

# Correlations in state space can cause sub-optimal adaptation of optimal feedback control models

Jonathan Aprasoff · Opher Donchin

Received: 21 November 2010 / Revised: 29 April 2011 / Accepted: 20 June 2011  
© The Author(s) 2011. This article is published with open access at Springerlink.com

**Abstract** Control of our movements is apparently facilitated by an adaptive internal model in the cerebellum. It was long thought that this internal model implemented an adaptive inverse model and generated motor commands, but recently many reject that idea in favor of a forward model hypothesis. In theory, the forward model predicts upcoming state during reaching movements so the motor cortex can generate appropriate motor commands. Recent computational models of this process rely on the optimal feedback control (OFC) framework of control theory. OFC is a powerful tool for describing motor control, it does not describe adaptation. Some assume that adaptation of the forward model alone could explain motor adaptation, but this is widely understood to be overly simplistic. However, an adaptive optimal controller is difficult to implement. A reasonable alternative is to allow forward model adaptation to ‘re-tune’ the controller. Our simulations show that, as expected, forward model adaptation alone does not produce optimal trajectories during reaching movements perturbed by force fields. However, they also show that re-optimizing the controller from the forward model can be sub-optimal. This is because, in a system with state correlations or redundancies, accurate prediction requires different information than optimal control. We find that adding noise to the movements that matches noise found in human data is enough to overcome this

problem. However, since the state space for control of real movements is far more complex than in our simple simulations, the effects of correlations on re-adaptation of the controller from the forward model cannot be overlooked.

**Keywords** Optimal control · Motor adaptation · Forward model · Reaching movements

## 1 Introduction

Behavioral (Catz et al. 2005; Morton and Bastian 2004, 2006; Frens and Donchin 2009; Rabe et al. 2009), anatomical (Dean et al. 2010), and physiological (Robinson 1995) evidence supports the view that the cerebellum plays the central role in adaptation, possibly by implementing an adaptive filter (Miall et al. 1993). While the traditional view has been that the cerebellum implements an inverse model—generating motor commands or supplementing the motor commands generated in other brain areas (Kawato and Gomi 1992)—an alternative view has recently gained popularity. It holds that the cerebellum calculates a forward model that predicts the sensory consequences of our motor commands (Miall et al. 1993). Forward model output is used to supplement information provided by sensory afferents, which may be either noisy or subject to feedback delays. It has long been hypothesized that both types of model are generated by different parts of the cerebellum for different purposes. However, some recent studies of visuomotor adaptation suggest that motor adaptation specifically reflects changes in a forward model (Mazzoni and Krakauer 2006; Tseng et al. 2007). This is consistent with the hypothesis that the function which is learned during cerebellar-dependent adaptation is a forward model (Shadmehr and Krakauer 2008).

Action Editor: Eberhard Fetz

**Electronic supplementary material** The online version of this article (doi:10.1007/s10827-011-0350-z) contains supplementary material, which is available to authorized users.

J. Aprasoff · O. Donchin (✉)  
Department of Biomedical Engineering,  
Ben-Gurion University of the Negev,  
Beersheba, Israel  
e-mail: donchin@bgu.ac.il  
URL: <http://www.bgu.ac.il/~donchin>

However, a forward model has only an indirect influence on motor commands. Will improved prediction necessarily lead to optimal motor control? Our research explores this question. We built an optimal feedback controller (OFC, Todorov and Jordan 2002) with an adaptive forward model in order to test the effects of forward model adaptation on the overall behavior of the controller. We use the OFC model because it has been successful in describing a number of psychophysical results associated with cerebellar-dependent adaptation. These include the curved trajectories humans make after adapting to force field perturbations in reaching movements (Shadmehr and Krakauer 2008) and synergetic control of bimanual movements (Diedrichsen 2007). The OFC framework has also been specifically mapped onto the circuitry of the motor system, with the cerebellum assigned the role of the forward model (Frens and Donchin 2009; Shadmehr and Krakauer 2008).

To explore the role of forward model adaptation in OFC, we extended the classical OFC framework to allow adaptation using gradient descent. We first demonstrated that adaptation of the forward model is not sufficient to produce optimal movements. This is widely accepted: unless the motor commands are changed to fit the new dynamics then prediction is insufficient to produce optimal movements. An alternative idea with more support is that adaptation of predictive estimates can be used simply to re-optimize the controller that generates movements. We tested this by allowing the controller to ‘learn from the cerebellar model,’ as suggested by Gomi and Kawato (1992). We show that this strategy can also fail to achieve optimal control if there are correlations in state space variables. This is specifically an issue in situations like human motor control where there are many redundancies in the effectors and where movement is often confined to a small subspace of the possible work space.

However, when we reduced correlations by increasing the noise in the model or adding more movement directions, the performance of the adapted forward model became very close to optimal performance. Thus, our modeling leaves open the question of whether the degree of variability in human movements is generally sufficient to resolve the true dynamics during forward model adaptation. What we do show is that this question is of central importance in the ability of the OFC framework to explain human motor adaptation.

## 2 Methods

### 2.1 Analysis of human data

A subset of the data previously collected for the study published in Donchin and Shadmehr (2004) was analyzed

in order to determine the amount of variance in human movements and the degree of correlation between different state variables. The methods are completely described in the earlier paper. Briefly, the subjects performed a standard curl field paradigm described making 10 cm movements in one direction (straight away from the body, the Y direction) with the robot assisting in passive return of the hand to the starting position after each movement. Data was analyzed from 4 subjects who each performed 3 sets of 150 movements with no catch trials. We calculated the correlation of X and Y position and X and Y velocity across all time steps of all movements for each subject. We also calculated, for each subject, the mean of the X and Y starting position across all 450 trials and the mean of the starting velocity taken across the first 30 ms of movement in each of the 450 trials. Finally, we calculated the covariance over all four of these variables.

