Motor cortical dynamics are shaped by multiple distinct subspaces during naturalistic behavior

Matthew G. Perich[1,2,†], Sara Conti[3,4], Marion Badi[5], Andrew Bogaard[5], Beatrice Barra[3], Sophie Wurth[5], Jocelyne Bloch[6,7,8], Gregoire Courtine[6,7,8], Silvestro Micera[4,5], Marco Capogrosso[3,9,*], Tomislav Milekovic[1,6,7,8*]

[1] Department of Fundamental Neuroscience, University of Geneva, Geneva, Switzerland
[2] Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, NY, USA
[3] Department of Medicine, University of Fribourg, Fribourg, Switzerland
[4] The BioRobotics Institute and Department of Excellent in Robotics and AI, Scuola Superiore Sant'Anna, Pisa, Italy
[5] Bertarelli Foundation Chair in Translational Neuroengineering, Center for Neuroprosthetics and Institute of Bioengineering, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
[6] Centre hospitalier universitaire Vaudois (CHUV), Lausanne, Switzerland
[7] Defitech Center for Interventional Neurotherapies (NeuroRestore), Lausanne, Switzerland
[8] Center for Neuroprosthetics and the Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
[9] Department of Neurological Surgery, University of Pittsburgh, Pittsburgh, PA, USA
[*] these authors jointly supervised this work

† correspondence to: mperich@gmail.com

ABSTRACT

Behavior relies on continuous influx of sensory information about the body and the environment. In primates, cortex integrates somatic feedback to accurately reach and manipulate objects. Yet, in many experimental regimes motor cortex seems paradoxically to operate as a feedforward, rather than feedback-driven, system. Here, we recorded simultaneously from motor and somatosensory cortex as monkeys performed a naturalistic reaching and object interaction behavior. We studied how unexpected feedback from behavioral errors influences cortical dynamics. Motor cortex generally exhibited robust feedforward dynamics, yet displayed feedback-driven dynamics surrounding correction of behavioral errors. We then decomposed motor cortical activity into orthogonal subspaces capturing communication with somatosensory cortex or behavior-generating dynamics. During error correction, the communication subspace became feedback-driven, while the behavioral subspace maintained feedforward dynamics. We therefore demonstrate that cortical activity is compartmentalized within distinct subspaces that shape the population dynamics, enabling flexible integration of salient inputs with ongoing activity for robust behavior.

INTRODUCTION

Dexterous control of the large number of degrees of freedom in the arm and hand is a daunting task, further complicated by transmission delays that make the immediate effect of a motor command uncertain[1]. The brain appears to have solved this problem by predicting the effect of its commands, and updating future output based on the sensory feedback it receives[1]. Consequently, somatic feedback from the body is critical for the execution of accurate and robust movements: patients suffering from the loss of cutaneous touch sensations or proprioception have great difficulty performing motor skills[2]. Understanding how the motor system incorporates somatosensory feedback is critical to understand how the brain generates behavior[1]. The motor cortex is an appealing hub for such integration. Motor cortical neurons are the primary output of the cortex to the spinal cord[3], but also broadcast commands to many parts of the brain including sensory regions like somatosensory cortex[1,4]. Motor cortical neurons respond rapidly to limb perturbations and other sensory inputs[5,6], though it remains unclear how sensory feedback is integrated within motor cortex throughout the generation of natural movements.

Despite the large numbers of neurons in motor cortex, the activity of individual motor cortical neurons largely reflects some underlying, lower-dimensional, population-wide dynamics[7–12]. The dynamics are captured within the *neural manifold*[8], which can be derived using dimensionality reduction techniques[13]. Recent work has shown that a large part of motor cortical activity within the manifold is well-modeled as a dynamical system[14,15] wherein the future state of the neural population activity evolves deterministically from the present state. From these studies, it may seem that motor cortical activity can be viewed as a feedforward generator of movement (i.e. a system that is not strongly influenced by external feedback inputs) with dynamics that are inherently autonomous or self-driven. However, motor cortex requires constant input during the generation of behavior[16] and its activity is shaped by somatosensory feedback[1]. Why, then, do the motor cortical dynamics appear to be feedforward and not driven by feedback inputs?

We propose that even a largely feedback-driven system may appear to be feedforward if all inputs are predictable from the current state of the population activity, such as during planned, practiced movements without errors or perturbations (Fig. 1a). Throughout these movements, the feedback inputs largely match what is expected from the outgoing motor commands, making it challenging to dissociate possible feedforward or feedback-driven components of motor cortical activity. The arrival of unexpected inputs – such as somatic feedback from behavioral errors that deviate from the intended motor output – should lead to more apparently feedback-driven dynamics in the same cortical population (Fig. 1b). The degree to which a dynamical system is behaving as feedforward or feedback-driven can be quantified using *trajectory tangling*[17], which occurs when identical neural states lead to diverging future states. A completely feedforward or autonomous dynamical system will have low tangling since the current activity deterministically predicts future activity, while high tangling indicates a system that is driven by unexpected inputs or is otherwise unstable. Prior analysis of motor cortical dynamics have used simple behavioral tasks with repetitive, well-practiced movements and found largely untangled, feedforward dynamics[17]. Such feedforward dynamics are advantageous as they allow cortex to produce

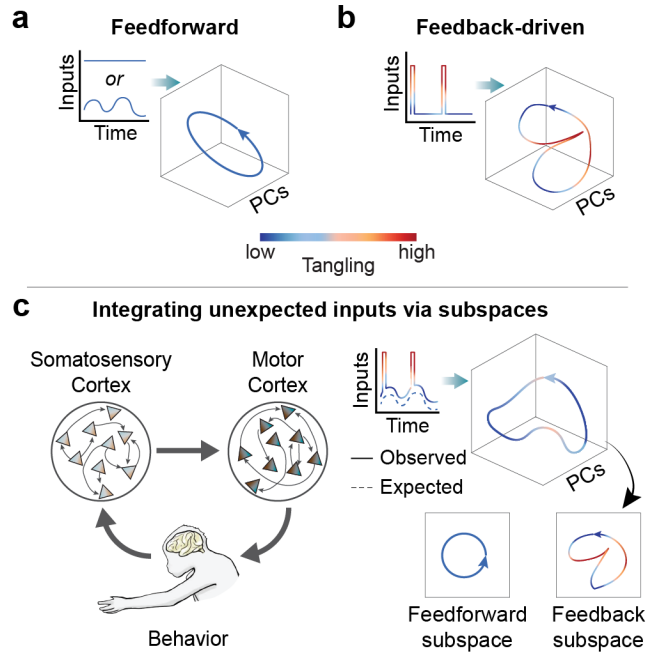

Figure 1. *Hypothesis for integrating unexpected inputs with ongoing population activity via subspaces.* **(a)** Cortical dynamics can appear to be a feedforward system as long as the observed inputs match the expectation of the current state of activity. Tangling quantifies the regime of a dynamical system (low tangling: feedforward or autonomous dynamics; high tangling: feedback-driven dynamics) and can measure the influence of unexpected inputs on cortical activity throughout the behavior. **(b)** Such unexpected inputs can increase tangling in the population dynamics, particularly around the time that the unexpected inputs arrive. **(c)** Schematic of interactions between the motor cortex, somatosensory cortex, and behavior. Motor cortex orchestrates the behavior, while somatosensory cortex receives sensory inputs about the state of the body. Sensory feedback is communicated to motor cortex from somatosensory cortex to monitor behavioral output and enable correction of behavioral errors. When all feedback inputs are expected (dashed line), cortex can behave as a feedforward dynamical system, as in Panel a. However, when the observed inputs (solid line) deviate from expectations (dashed line), the population can appear more feedback-driven. We hypothesized that cortex integrates unexpected, tangled sensory inputs while maintaining robust, untangled movement generation activity through orthogonal subspaces within the neural population.

a highly robust control signal that is resistant to noise[17], including the biological limitations of neural circuits[18] such as neuronal death[19] or unreliable synapses[20]. If motor cortex is indeed input-driven[16], more complex feedback-driven dynamics may only become apparent when studying naturalistic behaviors that include unexpected behavioral outcomes.

Can the motor cortex integrate relevant inputs while maintaining feedforward dynamics to robustly generate behavior, or is the population of neurons forced to switch between feedback-driven and feedforward regimes? Recent work has demonstrated that the neural manifold can be decomposed into *subspaces* capturing specific population-wide features. Subspaces can enable planning a movement[7,11], adjusting motor output during learning[21,22], sharing the control of a limb across hemispheres[23,24], and even communicating between cortical regions[25]. We hypothesized that subspaces of neural population activity also enable effective coexistence of feedforward and feedback-driven dynamics (**Fig. 1c**). This framework allows cortex to operate as a set of linked, but independently-driven, systems comprising both feedback-driven and feedforward dynamics. Together, these systems shape the global cortical dynamics observed in the full population.

To test our hypothesis, we simultaneously recorded the activity of neural populations from motor and somatosensory cortex of four monkeys as they reached for and interacted with objects[26]. Using trajectory tangling[17], we showed that the neural population dynamics substantially differed between the two cortical regions. Despite the complexity of the behavior, motor cortex exhibited untangled dynamics suggestive of a robust, feedforward system. In contrast, somatosensory cortex was comparatively more feedback-driven (tangled). As predicted, motor cortical dynamics became more feedback-driven (higher tangling) during behavioral errors and the subsequent correction.

Using a novel subspace decomposition algorithm, we showed that the motor cortex interacts with somatosensory cortex through a low-dimensional communication subspace, while an orthogonal low-dimensional *behavioral subspace* correlated with the behavior and exhibited feedforward dynamics. Intriguingly, the increased tangling in motor cortex during error correction was isolated only to its communication subspace, while the tangling of the behavioral subspace remained unchanged. Our work reimagines cortical activity as a composition of multiple function-specific systems isolated to distinct subspaces, and provides a novel framework for flexible integration of salient inputs with ongoing population activity.

RESULTS

A naturalistic reach and grasp task to study sensorimotor integration. The surprisingly feedforward nature of motor cortical dynamics observed in prior studies may be attributed to examining behavioral tasks performed by highly trained subjects that were allowed to fully plan movements before executing them[10,14,15]. To fully understand the effect of sensory feedback on motor cortical dynamics, we designed a more naturalistic task that resulted with occasional behavioral errors. In our task, monkeys reached for, grasped, and pulled objects mounted on a robotic arm[26] (**Fig. 2a**). During the pulling phase, the robot acted as a spring, applying a resistive force proportional to the displacement. This task had several unique properties: 1) the task began with a ballistic reaching phase, similar to many prior studies[11,14,27]; 2) grasping the object provides a salient sensory cue; and 3) the resistive pulling phase required interactions with an external object. Four monkeys (*macaca fascicularis*) effectively learned to complete this task.

