:**  $N = 4, K = 25$ , **B:**  $N = 6, K = 10$ , and **C:**  $N = 3, K = 80$ ; **D:** baseline-B model. In the baseline-B model, the robot was stopped while writing because of strong oscillations.

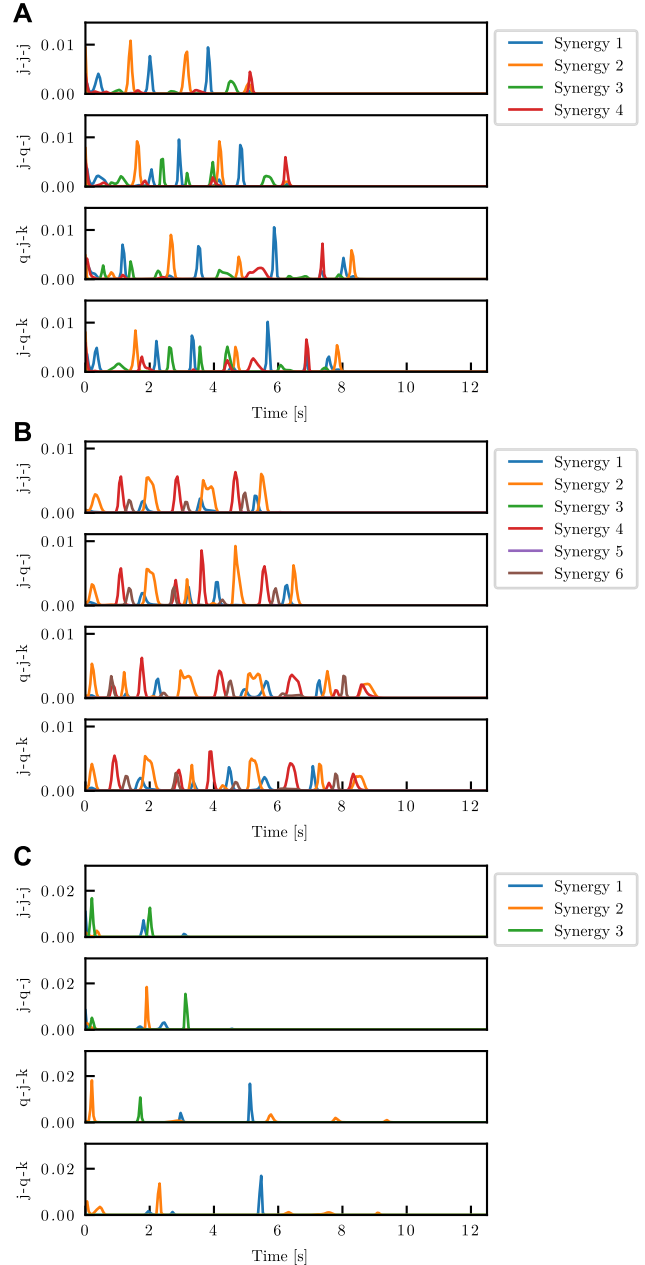


**FIGURE 13.** Snapshots of the robot writing *j-q-k* (an unlearned pattern) using the proposed method with  $N = 4, K = 25$ .

that the synergies could not represent the deviations in letter shapes. Even the same letter varies in shape each time. The shape also varies depending on the letters immediately before and after it. However, such variations cannot be represented when only a single synergy exists for each letter. Consequently, an undesirable synergy activity was obtained by attempting to fit variations in letter shapes, as observed in Fig. 8F.

**B. PERFORMANCE OF MOVEMENT GENERATION**

Time-varying synergies improved the performance of the neural networks with fewer parameters, as shown in Figs. 9–11 and Table 4. The time-varying synergies assisted NNs in learning movements by providing the spatiotemporal structures of the task. The proposed method decomposes demonstrations into movement primitives and combination patterns. Using these primitive patterns (synergies), NNs could generate unlearned patterns relatively more easily than learning demonstrations directly. However, without syner-



**FIGURE 14.** Synergy activities generated by the proposed model when the robot is writing. **A:**  $N = 4, K = 25$ . **B:**  $N = 6, K = 10$ . **C:**  $N = 3, K = 80$ .

gies, NNs would need more layers, parameters, and training data to capture these structures.

Moreover, the proposed models succeeded in the task in an actual robot against control deviations owing to the controller and disturbance by friction. This could also be due to synergies. Because the synergies deal with short-term movements, NNs could have focused on longer-term movements and have become robust against small deviations in the input-output relationship. Additionally, the small number of parameters and layers in the proposed model also contributed to stable movement generation; this was also realized owing to synergies.