### 2.2 Task

We created a model that was capable of accomplishing a simulated simplified reaching task in a two dimensional space (Todorov and Jordan 2002). The model end effector began movements either from the point (0,) or from a starting position and velocity that were chosen from a normal distribution with a mean and covariance structure matched to the average mean and covariance in starting movements and velocities in human movements. It was then required to move to a target located at  $Y=10$  cm. Each trial was composed of 100 discrete time steps (0.01 s each), in which the controller generated a motor command  $\mathbf{u}$  and the state of the plant was updated in response to the motor command and plant dynamics:

$$\mathbf{x}_{n+1} = \mathbf{A}\mathbf{x}_n + \mathbf{B}\mathbf{u}_n + \mathbf{w}_n, \mathbf{w}_n \sim N(0, \Sigma_w) \quad (1)$$

Where  $\mathbf{x}_n$  is the state vector at time step  $n$ ,  $\mathbf{u}_n$  is the command vector at time step  $n$  (a force that affects acceleration),  $\mathbf{A}$  is the dynamics matrix,  $\mathbf{B}$  is the command matrix, and  $\mathbf{w}_n$  is the process noise (distributed normally with covariance matrix  $\Sigma_w$ )

Simulated movements were made either in free space (null field) or in a perturbing curl force field (Shadmehr and Mussa-Ivaldi 1994). In the null field, the only external force acting on the plant was caused by the motor command. In the curl field, an external force perpendicular to the instantaneous velocity of the end effector is added to the motor command. Its magnitude is proportional to speed ( $\mathbf{F}_{ext} = \mathbf{C} \cdot \mathbf{v}$ ,  $\mathbf{C} = [0, -b; b, 0]$ ). Curl fields are commonly used in experiments on adaptation of reaching movements (Shadmehr et al. 2005), and human subjects are able to adapt to curl fields: trajectories gradually shift to curved movements which are optimal for the curl field (Izawa et al. 2008).

### 2.3 Cost

An analytic solution for the generation of control signals in an OFC controller can be found only when the cost function has a quadratic form. Hence, as is common in such simulations, we used a cost function that was a quadratic function of the state and command:

$$C = \sum_{n=1}^N \mathbf{x}_n^T \mathbf{Q}_n \mathbf{x}_n + \mathbf{u}_n^T \mathbf{R} \mathbf{u}_n \tag{2}$$

The state dependent cost matrix (Q) penalized distance of the end effector from the target position and velocity. The penalty was imposed only in the final 20% (200 msec) of the trial. To minimize state dependent cost, the controller must reach the target within 800 msec and bring the end effector to a stop. The action dependant cost matrix (R) is constant and diagonal, so it simply penalized effort. This meant that movements were made using as little force as possible. An optimal movement is one which reaches the target in 800 msec with as little effort as possible, and stays there. In total, without using matrix notation, the cost was

$$C = \alpha_{pos} (x_{pos} - x_{tar})^2 + \alpha_{vel} x_{vel}^2 + \alpha_{action} u^2 \tag{3}$$

with

$$\alpha_{vel} = \begin{cases} 100, & \text{if } T > 800\text{m sec} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

$$\alpha_{vel} = \begin{cases} 0.2, & \text{if } T > 800\text{m sec} \\ 0 & \text{otherwise} \end{cases}$$

$$\alpha_{action} = \frac{10^{-5}}{N}$$

### 2.4 Plant

The plant that was used for all simulations was a point mass. The plant’s dynamics were inertial; in addition, forces generated by both the motor command and the force field drove the inertial dynamics.

The state vector was defined as  $\mathbf{x} = [p_x \ p_y \ v_x \ v_y \ a_x \ a_y \ T_x \ T_y]^T$  (with  $p_x$  and  $p_y$  representing position in the  $x$  and  $y$  coordinates respectively; and  $v_\bullet$ ,  $a_\bullet$  and  $T_\bullet$  representing the coordinates of velocity, acceleration, and target position).

The state was calculated at every time step according to the dynamics equation (Eq. (1)), using the dynamics matrix  $\mathbf{A}$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & dt & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & dt & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & dt & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & dt & 0 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 0 & 0 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{5}$$

and the command matrix  $\mathbf{B}$

$$\mathbf{B} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{6}$$

Thus motor commands were applied as forces that affect acceleration. When the curl field is applied, the dynamics matrix changes to:

$$\mathbf{A}_{CF} = \begin{bmatrix} 1 & 0 & dt & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & dt & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & dt & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & dt & 0 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-2} & 0 & 0 & 0 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{2} & \mathbf{0} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{7}$$

The difference between the dynamics matrices in Eqs. (5) and (7) are in bold face. They describe the effects of position and velocity on acceleration.

Feedback,  $\mathbf{y}_n$ , was returned from the plant to the control system according to the equation:

$$\mathbf{y}_n = \mathbf{H} \cdot \mathbf{x}_n + \boldsymbol{\eta}, \boldsymbol{\eta} \sim N(\mathbf{0}, \Sigma_\eta) \tag{8}$$

Where  $\mathbf{H}$  is the observation matrix, and  $\boldsymbol{\eta}$  is the observation noise, normally distributed with zero mean and covariance  $\Sigma_\eta$ . The observation matrix,  $\mathbf{H}$  is diagonal in the first six rows, so the system receives feedback regarding the actual position, velocity and acceleration in both the  $x$  and  $y$  coordinates.

### 2.5 Simulations

All simulations were run using MATLAB 7.6.0 (MathWorks, Natick, MA). In order to implement adaptation in different parts of the OFC model, we had to construct our software in such a way that every part of the OFC model could be altered and manipulated independently. We used OFC code made publicly available by Emanuel Todorov (<http://www.cs.washington.edu/homes/todorov/software.htm>) as a framework but modified the code such that different parts of the model could be made adaptive. The modified code is available at <http://www.mll.org.il/AprasoffForwardModel>.

We first modified the code to allow adaptation of the forward model. We will denote the output of the forward model  $\hat{\mathbf{x}}_{n+1|n}$ , meaning the prediction of the state on trial  $n+1$  given sensory information from time  $n$ . Inputs to the forward model are the current estimated state  $\hat{\mathbf{x}}_{n|n}$  (the estimate of state  $n$

given sensory information from time  $n$ ; the output of the state estimator) and the current command  $\mathbf{u}_n$ . The forward model then calculates the a-priori estimated next state using its estimate of the dynamics matrix ( $\hat{\mathbf{A}}$ ) and the command matrix ( $\hat{\mathbf{B}}$ ):

$$\hat{\mathbf{x}}_{n+1|n} = \hat{\mathbf{A}}\hat{\mathbf{x}}_{n|n} + \hat{\mathbf{B}}\mathbf{u}_n \tag{9}$$

Forward model adaptation was based on two assumptions: (1) the forward model is implemented by the cerebellum (Shadmehr and Krakauer 2008), and (2) the cerebellum is involved in supervised learning (Doya et al. 2001). Therefore, learning can be approximated as a process of gradient descent. Indeed, some analytical models of behavior have suggested that adaptation is well modeled as a process of gradient descent (Donchin et al. 2003; Hwang et al. 2003). Thus, we used gradient descent learning on the prediction error in order to adapt the forward model in our simulation.

