

# Predicting visually-modulated precisely-timed spikes across a coordinated and comprehensive motor program

Usama Bin Sikandar  
*School of Electrical and  
Computer Engineering  
Georgia Institute of Technology*  
Atlanta, GA, USA  
usama@gatech.edu

Hannah Choi  
*School of Mathematics  
Georgia Institute of Technology*  
Atlanta, GA, USA  
hannahch@gatech.edu

Joy Putney  
*School of Biological Sciences  
Georgia Institute of Technology*  
Atlanta, GA, USA  
jputney3@gatech.edu

Hengye Yang  
*Sibley School of Mechanical and  
Aerospace Engineering  
Cornell University*  
Ithaca, NY, USA  
hy469@cornell.edu

Silvia Ferrari  
*Sibley School of Mechanical and  
Aerospace Engineering  
Cornell University*  
Ithaca, NY, USA  
ferrari@cornell.edu

Simon Sponberg  
*Schools of Physics and  
Biological Sciences  
Georgia Institute of Technology*  
Atlanta, GA, USA  
sponberg@gatech.edu

**Abstract**—Traditional models of motor control typically operate in the domain of continuous signals such as spike rates, forces, and kinematics. However, there is growing evidence that precise spike timings encode significant information that coordinates and causally influences motor control. Some existing neural network models incorporate spike timing precision but they neither predict motor spikes coordinated across multiple motor units nor capture sensory-driven modulation of agile locomotor control. In this paper, we propose a visual encoder and model of a sensorimotor system based on a recurrent neural network (RNN) that utilizes spike timing encoding during smooth pursuit target tracking. We use this to predict a nearly complete, spike-resolved motor program of a hawkmoth that requires coordinated millisecond precision across 10 major flight motor units. Each motor unit enervates one muscle and utilizes both rate and timing encoding. Our model includes a motion detection mechanism inspired by the hawkmoth’s compound eye, a convolutional encoder that compresses the sensory input, and a simple RNN that is sufficient to sequentially predict wingstroke-to-wingstroke modulation in millisecond-precise spike timings. The two-layer output architecture of the RNN separately predicts the occurrence and timing of each spike in the motor program. The dataset includes spikes recorded from all motor units during a tethered flight where the hawkmoth attends to a moving robotic flower, with a total of roughly 7000 wingstrokes from 16 trials on 5 hawkmoth subjects. Intra-trial and same-subject inter-trial predictions on the test data show that nearly every spike can be predicted within 2 ms of its known spike timing precision values. Whereas, spike occurrence prediction accuracy is about 90%. Overall, our model can predict the precise spike timing of a nearly complete motor program for hawkmoth flight with a precision comparable to that seen in agile flying insects. Such an encoding framework that captures visually-modulated precise spike timing codes and coordination can reveal how organisms process visual cues for agile movements. It can also drive the next generation of neuromorphic controllers for navigation in complex environments.

**Index Terms**—sensorimotor control, precise spike timings, recurrent neural network, insect flight

## I. INTRODUCTION

Animal locomotion poses many challenges to an animal’s sensorimotor processing, especially for goal-directed tasks in dynamic and uncertain environments. To understand these challenges, some neural network models can predict firing rates or other continuous representations of neural parameters or can classify sensory stimuli and motor responses. These models are integrated with traditional motor control models that typically operate in the domain of continuous signals such as muscle forces, muscle activation, and limb and body kinematics [1], [2]. However, they do not capture a key aspect of neural signals essential for movement: temporal precision of spikes and coordination of neural activity. There is growing and ubiquitous evidence that neurons across the sensorimotor cascade encode information in the precise timings of spikes – not just for sensory information encoding and processing, but for the generation of motor behavior as well [3]. To respond to external stimuli, especially during periodic gaits, the nervous system uses a control strategy making slight modulations to the temporally consistent motor spiking pattern coordinated across synergetic muscles [4], [5]. Together with spike rate, millisecond-precise spike timings have been shown to encode a significant (and sometimes greater) amount of information across insect flight control and turning maneuvers [6], [7], songbird acoustic structures [8], respiration [5], human movement [3], muscle coordination [7], and insect escape behaviors [9]. Beyond understanding motor control, such models can also potentially motivate the next generation of neuromorphic

spike-based controllers for goal-directed robot navigation in dynamic and uncertain environments.

The use of artificial neural networks (ANNs) is on the rise as predictive models of animal behavior and locomotion based on neural data, for their ability to capture complex and nonlinear relationships between neural activity and body mechanics [10]–[12]. Given the importance of incorporating neural spiking patterns into neuromechanical models, some recent algorithms have been developed to train spiking neuron models to spike either at rates that correlate with the given motor control outputs or within short time windows. For instance, ReSuMe and SPAN can learn temporally precise spike patterns to solve classification and decision-making tasks [13], [14]. Chronotron spiking neurons can learn to generate precisely timed output spikes in response to precisely-timed input spike patterns, and thus can process entirely temporally coded information [15]. Another learning algorithm is capable of training spiking neural networks (SNNs) to generate the desired sequence of spikes within precisely specified timing windows [16]. However, the capability of some of these algorithms to predict precise spike timings is yet to be shown beyond feedforward or two-layered networks trained on simulated sensory input data. Other algorithms that currently exist cannot capture the features of the nervous system that generate precisely-timed motor spikes modulated by varying sensory stimuli and coordinated across multiple motor units. Nonetheless, they serve as motivation for building an ANN-based model that can accurately predict spike timings across a motor program for agile movement and thereby interface with neuromorphic control models.

To begin with, what minimal ANN-based modeling framework might be sufficient to learn the time-dependent dynamics of the sensorimotor system and motor control for predicting underlying spiking patterns that enable movement? Learning these dynamics is very data intensive which makes neural network approaches ideal for this purpose. The presence of time dependencies in the data means that recurrence would be a useful feature to incorporate into the model. The use of recurrence is biologically motivated and recurrent neural networks (RNNs) are widely used to reliably model dynamical systems in computational neuroscience [17]. Therefore, using a bottom-up approach in this paper, we propose a model of a sensorimotor system based on a simple RNN that utilizes temporal encoding of sensory information on a moving target to precisely predict the spike-resolved coordinated motor output that tracks the target. Our model is based on a hawkmoth’s (*M. sexta*) visuomotor system in the context of floral target tracking. From spike-resolved electromyography (EMG) signals recorded from a hawkmoth generating steering maneuvers while tracking a floral target, our model uses an RNN to learn the patterns and precise timings of spikes that are coordinated across 10 major flight muscles, each acting effectively as a single motor unit, as opposed to 10 independent channels. It also features a motion detection mechanism inspired by the hawkmoth’s compound eye.

