Simulation of fronto-parietal neural population activity to probe dimensionality reduction methods.

Dissertation

for the award of the degree

"Doctor rerum naturalium"

of the Georg-August-Universität Göttingen

within the doctoral program *Sensory and Motor Neuroscience*

of the Georg-August University School of Science (GAUSS)

submitted by

Andrej Filippow

from Ekaterinenburg, formerly Sverdlovsk, Russia

Göttingen, 2021

Thesis Committee

Prof. Dr. Hansjörg Scherberger
Research Group Neurobiology
German Primate Center
Kellnerweg 4, 37077 Göttingen

Prof. Dr. Alexander Gail
Cognitive Neuroscience Laboratory
German Primate Center
Kellnerweg 4, 37077 Göttingen

Prof. Dr. Florentin Wörgötter
Department of Computational Neuroscience
Third Physical Institute/BCCN
Georg-August University
Friedrich-Hund-Platz 1, 37077 Göttingen

Members of the Examination Board

Prof. Dr. Annekathrin Schacht
Effective Neuroscience and Psychophysiology
Georg-Elias-Müller-Institute of Psychology
Georg-August University
Goßlerstraße 14, 37073 Göttingen

Dr. Caspar M. Schwiedrzik
Neural Circuits and Cognition
European Neuroscience Institute
Grisebachstr. 5, 37077 Göttingen

Prof. Dr. Tim Gollisch,
Department of Ophthalmology
School of Medicine
Georg-August University
Waldweg 33, 37073 Göttingen

Date of oral examination: September 28th, 2021

Herewith I declare that I have written this thesis independently and with no other aids and sources than quoted.

Göttingen, 30.07.2021 Andrej Filippow

Acknowledgments

First of all, I want to thank Hans Scherberger for providing me the opportunity to work in an open, relaxed and supportive laboratory. My project would not have been possible without his guidance and he provided me with feedback and reassurance whenever I needed it.

I would also like to thank the other two members of my thesis committee, Alexander Gail and Florentin Wörgötter, for the constructive criticism and support provided in our meetings.

I thank especially Benjamin Dann for collaborating with me, providing me with scientific ideas and data to build my project upon and teaching me how to write scientifically. His experience and knowledge greatly contributed to the success of this project.

I would furthermore like to thank the other members of the lab for fostering a positive and friendly atmosphere and especially Sabine Borchert and Nathalie Bobb for introducing me to animal training and handling, and Matthias Dörge for his technical support. Thank you, Andres Agudelo-Toro, for introducing me to Hans and to the lab, and for supporting me during my first time in the lab.

My thanks go out to my family, my parents Sergej and Luisa, and my aunt Klara for supporting and enabling me in my scientific career, and finally, thanks Léana, for being there for me and believing in me.

Contents

1 Introduction

In most animals, the brain is the center for the integration of a wide range of sensory inputs, the cognitive processing of these inputs, and the rich variety of behaviors generated in response. The most essential cell type to generate these mental processes are neurons forming a wide-spanning, densely connected network. Information is transmitted in the network by short electrical impulses called action potentials. And while the mechanisms and functions of each individual neuron can be understood and described in simple electromechanical terms, the immense complexity that arises from the interactions of large networks of neurons is still far beyond our current understanding.

However, ever since the recordings of the simultaneous activity of neural populations first became feasible, neuroscience has been searching for a basis of this complexity, for driving principles that would explain the dynamics of neural population activity. A major step towards that goal was the finding that the responses of single neurons are correlated and to a high degree redundant. The correlations and redundancy suggest that the activity of large, high-dimensional neural populations can be explained by the activity of a much smaller set of independent variables that can't be observed directly, the so-called latent variables. A way to extract these low-dimensional variables are dimensionality reduction methods, a diverse group of mathematical algorithms. However, many dimensionality reduction methods have not been extensively examined with respect to their applicability to neural population activity. Such an evaluation is difficult, because we do not know the latent variables structuring the population response of recorded brain activity. Therefore, we cannot directly evaluate the performance and accuracy of any dimensionality reduction method.