Forward model estimated dynamics  $\hat{\mathbf{A}}$  were changed after every trial. The prediction error in a certain trial was defined as:

$$E = \sum_{n=1}^N (\mathbf{y}_n - \mathbf{H}\hat{\mathbf{x}}_{n|n-1})^T \cdot (\mathbf{y}_n - \mathbf{H}\hat{\mathbf{x}}_{n|n-1}) \tag{10}$$

To minimize this error,  $\hat{\mathbf{A}}$  must be changed in order to go down the gradient of the error with respect to the estimated dynamics:

$$\begin{aligned} \frac{\partial}{\partial \hat{\mathbf{A}}} E &= \frac{\partial}{\partial \hat{\mathbf{A}}} \left( \sum_{n=1}^N (\mathbf{y}_n - \mathbf{H}\hat{\mathbf{x}}_{n|n-1})^T \cdot (\mathbf{y}_n - \mathbf{H}\hat{\mathbf{x}}_{n|n-1}) \right) = \\ &= \frac{\partial}{\partial \hat{\mathbf{A}}} \sum_{n=1}^N \left( \mathbf{y}_n^T \mathbf{y}_n - \mathbf{y}_n^T \mathbf{H}\hat{\mathbf{x}}_{n|n-1} - \hat{\mathbf{x}}_{n|n-1}^T \mathbf{H}^T \mathbf{y}_n + \hat{\mathbf{x}}_{n|n-1}^T \mathbf{H}^T \mathbf{H}\hat{\mathbf{x}}_{n|n-1} \right) \\ &= \frac{\partial}{\partial \hat{\mathbf{A}}} \sum_{n=1}^N \left( \mathbf{y}_n^T \mathbf{y}_n - \mathbf{y}_n^T \mathbf{H}(\mathbf{A}\hat{\mathbf{x}}_{n-1} + \mathbf{B}\mathbf{u}_{n-1}) - (\mathbf{A}\hat{\mathbf{x}}_{n-1} + \mathbf{B}\mathbf{u}_{n-1})^T \mathbf{H}^T \mathbf{y}_n + \right. \\ &\quad \left. (\mathbf{A}\hat{\mathbf{x}}_{n-1} + \mathbf{B}\mathbf{u}_{n-1})^T \mathbf{H}^T \mathbf{H}(\mathbf{A}\hat{\mathbf{x}}_{n-1} + \mathbf{B}\mathbf{u}_{n-1}) \right) \\ &= 2 \left( \mathbf{H}^T \sum_{n=1}^N \mathbf{y}_n \hat{\mathbf{x}}_{n-1|n-1}^T - \mathbf{H}^T \mathbf{H} \mathbf{A} \sum_{n=1}^N \hat{\mathbf{x}}_{n-1|n-1} \hat{\mathbf{x}}_{n-1|n-1}^T - \mathbf{H}^T \mathbf{H} \mathbf{B} \sum_{n=1}^N \mathbf{u}_{n-1} \hat{\mathbf{x}}_{n-1|n-1}^T \right) \end{aligned} \tag{11}$$

Hence, after every trial, we calculated the gradient and changed the estimated dynamics matrix,  $\hat{\mathbf{A}}$ , by a fraction of the gradient in the direction that reduced the error. The steps were always 0.005 of the gradient. We let the forward model make the a priori assumption that only the forces acting on the plant were likely to change and that the Newtonian laws relating acceleration to velocity and velocity to position were likely to remain fixed. This meant that we confined adaptation to the rows of the dynamics matrix that calculate the force as a function of the other state variables (indicated with a bold font in Eqs. (5) and (7)).

We then modified the state estimator to make it partially adaptive. The altered state estimator is the optimal adaptive Kalman filter (see Eq. (1) in [online resource 1](#)), as suggested in Todorov (2005). The optimal filter needs to be adaptive because the size of the signal generated on a specific movement affects the amount of noise in the state

and, thus, the variance of our state estimate. However, the adaptive Kalman filter still relies on the known system dynamics for the estimation of the variances. Hence, adaptation of the Kalman filter will not prevent unknown dynamics for biasing state estimates.

The controller remained the same as in the original Todorov formulation. Its input is the current estimated state of the plant, as estimated by the Kalman filter:

$$\hat{\mathbf{x}}_{n+1|n+1} = \hat{\mathbf{x}}_{n+1|n} + \hat{\mathbf{A}} \cdot \mathbf{K}_n (\mathbf{y}_n - \mathbf{H} \cdot \hat{\mathbf{x}}_{n|n}) \tag{12}$$

Note that we altered the Kalman filter somewhat, by taking the dynamics matrix multiplication out of the Kalman gain and writing it separately. This is purely notational, but it enables us to change the estimated dynamics used by the Kalman filter without re-calculating the Kalman gains. The controller’s output is the command to the plant. The command is calculated linearly from the state according to the equation:

$$\mathbf{u}_n = -\mathbf{L} \cdot \hat{\mathbf{x}}_n \tag{13}$$

The controller gain matrix  $\mathbf{L}$  is the optimal feedback control gains as calculated in Todorov (2005), using the estimated dynamics  $\hat{\mathbf{A}}$  and  $\hat{\mathbf{B}}$  in Eq. (8).

We ran three types of adaptation simulation. All simulations were run for 800 trials. In the first type, only the forward model was adapted. At the beginning of the simulation, the OFC model was optimized to work in a null field (Eq. (5)) although the actual dynamics were curl force field dynamics (Eq. (7)). Throughout the simulation, the dynamics driving the controller and state estimator were kept constant and only the forward model was allowed to adapt.

In a second type of simulation, we used what we called a shared internal model. We began, as before, with a model optimized for the null field dynamics. Again the real dynamics were the curl force field dynamics, and the forward model adapted using gradient descent. However, in the shared internal model simulations the controller and state estimators are re-optimized every 15 trials, according to the dynamics learned by the forward model ( $\hat{\mathbf{A}}, \hat{\mathbf{B}}$  in Eq. (8)). Thus, the controller and state estimator share the internal model learned by the forward model.

We ran three variations of this simulation. The first variation had a fixed starting point and 0 velocity at the beginning of movements, and all movements were made towards the same target. In the second variation, the starting point and initial velocity were distributed normally with a mean and covariance matched to the mean and covariance in human movements, as described above. The third variation returned to the fixed starting point and 0 initial velocity, but there were nine different targets distributed uniformly along an arc with a radius of 0.1 m and running from the positive x axis, 180° to the negative x axis. The

targets thus covered the range of 0–180°. Target order was selected randomly.

In the third type of simulation, we used a linear approximation of controller adaptation, in order to examine the behavior of the system when the controller and forward model are adapted independently. In OFC theories of motor control, the controller is thought to be located outside of the cerebellum and to learn through unsupervised or reinforcement learning (Shadmehr and Krakauer 2008). In order to avoid the complexities of such adaptive controllers, we had the controller make a linear transition from the optimal controller and state estimator for the null field condition to the optimal controller and state estimator for the curl field:

$$\mathbf{L}_k = \frac{(K - k) \cdot \mathbf{L}_{Null} + k \cdot \mathbf{L}_{FF}}{K} \quad (14)$$

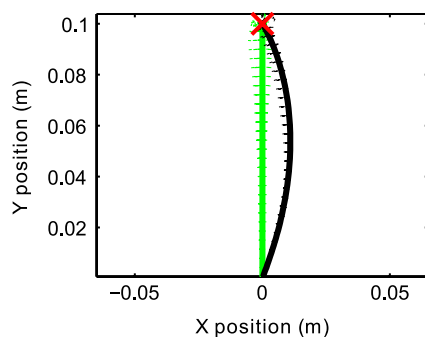
Here,  $k$  is the trial number and  $K$  is the number of trials. The forward model had been adapted using gradient descent, as in the other simulations. The transition from one optimal controller to the other was made using simple linear interpolation between the two sets of parameters.