Our model organism, hawkmoth *M. sexta*, is a remarkably

agile flier that can not just navigate quickly through dense and cluttered environments at low ( $< 1$  lux) light levels [18], but sustain long bouts of hovering while feeding from flowers swaying in the wind [19]. The spike timing information on its steering response to a floral target is thrice the amount of information in the spike count [7]. The spiking patterns of its 10 major flight muscles are periodic and contain shared near-complete information on coordination for flight control [7]. On top of the spike pattern periodicity, its nervous system’s spike time coding strategy for motor control is sub-millisecond precise and consistent across muscles which actuate wing movement to generate aerodynamic and inertial forces for rapid steering maneuvers [6], [7], [20]. A recently developed method can accurately decode the hawkmoth’s instantaneous forces and torques from its comprehensive spike-resolved motor output [21]. Hence, given the time-dependencies, precision, completeness and coordination in the hawkmoth’s flight motor program, in this paper, we explore whether a simple RNN-based model is sufficient to encode the visual modulation of the hawkmoth’s comprehensive motor program. Beyond predicting a coordinated set of motor spikes for an agile sensorimotor task, such a model can potentially be a critical component of a neuromechanical framework that captures precise timings of neural signals not just to encode information and accomplish coordination but also to make causal and context-dependent predictions about locomotion.

## II. METHODS

The modeling task at hand is to predict the visuomotor response of the hawkmoth, *M. sexta*, in terms of precisely-timed spike patterns of its nearly complete set of spike-resolved motor signals for flight control. A moving floral target drives variation in spike numbers and timings. The complete modeling framework as a transformation from visual stimulus to the comprehensive spike-resolved motor program is shown in Fig. 1a. It consists of four functional blocks. The first block, the ‘compound eye model’ recreates the visual scene that the photoreceptors of the hawkmoth’s compound eye perceive in the form of a sequence of frames sampled at the eye’s natural resolution. The second block performs event-based motion detection in the visual scene as a series of binary (or trinary) event codes at each pixel similar to neuromorphic cameras. The third block transforms the spatiotemporal features of the detected motion to a lower-dimensional latent space using a convolutional autoencoder. Then, from a sequence of these latent sensory signals, the fourth and final block uses an RNN to predict the corresponding precisely-timed spike patterns of the 10 major flight muscles that represent the hawkmoth’s flight motor control for floral target tracking.

### A. Dataset

The dataset used to model and test this framework was previously recorded from multiple independent trials on 5 *M. sexta* subjects flying on a fixed tether and responding to a sinusoidally oscillating robotic flower [7]. The flower was oscillating laterally at 1 Hz. Since this oscillation is much

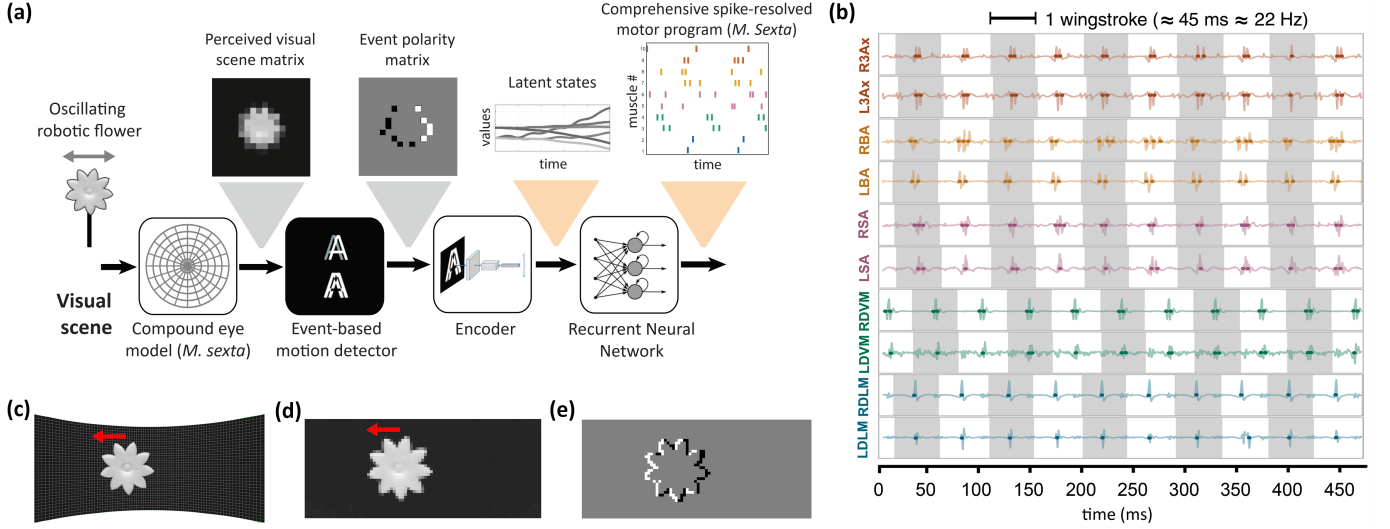


Fig. 1. (a) Our model of a hawkmoth’s visuomotor system utilizes the visual encoding of a moving floral target to sequentially predict precise timings of spike-resolved motor output coordinated across 10 major flight muscles. (b) Spike patterns in EMG data from the 10 major flight muscles: Left (L) and right (R) dorsolongitudinal (DLM), dorsoventral (DVM), subalar (SA), basalar (BA) and third axillary (3Ax). The time series of spike data has been split into wingstrokes (shown as alternating shaded and unshaded rectangles) in a way to preserve contiguous spike bursts. Hence, the rectangles do not coincide in time. But the duration of all the temporally coinciding wingstrokes is the same across all muscles. (c) Flower in the reconstructed animated scene to be projected on the ommatidial grid on the compound eye’s surface (d) Reconstruction of flower’s image in hawkmoth’s compound eye shown in a  $36 \times 84$  pixels grid of model photoreceptors. The grayscale value at each pixel represents the luminance intensity detected by the corresponding model photoreceptor. (e) Event-based flower motion detection is implemented where each pixel represents the motion-detection polarity (white +1, gray 0, black -1).

slower than the wingstroke (22 Hz), it staples a large variety of turns. A hawkmoth’s tracking response to the flower motion was measured as forces and torques through a transducer but here we only use these values to determine the flight bouts where the hawkmoth was responding. Each 20-second trial in the dataset captures roughly 450 wingstrokes and consists of a time series of 1-D flower position, precise (0.1 ms) timings of roughly 7000 spikes from all the 10 major flight muscles, and force and torque measurements from the force transducer. The precise spike timings were previously extracted from spike-resolved EMG recordings. A sample 500 ms recording is shown in Fig. 1b. The entire dataset consists of 16 trials of 20-second each from experiments on 5 hawkmoth subjects. This makes up a total of roughly 7000 wingstrokes and more than 100,000 EMG spikes from the 10 major flight muscles.