Due to the unconstrained nature of the task, each monkey exhibited an idiosyncratic behavioral strategy (**Fig. 2c**, left; **Fig. S1**). However, all monkeys occasionally made behavioral errors (**Fig. 2c**, pie charts; 18%, 29%, 27%, and 18% of trials for Mk-Cs, Mk-Br, Mk-Ol, and Mk-Jo, respectively) before ultimately succeeding in pulling the object. For Mk-Br, most errors occurred after grasping the object in order to adjust the posture of the hand to enable the pulling movement. For the remaining monkeys, errors were typically due to the hand missing the object or slipping off the object as it was pulled. These monkeys compensated by adjusting the position of the hand and grasping and pulling again, as shown by the second, smaller positive peak in hand velocity (**Fig. 2c**, right). These error trials provided an opportunity to study the effect of new, unexpected inputs on cortical activity.

As the monkeys performed the task, we simultaneously recorded kinematics of the left arm (Vicon, Oxford, UK), three-dimensional pulling force on the object, and the spiking activity of motor and somatosensory cortical neural populations (**Fig. 2b**) using chronically-implanted electrode arrays (Blackrock Microsystems, Salt Lake City, UT, USA). We isolated 78, 126, 62, and 88 motor and 82, 59, 44, and 42 somatosensory cortical neurons from Mk-Cs, Mk-Br, Mk-Ol, and Mk-Jo, respectively. Throughout the task, we observed striking similarities in the firing patterns of the somatosensory and motor cortical populations during the normal trials (**Fig. 2d**). On the error trials, we saw a transient reorganization of activity in both motor and somatosensory cortex around the time of error correction. For example, Mk-Cs (**Fig. 2d**, top row) had a prominent burst in activity in both

regions around the time of error correction and subsequent pulling phase. The new activity patterns likely result from unexpected inputs arriving to both motor and somatosensory cortex reflecting the occurrence of the error, as well as altered activity driving the behavioral correction.

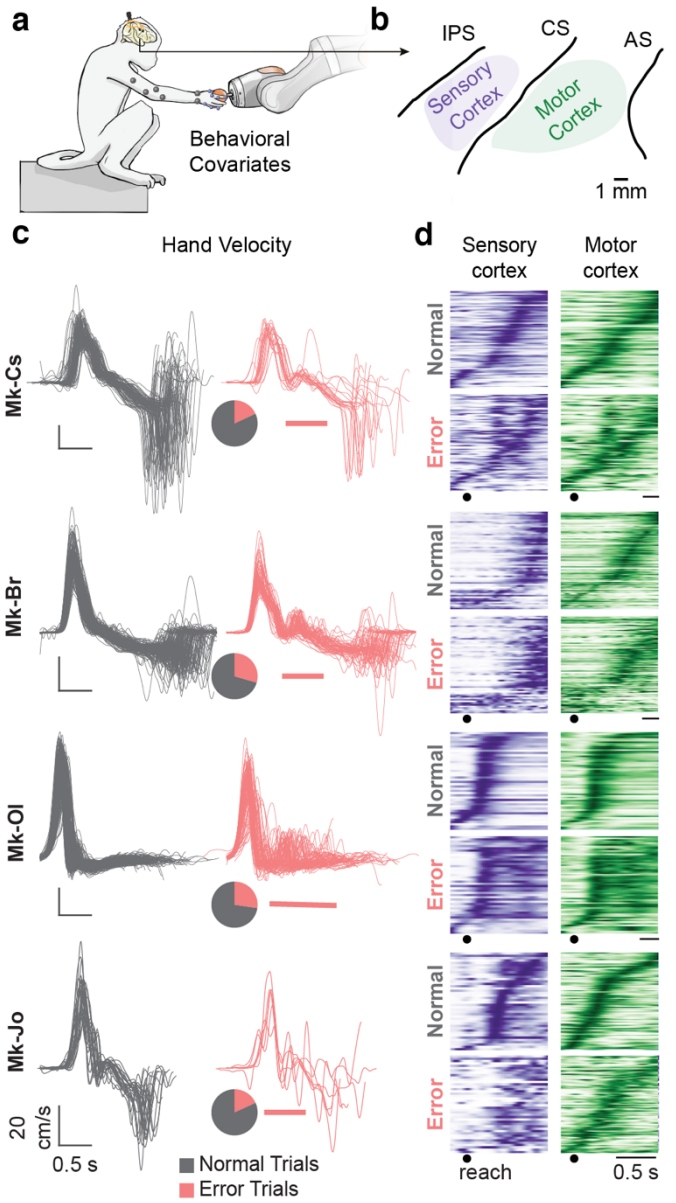

Figure 2. *Behavioral task and recordings.* **(a)** Monkeys reached for, grasped, and pulled an object mounted on a haptic robot while we recorded limb kinematics, pulling forces, and neural population activity from motor and somatosensory cortex simultaneously. **(b)** Approximate recording locations for all four monkeys spanned between the central sulcus (CS) and intraparietal sulcus (IPS) for somatosensory cortex (purple), and the CS and arcuate sulcus (AS) for motor cortex (green). **(c)** Plots show the hand velocity in the sagittal plane for all normal (gray, left) and error (pink, right) trials for each of the four monkeys. Pie charts show the proportion of each trial type. Each behavioral error required a behavioral correction characterized by a second period of forward velocity to either reach for the robot again or adjust the position of the hand. Pink bars beneath the trajectories highlight the approximate window of error correction for each monkey. **(d)** Average neural activity for all somatosensory (purple) and motor (green) cortical neurons for each monkey for normal (top) and error (bottom) trials. Activity was aligned on reach and averaged across trials. Neurons are sorted based on their time of maximal activity in normal trials.

Motor and somatosensory cortex exhibit distinct dynamics during reach to grasp. To understand the change in neural responses observed on the error trials, we first characterized the dynamics of the somatosensory and motor cortical population activity. We applied Principal Component Analysis (PCA) to identify 20-dimensional manifolds[8] that captured between 68-86% of neural population variance in each cortical region. We also applied PCA on the recorded arm kinematics and pulling force (behavioral covariates) to identify a 20-dimensional "behavioral manifold" capturing at least 95% of the variance. To quantify the local dynamical regime of each manifold, we used trajectory tangling[17]. High tangling values occur when the current state (e.g. a 20-dimensional vector in a neural or a behavioral manifold) leads to an array of vastly different future states. This is often true when the population is driven by unexpected inputs that change the temporal evolution of the state independently of its history. Low tangling values, in contrast, imply that the current state of the system accurately predicts future states. Importantly, even a feedback-driven system may appear to be feedforward if the inputs occur as expected in a manner that is consistent across trials. Since somatosensory cortex is driven by inputs from receptors in the limb[28], we expected that its neural population dynamics should exhibit higher tangling than motor cortex, reflecting unpredictable changes of the somatosensory feedback[17].

For the trajectories in the behavioral manifold, tangling was consistently very high for all four monkeys (**Fig. 3a**; **Fig. 3b**, left). The trajectories typically exhibited a cross-over point at reach onset and grasp, where most of behavioral variables are similar (hand at rest). Although somatosensory cortical activity is largely driven by feedback about the behavioral output, the trajectories of the somatosensory manifold (**Fig. 3a**, purple; **Fig. 3b**, middle) were significantly less tangled than in the behavioral manifold. Notably, the cross-over point was absent - the somatosensory cortical manifold trajectories during the movements followed a single path without prominent bifurcations. Nonetheless, the trajectories made sharp turns at different points, thereby causing the paths to frequently diverge and converge resulting in increased tangling. Despite the apparent similarities in the firing patterns of somatosensory and motor cortical neurons (**Fig. 2d**; **Fig. S1d-e**), trajectories in the motor cortical manifold were circular and smooth (**Fig. 3a**, green; **Fig. 3b**, right) resulting in significantly lower tangling than in both behavioral and somatosensory cortical manifolds[17].

We then analyzed the manifold trajectories on error trials. Previous studies have shown that some motor cortex neurons respond rapidly to mechanical perturbations of the limb, despite the apparently feedforward nature of its population

activity on normal trials[1]. We thus predicted that tangling of the motor cortex manifold trajectory increases when sensory inputs indicate a need to alter the current behavior, such as following the behavioral errors. In our task, each error was followed by a behavioral correction in order to complete the trial (Fig. 2b). These corrections corresponded to loops or distortions in the motor cortical manifold trajectories (Fig. 3c). We saw that motor cortical manifold tangling increased transiently surrounding the error correction (Fig. 3c-e), indicating that the motor cortex may then be driven by unexpected inputs. We observed similar increases in tangling for the somatosensory cortical manifold trajectories following behavioral errors (Fig. S3). We hypothesized that communication between motor and somatosensory cortex could account for the increased tangling in motor cortical activity surrounding errors.