For a full list of parameters and values used in the simulations, please refer to [online resource 2](#).

### 3 Results

#### 3.1 Optimal trajectories

In order to have some optimal performance to which we can compare the results obtained from the adaptive models, we simulated the reaching movements using an OFC model which was optimized with full knowledge of the dynamics. As depicted in (Fig. 1) the optimal trajectory in a null-field environment is a straight line, and the optimal trajectory in



**Fig. 1** System behavior with perfect knowledge of the dynamics. The optimal trajectory, as calculated by the OFC model, without a force field (*solid green*) is straight, as are the noisy simulations (*dotted green*). However, the optimal trajectory (*solid black*) and noisy simulations (*dotted black*), as calculated by the OFC model in a curl force field are curved to the opposite direction of the force

the curl field is curved to the side opposite the direction of the force, as shown previously (Izawa et al. 2008).

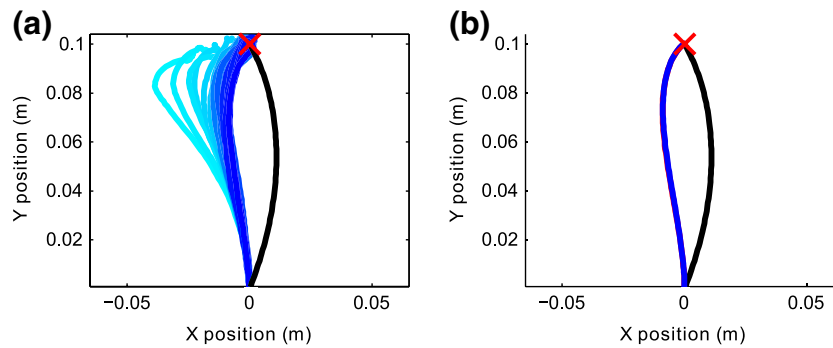
#### 3.2 First simulation: forward model adaptation

In the first simulation, we added an adaptive forward model to a standard OFC system. The controller and state estimator were optimized for null field dynamics, and the actual dynamics of the plant included a curl field. The forward model adapted by using gradient descent in order to minimize its prediction error. Initially, the model was unable to control the plant correctly and did not reach the target (Fig. 2(a)). As the forward model adapted, the model’s performance improved. After about 150 trials, movements started to reach the target. With improved estimation of the dynamics, prediction errors were reduced, adaptation slowed, and the forward model converged. However, the resulting trajectory was far from optimal (Fig. 2(a), black line). Adaptation of the forward model produced a trajectory that was similar to one produced using an OFC framework where the controller has been optimized for the null field and the forward model has perfect knowledge of dynamics (Fig. 2(b), red line). Hence, the adaptive forward model, coupled with an ignorant controller, reached the same level of performance as a perfect forward model, with the same controller. The resulting trajectory and the fact that prediction error was reduced to noise levels (Fig. 3) led us to conclude that the adaptive forward model provided adequate state predictions, and the sub-optimal result must have a different explanation. This result is expected. Since the controller has not been optimized, it continues to produce non-optimal commands regardless of the accuracy of its state information. In order to adapt an OFC model to force field perturbations, one must also permit change in the controller and state estimator.

#### 3.3 Second simulation: shared internal model

Our second simulation explored the possibility that the controller is updated according to the dynamics learned by the forward model. We did this by re-optimizing the controller and state estimator every 15 trials according to the dynamics learned by the forward model. The forward model adapted using gradient descent, as before. As in the first simulation, the model was initially unable to reach the target, and performance improved as a result of adaptation (Fig. 4(a)). Ongoing controller reoptimization led to a curved trajectory, resembling the optimal trajectory. However, the final trajectory was not the same as the optimal trajectory (Fig. 4(b)). The trajectory remained suboptimal even when learning was continued for many more movements.

This seems to show that the dynamics learned by the forward model are incorrect. Had the forward model learned the dynamics perfectly, there would be no differ-



**Fig. 2** Simulation 1, adapting to force field perturbation by altering the forward model. **(a)** Simulated trajectories during force field adaptation. Force field pushes the plant to the left. First trial (*light blue*) does not reach the target (*red X*). Last trial (no. 800—*dark blue*) reaches the target, yet is far from the optimal trajectory (*black*). **(b)** Average

trajectories—without noise. The adaptive model’s trajectory (*blue*) is very close to the optimal trajectory of a system using a controller that is ignorant of the force field, and a forward model with perfect knowledge of the force field (*red*, almost completely obscured by the overlying *blue*). However, both are far from the optimal trajectory (*black*)

ence between the adapted model and the optimal force field model. However, this contradicts our earlier conclusion that low prediction errors imply that the learned dynamics were correct. In order to resolve the issue, we looked at the rows of the dynamics matrix that had been learned by the forward model.

$$\hat{\mathbf{A}}_{CF,Partial} = \begin{bmatrix} 0.02 & -0.15 & -0.14 & -1.9 \\ 0.004 & -0.06 & 0.49 & 0.08 \end{bmatrix} \quad (15)$$

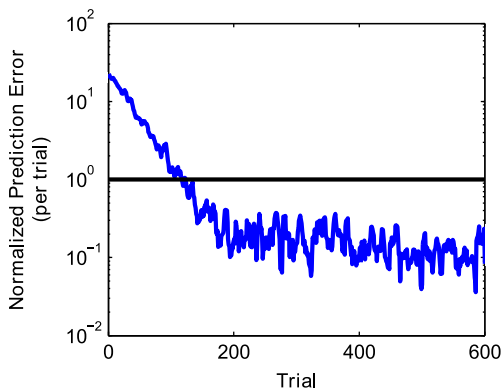
If we compare the learned dynamics to the relevant rows of the actual dynamics matrix (bold in Eq. (7)) we can see that several of the terms have been mis-estimated. This is especially true for the dependence of the force in the y direction on x velocity (0.43 instead of 2.0) and the dependence of the force in the x direction on y position and x velocity (−0.15 and −0.14 respectively instead of 0 for both). Thus, learned dynamics were incorrect. Correlation between state variables in our task led the adaptive forward model to attribute force incorrectly to both position and velocity. The first clue to this came from testing for

correlations in the forward model’s estimates of the dynamics. We ran the second simulation fifty times and looked at the distribution of values in the adapted terms of the dynamics matrix. We found correlations in the estimates of several quantities, as shown in Fig. 5. The two correlations that were significant after a Holm correction for 28 tests were between the estimates of the effect of x-position and x-velocity on force in the y direction ( $r=-0.48$ ;  $p=4e-4$ ) and similarly between the effects of x-position and y-velocity on force in the direction ( $r=0.41$ ;  $p=0.002$ ). We found that there was a clear relation between the estimated effect of the x-position and y-velocity on the force (Pearson’s  $r: -0.35$ ,  $p=0.012$ ). Similarly, a correlation exists between the effects of x-velocity and x-position on the force in the x direction (Pearson’s  $r: 0.31$ ,  $p=0.026$ ).