## B. Visual system model

1) *Scene reconstruction and the compound eye model:* For each 20-second trial, we track the flower and extract its trajectory by applying DLTdv8 [22] on a recorded video captured from above the hawkmoth and the flower (top view). Then using the measurements of the 3-D relative position of the moth’s tether with respect to the flower rig, we reconstruct the motion of the flower as an animation in the hawkmoth’s frame of reference. Once this visual scene is created, we compute its projection onto an array of photoreceptors in the hawkmoth’s compound eye model. The compound eye is assumed as a simple hemispherical eye with a  $180^\circ$  field of view along both azimuthal and elevation angles. Each facet is connected to one photoreceptor projected by an interommatidial angle of

$0.96^\circ$  [23], [24]. The visual scene of the flower on a plane outside the eye is then projected onto the hemispherical array of photoreceptors using Gnomonic projection [25],

$$x = x_0 - \frac{mr_e}{\cos \theta} (\cos \phi_x \sin \phi_{x_0} - \sin \phi_x \cos \phi_{x_0}) \cos(\phi_{y_0} - \phi_y), \quad (1)$$

$$y = y_0 + \frac{mr_e}{\cos \theta} \cos \phi_{x_0} \sin(\phi_{y_0} - \phi_y), \quad (2)$$

where  $x$  and  $y$  are the coordinates of points on the image plane that is being projected from outside the eye,  $x_0$  and  $y_0$  are the coordinates of the center of focus of the hawkmoth’s eye on the image plane (assumed  $(0, 0)$ ),  $\theta = |\cos^{-1}(\sin \phi_x \sin \phi_{x_0} + \cos \phi_x \cos \phi_{x_0} \cos(\phi_{y_0} - \phi_y))|$ ,  $\phi_x$  and  $\phi_y$  are the angular locations of a photoreceptor along the horizontal and vertical directions across the hemispherical photoreceptor array,  $\phi_{x_0}$  and  $\phi_{y_0}$  are values at the center of the array (assumed  $(0, 0)$ ),  $r_e = 0.00205$  m is the radius of hawkmoth’s eye [23] and  $m = d_{mf}/r_e + 1$  is the magnification ratio where  $d_{mf}$  is the distance between hawkmoth and flower. Because there is no moving object in the data except the robotic flower, the hawkmoth’s field of view is restricted between  $-80^\circ < \phi_x < 80^\circ$  and  $-35^\circ < \phi_y < 35^\circ$  to discard the pixels with no information. The resulting grid shown in Fig. 1c is 36 pixels  $\times$  84 pixels, which makes a total of 3024 photoreceptors – a small subset of  $\sim 25000$  facets on *M. sexta*’s compound eye [24]. The frames are sampled at 125 FPS, which is approximately the flicker fusion frequency of hawkmoths [26]. The luminosity level  $L(p_x, p_y, n)$  of each pixel is converted to grayscale range 0-255, where  $0 \leq p_x \leq 83$  and  $0 \leq p_y \leq 35$  are  $x$  and  $y$  pixel indices respectively, and  $n$  is the frame number. A sample

frame of the flower's image projection inside the compound eye model is shown in Fig. 1d.

2) *Event-based motion detection*: To detect the flower motion, an event-based motion detection algorithm is implemented [27]. This is a simplified neuromorphic sensing technique inspired by elementary motion detection in optic lobes of insects [28]. At each pixel, it evaluates the difference in luminosity between consecutive frames. If the absolute difference at a pixel  $(p_x, p_y)$  between  $n$ th and  $(n-1)$ th frame is greater than a certain threshold, an event  $E(p_x, p_y, n)$  is said to be detected. A polarity of either  $+1$  or  $-1$  is assigned to that pixel depending on whether the luminosity difference was positive or negative respectively. Otherwise, a value of 0 is assigned. A sample frame of the event polarity matrix (EPM) of flower motion is shown in Fig. 1e where the flower is moving to the left of the observer. Visual scene reconstruction, compound eye modeling and event-based motion detection are all performed using custom code in MATLAB.

3) *EPM compression*: The EPM is a high-dimensional but sparse representation of the flower motion. So it can be compressed to reduce its dimensionality with little or no loss of information content. Indeed, such extraction of task-relevant variables from large, sparse, event-based sensory blocks is one way the brain is thought to process information for controlled motor tasks [29]. We train a convolutional autoencoder, with 6 encoding and 6 decoding layers in series, and then extract a 64-dimensional array of latent features from the output of the encoding layers. The input to the autoencoder and the autoencoded output are stacked EPMs where  $(-1, 0, +1)$  polarities are transformed to  $(0, 1)$  polarities for simplification of the learning task. Pixels in the matrix with  $+1$  polarities occupy the upper stack while those with  $-1$  polarities occupy the lower one. Stacked EPM is also a closer representation of motion detection in the optic lobe of the hawkmoth because separate cells are responsible for detecting an abrupt rise or fall in luminosity [30]. For each frame or at each time-step, a 6048-element event polarity matrix  $E(n)$  of flower motion is compressed to a 64-element latent feature array  $g_{i,n}$ , where  $i \in \{1, 2, 3, \dots, 64\}$  is the feature dimension and  $n$  is the frame number. The autoencoder is modeled and trained using Keras API in Python, and the optimizer used for training is Adam with binary cross-entropy loss function.