In this thesis, we therefore present a detailed simulation of neural population activity to evaluate the performance and accuracy of dimensionality reduction methods using this type of data. We model a real recording of the neural population response of two macaques to a visually cued delayed grasping task. First, we matched a large variety of population features of recorded data as accurately as possible in our model to maximize relevance to recorded data. Second, we use the model to evaluate the most widespread dimensionality reduction method, principal component analysis (PCA). Based on our model, the application of PCA to

neural data can result in (1) a mis-estimation of the true dimensionality of the dataset, (2) failure to capture latent variables of small amplitude and (3) principal components that consist of mixtures of individual latent variables and noise. For these reasons, prior information about the dynamics and features of the analyzed dataset is necessary to interpret the dimensionality, or the individual components recovered by PCA.

1.1   History of recordings of neural population activity

To understand the dynamics and interactions of a neural population, we first need to simultaneously record the activity of large numbers of neurons. Over the history of neuroscience, a number of different methods to capture the activity of neurons has been developed. In general, the approaches range from fine-scale recordings of individual neuronal activity to large-scale recordings of the activity of entire areas of the brain.

The signal transfer between neurons is a complex electrochemical process, of which electrophysiological recording methods capture the electric component. Neuronal function is based on the variable voltage across the membrane of the neuron, which is regulated by the flow of electric current, in the form of ions, through protein channels and pumps that cross the membrane (Schwartz et al., 1991). Individual neurons receive excitatory or inhibitory inputs through chemical or electric synapses, typically situated at the dendrites and soma of a neuron. The inputs result in changes in the electric potential across the membrane of the cell. Once this post-synaptic potential crosses a certain threshold, the neuron "fires". A strong depolarization of the neuronal membrane potential, called action potential, is evoked and travels in a self-propagating manner down the axon of a neuron, where it is passed through synapses to the next neuron (It is thus called the pre-synaptic potential). Action potentials are short, the membrane returns to its resting potential within one to two milliseconds (Whittingstall and Logothetis, 2013). The amplitude of the action potential is constant, and stronger excitation of the neuron is reflected in a higher firing frequency.

The changes in the potential of the cell membrane result in currents flowing through both the cell and the extracellular medium. The currents within the cell can be measured by inserting an electrode directly into the cell, or attaching a patch clamp to the cell membrane. Such intra-cellular measurements can capture both presynaptic action

potentials and postsynaptic fluctuations of the membrane potential. However, intra-cellular recordings are more difficult to conduct than other recording methods, because the electrode needs to pierce the membrane and patch clamps need to be placed onto the membrane surface. Intra-cellular recordings are therefore mostly performed in vitro on single cell preparations (Perin et al., 2011).

Neural activity can also be measured without piercing or attaching to a cell membrane, by placing an electrode in the proximity of a neuron. Such extracellular electrode recordings are significantly easier to conduct in vivo and multiple electrodes can be used simultaneously and chronically in an electrode array (Nicolelis et al., 1997; Musallam et al., 2007). Each electrode captures the electric activity of multiple nearby neurons (multi-unit activity, MUA), however, with a much worse signal-to-noise ratio than intracellular recordings. The action potential of each neuron has a characteristic temporal and spatial (due to the position of the electrode relative to the neuron) profile that is different even between neurons of the same type. It is therefore possible to sort, to a degree, multi-unit activity of a single electrode into action potentials attributable to individual neurons and from background noise. This process of spike sorting is a difficult and computationally intensive process, that is often done at least in part manually (Wood et al., 2004; Pachitariu et al., 2016).

The flow of ions through the membrane also results in a change of the potential outside of the neuron, called the Local Field Potential (LFP). Extracellular electrode recordings capture both LFP and MUA simultaneously. Changes in LFP are predominantly attributed to postsynaptic activity and LFP is assumed to capture the combined postsynaptic activity on a larger scale than MUA (Katzner et al., 2009; Whittingstall and Logothetis, 2013), in other words, more neurons contribute to the LFP recorded by an electrode than to the MUA recorded by the same electrode. LFP and spiking activity occupy different ranges of electromagnetic frequency band; MUA is high-frequency, while LFP activity is of lower frequency. Therefore, LFP and MUA can be separated by filtering the recorded signal with a high-pass or a low-pass frequency filter. When considering the LFP of a population of neurons, LFP reflects predominantly activity that is positively correlated across single neurons rather than uncorrelated activity (Einevoll et al., 2013). Therefore, we can't record the part of the neural signal that is not correlated across neurons.