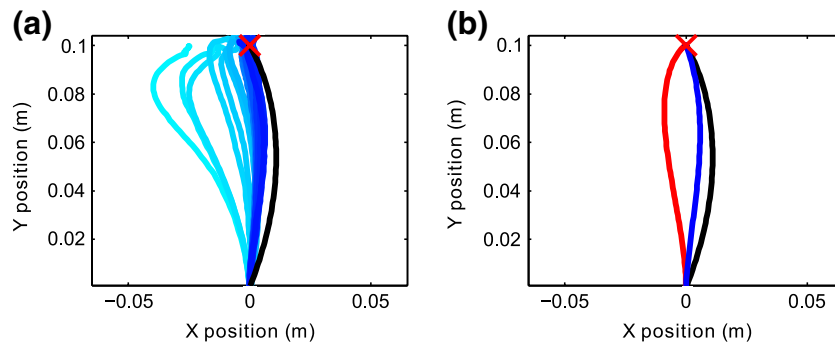
The correlations in the forward model estimates of the dynamics arise because there is no way to resolve which of two correlated state variables is actually related to force. The first four rows of Table 1 show the correlation among state variables in the movements made by the simulation. X position is highly correlated with x velocity ( $r=0.57$ ) and y position is negatively correlated with x velocity ( $r=-0.51$ ). Smaller correlations also exist for other variables.

These correlations, in turn, reflect the fact that our simulations only explored a very limited area of the state space. The simulations, like many real experiments on reaching movements, were confined to repetitive movements with similar trajectories. For instance, data from human subjects performing movements in one direction while adapting to a curl field are shown in the last four rows of Table 1. It can be seen that the correlations among these variables can be just as high or even higher than the correlations in our simulated data.

The important point is that forward estimation can be correct even though the estimated dynamics are incorrect. In contrast, control optimization is sensitive to the discrepancy between real and estimated dynamics, because



**Fig. 3** Adaptive forward model converges and minimizes prediction error. Normalized average prediction error per trial (*blue*, Eq. (2) in [online resource 1](#)) decreases as the simulation progresses, reaching noise level (*black*) errors in the end of the simulation



**Fig. 4** Simulation 2, using forward model learned dynamics to re-optimize controller results. **(a)** Simulated trajectories during force field adaptation. First trial (*light blue*) does not reach the target (*red X*). Last trial (no 800, *dark blue*) reaches the target and comes closer to the optimal trajectory (*black*) than the first simulation, yet is still

quite different. **(b)** Average trajectories without noise. The adaptive model's trajectory (*blue*) is closer to the optimal trajectory (*black*) than the optimal trajectory of a system using a naïve controller and a fully adapted forward model (*red*). However, the final trajectory is still sub-optimal

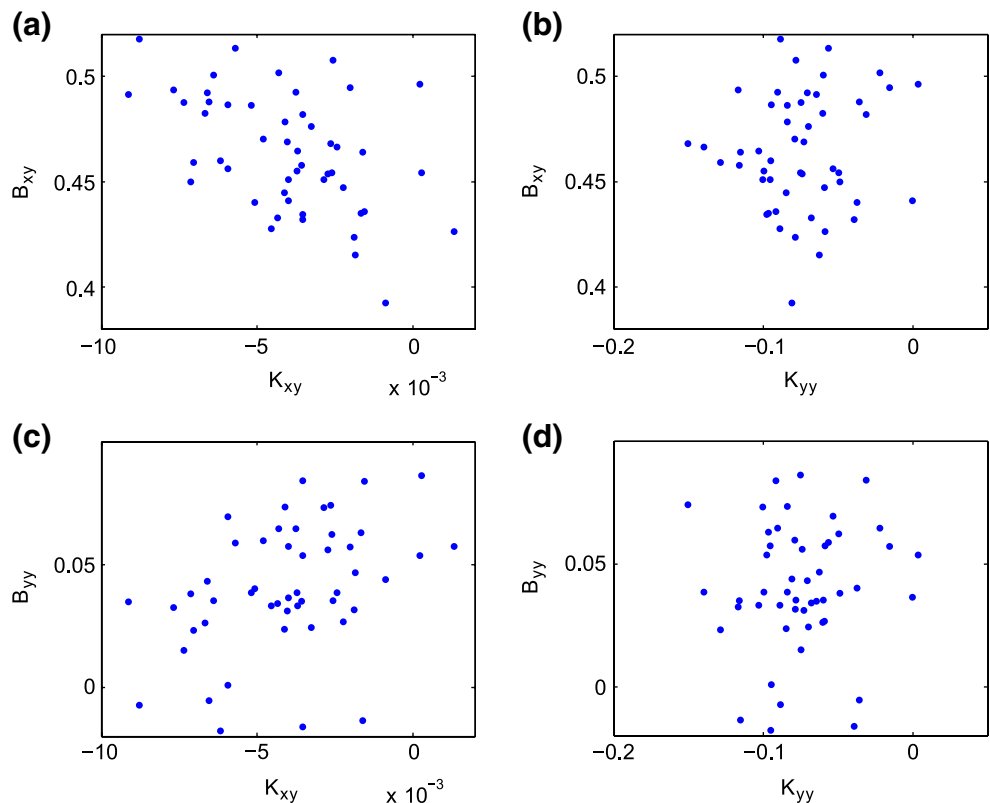
the calculation of the optimal controller relies on the dynamics of the plant. Thus, accurate prediction and optimal control are sensitive to limited information in different ways.