### C. Predicting precisely-timed spike patterns

The prediction part of our model consists of a neural network with one input layer, one hidden RNN layer and two output layers as shown in Fig. 2a. The latent features  $g_{i,n}$  are normalized and downsampled to wingbeat frequency and are then input to the neural network model as  $\tilde{g}_{i,k}$ , where  $k$  is the wingstroke number. Because there are 64 latent features, the input layer has 64 neurons. The hidden layer consists of 1000 recurrent neurons. Our choice of the number of neurons was made after trying networks of different sizes and then observing that at least roughly 1000 neurons are necessary to capture the spike timing variance observed in the data. The time-step of the recurrence equals the wingstroke time

period of the hawkmoth in the data because its flight muscles spike in periodic bursts. As shown in Fig. 1b, each motor unit corresponds to one muscle and most of the muscles fire in bursts nearly every wingstroke. The burst of spikes activating a muscle occurs in a narrow phase range that varies across muscles. Moreover, in our dataset, each muscle has a different maximum number of spikes in a single burst as shown in Fig. 2b. The observed maximum number of spikes per burst summed across all muscles in our dataset is 41. Because each spike in a burst can be viewed as occurring at most once per wingstroke, we assign features to these 41 'spiking units' separately. These features include 1) binary occurrence of the spike or spike event  $e_{j,k}$  i.e. whether  $j$ th spiking unit has fired ( $e_{j,k} = 1$ ) or not fired ( $e_{j,k} = 0$ ) in the  $k$ th wingstroke, and 2) spike timing  $t_{j,k}$  in milliseconds relative to the start of the wingstroke. In case a spiking unit does not fire in the  $k$ th wingstroke, its spike time is assumed to be  $t_{j,k} = t_{j,k-1}$ . Spikes in a burst are assigned to each spiking unit ordinaly i.e. based on the sequence of their appearance in the burst. Extracted spike event  $e_{j,k}$  and spike timing  $t_{j,k}$  can always be recombined into spike trains because we know both the specific muscle each spiking unit corresponds to and the time instance when the current wingstroke started. To predict these two features for each spiking unit sequentially every wingstroke, there are two output layers with 41 neurons each. One output layer makes binary predictions on spike events  $e_{j,k}$ , whether  $j$ th spiking unit has fired in the  $k$ th wingstroke. The other output layer predicts normalized spike timings  $\tilde{t}_{j,k}$  relative to the wingstroke onset. Hence, the task for the neural network is to accurately predict for the  $k$ th wingstroke the spike events  $e_{j,k}$  and normalized timings  $\tilde{t}_{j,k}$  for  $j \in \{1, 2, 3, \dots, 41\}$  spiking unit, given the normalized latent features  $\tilde{g}_{i,k}$  of the flower motion for  $i \in \{1, 2, 3, \dots, 64\}$ . In our dataset, the maximum number of spiking units turned out to be 41. However, more generally in other datasets, the number of total spiking units could vary. Hence more generally, the maximum number of spikes per burst for a muscle or a motor unit can be chosen to be a large number (maximum number of possible spikes in a wingstroke given spike duration and refractory period) and then the total number of spiking units can be calculated similarly by summing over all the muscles or motor units. Among the spike patterns of 10 major flight muscles shown in Fig. 1b, RDLM and LDLM consistently spike once every wingstroke. Hence, their mean spike time is used as the reference time, relative to which  $t_{j,k}$  is predicted. In other words, the EMG data is split into wingstrokes based on the mean spike timings of the DLM muscles [31] but a small constant time shift (as wingstroke fraction) is applied for each muscle to center the spike bursts in order to keep them contiguous as shown in Fig. 1b.

1) *Network training*: The entire prediction part is modeled and trained using Keras. The recurrent layer has a time step of discrete 1 unit which means that its inputs and outputs from  $(k-1)$ th wingstroke were used for making predictions for the  $k$ th wingstroke. For both output layers of the neural network, the loss function is the mean-squared error but the overall loss

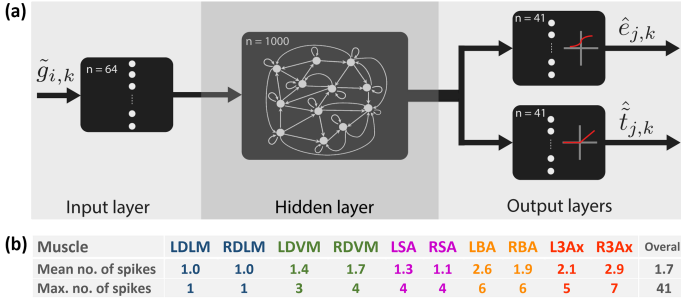


Fig. 2. (a) Structure of the neural network that is trained to predict the spike events and timings of each of the 41 spiking units across 10 muscles. The input layer receives a sequence of normalized latent features  $\tilde{g}_{i,k}$  from the convolutional autoencoder. The RNN in the hidden layer processes these values. During the  $k$ th wingstroke and for the  $j$ th spiking unit, the first output layer predicts spike events  $e_{j,k}$  (followed by a rounding operator), and the second output layer predicts normalized (by wingstroke time period) spike timings  $\tilde{t}_{j,k}$ . Sigmoid and ReLU activation functions are applied at the first and the second output layers respectively. (b) The mean and the maximum number of spikes per wingstroke for each muscle observed across the entire dataset of roughly 7000 wingstrokes

function is the sum of the two given by,

$$J = \frac{w_e}{M} \sum_k \sum_j (e_{j,k} - \hat{e}_{j,k})^2 + \frac{w_t}{M} \sum_k \sum_j (\tilde{t}_{j,k} - \hat{\tilde{t}}_{j,k})^2, \quad (3)$$

where  $w_e = 1$  and  $w_t = 0.25$  are the weights assigned to each loss function to make their magnitudes comparable,  $M$  is the product of the number of wingstrokes and the number of spiking units,  $\tilde{t}_{j,k}$  is the normalized value of  $t_{j,k}$ , and  $\hat{e}_{j,k}$  and  $\hat{\tilde{t}}_{j,k}$  are the predicted values of  $e_{j,k}$  and  $\tilde{t}_{j,k}$  respectively. We train our model to make intra-trial (same trial, same moth subject) as well as inter-trial (different trial, same moth subject) predictions. Predictions are only limited to the same subject because motor control patterns of different hawkmoth subjects can vary despite the same sensory stimulus [7], [32]. Our dataset consists of 16 experimental trials of 20 seconds each from 5 hawkmoth subjects, with 2 to 5 trials per subject. The first half of each 20-second trial is the training set (80-20% validation split) for intra-trial predictions, and the second half of the same trial is used as the test set. For inter-trial predictions, one randomly chosen trial from each subject is selected as the training set while the remaining trials from the same subject are placed in the test set. This makes the total number of test trials for inter-trial predictions to be 11. The total number of wingstrokes in the test set of intra-trial and inter-trial predictions are 3506 and 4914 respectively. Only the first half of a 20-second trial is used as the training set so that the same trained weights can be used to make both inter-trial and intra-trial predictions. With the Adam optimizer used for network training for all training trials, the total loss function converges to a value of about 0.2 within 500 epochs. Beyond this point, the validation loss starts to increase.