### C. LIMITATIONS AND FUTURE WORKS

Although time-varying synergies can help imitation learning, there are some limitations. First, providing guidelines for designing  $N$  and  $K$  of synergy extraction is beneficial. As shown in Fig. 11 and Table 4, the performance varied depending on these hyperparameters. Although they were not very sensitive, a technique to find the best ones for performing the task would be useful. Second, the proposed method does not have a function to halt synergy activation once it starts, causing delays in reacting to unexpected disturbances. However, it has been reported that time-varying synergies can account for human reaching movements when the target positions suddenly change during a task [45]. Therefore, we expect a certain amount of adaptability to be realized without a synergy-activity interruption function. This evaluation is a topic for future research. Third, extending the proposed method to contact-rich tasks, such as assembly tasks, by combining force and impedance controls is desirable.

### VI. CONCLUSION

This study proposes an imitation-learning method using time-varying synergies to decompose human demonstrations into linear combinations of a few primitives activated at various times. Owing to the spatiotemporal representation ability of the synergies, even compact NNs could learn movements comprising several sub-behaviors.

We evaluated the proposed method with the task of writing cursive-script letters. Consequently, the NN with the proposed method could generate movements even for new patterns using a small number of model parameters. In addition, the proposed method generated movements even when applied to a robot. The proposed model focuses on temporal structures (synergy activities) because synergies encode primitive temporal patterns, making it relatively easy for NNs to generate new movements by varying the onset times and amplitudes of the synergies. This contrasts with the baseline models, which need to vary the entire behavior without knowing the temporal structures of the demonstrations. Therefore, the time-varying synergy is a promising representation method for improving the efficiency and generalization ability of imitation learning.