### 3.4 Second simulation, variant 2: varying start positions

One way to reduce the correlations among the state variables is to increase variability. In subject movements, start positions and velocities were noisy, while our simulation always started movements at the same point

and had little noise in the initial velocity. We calculated the mean and covariance matrix for the initial values of the position and velocity in the data and began our simulations at random positions and with random initial velocities drawn from a Gaussian with matched mean and covariance. The middle rows of Table 1 show that this reduced the most extreme correlations. Correlation of x position with y velocity, for instance, was reduced to 0.48 and correlation of y position with y velocity was reduced to -0.48. However, some correlations also increased. The correlation of x and y velocity actually increased from 0.14 to 0.31,

**Fig. 5** Some elements in the learned dynamics matrix are correlated, suggesting underlying state space correlations. Elements of the dynamics matrices learned by the adaptive forward model in fifty separate simulations are correlated. **(a)** The element  $K_{xy}$  (the effect of x position on y force) is negatively correlated with  $B_{xy}$  (the effect of x velocity on y force),  $r=0.47$ ;  $p=4 \times 10^{-4}$ ; **(b)** No correlation between  $K_{yy}$  and  $B_{xy}$ ,  $r=0.11$ ;  $p>0.2$ ; **(c)**  $K_{xy}$  is also correlated with  $B_{yy}$ ,  $r=0.42$ ;  $p=0.003$ ; **(d)**  $K_{yy}$  is not correlated with  $B_{yy}$ ,  $r=0.08$ ;  $p>0.2$



**Table 1** Correlation in state variables

	$x$	$y$	$v_x$	$v_y$
Simulation: fixed start				
$x$	1.00	-0.10	0.57	0.12
$y$	-0.10	1.00	-0.51	0.07
$v_x$	0.57	-0.51	1.00	0.14
$v_y$	0.12	0.07	0.14	1.00
Simulation: noisy start				
$x$	1.00	-0.12	0.49	0.18
$y$	-0.12	1.00	-0.48	-0.20
$v_x$	0.49	-0.48	1.00	0.31
$v_y$	0.18	-0.20	0.31	1.00
Data				
$x$	1.00	0.10 (0.04)	0.10 (0.02)	0.14 (0.15)
$y$	0.10 (0.04)	1.00	-0.32 (0.08)	-0.49 (0.09)
$v_x$	0.10 (0.02)	-0.32 (0.08)	1.00	0.62 (0.02)
$v_y$	0.14 (0.14)	-0.49 (0.09)	0.62 (0.02)	1.00

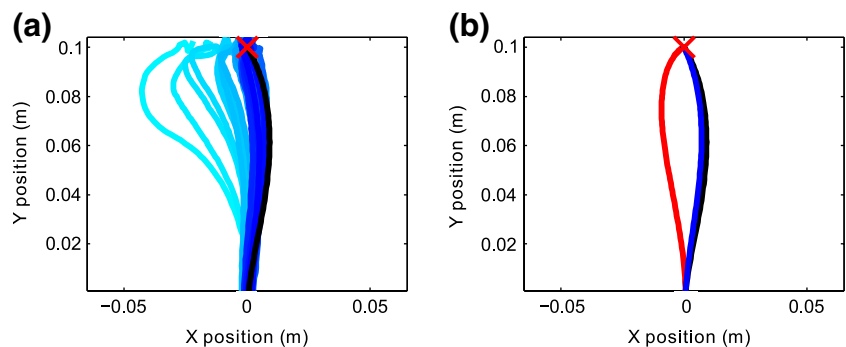
The first four rows show the matrix of correlation coefficients for the position ( $x, y$ ) and velocity ( $v_x, v_y$ ) in the simulated movements. The second set of four rows shows the same correlation coefficients calculated for the second variation of the second simulation, the bottom four rows show the mean and standard deviation (across subjects) for the same correlation matrix calculated on data generated by four human subjects

which is somewhat closer to the figure in human data, 0.62. In any case, Fig. 6 shows that the adding noise to the initial state led to movements that were much closer to the optimal movements.

### 3.5 Second simulation, variant 3: multiple targets

As a further test of our hypothesis regarding the effects of state space correlations on the learning of the plant’s dynamics, we performed a simulation that was identical to the second simulation, except that we used targets in multiple directions. By performing movements to different targets, we ‘forced’ exploration of the state space and reduced correlations. Correlations among state space variables in this simulation were all less than 0.01. As a

**Fig. 6** Simulation 2, variant 2, using noisy starting position and velocity. Format as in Fig. 4. (a) The later adapted trajectories (dark blue) are now closer to the optimal trajectory (black). (b) Average trajectories without noise. The adaptive model’s trajectory (blue) is very close to the optimal trajectory (black), although a small difference can still be discerned



result, the learned dynamics were closer to the real dynamics, although there are still discrepancies:

$$\hat{A}_{CF,Partial} = \begin{bmatrix} -0.02 & -0.32 & -0.07 & -1.94 \\ 0.29 & -0.06 & 1.93 & -0.05 \end{bmatrix}$$

As shown in Fig. 7, the resulting trajectories are now nearly optimal.

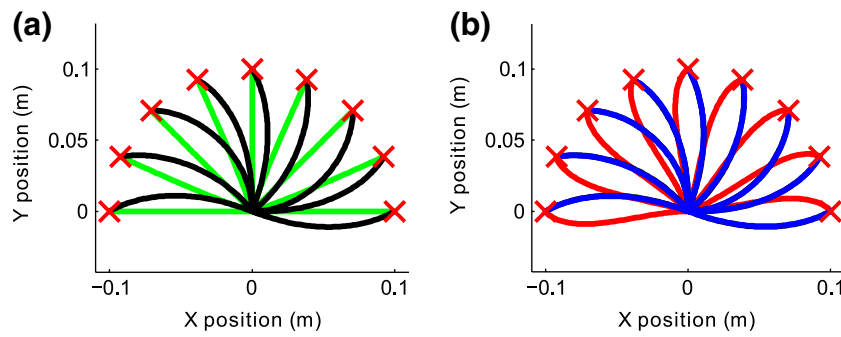
### 3.6 Third simulation: linear approximation of controller adaptation

Our final simulation simulated an independently adaptive controller. The parameters of the controller and state estimator shifted from being optimal under the null field dynamics, to being optimal under the force field dynamics (from the dynamics of Eq. (5) to the dynamics of Eq. (7)). Meanwhile, the forward model adapted to the new dynamics as before (Eq. (11)). As in other simulations, initial movements did not reach the target (Fig. 8(a)). As the forward model learned the dynamics, and the controller shifted closer to the force field optimal state, control improved. Finally, the trajectories converged to the optimal trajectory (Fig. 8(b)). In this simulation, as in the Simulation 2, the forward model did not converge to the correct estimate of the dynamics.

$$\hat{A}_{CF,Partial} = \begin{bmatrix} -0.03 & -0.13 & -0.10 & -1.98 \\ 0.001 & -0.02 & 0.60 & -0.02 \end{bmatrix}$$

However, it did produce accurate predictions that allowed the optimized controller to produce optimal motor commands. Thus, the adapted model behaved just like an optimal OFC system.