### III. RESULTS

To test the success of our visual perception and RNN-based model, we analyze how accurately it can learn and predict

patterns of the hawkmoth's motor spikes coordinated across its 10 major flight muscles. We break down the predictions into two fundamental features of the spiking patterns: spike counts measured in terms of spike events  $e_{j,k}$ , and spike timings  $t_{j,k}$  measured in milliseconds. From the predicted values of these features, a spike train for each muscle can be reconstructed. A sample spike train from intra-trial and inter-trial predictions is shown in Fig. 3a.

#### A. Spike event predictions

For intra-trial predictions on the test data comprising 3506 wingstrokes, the RNN correctly predicts spike events with 84-100% accuracy of the wingstrokes across all muscles as shown in Fig. 3b. Spike events for both DLM muscles are predicted with 100% accuracy because these muscles consistently spike once every wingstroke. Therefore, the RNN is able to learn their spike events perfectly. The R3Ax muscle has the least spike prediction accuracy at 84%. Overall 90% accuracy is achieved across the entire motor program. In the data, the range of the maximum percentage of one particular number of spike counts per wingstroke across all muscles is 40-68%. Overall at 90%, our predictor is performing considerably well compared to if it just predicted these maximum numbers for each muscle instead of learning. Results on spike events that are correctly predicted, missed (false negatives) and undesired (false positives) are also plotted and compared with the observed ones in Fig. 3c. The number of undesired spikes as a percentage of actual spikes is also less than 14%. For the first spike in a burst, the number of missed spikes and undesired spikes as a percentage of actual spikes is always less than 1% and 8% respectively.

For inter-trial predictions on the test data comprising 4914 wingstrokes, the RNN correctly predicts spike events with 80-100% accuracy of the wingstrokes across 9 muscles, while the prediction accuracy for RDVM is relatively low at 69.6% (Figs. 3b and 3d). The reason for the low prediction accuracy of RDVM is one particular hawkmoth subject that showed a large variation in the percentage of times RDVM muscle spiking exactly twice. This percentage varies between 16% and 95% in the test data of the subject while the percentage in training data is about 40%. Even though spike prediction accuracy for SA, BA and L3Ax muscles seems to have increased for inter-trial predictions, there are roughly 1.5 times as many undesired spikes compared to the intra-trial predictions. Overall, across all muscles, an average inter-trial accuracy of 87% is achieved in spike event predictions. The muscles that have relatively low spike event prediction accuracy show a higher number of spikes per wingstroke burst and have more uniformly distributed spike count numbers. This is also the reason why R3Ax has the largest spike prediction error as it has the highest average number of spikes per wingstroke (or burst) and the lowest peak of the mode of the spike count distribution.