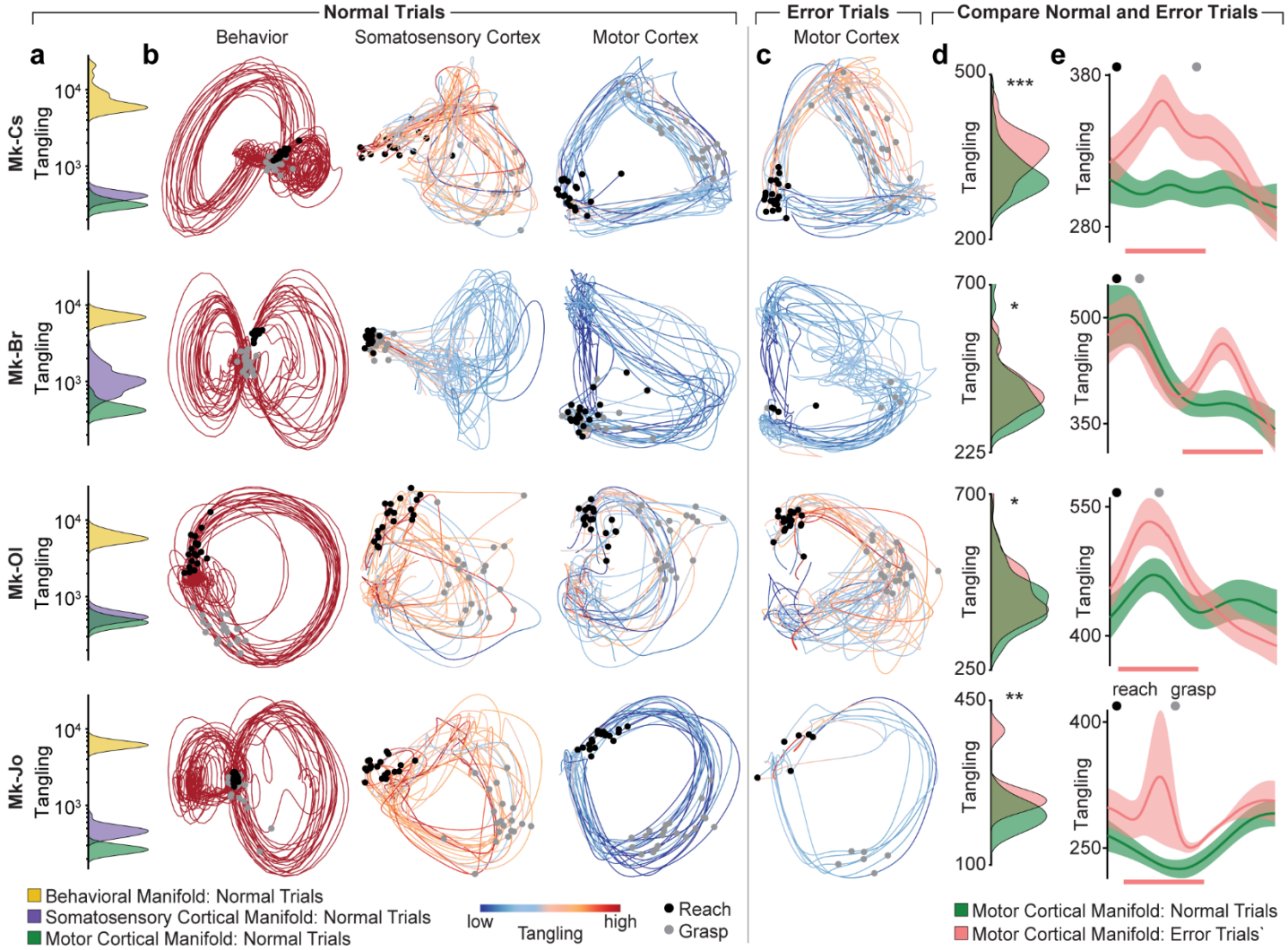

Figure 3. *Motor cortical population activity is driven by unexpected inputs on error trials.* **(a)** We identified 20-D manifolds capturing the behavioral signals (yellow), as well as the somatosensory (purple) and motor (green) cortical population activity. We then computed the tangling of the trajectories of behavior, somatosensory cortical and motor cortical manifold states. The distributions show average tangling throughout each trial in these three manifolds on a log scale. All distributions are significantly different at $p < 0.001$ (ranksum). **(b)** Each subplot shows the trajectories in the first two principal components of the behavioral (left), somatosensory cortical (middle), and motor cortical (right) manifolds for 20 normal trials. Dark blue indicates the minimum tangling value and dark red indicates the maximum value observed in somatosensory or motor cortical activity for each monkey (allowing behavior manifold values to saturate). Black circles: reach onset; gray circles: object grasp. **(c)** Tangling in the motor cortical manifold increases in the error trials. Plots show motor cortical manifold trajectories and tangling for all error trials for each of the four monkeys. Color scale for each monkey is the same as in Panel b. Note that the states at the time of object grasp are similar on the normal and error trials, indicating that error correction required a return to a similar neural state to complete the trial. **(d)** Single-trial error-related motor cortical manifold tangling is significantly higher in error trials compared to the normal trials. Tangling is computed in windows shown by pink lines in Panel e. Plots show single-trial tangling distributions in the motor cortical manifold for all normal (green) and error (pink) trials. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$; ranksum. **(e)** Motor cortical manifold tangling acutely increases surrounding the error correction. Plots show tangling in the motor cortical manifold averaged across all normal and error trials after aligning on reach onset and time-warping to match the length across all trials (see Methods). Error bars: mean ± s.e.m. See also: Supplementary Video 1.

Decomposing motor cortical population activity into subspaces. We sought to understand the dynamics of the interactions between the two cortical regions to determine whether communication with somatosensory cortex might account for the increased tangling observed in motor cortex on error trials. We hypothesized that communication between somatosensory and motor cortex could be achieved using a specific subspace of population activity (Semedo et al., 2019). This communication subspace would function independently from an orthogonal set of dimensions that facilitates the generation of behavior, which we call the behavioral subspace. By being independently-driven, these two subspaces could exhibit correspondingly different dynamics, allowing new, unexpected sensory inputs to drive motor cortical activity without directly perturbing the ongoing, behavior-generating dynamics (Fig. 1c).

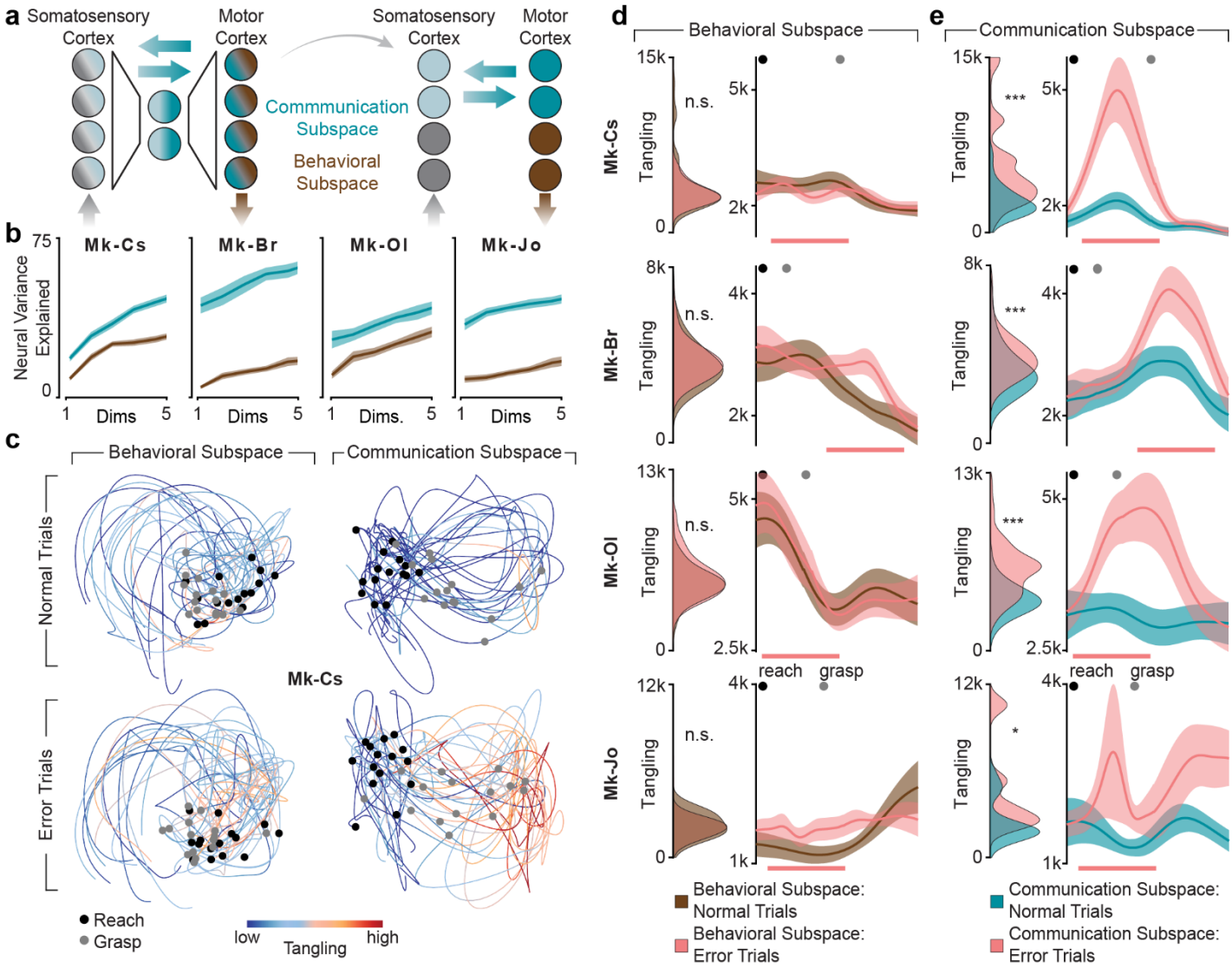

Figure 4. *The increase in motor cortical tangling during behavioral errors is confined to its communication subspace.* **(a)** Schematic of the subspace decomposition. The neural manifold identified by PCA comprises a mixture of activity related to cortico-cortical communication (brown) and behavioral output (blue). We developed the DOS algorithm to identify communication and behavioral subspaces that isolate dynamics related to somatosensory cortical activity or behavioral output, respectively. **(b)** The dominant components of motor cortical activity captures communication with the somatosensory cortex. The plots show the percentage of motor cortical neural variance explained by the communication and behavioral subspaces as a function of the chosen subspace dimensionality. **(c)** Communication subspace tangling substantially increases in error trials. Plots show trajectories within the behavioral (left) and communication (right) subspaces on normal (top) and error (bottom) trials for Mk-Cs. Data plotted as in Fig. 3a. **(d)** Behavioral subspace tangling remains unchanged between normal and error trials. Panels show tangling in the behavioral subspace for all normal (brown) and error (pink) trials for the four monkeys. Left column data are plotted as in Fig. 3d. Right column data are plotted as in Fig. 3e. **(e)** Communication subspace tangling substantially increases in the error trials surrounding the error correction. Panels show tangling in the communication subspace for all normal (blue) and error (pink) trials for the four monkeys. Pink bars show approximate window of error correction. Data presented as in Panel d. Error bars: s.e.m. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$; ranksum. See also: Supplementary Video 2.

We developed an algorithm – Decomposition into Optimal Subspaces (DOS) – to identify these subspaces within the population activity (see Methods). This algorithm optimally decomposes neural population activity into orthogonal subspaces based on the covariance of the neural population activity with other variables: here, the population activity of somatosensory cortex (communication subspace) and the behavioral covariates (behavioral subspace).

We designed a simulation to validate DOS in a dataset with known ground truth dynamics (see Methods). We determined that DOS isolates specific dynamics with high accuracy on simulated data (Fig. S4). We then applied DOS to our behavioral, motor cortical and somatosensory cortical manifolds to identify communication and behavioral subspaces in motor cortical activity (Fig. 4a and S5a). Using cross-validation across all normal trials, we computed the neural variance explained by the communication and behavioral subspaces for assumed dimensionalities between one and five (Fig. 4b). For all four monkeys, the communication subspace explained the majority of neural population variance at all dimensionalities. Since the cost function of DOS does not optimize to explain variance in the motor cortical population, only covariance with somatosensory cortex, cortico-cortical communication appears to be a dominant feature of motor cortical population activity during movement. This observation is particularly surprising given the strikingly different dynamics of the full motor and somatosensory cortical manifolds (Fig. 3a-b).