The simulation results suggest that the sub-optimality we encountered in the first two simulations was not a result of poor forward model adaptation. The sub-optimality can be accounted for by the inability of forward model adaptation to provide the changes needed for optimal control.



**Fig. 7** Simulation 2, variant 3, reaching to multiple targets. Reaching to multiple targets improves adaptation. **(a)** The optimal trajectory as calculated by the OFC model, without a force field (*green*) and with a curl force field perturbation (*black*). **(b)** Noise free trajectory after

adaptation (*blue*) is indistinguishable from the optimal trajectory (*black*), and far from the optimal trajectory calculated with a naive controller, and a fully adapted forward model of the dynamics (*red*)

### 4 Discussion

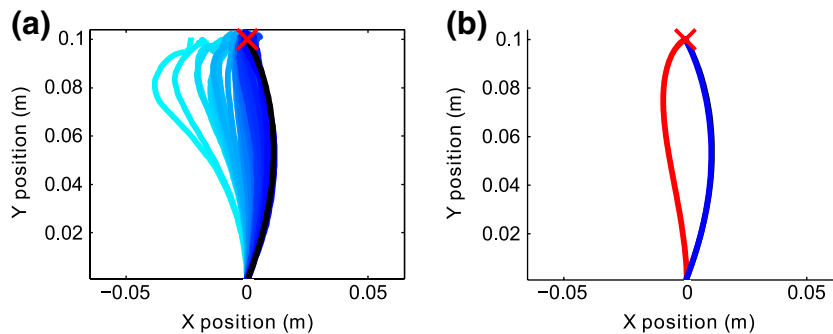
In this study we used simple simulations in order to explore issues of motor adaptation under the OFC framework. We simulated a two dimensional reaching movement, perturbed by a curl force field. We first simulated an OFC framework with an adaptive forward model. Adaptation of the forward model improved its predictions by minimizing the prediction error. As learning progressed, the forward model’s estimation errors decreased, finally reaching noise-level. However, even though the model’s trajectories converged, the resulting trajectory was far from the trajectories produced by a model that had perfect knowledge of the dynamics (Fig. 2). As expected, an adaptive forward model is not sufficient for adaptation of a controller in an OFC framework. This is consistent with recent simulations of the adaptation to target jumps in saccades (Chen-Harris et al. 2008).

We then modified the controller and state-estimator so that that they were re-optimized periodically to match the dynamics learned by the forward model. The resulting trajectories came closer to the optimal trajectory, yet they were still clearly different (Fig. 4(a)). We hypothesized that the difference was due to inadequate learning of the dynamics by the forward

model. Because position and velocity were correlated in the small region of state space explored by the model, an infinite number of dynamics matrices could minimize prediction error. This was emphasized by introducing variability in the starting position and velocity of the movements that match the starting position and velocity in human movements. This increased variability reduced correlations in the state space and led to movements that were much closer to optimal (Fig. 6).

The problem we observe is a credit assignment problem. The forward model has a hard time distinguishing whether the force is related to position or velocity. This problem has also been demonstrated in human experiments (Sing et al. 2009). In their study, Sing et al. show that in the early stages of force field adaptation, human subjects relate the force field to both position and velocity, only converging to the correct solution later in the adaptation process. Even though Sing et al. have a different explanation for this phenomenon, their results show that credit assignment is not a problem that can be overlooked when modeling motor adaptation processes.

Even though the learned dynamics suffered from wrong credit assignment, they were suitable for state prediction.



**Fig. 8** Independent adaptation of the forward model and controller reaches optimal performance. **(a)** Simulated trajectories during force field adaptation. As in the first two simulations, the first trial (*light blue*) does not reach the target (*red X*). However, in this simulation the last

trial (No. 800, *dark blue*) is identical to the optimal trajectory (*black*). **(b)** Average trajectories without noise. The adaptive model’s trajectory (*blue*) is exactly the optimal trajectory (*black*), and is now very far from the optimal trajectory calculated using a naive controller (*red*)

This is because in the region of the state space explored by the model, correlations really did exist. However, control and state estimation depend on different aspects of the dynamics than prediction. This means that even though the forward model has been learned perfectly, the controller that is built from it may perform sub-optimally. Difficulties arise because the learned forward model dynamics are extrapolated by the controller, and this extrapolation may cause the system to converge to a sub-optimal trajectory.

We should emphasize that this convergence to a sub-optimal trajectory is unlikely to be a result of the learning algorithms we have used. While gradient descent learning is the simplest of the error-based learning algorithms, other error-based algorithms differ primarily in terms of speed of convergence or robustness in the face of local minima. Since Figs. 3 and 8 show successful forward model prediction, it does not seem that the problem arises because our forward model got stuck in a local error minimum. It is the structure of the OFC architecture—with a forward model that is independent of the controller and the task constraints—which leads to the difficulty. However, it is precisely the relative simplicity of this architecture and its apparent connection to motor system physiology which has made it so attractive.

We choose to treat this sub-optimality as a problem. We assume that the model should reach optimal control, and that humans tend to do the same. However, one might choose to look at this sub-optimality as a solution. In their paper, Izawa et al. (2008) show that during force field adaptation, humans do not reach the optimal solution. The human movements are better modeled by an OFC model calculated with only 80% of the actual force field's amplitude. Izawa et al. show that this sub-optimality exists, however, they present no explanation for this sub-optimality. Here we may have stumbled upon an explanation. Our model shows that this sort of sub-optimality might be an inherent characteristic of the OFC model in a repetitive reaching experiment.

However, other models of OFC adaptation manage to dodge the sub optimality problem. Our results seem to contradict results showing successful adaptation of an OFC framework where the controller was reoptimized to incorporate forward model adaptation (Mitrovic et al. 2008). However, that study differed from ours in a number of important respects. First, the earlier study used a non-linear model of the arm and a relatively sophisticated iterative optimization algorithm (Todorov and Li 2005) to simulate trial-by-trial adaptation of the controller. Our own model is linear so we could simply re-optimize the controller on the basis of the changes in the forward model. Second, Mitrovic et al. simulated movements to three alternating targets where we only used a single target. Third, their force field was constant while ours was velocity dependent. The first possibility seems unlikely to have created the difference in

our results. Rather, we believe that the other differences explain the divergent results differences. Adaptation to a constant force field, independent of state, trivializes a key issue: the credit assignment problem. Multiple targets cause further decorrelation of the state variables and lead to more successful adaptation in a framework where forward model changes are used to reoptimize the controller (Fig. 7).