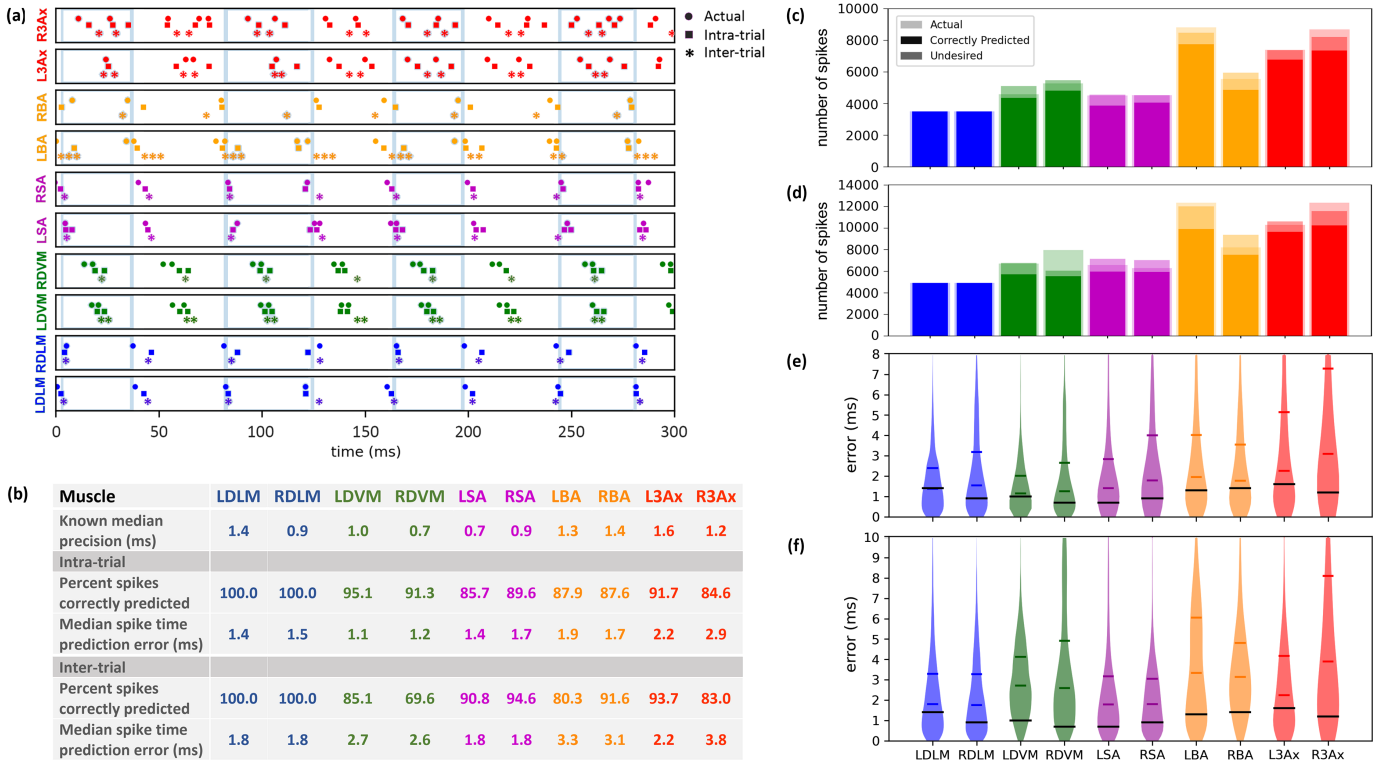


Fig. 3. (a) Spanning roughly 7 wingstrokes in time, a sample of actual and predicted spike trains for the 10 flight muscles. Circles correspond to the actual spike timings, squares to the intra-trial predictions and asterisks to the inter-trial predictions. Time intervals between transparent blue bars are wingstroke time periods. (b) Summary of the intra-trial (3506 wingstrokes) and inter-trial (4914 wingstrokes) prediction errors for each muscle. Known values of precision for each muscle [33] are also given for comparison with spike time prediction errors. (c)-(d) Bar plots of spike event prediction errors for the 10 muscles (c) intra-trial (3506 wingstrokes) (d) inter-trial (4914 wingstrokes). Opaque bars show correctly predicted spikes. Bars with the most transparency show the actual number of spikes. The difference between the heights of a most transparent bar and its corresponding opaque bar gives the number of missed spikes (false negatives). Lengths of bars with lesser transparency that are stacked on top of opaque bars show the number of undesired spikes (false positives). (e)-(f) Violin plots show the distributions of (e) intra-trial (3506 wingstrokes) and (f) inter-trial (4914 wingstrokes) spike time prediction errors across 10 muscles. Horizontal colored bars represent median values and 75% quantiles. Black horizontal bars correspond to the median values of known spike timing precision of each muscle [33].

### B. Spike timing predictions

Accurately predicting muscle spike timings down to the millisecond scale is important because timings contain nearly the entire information on coordination between flight muscles which drive and adjust wingstrokes with time periods of about 45 ms. [7]. These spike timings also contain most of the information on how the spikes are being modulated by the sensory stimulus. To measure the amount of modulation in the model, we raise the magnitude of RNN's input (latent feature array  $g_{j,k}$ ) step-wise by scaling it from 0 to 2 and evaluating the corresponding variation in  $\hat{t}_{j,k}$ . On average, the standard deviations of  $\hat{t}_{j,k}$  distributions gradually increase from  $0.006 \pm 0.002$  to  $0.078 \pm 0.008$  on scaling the RNN-input from 0 to 0.5. Then from 0.5 to 1, they stay nearly steady. But beyond this point, the deviations increase at a constant rate of about 0.1 wingstroke time periods per unit increase in the RNN-input scaling. This shows that the scaling  $g_{j,k}$  between 0.5 and 1 is enough to capture the actual magnitude of spike timing modulation. However, to determine the temporal precision of our model, we need to calculate the spike timings prediction errors at the actual values of  $g_{j,k}$ .

For the correctly predicted spikes, we calculate the median spike time prediction errors and show them in Fig. 3b. The prediction error distributions are shown in Fig. 3e. For intra-trial predictions, the errors were less than 2.2 ms in all cases except the R3Ax muscle where the error was 2.9 ms. These values of spike timing precision are close to the known precision for each muscle which varies from 0.7 to 1.6 ms (Fig. 3e – black horizontal bars) for comparison [33]. The known precision values are based on the estimates of the error tolerance of the continuous mutual information between spike timings of each muscle and yaw torque generated by the hawkmoth to respond to the stimulus [33]. The median errors for nine muscles are within 0.8 ms of the known median spike timing precision, while the R3Ax muscle is an outlier with a median prediction error of 1.7 ms more than the known precision. The mode of the error distribution of R3Ax is approximately 1.5 ms though, which is within 0.3 ms of the measured precision of R3Ax.

For inter-trial predictions, the median errors ranged between 1.8 and 3.8 ms across all muscles (Fig. 3b). These values of spike timing precision are within 2.6 ms of the known



precision values for each muscle (Fig. 3f – black horizontal bars) [33]. Inter-trial predictions are less accurate than intra-trial predictions because the animal’s sensorimotor state can change over time between trials [7], [32]. Overall, our spike timing prediction error distribution considerably overlaps with the known precision of the hawkmoth’s flight muscles [33]. Moreover, our timing predictions are also able to capture the distributions of timing differences in each contralateral pair of muscles, some of which have been shown to contribute to torque production during visuomotor yawing responses [6], [31] and also to compensate for wing damage [20].

#### IV. DISCUSSION

Our model captures spike count and timing codes observed in our hawkmoth flower tracking dataset. Intra-trial spike timing prediction errors are roughly 1-2 ms, while inter-trial errors are about a millisecond larger. Yet these errors closely coincide with the known precision values of each of the major flight muscles [33]. Same-trial predictions involve learning the motor state of the animal and making predictions within a 10-second window. However, between trials, errors might increase because predictions are made about the animal’s motor response with a gap of minutes or even a few hours. Over longer periods, the animal’s motor state may transition due to physiological changes, such as muscle fatigue. Therefore, for an interrupted execution of the behavior, adjustments to either the firing rates or the overall motor coordination across muscles might be required [34], [35].

some existing frameworks for generating precisely-timed spikes use SNNs whose outputs more closely represent neural signals. However, their capabilities are limited to feedforward or two-layered networks trained on simulated input data [13], [14]. We do not yet use a fully spiking network even though our output layers predict spike events as well as spike timings. This is sufficient to extend ANN predictability to the comprehensive motor program of moths but a fully spiking framework would better reflect the neural processing used throughout the animal’s sensorimotor circuits. This is an important future direction for our work. Advancing such models can deepen our understanding of how movement is maintained from cycle to cycle by modulations in rates and timings of these spikes based on sensory feedback from external as well as self-motion cues.