Increased tangling of the motor cortex during error correction is isolated to its communication subspace. We next analyzed the dynamics within the motor cortical subspaces using tangling. Since the tangling metric is sensitive to dimensionality of the subspaces, we assumed both subspaces to be five-dimensional to facilitate direct comparison between them. If motor cortex receives unexpected inputs from somatosensory cortex during error trials, we hypothesized that the increase in tangling observed in motor cortical activity is isolated to its communication subspace. To ensure our results were not biased, we projected the activity on error trials into the subspaces previously identified using only the normal trials (Fig. 4c, Fig. S5).

Despite the increase in tangling observed within the full motor cortical population on error trials (Fig. 3c-e), we observed no significant difference in tangling within the behavioral subspace (Fig. 4d). In contrast, trajectories within the communication subspace increased in tangling around the time of error correction (Fig. 4e), similar to the full motor (Fig. 3c-e) and somatosensory (Fig. S3) cortical population activity. This increase in tangling was not observed in the remaining dimensions of population activity that were not captured by either the communication or behavioral subspaces (the *null subspace*; see Methods), indicating that the effect of unexpected somatosensory inputs related to the behavioral errors is confined only to the communication subspace (Fig. S6). These results together demonstrate that the increased tangling in motor cortex on error trials, indicative of a feedback-driven regime, is mediated using a communication subspace shared with somatosensory cortex.

DISCUSSION

We showed that during a naturalistic movement task composed of reaching, grasping, and pulling a loaded object, somatosensory cortical population activity appeared to be predominantly feedback-driven (Fig. 3), reflecting the tangled nature of the behavioral output. Despite the complexity of the task, motor cortical population activity exhibited smooth and untangled dynamics, characteristic of a dynamical system that is behaving in a feedforward manner. However, when the monkeys made a behavioral error, the motor cortical tangling acutely increased as if it were transiently more feedback-driven (Fig. 3c-e). The dynamics of the neural population activity in somatosensory cortex was more tangled than the motor cortical activity in both normal and error trials (Fig. 3a, Fig. S2, Fig. S3). This result was surprising given the apparent similarity of single neuron activity in the two regions (Fig. 2d). To study the interactions between the two cortical regions, we developed DOS, a subspace decomposition algorithm that isolated dimensions capturing cortico-cortical communication from those capturing behavioral output. In agreement with our hypothesis, the increase of tangling in motor cortex was confined to the communication subspace linking motor cortical activity to somatosensory cortex (Fig. 4). The remaining orthogonal components of motor cortical population activity (both the behavioral subspace and the null subspace) showed no change in tangling during the error trials (Fig. 4d, Fig. S6). These observations demonstrate that motor cortical activity is composed of multiple, independently-driven systems that together comprise the large-scale dynamics of the region.

Subspaces enable inputs to co-exist with ongoing population activity. Although inputs, whether sensory[5,29] or cognitive such as during changes of mind[30,31], influence motor cortical activity, the mechanisms that allow these inputs to coexist with ongoing population activity driving behavior have remained unclear. Intriguingly, motor cortex is dominated by apparently feedforward dynamics during the

execution of planned or repetitive movements[14,17,18]. Modeling provides strong evidence that feedforward dynamics enables motor cortical populations to robustly generate motor output despite biological limitations of the neurons, such as neuronal noise, unreliable synapses or spontaneous death of neurons[18]. Nonetheless, motor cortex has to respond to external inputs when those inputs require unplanned changes of behavior. Here, we show how this apparent contradiction is resolved during correction of behavioral errors. The arrival of new information is compartmentalized into lower-dimensional subspaces of the neural manifold. This result has been challenging to observe in prior experiments focusing on highly trained, fully planned movements. This principle can only be uncovered by studying population activity during more complex behaviors where the possible arrival of inputs can be dissociated from the ongoing movements, such as the error trials in our naturalistic reaching and grasping task.

We studied the interactions between two brain regions, the somatosensory cortex and the motor cortex. However, a region such as the motor cortex receives input from many other regions in the brain. Through subspaces, our framework could allow any number of different brain regions to independently drive motor cortex. An interesting open question is how the independently-driven subspaces are linked to allow flexible sensorimotor integration and generate the full, population-wide dynamics. One intriguing possibility is that the link between the subspaces is direct but transient. Another possibility is that the subspaces are perpetually independent, but communication between them is mediated using additional dimensions within the neural manifold[32] or even through other brain areas. Future theoretical and experimental work is needed to elucidate possible mechanisms underlying this integration.

Tangling as a means to quantify differences in dynamical regimes. The error trials demonstrate how independently-driven subspaces can allow motor cortex to adjust behavior. We found that motor cortex exhibited feedback-driven dynamics (quantified by an increase in tangling) transiently surrounding the error correction (**Fig. 3c-e**). Trajectory tangling is a compact and data-driven metric to quantify the local dynamical behavior of the neural population, i.e. to measure where the current state of a system is on a scale between a fully feedback-driven and a fully feedforward. The tangling value, however, only offers a relative measure of the dynamics of the neural and behavioral signals. To ensure that the tangling values are interpretable, we performed all comparisons between trajectories that were matched for dimensionality (20 dimensions for the manifolds, and 5 dimensions for the subspaces).

Surprisingly, the increase in tangling in motor cortex was contained to the communication subspace (**Fig. 4d-e**; **Fig. S6**). This result cannot be trivially explained by our methods. First, we derived the subspaces using only normal trials, ensuring that there was no bias enforcing that only the communication subspace captures the feedback-driven dynamics on error trials. Second, the tangling measure was not considered when defining the subspaces, ensuring that tangling remains an unbiased metric quantifying the dynamics. Third, the increased tangling is not a trivial consequence of covarying motor cortex with a more tangled cortical region like somatosensory cortex. There was no apparent bias for the communication subspace to have higher tangling on normal trials, and the behavioral covariates used to define the behavioral subspace had the highest tangling overall. Note that although we defined the communication subspace using covariance between motor and somatosensory cortex, our current experiment cannot causally implicate somatosensory cortex in driving these new motor cortical dynamics. Indeed, we would expect similar results if both cortical regions were driven simultaneously by a third region such as the thalamus[16]. Modern experimental tools that facilitate precise, real-time manipulation of neural circuits could help in future work to causally demonstrate the role of subspaces for cortico-cortical communication.

Decomposing population activity into subspaces. Several recent studies have leveraged population recordings from multiple areas to identify subspaces capturing putative cortico-cortical communication[11,21,25], while others have employed targeted analysis methods to identify specific subspaces that capture covariance across conditions[33], task epochs[7], or with behavioral output[34]. The DOS algorithm expands upon many of these approaches by providing a flexible framework that jointly identifies multiple subspaces within the neural manifold capturing specific patterns of covariance with external signals. Here, we apply the algorithm to identify two subspaces within the motor cortical population activity, though it can be readily scaled to larger numbers of brain regions, as well as behavioral or environmental signals.

Activity of many cortical regions can be accurately represented by low-dimensional dynamics, despite the large number of neurons involved in their function[8,21,35–38]. Prior work has shown that untangled population activity allows for robust control signals within motor cortex[17], but is not necessarily an ubiquitous phenomenon across all cortical regions. Here, we have shown in four monkeys that sensory cortical areas are more tangled than motor areas, as has been

previously reported[17]. Yet, our subspace analysis shows that even within a single cortical region the neural population can be decomposed into independently-driven subspaces operating within distinct dynamical regimes that together shape the full population dynamics. In our experiments studying the control of reaching and grasping, motor cortical populations can leverage these subspaces to smoothly integrate new inputs. Broadly speaking, the high possible dimensionality of cortical populations enables a large number of inputs to be simultaneously integrated with ongoing computations by compartmentalizing them into distinct subspaces. Further experiments that simultaneously record from several interconnected cortical regions are needed to confirm whether cortex utilizes this approach to integrate multiple inter-regional inputs.

ACKNOWLEDGEMENTS

We thank Maude Delacombaz and Melanie Kaeser for assistance with behavioral training of the animals. M.G.P. thanks Dr. Juan A. Gallego and Dr. Raeed H. Chowdhury for their valuable discussions in the preparation of this manuscript. Funding: Whitaker International Scholars Program fellowship (M.G.P.), Wyss Center for Bio and Neuroengineering (WCP-008, G.C., M.C. and T.M), GTX medical (G.C, M.C), Bertarelli Foundation Catalyst Fund (BC1709, T.M. and M.C.), the Bertarelli Foundation, and the Swiss National Science Foundation through Ambizione awards (PZ00P2_168103, T.M.; PZ00P2_167912, M.C.), the NeuGrasp project (170032, S.M.), and the Swiss National Centre of Competence in Research (NCCR) Robotics.

AUTHOR CONTRIBUTIONS

Conception and theory: M.G.P.
Experimental Design: M.G.P., S.C., M.B., A.B., B.B., J.B., G.C., S.M., M.C., T.M.
Surgical implantation: M.G.P., A.B., J.B., G.C.
Animal care: M.G.P., S.C., M.B., A.B., S.W., M.C.
Data collection and processing: M.G.P., B.B., S.C., M.B., A.B., S.W., M.C.
Data analysis: M.G.P.
Figure preparation: M.G.P.
Manuscript preparation: M.G.P.
Manuscript and figure revisions: M.G.P., M.C., T.M.