### REFERENCES

- [1] Y. Lee, E. S. Hu, and J. J. Lim, "IKEA furniture assembly environment for long-horizon complex manipulation tasks," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2021, pp. 6343–6349.
- [2] W. A. Lee, "Neuromotor synergies as a basis for coordinated intentional action," *J. Motor Behav.*, vol. 16, no. 2, pp. 135–170, Jun. 1984.
- [3] M. Santello, M. Flanders, and J. F. Soechting, "Postural hand synergies for tool use," *J. Neurosci.*, vol. 18, no. 23, pp. 10105–10115, Dec. 1998.
- [4] A. d'Avella, A. Portone, L. Fernandez, and F. Lacquaniti, "Control of fast-reaching movements by muscle synergy combinations," *J. Neurosci.*, vol. 26, no. 30, pp. 7791–7810, Jul. 2006.
- [5] A. d'Avella and F. Lacquaniti, "Control of reaching movements by muscle synergy combinations," *Frontiers Comput. Neurosci.*, vol. 7, p. 42, Apr. 2013.
- [6] T. Takei, J. Confais, S. Tomatsu, T. Oya, and K. Seki, "Neural basis for hand muscle synergies in the primate spinal cord," *Proc. Nat. Acad. Sci. USA*, vol. 114, no. 32, pp. 8643–8648, 2017.
- [7] T. Geng, M. Lee, and M. Hülse, "Transferring human grasping synergies to a robot," *Mechatronics*, vol. 21, no. 1, pp. 272–284, 2011.
- [8] H. B. Amor, O. Kroemer, U. Hillenbrand, G. Neumann, and J. Peters, "Generalization of human grasping for multi-fingered robot hands," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2012, pp. 2043–2050.
- [9] F. Ficuciello, G. Palli, C. Melchiorri, and B. Siciliano, "Postural synergies and neural network for autonomous grasping: A tool for dextrous prosthetic and robotic hands," *Biosyst. Biorobotics*, vol. 1, pp. 467–480, Jan. 2013.
- [10] A. Furui, S. Eto, K. Nakagaki, K. Shimada, G. Nakamura, A. Masuda, T. Chin, and T. Tsuji, "A myoelectric prosthetic hand with muscle synergy-based motion determination and impedance model-based biomimetic control," *Sci. Robot.*, vol. 4, no. 31, Jun. 2019, Art. no. eaaw6339.
- [11] A. d'Avella and M. C. Tresch, "Modularity in the motor system: Decomposition of muscle patterns as combinations of time-varying synergies," in *Proc. Adv. Neural Inf. Process. Syst.*, 2002, pp. 1–8.
- [12] A. d'Avella, P. Saltiel, and E. Bizzi, "Combinations of muscle synergies in the construction of a natural motor behavior," *Nature Neurosci.*, vol. 6, pp. 300–308, Mar. 2003.
- [13] Y. Wang, C. C. Beltran-hernandez, W. Wan, and K. Harada, "Robotic imitation of human assembly skills using hybrid trajectory and force learning," 2020, *arXiv:2103.05912*.
- [14] K. Hakhamaneshi, R. Zhao, A. Zhan, P. Abbeel, and M. Laskin, "Hierarchical few-shot imitation with skill transition models," 2021, *arXiv:2107.08981*.
- [15] G. Konidaris, L. P. Kaelbling, and T. Lozano-Perez, "Constructing symbolic representations for high-level planning," in *Proc. AAAI Conf. Artif. Intell.*, vol. 3, 2014, pp. 1932–1938.
- [16] J. Choi and K. E. Kim, "Hierarchical Bayesian inverse reinforcement learning," *IEEE Trans. Cybern.*, vol. 45, no. 4, pp. 793–805, Apr. 2015.
- [17] O. Kroemer, C. Daniel, G. Neumann, H. van Hoof, and J. Peters, "Towards learning hierarchical skills for multi-phase manipulation tasks," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2015, pp. 1503–1510.
- [18] C. Lynch, M. Khansari, T. Xiao, V. Kumar, J. Tompson, S. Levine, and P. Sermanet, "Learning latent plans from play," in *Proc. 3rd Conf. Robot Learn.*, Mar. 2019, pp. 1113–1132.
- [19] Z. Wang, J. Merel, S. Reed, G. Wayne, N. de Freitas, and N. Heess, "Robust imitation of diverse behaviors," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 5320–5329.
- [20] K. Kutsuzawa, S. Sakaino, and T. Tsuji, "Trajectory adjustment for nonprehensile manipulation using latent space of trained sequence-to-sequence model," *Adv. Robot.*, vol. 33, no. 21, pp. 1144–1154, Oct. 2019.
- [21] T. Li, R. Calandra, D. Pathak, Y. Tian, F. Meier, and A. Rai, "Planning in learned latent action spaces for generalizable legged locomotion," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 2682–2689, Apr. 2021.
- [22] K. Kase, K. Suzuki, P.-C. Yang, H. Mori, and T. Ogata, "Put-in-box task generated from multiple discrete tasks by humanoid robot using deep learning," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 6447–6452.
- [23] K. Suzuki, H. Mori, and T. Ogata, "Motion switching with sensory and instruction signals by designing dynamical systems using deep neural network," *IEEE Robot. Autom. Lett.*, vol. 3, no. 4, pp. 3481–3488, Oct. 2018.
- [24] A. Gupta, V. Kumar, C. Lynch, S. Levine, and K. Hausman, "Relay policy learning: Solving long-horizon tasks via imitation and reinforcement learning," Oct. 2019, *arXiv:1910.11956*.
- [25] K. Hayashi, S. Sakaino, and T. Tsuji, "An independently learnable hierarchical model for bilateral control-based imitation learning applications," *IEEE Access*, vol. 10, pp. 32766–32781, 2022.
- [26] M. Tagliabue, A. L. Ciancio, T. Brochier, S. Eskiizmirli, and M. A. Maier, "Differences between kinematic synergies and muscle synergies during two-digit grasping," *Frontiers Human Neurosci.*, vol. 9, pp. 1–17, Mar. 2015.
- [27] M. L. Latash, "The organization of quick corrections within a two-joint synergy in conditions of unexpected blocking and release of a fast movement," *Clin. Neurophysiol.*, vol. 111, no. 6, pp. 975–987, Jun. 2000.
- [28] J. Chai and M. Hayashibe, "Motor synergy development in high-performing deep reinforcement learning algorithms," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 1271–1278, Apr. 2020.
- [29] M. Chhabra and R. A. Jacobs, "Properties of synergies arising from a theory of optimal motor behavior," *Neural Comput.*, vol. 18, no. 10, pp. 2320–2342, Oct. 2006.