Compared to intra-trial predictions, our inter-trial spike timing prediction errors for different muscles increase by different amounts (larger increase in DVM and BA muscles) but the increase for each muscle within a contralateral pair is roughly the same. This means that in case a change in motor control strategy elicits slight changes in spike timings, for effective coordination the timing changes must be shared across multiple muscle pairs. Moreover, in each trial, each contralateral muscle pair has a distinct spiking pattern and timing precision value which slightly varies even within a pair. This is because each contralateral pair has a distinct functional role, each muscle is innervated by a different neuron [36] and the experimental preparation might now be perfectly

symmetric. The DLMs and the DVMs are predominantly wingstroke power muscles while the SAs, the BAs and the 3Axs are primarily used for steering [37], [38]. Within a wingstroke, the DLMs and the DVMs power the downstroke and the upstroke respectively so they all must spike every wingstroke to continuously drive the flapping of the wings. For this, they must encode sub-millisecond scale information individually and not just relative to one another [33]. This might be the reason why our intra-trial spike timing prediction errors for the power muscles are smaller than the three steering muscles. On the other hand, our intra-trial prediction errors for the 3Ax muscles are the largest. On average, their spike counts per wingstroke are also the highest. But they carry the same amount of information in their spike timings as the DVMs, the SAs and the BAs [7]. Having more spikes to carry the same amount of information reduces the timing precision requirement for the 3Ax muscles and thus they can still carry the essential information despite larger errors. Moreover, prediction errors later in the burst do not matter because physiologically a single spike error might be less likely to make a difference in a long burst as the calcium activation begins to saturate, although there can be exceptions based on biomechanical context [3], [39], [40].

Placed in the context of the dataset used, our model learns the motor program underlying a visuomotor response of the hawkmoth subject to a slowly moving floral target along a single axis – an apparently simple task involving yaw steering only. However, flight steering control is not a simple task for the nervous system because of the unstable dynamics of flight [41]. Control is executed using all muscles, not just a single or relatively few motor channels [7]. Information is processed in precise spike timing that does not simply linearly translate into forces and torques [7], [21]. Moreover, there is evidence that a hawkmoth’s muscle coordination patterns are conserved at the level of motor neuronal timings to the point that they can predict functionally distinct behaviors across different visual stimulus conditions [32]. Despite this, in the next steps, we plan to progressively shift towards recording the hawkmoth’s visuomotor response to complex motion patterns of targets in scenes with multiple moving objects where some objects may have more causal influences on the motor output than others.

Once combined with existing decoders [21] that translate coordinated motor spikes into forces, our model paves the path to deepen our understanding of neuromechanics in terms of causal effects of comprehensive and ubiquitous precise timing codes on motor control and locomotion. Incorporating temporally precise spiking mechanisms into neural network models provides an opportunity to understand control and coordination in neuronal networks. Moreover, a spike-based control framework that captures temporal precision and coordination may also reveal how such architectures expand the control capabilities of agile animals beyond simple tasks. They might help us discover how mathematically simple biomechanical control laws emerge from underlying complex spiking patterns as well as drive the next generation of neuromorphic controllers to navigate in dynamic and uncertain environments.

## V. CONCLUSION

We show that a simple RNN interfaced with an elementary model of an animal's vision is sufficient to predict a comprehensive, coordinated and temporally precise spike-resolved motor response that was experimentally measured. It is comprehensive in terms of the encoded information for yaw steering motor control in a hawkmoth and is coordinated across 10 major flight muscles using millisecond-precise motor spikes modulated by sensory processing in a visual system. In addition, our predictions match the natural spike timing precision observed for each muscle [33].

## FUNDING AND ACKNOWLEDGMENTS

Funding was provided by AFOSR grants FA9550-19-1-0396 to S.F. and S.S., and FA9550-22-1-0315 to S.S., S.F. and H.C. We would like to thank Varun Sharma, Leo Wood, Ethan Wold, Ellen Liu and Zach Mobbille for their invaluable feedback.