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

REFERENCES

1. Scott, S. H., Cluff, T., Lowrey, C. R. & Takei, T. Feedback control during voluntary motor actions. *Curr. Opin. Neurobiol.* **33**, 85–94 (2015).
2. Sainburg, R. L., Ghilardi, M. F., Poizner, H. & Ghez, C. Control of limb dynamics in normal subjects and patients without proprioception. *J. Neurophysiol.* **73**, 820–835 (1995).
3. Rathelot, J.-A. & Strick, P. L. Subdivisions of primary motor cortex based on cortico-motoneuronal cells. *Proc. Natl. Acad. Sci.* **106**, 918–923 (2009).
4. Dum, R. & Strick, P. Motor areas in the frontal lobe of the primate. *Physiol. Behav.* **77**, 677–682 (2002).
5. Pruszynski, J. A. *et al.* Primary motor cortex underlies multi-joint integration for fast feedback control. *Nature* **478**, 387–390 (2011).
6. Cheney, P. D. & Fetz, E. E. Corticomotoneuronal cells contribute to long-latency stretch reflexes in the rhesus monkey. *J. Physiol.* **349**, 249–272 (1984).
7. Elsayed, G. F., Lara, A. H., Kaufman, M. T., Churchland, M. M. & Cunningham, J. P. Reorganization between preparatory and movement population responses in motor cortex. *Nat. Commun.* **7**, 13239 (2016).
8. Gallego, J. A., Perich, M. G., Miller, L. E. & Solla, S. A. Neural Manifolds for the Control of Movement. *Neuron* **94**, 978–984 (2017).
9. Gallego, J. A. *et al.* Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. *Nat. Commun.* **9**, 4233 (2018).
10. Gallego, J. A., Perich, M. G., Chowdhury, R. H., Solla, S. A. & Miller, L. E. Long-term stability of cortical population dynamics underlying consistent behavior. *Nat. Neurosci.* **23**, 260–270 (2020).
11. Kaufman, M. T., Churchland, M. M., Ryu, S. I. & Shenoy, K. V. Cortical activity in the null space: permitting preparation without movement. *Nat. Neurosci.* **17**, 440–448 (2014).
12. Sadtler, P. T. *et al.* Neural constraints on learning. *Nature* **512**, 423–426 (2014).
13. Cunningham, J. P. & Yu, B. M. Dimensionality reduction for large-scale neural recordings. *Nat. Neurosci.* **17**, 1500–1509 (2014).
14. Churchland, M. M. *et al.* Neural population dynamics during reaching. *Nature* **487**, 51–56 (2012).
15. Michaels, J. A., Dann, B. & Scherberger, H. Neural Population Dynamics during Reaching Are Better Explained by a Dynamical System than Representational Tuning. *PLOS Comput. Biol.* **12**, e1005175 (2016).
16. Sauerbrei, B. A. *et al.* Cortical pattern generation during dexterous movement is input-driven. *Nature* **577**, 386–391 (2020).
17. Russo, A. A. *et al.* Motor Cortex Embeds Muscle-like Commands in an Untangled Population Response. *Neuron* **97**, 953-966.e8 (2018).
18. Sussillo, D., Churchland, M. M., Kaufman, M. T. & Shenoy, K. V. A neural network that finds a naturalistic solution for the production of muscle activity. *Nat. Neurosci.* **18**, 1025–1033 (2015).
19. Pakkenberg, B. Aging and the human neocortex. *Exp. Gerontol.* **38**, 95–99 (2003).
20. Tsodyks, M. V. & Markram, H. The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci.* **94**, 719–723 (1997).
21. Perich, M. G., Gallego, J. A. & Miller, L. E. A Neural Population Mechanism for Rapid Learning. *Neuron* **100**, 964-976.e7 (2018).
22. Stavisky, S. D., Kao, J. C., Ryu, S. I. & Shenoy, K. V. Motor Cortical Visuomotor Feedback Activity Is Initially Isolated from Downstream Targets in Output-Null Neural State Space Dimensions. *Neuron* **95**, 195-208.e9 (2017).
23. Ames, K. C. & Churchland, M. M. Motor cortex signals for each arm are mixed across hemispheres and neurons yet partitioned within the population response. *eLife* **8**, e46159 (2019).
24. Heming, E. A., Cross, K. P., Takei, T., Cook, D. J. & Scott, S. H. Independent representations of ipsilateral and contralateral limbs in primary motor cortex. *eLife* **8**, e48190 (2019).
25. Semedo, J. D., Zandvakili, A., Machens, C. K., Yu, B. M. & Kohn, A. Cortical Areas Interact through a Communication Subspace. *Neuron* **102**, 249-259.e4 (2019).

26. Barra, B. *et al.* A versatile robotic platform for the design of natural, three-dimensional reaching and grasping tasks in monkeys. *J. Neural Eng.* (2019) doi:10.1088/1741-2552/ab4c77.
27. Perich, M. G. & Miller, L. E. Altered tuning in primary motor cortex does not account for behavioral adaptation during force field learning. *Exp. Brain Res.* **235**, 2689–2704 (2017).
28. Scott, S. H. Optimal feedback control and the neural basis of volitional motor control. *Nat. Rev. Neurosci.* **5**, 532–545 (2004).
29. Hatsopoulos, N. G. & Suminski, A. J. Sensing with the Motor Cortex. *Neuron* **72**, 477–487 (2011).
30. Dekleva, B. M., Kording, K. P. & Miller, L. E. Single reach plans in dorsal premotor cortex during a two-target task. *Nat. Commun.* **9**, 3556 (2018).
31. Kaufman, M. T., Churchland, M. M., Ryu, S. I. & Shenoy, K. V. Vacillation, indecision and hesitation in moment-by-moment decoding of monkey motor cortex. *eLife* **4**, e04677 (2015).
32. Kaufman, M. T. *et al.* The Largest Response Component in the Motor Cortex Reflects Movement Timing but Not Movement Type. *eneuro* **3**, ENEURO.0085-16.2016 (2016).
33. Kobak, D. *et al.* Demixed principal component analysis of neural population data. *eLife* **5**, e10989 (2016).
34. Sani, O. G., Pesaran, B. & Shanechi, M. M. *Modeling behaviorally relevant neural dynamics enabled by preferential subspace identification (PSID)*. http://biorxiv.org/lookup/doi/10.1101/808154 (2019) doi:10.1101/808154.
35. Lara, A. H., Cunningham, J. P. & Churchland, M. M. Different population dynamics in the supplementary motor area and motor cortex during reaching. *Nat. Commun.* **9**, 2754 (2018).
36. Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013).
37. Stopfer, M., Jayaraman, V. & Laurent, G. Intensity versus Identity Coding in an Olfactory System. *Neuron* **39**, 991–1004 (2003).
38. Wang, J., Narain, D., Hosseini, E. A. & Jazayeri, M. Flexible timing by temporal scaling of cortical responses. *Nat. Neurosci.* **21**, 102–110 (2018).
39. Capogrosso, M. *et al.* Configuration of electrical spinal cord stimulation through real-time processing of gait kinematics. *Nat. Protoc.* **13**, 2031–2061 (2018).
40. Boumal, N., Mishra, B., Absil, P. & Sepulchre, R. Manopt, a Matlab Toolbox for Optimization on Manifolds. *J. Mach. Learn. Res.* **15**, 1455–1459 (2014).

METHODS

Behavioral task. The monkeys were seated in a custom primate chair that enabled unconstrained interaction with a workspace in front of their body. The experimental platform and task has been previously described in detail[26]. In brief, a small, spherical, custom-molded, silicon object was affixed to the end of a seven-degree-of-freedom robotic arm (Intelligent Industrial Work Assistant, IIWA – KUKA, Augsburg, Germany). Custom control software enabled positioning of the robot in space. For Mk-Jo, only one position was provided, roughly centered in front of the body. For Mk-Cs, Mk-Br, and Mk-Ol, three positions were used: left, middle, and right. All positions were in the same coronal plane. The monkeys were trained to freely reach for the object, grasp it using only their left hand, and then pull it towards the body. In each trial, the robot brought the object to one of three randomly-selected positions (except Mk-Jo, where only one position was used). While the object was displaced from its starting position, the robot provided a force towards the starting position that was proportional to the horizontal displacement. Therefore, the monkeys had to resist this force to pull the object and, if they released it, the object quickly returned to its starting position. The system recorded the interaction forces produced by the robot.

Each trial was deemed successful after the position of the object passed a pre-determined distance threshold. Pulling movements that failed to reach this threshold were excluded. Upon success, Mk-Cs and Mk-Jo manually received a food reward, while Mk-Br and Mk-Ol received an automated liquid reward through a sipper tube. Error trials were manually identified by assessing the videos and kinematic trajectories. Most of the errors occurred when the monkey missed the robot in the initial reach or when their hand slipped off the robot while pulling. Typically, this resulted in errors preceding the final successful grasp, though for Mk-Br most errors occurred after grasping the object in order to adjust the posture of the hand to enable the pulling movement. While performing the task, the monkeys were affixed with up to 15 reflective markers to track kinematics of the joints of the arm and hand (Vicon Motion Systems, Oxford, UK).

Surgical procedures. All surgical procedures were performed using aseptic technique under general anaesthesia (induction: 0.1 mg/kg midazolam and 10 mg/kg ketamine; maintenance: 5 ml/kg/h propofol and 0.2-1.7 ml/kg/h fentanyl IV). A certified neurosurgeon (Dr. Jocelyne Bloch, CHUV, Lausanne, Switzerland) supervised all procedures. Each monkey was implanted with two Utah electrode arrays (Blackrock Microsystems, Salt Lake City, UT, USA). The first array targeted the arm region of motor cortex near the central sulcus (primary motor cortex; 1.5 mm shaft length) and the other targeted somatosensory cortex (Brodmann's Area 2; 1 mm shaft length). The two arrays were connected to a single pedestal mounted to an implanted titanium that covered a portion of the skull. The mesh was molded to the shape of the skull of each monkey using a plastic skull 3-D printed from an individualized MRI scan. The implants in Mk-Cs and Mk-Br targeted arm areas identified by intraoperative surface stimulation (biphasic pulses, 3mA), while the implants in Mk-Jo targeted the hand areas. All arrays had 64 channels in an 8x8 configuration except the somatosensory cortical array of Mk-Br, which was 32 channels in an 8x4 configuration. Mk-Br and Mk-Ol each additionally received a third array (32-channel, 1mm shaft) anterior to the first array near to ventral premotor cortex. We saw qualitatively similar dynamics in this and the array implanted near the Central Sulcus. We therefore combined the signals recorded from both arrays in each monkey and for

simplicity refer to this combined signal as "motor cortex". In the same surgery, Mk-Br also received bipolar EMG electrodes in the left arm to record the activation of eight muscles of the arm and hand. Detailed surgical and post-operative care procedures have been described previously[39].

Data acquisition. As the monkeys performed the behavioral task, we recorded infrared video from 12 cameras to track the kinematic markers at 100 Hz (Vicon Motion Systems, Oxford, UK) and neural activity at 30kHz using a Cerebus system (Blackrock Microsystems, Salt Lake City, UT, USA). To process the neural data, we bandpass-filtered each channel at 750 to 5000 Hz and set a threshold between -5x and -6.25x the RMS value to extract spike events. We sorted these spikes using Offline Sorter (Plexon, Dallas, TX, USA) to identify putative single neurons. We counted the number of spikes occurring in 10ms bins matched to the 100 Hz kinematic data. We used these binned spike counts in the Generalized Linear Model analysis (see below). For all other analyses, we square-root transformed each spike train to stabilize the variance[13] and then converted them to an instantaneous firing rate by convolution with a Gaussian kernel with a 75 ms standard deviation.