To what extent do correlations between state variables represent a problem for the brain? Of course, we do not know what the state variables are in the real motor system, or whether they are correlated. Recent research suggests that adaptation to force fields actually depends on a representation of space in which position and velocity tend to be correlated (Hwang et al. 2003). It is interesting to note that this research not only implied that the system learns in the face of an assumed correlation between position and velocity, but also that it does so while adapting the motor command rather than the forward model prediction. This is in line with recent work from another group (Franklin et al. 2008) that has shown that adaptation to force fields is well described as direct adaptation of muscle activation driven by error in that activation. Thus, it is possible that motor error drives adaptation to force fields and not sensory prediction error. In any case, it is well accepted that the dimensionality of the state space in motor control is very large and that we confine our movements to a small subspace of this large space (Bernstein 1996). Thus, it is likely that our simplistic model highlights an important issue. That is, 'performance irrelevant' variability is likely to be different for the forward model and the controller.

**Acknowledgements** We want to thank Ilana Nisky for helpful comments on the manuscript. This research was supported by the Israeli Science Foundation, grant number 0624/06 and by a grant from Israel/Lower Saxony fund.

**Author contributions** J.A.: Had the original idea, designed and implemented the model, analyzed the data, and wrote the paper.

O.D.: Helped in the design of the simulations, contributed to the mathematical derivations, did part of the programming, and edited the paper.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Bernstein, N. A. (1996). On dexterity and its development. In M. L. Latash & M. T. Turvey (Eds.), *Dexterity and its development* (pp. 3–244). Mahwah: Lawrence Erlbaum Associates.
- Catz, N., Dicke, P. W., & Thier, P. (2005). Cerebellar complex spike firing is suitable to induce as well as to stabilize motor learning. *Current Biology*, *15*(24), 2179–2189.

- Chen-Harris, H., Joiner, W. M., Ethier, V., Zee, D. S., & Shadmehr, R. (2008). Adaptive control of saccades via internal feedback. *Journal of Neuroscience*, *28*(11), 2804–2813.
- Dean, P., Porrill, J., Ekerot, C. F., & Jorntell, H. (2010). The cerebellar microcircuit as an adaptive filter: experimental and computational evidence. *Nature Reviews. Neuroscience*, *11*(1), 30–43.
- Diedrichsen, J. (2007). Optimal task-dependent changes of bimanual feedback control and adaptation. *Current Biology*, *17*(19), 1675–1679.
- Donchin, O., & Shadmehr, R. (2004). Change of desired trajectory caused by training in a novel motor task. *Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE*, doi:10.1109/IEMBS.2004.1404249.
- Donchin, O., Francis, J. T., & Shadmehr, R. (2003). Quantifying generalization from trial-by-trial behavior of adaptive systems that learn with basis functions: theory and experiments in human motor control. *Journal of Neuroscience*, *23*(27), 9032–9045.
- Doya, K., Kimura, H., & Kawato, M. (2001). Neural mechanisms of learning and control. *IEEE Control Systems Magazine*, *21*(4), 42–54.
- Franklin, D. W., Burdet, E., Tee, K. P., Osu, R., Chew, C. M., Milner, T. E., et al. (2008). CNS learns stable, accurate, and efficient movements using a simple algorithm. *Journal of Neuroscience*, *28*(44), 11165–11173.
- Frens, M. A., & Donchin, O. (2009). Forward models and state estimation in compensatory eye movements. *Frontiers in Cellular Neuroscience*, *3*, 131.
- Gomi, H., & Kawato, M. (1992). Adaptive feedback control models of the vestibulocerebellum and spinocerebellum. *Biological Cybernetics*, *68*(2), 105–114.
- Hwang, E. J., Donchin, O., Smith, M. A., & Shadmehr, R. (2003). A gain-field encoding of limb position and velocity in the internal model of arm dynamics. *Public Library of Science: Biology*, *1*(2), E25.
- Izawa, J., Rane, T., Donchin, O., & Shadmehr, R. (2008). Motor adaptation as a process of reoptimization. *Journal of Neuroscience*, *28*(11), 2883–2891.
- Kawato, M., & Gomi, H. (1992). A computational model of four regions of the cerebellum based on feedback-error learning. *Biological Cybernetics*, *68*(2), 95–103.
- Mazzoni, P., & Krakauer, J. W. (2006). An implicit plan overrides an explicit strategy during visuomotor adaptation. *Journal of Neuroscience*, *26*(14), 3642–3645.
- Miall, R. C., Weir, D. J., Wolpert, D. M., & Stein, J. F. (1993). Is the cerebellum a smith predictor? *Journal of Motor Behavior*, *25*(3), 203–216.
- Mitrovic, D., Klanke, S., & Vijayakumar, S. (2008). Adaptive optimal control for redundantly actuated arms. In M. Asada, J. Hallam, J. Meyer, & J. Tani (Eds.), *From animals to animats 10* (pp. 93–102). Berlin: Springer.
- Morton, S. M., & Bastian, A. J. (2004). Prism adaptation during walking generalizes to reaching and requires the cerebellum. *Journal of Neurophysiology*, *92*(4), 2497–2509.
- Morton, S. M., & Bastian, A. J. (2006). Cerebellar contributions to locomotor adaptations during splitbelt treadmill walking. *Journal of Neuroscience*, *26*(36), 9107–9116.
- Rabe, K., Livne, O., Gizewski, E. R., Aurich, V., Beck, A., Timmann, D., et al. (2009). Adaptation to visuomotor rotation and force field perturbation is correlated to different brain areas in patients with cerebellar degeneration. *Journal of Neurophysiology*, *101*(4), 1961–1971.
- Robinson, F. R. (1995). Role of the cerebellum in movement control and adaptation. *Current Opinion in Neurobiology*, *5*(6), 755–762.
- Shadmehr, R., & Krakauer, J. W. (2008). A computational neuroanatomy for motor control. *Experimental Brain Research*, *185*(3), 359–381.
- Shadmehr, R., & Mussa-Ivaldi, F. A. (1994). Adaptive representation of dynamics during learning of a motor task. *Journal of Neuroscience*, *14*(5 Pt 2), 3208–3224.
- Shadmehr, R., Donchin, O., Hwang, E., Hemminger, S., & Rao, A. (2005). Learning dynamics of reaching. In: A. Riehle & E. Váadia (Eds.), *Motor cortex and voluntary movements* (pp. 297–328-11). CRC Press.
- Sing, G. C., Joiner, W. M., Nanayakkara, T., Braynov, J. B., & Smith, M. A. (2009). Primitives for motor adaptation reflect correlated neural tuning to position and velocity. *Neuron*, *64*(4), 575–589.
- Todorov, E. (2005). Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. *Neural Computation*, *17*(5), 1084–1108.
- Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, *5*(11), 1226–1235.
- Todorov, E., & Li, W. (2005). A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems. *American Control Conference, 2005. Proceedings of the 2005*, doi:10.1109/ACC.2005.1469949.
- Tseng, Y. W., Diedrichsen, J., Krakauer, J. W., Shadmehr, R., & Bastian, A. J. (2007). Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *Journal of Neurophysiology*, *98*(1), 54–62.