- [30] M. Al Borno, J. L. Hicks, and S. L. Delp, "The effects of motor modularity on performance, learning and generalizability in upper-extremity reaching: A computational analysis," *J. Roy. Soc. Interface*, vol. 17, no. 167, Jun. 2020, Art. no. 20200011.
- [31] J. Han, J. Chai, and M. Hayashibe, "Synergy emergence in deep reinforcement learning for full-dimensional arm manipulation," *IEEE Trans. Med. Robot. Bionics*, vol. 3, no. 2, pp. 498–509, May 2021.
- [32] A. d'Avella and E. Bizzi, "Shared and specific muscle synergies in natural motor behaviors," *Proc. Nat. Acad. Sci. USA*, vol. 102, pp. 3076–3081, Feb. 2005.
- [33] A. Scano, L. Dardari, F. Molteni, H. Giberti, L. M. Tosatti, and A. d'Avella, "A comprehensive spatial mapping of muscle synergies in highly variable upper-limb movements of healthy subjects," *Frontiers Physiol.*, vol. 10, p. 1231, Sep. 2019.
- [34] K. Kutsuzawa and M. Hayashibe, "Motor synergy generalization framework for new targets in multi-planar and multi-directional reaching task," *Roy. Soc. Open Sci.*, vol. 9, no. 5, May 2022, Art. no. 211721.
- [35] J. Chai, D. Owaki, and M. Hayashibe, "Deep reinforcement learning with gait mode specification for quadrupedal trot-gallop energetic analysis," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Nov. 2021, pp. 4583–4587.
- [36] E. Ruckert and A. d'Avella, "Learned parametrized dynamic movement primitives with shared synergies for controlling robotic and musculoskeletal systems," *Frontiers Comput. Neurosci.*, vol. 7, p. 138, Oct. 2013.
- [37] J. Chen and H. Qiao, "Muscle-synergies-based neuromuscular control for motion learning and generalization of a musculoskeletal system," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 6, pp. 3993–4006, Jun. 2021.
- [38] A. J. Ijspeert, J. Nakanishi, and S. Schaal, "Movement imitation with nonlinear dynamical systems in humanoid robots," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, vol. 2, May 2002, pp. 1398–1403.
- [39] A. J. Ijspeert, J. Nakanishi, H. Hoffmann, P. Pastor, and S. Schaal, "Dynamical movement primitives: Learning attractor models for motor behaviors," *Neural Comput.*, vol. 25, no. 2, pp. 328–373, Feb. 2013.
- [40] P. Kormushev, S. Calinon, and D. G. Caldwell, "Robot motor skill coordination with EM-based reinforcement learning," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2010, pp. 3232–3237.
- [41] J. Kober and J. Peters, "Policy search for motor primitives in robotics," *Mach. Learn.*, vol. 84, no. 1, pp. 83–117, 2014.
- [42] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.
- [43] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [44] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," *Neural Comput.*, vol. 12, no. 10, pp. 2451–2471, 2000.
- [45] A. d'Avella, A. Portone, and F. Lacquaniti, "Superposition and modulation of muscle synergies for reaching in response to a change in target location," *J. Neurophysiol.*, vol. 106, no. 6, pp. 2796–2812, Dec. 2011.



**KYO KUTSUZAWA** (Member, IEEE) received the B.E. degree in electrical and electronic systems and the M.E. and Dr.Eng. degrees in science and engineering from Saitama University, Japan, in 2015, 2017, and 2020, respectively. He is currently an Assistant Professor with Tohoku University. His research interests include robotics, motion generation, force signal processing, and neural networks. He received the Advanced Robotics Excellent Paper Award, in 2020, and the Young Investigation Excellence Award from the Robotics Society of Japan, in 2018.



**MITSUHIRO HAYASHIBE** (Senior Member, IEEE) received the B.S. degree from the Tokyo Institute of Technology, in 1999, and the M.S. and Ph.D. degrees from the Graduate School of Engineering, The University of Tokyo, in 2001 and 2005, respectively. From 2001 to 2006, he was an Assistant Professor with the Department of Medicine, Jikei University School of Medicine. In 2007, he was a Postdoctoral Fellow with the Institute National de Recherche en Informatique et en Automatique (INRIA) and the Laboratoire d'Informatique, de Robotique, de Microelectronique de Montpellier (LIRMM), CNRS/University of Montpellier, France. Since 2008, he has been a tenured Research Scientist with INRIA and the University of Montpellier. Since 2012, he has also been a Visiting Researcher with the RIKEN Brain Science Institute and the Toyota Collaboration Center. Since 2017, he has also been a Professor with the Department of Robotics, Tohoku University, and the Founder of the Neuro-Robotics Laboratory. He is a Senior Member of the IEEE Engineering in Medicine and Biology Society. He serves as a Technical Activity Board Member for the International Foundation of Robotics Research (IFRR). He was a recipient of the 15th Annual Delsys Prize for Innovation in Electromyography from the De Luca Foundation, USA, in 2017.

• • •