Calculation of neural and behavioral trajectories. We applied a soft normalization procedure[14] to all of the smoothed single neuron firing rates. We then used Principal Components Analysis (PCA) to reduce the dimensionality of each brain region independently and compute a neural manifold[8]. We selected the first 20 dimensions, though in practice our results were not impacted by the exact choice of dimensionality within reasonable ranges (10-30 dimensions; data not shown). We then computed a "behavioral manifold" using a similar procedure. First, we z-scored each of the position, velocity, acceleration, and force signals for the limb and robot. For Mk-Br, we also added the EMG recordings to this behavioral manifold to most accurately estimate the behavioral dynamics. We then applied PCA to this set of signals and selected the first 20 dimensions to correspond to the 20 dimensions of the neural manifold. The match in dimensionality was necessary for the subspace decomposition analysis described below, as well as to compare the trajectory tangling between manifolds.

Computing trajectory tangling. We employed trajectory tangling to quantify the relative dynamical behavior of the neural and behavioral dynamics. Tangling values should be highest at points in time where the current state could lead to multiple future states. We computed the instantaneous tangling (T) at each time t of the neural and behavioral trajectories using the formulation of [17].

$$T(t) = \max_{t'} \frac{\|\dot{x}_t - \dot{x}_{t'}\|^2}{\|x_t - x_{t'}\|^2 + \varepsilon}$$

The metric looks over all time points for all trials for each position, and is maximal when a similar position in state space (x) – which gives a small denominator value – corresponds to different state space velocities ($\dot{x}$) – which gives a large value in the numerator. The value of ε was chosen to be very small (10^{-6}) and served only to keep the value from becoming undefined if the denominator was otherwise equal to zero. Note that we computed the tangling independently for all trials for each position, though we obtained similar results computing the tangling across all trials together. The formulation above, taken from Russo et al., was applied to trial-averaged data. Thus, noise and outliers were very unlikely. We modified the approach to take the 99.99th percentile across all time points in each position instead of the max in order to ignore a very small number of outliers (e.g. artifacts) in our single trial data.

Tangling values are sensitive to the dimensionality of the state space and, therefore, can be directly compared only between spaces with matched dimensionality. To facilitate easy comparison across signals or conditions, we ensured that dimensionalities were matched (20 dimensions for manifolds, 5 dimensions for subspaces). To study the time course of tangling changes, we interpolated to time-warp each trial to be the same length, preserving the proportion of time spent on average in the reaching and pulling phases. We then averaged across normal or error trials (e.g., **Fig. 3d**) To compute an aggregate single-trial tangling value, we computed the root-mean-square tangling through the trial (e.g., **Fig. 3b**) or throughout the specific period of behavioral error correction (e.g., **Fig. 3e**).

Decomposition into Optimal Subspaces (DOS) algorithm. We based DOS on a published method that jointly optimizes subspaces separating preparatory and movement activity within motor cortex[7]. We designed DOS to identify two orthogonal subspaces within the neural manifold, one capturing activity that covaried with the neural activity of another brain region (*communication subspace*) and the other capturing activity that covaried with behavior (*behavioral subspace*). Our method looks for eigenvectors spanning subspaces of the neural manifold that maximize the covariance with external covariates. The relevant value to be optimized is defined as:

$$\frac{Tr(Q^T C Q)}{\sum_{i=1}^{d} \sigma_i}$$

Where:

Q: basis vectors of the subspace

C: covariance of neural activity and external covariates

d: dimensionality of subspace

σ_i: variance explained in the neural population by each subspace dimension

The numerator effectively measures the amount of covariance between neural activity and external covariates captured by the subspace. We defined C as the covariance between a matrix of N-dimensional covariates and N-dimensional neural population activity, which we can express as Y^TX. This formulation ignores a global scaling term for covariance since we are deriving unit basis vectors.

Using this metric, a cost function can be defined that optimizes for any arbitrary number of subspaces, with any value of d for each subspace, so long as the sum of their dimensionality is less than or equal to N. In our analysis, we defined two subspaces. For the behavioral subspace, Y was the activity within the 20-dimensional behavioral manifold identified from the behavioral covariates described above. For the communication subspace, Y was the activity within the somatosensory cortical manifold. These were referenced against X, the activity within the motor cortical manifold.

We defined the cost function using the above metric:

$$[Q_{com}, Q_{beh}] = argmax_{[Q_{com},Q_{beh}]} \frac{1}{2}\left(\frac{Tr(Q_{com}^T C_{com} Q_{com})}{\sum_{i=1}^{d_{com}} \sigma_{com,i}} + \frac{Tr\left(Q_{beh}^T C_{beh} Q_{beh}\right)}{\sum_{i=1}^{d_{beh}} \sigma_{beh,i}} - \alpha\left(\sum_{i=1}^{d_{com}} rms(\|Q_{com,i}\|) + \sum_{i=1}^{d_{beh}} rms(\|Q_{beh,i}\|)\right)\right)$$

Subject to the constraints:

$$Q_{com}^T Q_{beh} = 0,\ Q_{com}^T Q_{com} = I, \qquad Q_{beh}^T Q_{beh} = I$$

Where:

Q_{com}: basis of communication subspace

Q_{beh}: basis of behavioral subspace

X: local neural population activity (e.g. motor cortex)

Y_{neur}: neural population activity of other brain region (e.g. somatosensory cortex)

Y_{kin}: kinematic and behavioral covariates

d_{com}: dimensionality of the communication subspace

d_{beh}: dimensionality of the behavioral subspace

$C_{com} = Y_{neur}^T X$: covariance of local neural population activity with other brain region

$C_{beh} = Y_{kin}^T X$: covariance of local neural population activity with behavioral covariates

The first two terms of the cost function maximize the ability of the communication subspace to explain the covariance of the neural population with the other brain region, as well as the ability of the behavioral subspace to explain the covariance of the neural population with the kinematic signals. The last term penalizes sparsity in the subspace weights. This term is maximized when the weight terms are evenly distributed across the dimensions. In practice, our results on the neural datasets were qualitatively similar without this sparsity constraint, however from our simulation experiments described below, this term was important to avoid edge cases. For example, if a high-variance PC can explain a lot of behavioral covariance, the algorithm may settle into local minima by attributing a high weight to one dimension. Instead, we want to bias towards finding distributed representations across dimensions. Typically, such sparsity constraints are achieved using the L1-norm. However, since the algorithm optimizes for a basis set, the L1-norm of Q is always 1 regardless of the weights. To capture the same intuition of penalizing sparse weightings to a basis set, we opted to use the root-mean-square of the weights. The α term helps to balance the contribution of the different terms.

The cost function used here can be viewed as analogous to that of Reduced Rank Regression, which specifically finds dimensions within the full population space that covary with other signals[25]. It is important to note that there is no constraint in the cost function to explain neural variance, only to explain the *co*variance with external signals. In theory, one of the subspaces could actually explain very little neural variance. The optimization was performed using a Matlab-based manifold optimization framework[40]. Thus, our algorithm can jointly optimize multiple orthogonal subspaces within the same neural population, each subspace designed to explain the covariation of the full population with external signals. A key assumption in the above algorithm is that the dimensionality is matched across all sets of signals, i.e., that C_{beh} and C_{com} are square and of identical size. We achieved this by performing PCA on the population activity and kinematic signals before running this algorithm and selecting the same number of dimensions for all sets of signals. We used DOS to identify motor cortical communication and behavioral subspaces from neural activity and behavior covariates from normal trials only. To analyze the error trials, we projected the neural activity from the error trials onto these subspaces identified from normal trials. This ensured that the error analysis results were not biased by the optimization algorithm.

We quantified the amount of variance captured by each of the behavioral and communication subspaces using the

eigenvalues of the covariance matrix between the subspaces and the activity in the motor cortical manifold.

Simulation to validate the DOS algorithm. We simulated population of spiking motor and somatosensory cortical neurons where the ground truth latent dynamics are known to confirm that the above decomposition algorithm can correctly reconstruct the latent dynamics in the intended subspaces. We generated two sets of latent dynamics from a one-dimensional random walk, termed behavioral and communication. We then smoothed these random dynamics with a Gaussian kernel of 50 ms width to obtain smoothness that is comparable to that of the neural data. We developed DOS with intention to identify specific patterns of covariance, even in the face of confounding correlated signals. To test the accuracy of the algorithm, we added two confounds. First, both the motor and somatosensory cortical populations received correlated sinusoidal dynamics to simulate the effect of common input to both regions (referred to as null dynamics). Second, the somatosensory cortical population received the precise communication dynamics as well as a time-shifted version of the behavioral dynamics meant to simulate the real-world delays between intended motor output (here assumed to be 50ms) and the corresponding sensory feedback (assumed to be 50ms).

The motor and somatosensory cortical population activity was generated as a Poisson process whose input was randomly weighted combinations of the null, communication, and behavioral dynamics. We then processed the simulated spiking activity with the same methods as the real neural data, including normalization, smoothing, and dimensionality-reduction through PCA. We identified the subspaces within the simulated motor cortical data using the covariance between motor and somatosensory cortical PCs, as well as the covariance between motor cortex and the known behavioral dynamics. We quantified the performance by computing the variance explained (R^2) between the ground truth dynamics and the extracted subspace dynamics. We repeated this random simulation 1000 times to generate the performance distributions. On each simulation, we compared the performance of DOS to a null distribution found by taking the max correlation across all individual motor cortical PCs. This null distribution served to confirm that the reconstructed dynamics did not trivially arise from covariance across neurons found by PCA, but instead required a targeted approach like DOS.