## REFERENCES

- [1] S. D. Prentice, A. E. Patla, and D. A. Stacey, "Artificial neural network model for the generation of muscle activation patterns for human locomotion," *J. Electromyogr. Kinesiol.*, vol. 11, no. 1, pp. 19–30, 2001.
- [2] D. H. Edwards, "Neuromechanical simulation," *Front. Behav. Neurosci.*, vol. 4, no. JUL, pp. 1–7, 2010.
- [3] S. J. Sober, S. Sponberg, I. Nemenman, and L. H. Ting, "Millisecond Spike Timing Codes for Motor Control," *Trends Neurosci.*, vol. 41, pp. 644–648, 2018.
- [4] K. Nishikawa, A. A. Biewener, P. Aerts, A. N. Ahn, H. J. Chiel, M. A. Daley, T. L. Daniel, R. J. Full, M. E. Hale, T. L. Hedrick, A. K. Lappin, T. R. Nichols, R. D. Quinn, R. A. Satterlie, and B. Szymik, "Neuromechanics: An integrative approach for understanding motor control," *Integr. Comp. Biol.*, vol. 47, no. 1, pp. 16–54, 2007.
- [5] K. H. Srivastava, C. M. Holmes, M. Vellema, A. R. Pack, C. P. Elemans, I. Nemenman, and S. J. Sober, "Motor control by precisely timed spike patterns," *Proc. Natl. Acad. Sci.*, vol. 114, no. 5, pp. 1171–1176, 2017.
- [6] S. Sponberg and T. L. Daniel, "Abdicating power for control: A precision timing strategy to modulate function of flight power muscles," *Proc. R. Soc. B Biol. Sci.*, vol. 279, no. 1744, pp. 3958–3966, 2012.
- [7] J. Putney, R. Conn, and S. Sponberg, "Precise timing is ubiquitous, consistent, and coordinated across a comprehensive, spike-resolved flight motor program," *PNAS*, vol. 116, no. 52, pp. 26951–26960, 2019.
- [8] C. Tang, D. Chehayeb, K. Srivastava, I. Nemenman, and S. J. Sober, "Millisecond-scale motor encoding in a cortical vocal area," *PLOS Biology*, vol. 12, pp. 1–13, 12 2014.
- [9] C. R. Von Reyn, P. Breads, M. Y. Peek, G. Z. Zheng, W. R. Williamson, A. L. Yee, A. Leonardo, and G. M. Card, "A spike-timing mechanism for action selection," *Nat. Neurosci.*, vol. 17, no. 7, pp. 962–970, 2014.
- [10] M. Enquist and S. Ghirlanda, *Neural Networks and Animal Behavior*. Princeton University Press, 2006.
- [11] K. Kumarasinghe, N. Kasabov, and D. Taylor, "Brain-inspired spiking neural networks for decoding and understanding muscle activity and kinematics from electroencephalography signals during hand movements," *Sci. Rep.*, vol. 11, no. 1, pp. 1–15, 2021.
- [12] M. Gharagozloo, A. Amrani, K. Wittingstall, A. Hamilton-Wright, and D. Gris, "Machine Learning in Modeling of Mouse Behavior," *Front. Neurosci.*, vol. 15, no. September, pp. 1–11, 2021.
- [13] F. Ponulak and A. Kasiński, "Supervised learning in spiking neural networks with ReSuMe: Sequence learning, classification, and spike shifting," *Neural Comput.*, vol. 22, no. 2, pp. 467–510, 2010.
- [14] A. Mohammed, S. Schliebs, and N. Kasabov, "SPAN: A neuron for precise-time spike pattern association," in *Neural Information Processing*, vol. 7063 LNCS, pp. 718–725, Springer Berlin Heidelberg, 2011.
- [15] R. V. Florian, "The chronotron: A neuron that learns to fire temporally precise spike patterns," *PLoS One*, vol. 7, pp. 1–27, 08 2012.
- [16] R. M. Memmesheimer, R. Rubin, B. P. Ölveczky, and H. Sompolinsky, "Learning Precisely Timed Spikes," *Neuron*, vol. 82, no. 4, pp. 925–938, 2014.
- [17] D. Sussillo, "Neural circuits as computational dynamical systems," *Curr. Opin. Neurobiol.*, vol. 25, pp. 156–163, 2014.
- [18] S. Sponberg, J. P. Dyhr, R. W. Hall, and T. L. Daniel, "Luminance-dependent visual processing enables moth flight in low light," *Science* (80-. ), vol. 348, no. 6240, pp. 1245–1248, 2015.
- [19] J. D. Sprayberry and T. L. Daniel, "Flower tracking in hawkmoths: Behavior and energetics," *J. Exp. Biol.*, vol. 210, no. 1, pp. 37–45, 2007.
- [20] M. J. Fernandez, D. Springthorpe, and T. L. Hedrick, "Neuromuscular and biomechanical compensation for wing asymmetry in insect hovering flight," *J. Exp. Biol.*, vol. 215, no. 20, pp. 3631–3638, 2012.
- [21] H. Yang, J. Putney, U. B. Sikandar, P. Zhu, S. Sponberg, and S. Ferrari, "A relative spike-timing approach to kernel-based decoding demonstrated for insect flight experiments," in *2022 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, 2022.
- [22] T. L. Hedrick, "Software techniques for two- and three-dimensional kinematic measurements of biological and biomimetic systems," *Bioinspir. Biomim.*, vol. 3, no. 3, 2008.
- [23] J. C. Theobald, E. J. Warrant, and D. C. O'Carroll, "Wide-field motion tuning in nocturnal hawkmoths," *Proc. R. Soc. B Biol. Sci.*, vol. 277, no. 1683, pp. 853–860, 2009.
- [24] A. L. Stöckl, W. A. Ribi, and E. J. Warrant, "Adaptations for nocturnal and diurnal vision in the hawkmoth lamina," *J. Comp. Neurol.*, vol. 524, no. 1, pp. 160–175, 2016.
- [25] E. W. Weisstein, "Gnomonic projection." <https://mathworld.wolfram.com/GnomonicProjection.html>. Accessed: Jan 2023.
- [26] P. Chatterjee, U. Mohan, A. Krishnan, and S. P. Sane, "Evolutionary constraints on flicker fusion frequency in Lepidoptera," *J. Comp. Physiol. A Neuroethol. Sensory, Neural, Behav. Physiol.*, vol. 206, no. 5, pp. 671–681, 2020.
- [27] J. Gemerek, "Active vision and perception," *Ph.D. Thesis*, 2020.
- [28] A. Bouzerdoun, "The elementary movement detection mechanism in insect vision," *Philos. Trans. R. Soc. London. Ser. B Biol. Sci.*, vol. 339, no. 1290, pp. 375–384, 1993.
- [29] F. G. Barth, J. A. C. Humphrey, and M. V. Srinivasan, *Frontiers in Sensing, From Biology to Engineering*. SpringerNewYork, 2012.
- [30] E. J. Warrant, "The remarkable visual capacities of nocturnal insects: Vision at the limits with small eyes and tiny brains," *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 372, no. 1717, 2017.
- [31] S. Sponberg, T. L. Daniel, and A. L. Fairhall, "Dual dimensionality reduction reveals independent encoding of motor features in a muscle synergy for insect flight control," *PLoS Comput. Biol.*, vol. 11, no. 4, pp. 1–23, 2015.
- [32] J. Putney, M. Angelichinoski, R. Ravier, S. Ferrari, V. Tarokh, and S. Sponberg, "Consistent coordination patterns provide near perfect behavior decoding in a comprehensive motor program for insect flight," *bioRxiv*, 2021.
- [33] J. Putney, T. Niebur, R. Conn, and S. Sponberg, "An information theoretic method to resolve millisecond-scale spike timing precision in a comprehensive motor program," *bioRxiv*, 2021.
- [34] S. Binder-Macleod and T. Kesar, "Catchlike property of skeletal muscle: Recent findings and clinical implications," *Muscle and Nerve*, vol. 31, no. 6, pp. 681–693, 2005.
- [35] P. Contessa, J. Letizi, G. De Luca, and J. C. Kline, "Contribution from motor unit firing adaptations and muscle coactivation during fatigue," *J. Neurophysiol.*, vol. 119, no. 6, pp. 2186–2193, 2018.
- [36] M. B. Rheuben, "Quantitative comparison of the structural features of slow and fast neuromuscular junctions in Manduca," *J. Neurosci.*, vol. 5, no. 7, pp. 1704–1716, 1985.
- [37] A. E. Kammer, "Motor output during turning flight in a hawkmoth, *Manduca sexta*," *J. Insect Physiol.*, vol. 17, no. 6, pp. 1073–1086, 1971.
- [38] A. E. Kammer, "Flying," in *Comprehensive insect physiology, biochemistry and pharmacology*, pp. 491–552, Oxford, UK: Pergamon Press, 1985.
- [39] S. Sponberg, A. J. Spence, C. H. Mullens, and R. J. Full, "A single muscle's multifunctional control potential of body dynamics for postural control and running," *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 366, no. 1570, pp. 1592–1605, 2011.
- [40] L. H. Ting and H. J. Chiel, "Muscle, Biomechanics, and Implications for Neural Control," in *Neurobiology of Motor Control: Fundamental Concepts and New Directions*, ch. 12, pp. 365–416, John Wiley & Sons, Inc, 2017.
- [41] M. Sun, J. Wang, and Y. Xiong, "Dynamic flight stability of hovering insects," *Acta Mech. Sin.*, vol. 23, no. 3, pp. 231–246, 2007.