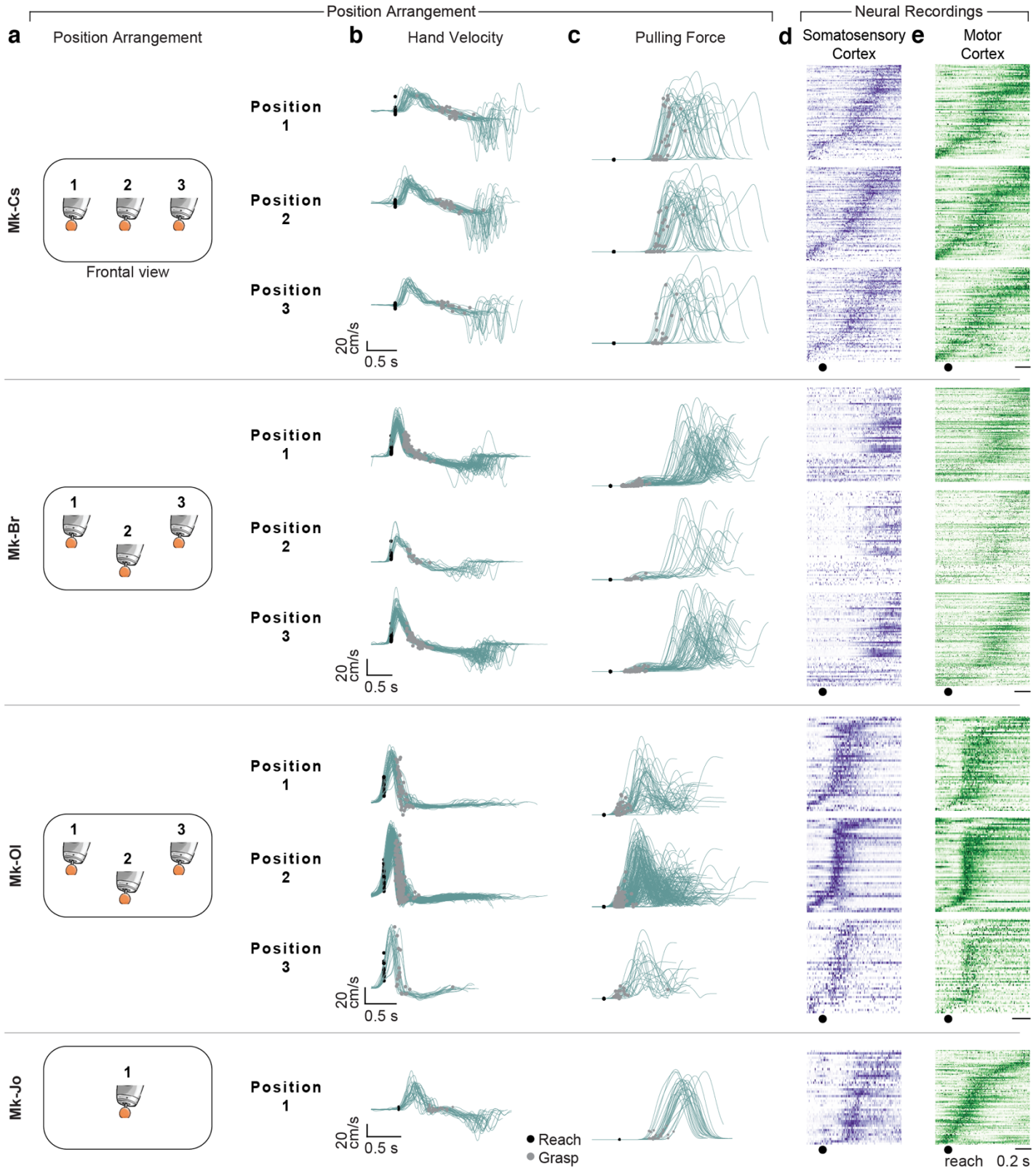

Figure S1. *Detailed description of behavioral and neurophysiological dataset.* **(a)** Arrangement of object positions in the coronal plane for each monkey. Mk-Jo reached to only one position in the center of the workspace, while the other three monkeys reached to three positions. **(b)** Plots show hand velocity traces for each position for all monkeys. Trajectories are duplicated from Fig. 2c, but separated by position. Black circles denote the velocity at the time of reach onset marked from the video recordings, and gray circles denote velocity at the time of object grasp. **(c)** Plots show the magnitude of pulling force registered by the force transducer in the robot for each position. **(d)** Peri-event histograms for somatosensory cortex spiking activity (spikes counted in 10ms bins aligned on reach onset). Each row represents the sum across trials of one neuron. **(e)** Data presented as in Panel d, but for motor cortical neurons.

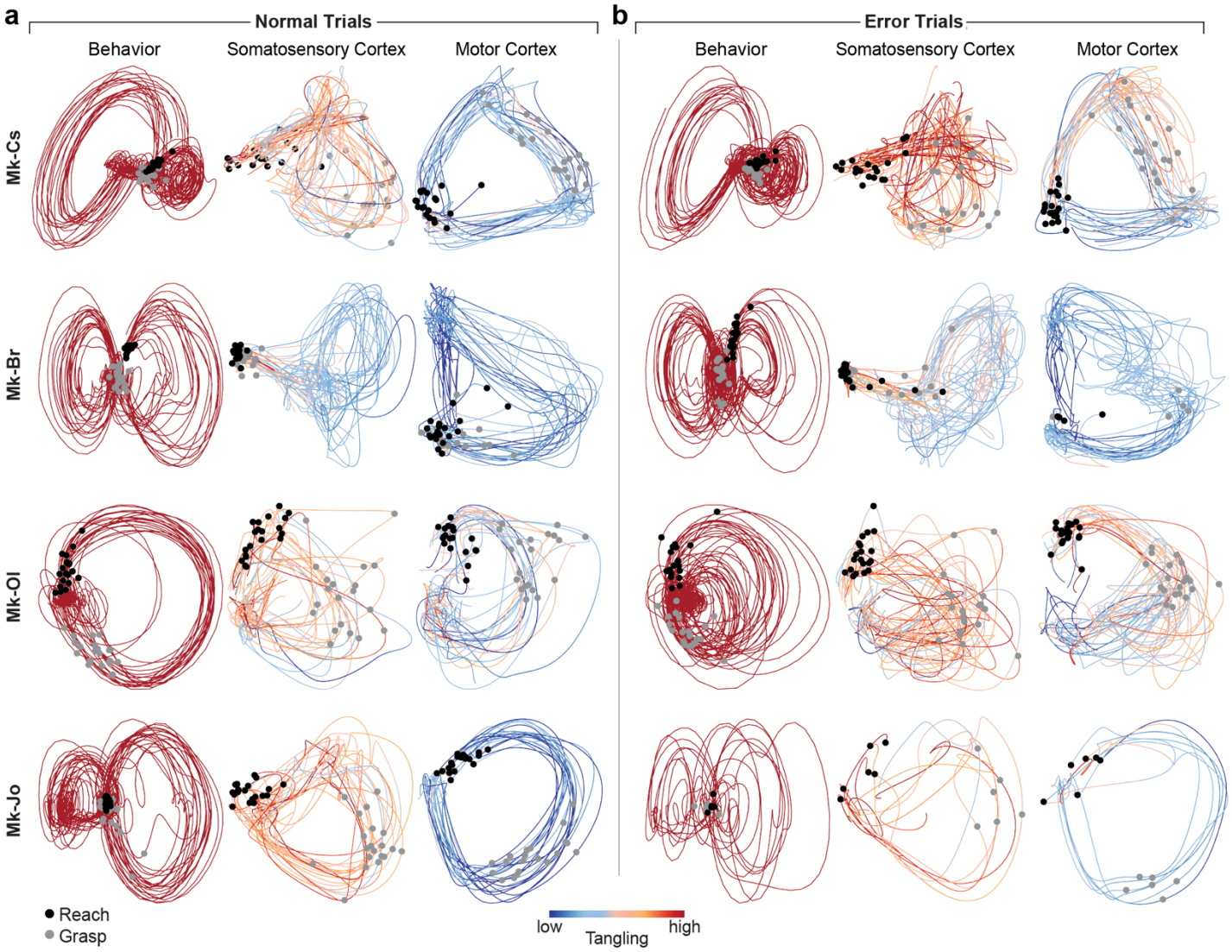

Figure S2. *Comparison of the trajectories and tangling between behavioral, somatosensory cortical and motor cortical manifolds during normal and error trials.* **(a)** Trajectories of the behavioral manifold are substantially more tangled than the somatosensory cortical manifold trajectories, which are substantially more tangled than the motor cortical manifold trajectories on normal trials. Plots show trajectories on 20 normal trials for behavioral (left), somatosensory cortical (middle), and motor cortical manifolds (right) shown in the two leading principal components for the four monkeys (four rows). Colors indicate tangling value at each time point. Data are reproduced from Fig. 3a. **(b)** Tangling in both somatosensory and motor cortical manifolds increases during the error trials. Plots show trajectories for up to 20 error trials for behavioral (left), somatosensory cortical (middle), and motor cortical manifolds (right) for the four monkeys. Color scale is identical for each monkey to the scale in Panel a.

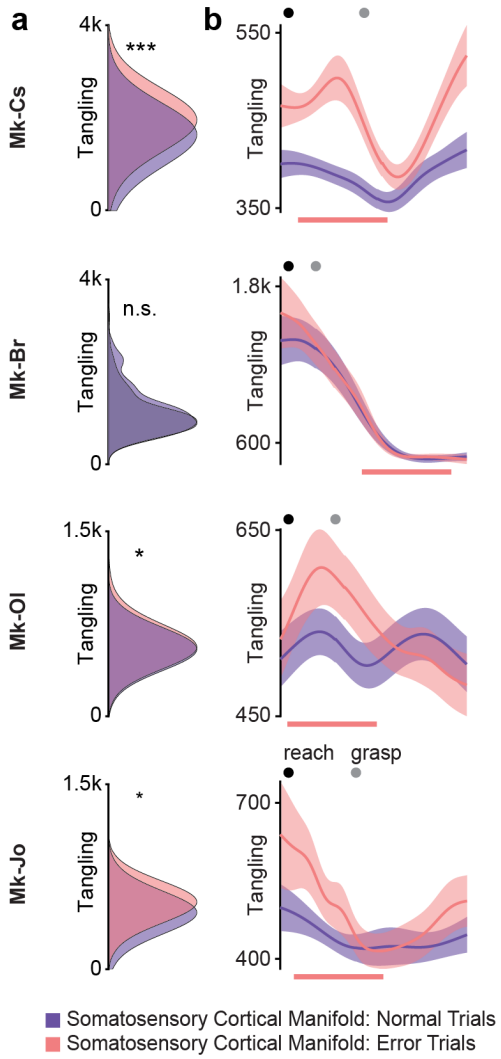

Figure S3. *Errors lead to a slight increase in the tangling of somatosensory cortical manifold.* **(a)** Plots show single-trial tangling distributions in the somatosensory cortical manifold for all normal (purple) and error (pink) trials. All monkeys except Mk-Br showed a significant increase in tangling on error trials, though Mk-Br did occasionally show small, transient increases in tangling (see Fig. S2b). *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$; ranksum. **(b)** Somatosensory cortical tangling increases surrounding the error correction in all but one monkey. Tangling in the somatosensory cortical population averaged across all normal (purple) and error (pink) trials after aligning on reach onset and time-warping to match the length across all trials. Black and grey circles indicate the reach onset and object grasp, respectively. Pink bars indicate approximate windows of error correction or each monkey. Error bars: mean ± s.e.m.

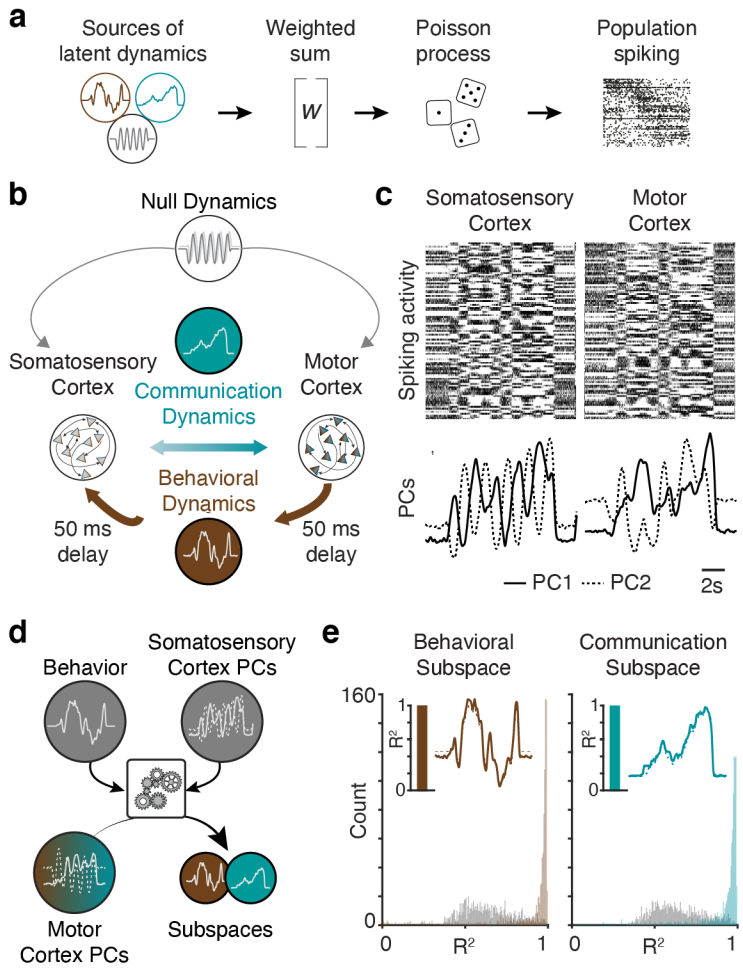

Figure S4. *Our Decomposition into Optimal Subspaces (DOS) algorithm accurately isolates subspaces from simulated neural activity.* **(a)** We modeled neural activity of 100 motor and somatosensory cortical neurons whose activity was generated by randomly weighted combinations of sources of latent dynamics. **(b)** We defined three types of latent dynamics: (i) communication dynamics shared between both regions, (ii) behavioral dynamics that were also correlated between regions but delayed in somatosensory cortex to account for sensory feedback delays, and (iii) oscillatory null dynamics intended to confound the algorithm. **(b)** We performed the same processing on the simulated neurons as the real neurons in Fig. 3 to reduce the dimensionality of the spiking activity (top) using PCA (bottom shows the dominant two principal components (PCs). Each PC reflects a mix of the three sources of dynamics, thus PCA alone is not sufficient to separate the dynamics. **(c)** The DOS algorithm aimed to isolate the communication and behavioral dynamics into orthogonal subspaces by referencing the motor cortical activity against the behavioral dynamics (assumed to be measurable as in the experimental data), and the somatosensory cortical manifold. **(d)** For the example simulation shown in Panels a-c, the algorithm achieved near-perfect reconstruction ($R^2 > 0.95$) of the behavioral and communication dynamics in the two orthogonal subspaces (insets). We performed 1000 such simulations and the algorithm consistently performed well (median R^2: 0.97), significantly better than the max across all principal components in gray ($p < 0.001$, ranksum).

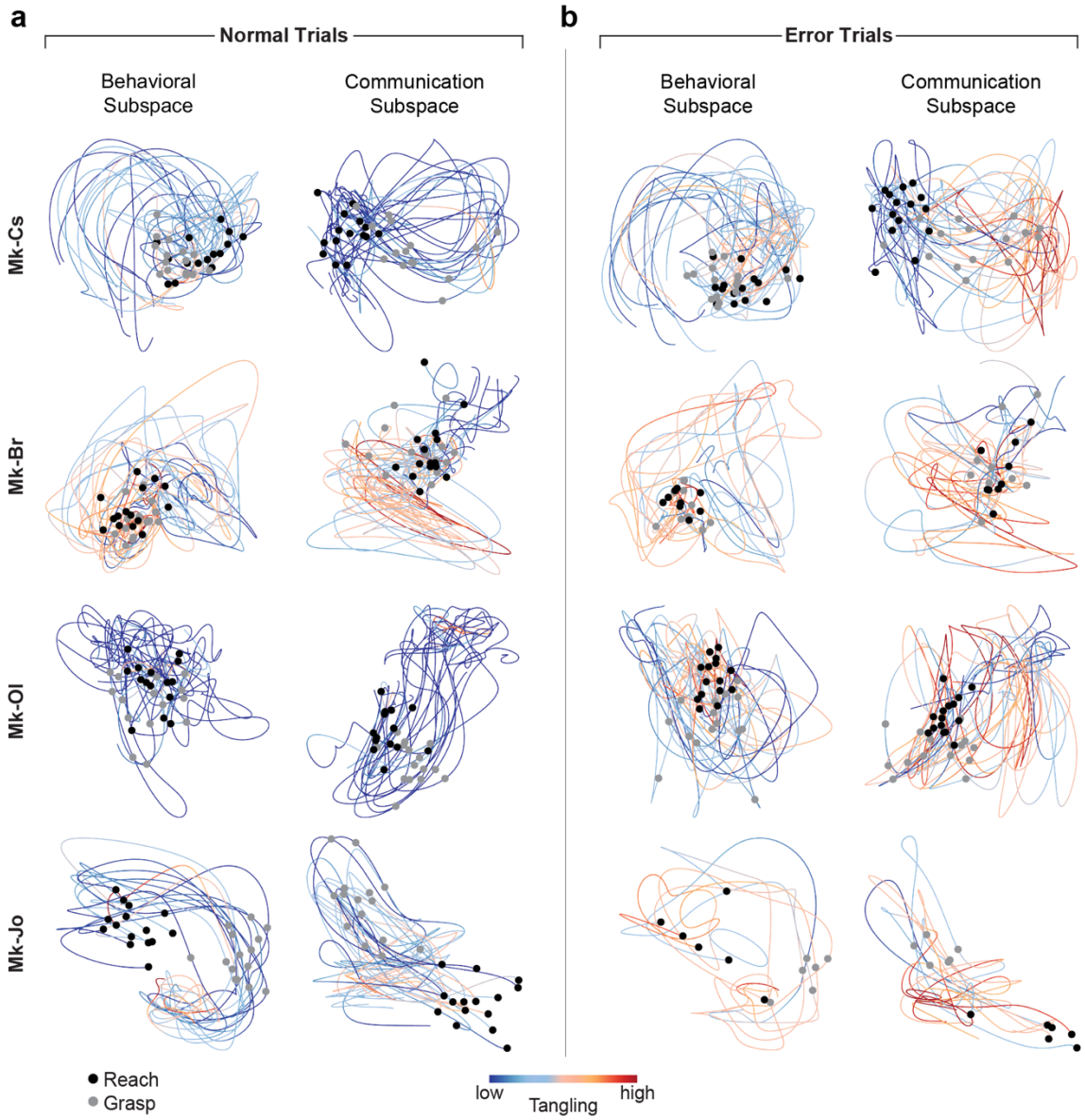

Figure S5. *Trajectories in the communication subspace increase in tangling during errors, while the tangling of behavioral subspace trajectories does not change.* **(a)** Neural trajectories within the behavioral (left) and communication (right) subspaces for 15 normal trials for the four monkeys. Data plotted as in Fig. 6a; Mk-Cs trajectories reproduced from that figure. **(b)** Neural trajectories in the behavioral (left) and communication (right) subspaces for up to 15 error trials for the four monkeys. Color scale is consistent with Panel a. In the error trials, the tangling of behavioral subspace trajectories remains similar to normal trials, while the tangling of communication subspace trajectories substantially increases.

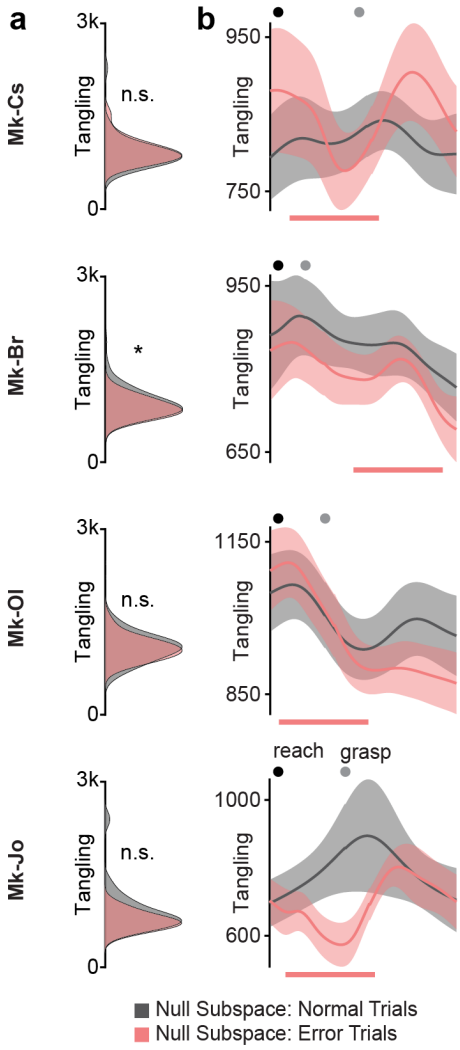

Figure S6. *Errors lead to no increase in the tangling of the null subspace of the motor cortical manifold.* **(a)** Single-trial tangling distributions in the null subspace for all normal (gray) and error (pink) trials. No significant difference was observed for Mk-Cs, Mk-Ol, or Mk-Jo, though Mk-Br showed a significant decrease on the error trials. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$; ranksum. **(b)** Plots show tangling in the null subspace (see Methods) averaged across all normal (gray) and error (pink) trials after aligning on reach onset and time-warping to match the length across all trials. Black and grey circles indicate the average time of reach onset and object grasp, respectively. Pink bars indicate approximate windows of error correction or each monkey. Error bars: mean ± s.e.m.