Another field of methods that is capable of simultaneously recording large numbers of single neurons is optical imaging technics. The changes in voltage across the cell membrane caused by changes in ion concentration can be made visible under a microscope, by using dyes that are sensitive to changes in voltage (Chemla and Chavane, 2012) or to calcium ion concentrations (two-photon calcium microscopy) (Stosiek et al., 2003). In either case, the dyes react to external excitation by photons of specific wavelength and emit photons with a wavelength that depends on the voltage across the membrane or the concentration of calcium ions. This fluorescence can be recorded by methods of optical microscopy, with high temporal and spatial resolution. It should be noted that the dyes used can be phototoxic and, together with the intensity of illumination needed for high-resolution microscopy, can be damaging to the recorded cells in the long term.

Non-invasive methods are also widely used. Electroencephalography (EEG) measures the electric field of the brain through the skin, skull and dura mater with a high temporal resolution, but the spatial resolution is very poor and the signal-to-noise ratio is very low. Functional imaging methods, like positron emission tomography and functional magnetic resonance imaging, capture brain activity indirectly by measuring changes in blood flow in response to changing neural activity (Logothetis et al., 2001). The spatial resolution of functional methods can reach values of around a millimeter, but the temporal resolution is limited by the slow response of the blood flow to increased neural activity. Nevertheless, these methods can be used to record the entire brain at the same time (at least in rodents).

Figure 1: The Neuropixels probe is one of several novel recording methods that allow for a simultaneous recordings of large numbers of neurons. A: Illustration of the >900 recording sites in relation to the probe tip. B: Microscope image of the tip. C: The entire probe with headstage, cables and electrode shank. Adapted from (Jun et Al. 2017).

In general, recording methods that capture the activity of larger neural populations do so with a worse resolution than single-neuron recording methods. An important effect was touched upon in the paragraph about LFP; The combination of the activity of multiple neurons into one measurement

is effectively an average over the signals of individual neurons (Einevoll et al., 2013). Low and high neural signals measured simultaneously cancel each other out on average, thus, we lose the uncorrelated or negatively correlated signals of individual neurons. If we want to fully understand the dynamics of the brain, in single-spike resolution recordings of single-neuron spiking activity are preferable.

Recent developments of new single-neuron recording methods and improvements of existing methods allow the simultaneous recording of very large numbers of neurons with high temporal and spatial resolutions. Multielectrode arrays allow the simultaneous recording of large neural populations with an excellent temporal and spatial resolution (Rousche and Normann, 1998; Nicolelis et al., 2003; Musallam et al., 2007). By integrating more than 900 electrodes on a single silicon probe shaft, recording devices like the Neuropixel probe (Jun et al., 2017) (Fig. *1*) or the NeuroSeeker (Raducanu et al., 2017) allow simultaneous recordings of hundreds to thousands of neurons. Resonance-scanning two-photon calcium microscopy was used to record ~10,000 cells simultaneously in vivo (Stringer et al., 2018, 2019), but the temporal resolution was low, at only 0.7 Hz. Now that recordings of the activity of large neuronal populations are possible, the important question is how to analyze and comprehend the dynamics of the activity of such large populations.

1.2   The neuron doctrine and the representational perspective

The theory that individual neurons are the structural and functional units of the nervous system was formulated by Santiago Ramón y Cajal (y Cajal, 1888) which is now known as the neuron doctrine (Yuste, 2015). Research under the neuronal doctrine was the focused on biological, structural, and physiological and computational mechanisms of individual neurons, eventually resulting in a comprehensive mathematical model of how action potentials are generated and propagated (Hodgkin and Huxley, 1952). While such a model could explain how and when a neuron generates action potentials, it could not explain how information of many neurons is combined to perform computations relevant behavior and cognition .

A first insight into the function of neurons was gained with the discovery of receptive fields. Studies of single neurons in the skin of dogs (Sherrington, 1906) and later in the visual cortex of cats (Hubel and Wiesel, 1962) revealed that their firing activity changed in

response to stimuli, but only if the stimulus was localized in specific areas, the receptive fields of the skin and the visual field of the eye. Furthermore, neurons in the visual cortex also responded specifically to parameters of the stimulus that are not captured by single specific receptors, like movement direction. The activity of single neurons was said to be tuned to a specific stimulus parameter, or conversely, a neuron represented some external sensory stimulus or behavioral parameter. Later studies even discovered neurons in the visual cortex in the monkey that responded to facial images of specific individuals (Desimone et al., 1984; Tanaka, 1996). Studies of motor-related areas also revealed representations of different movement parameters, like muscle activity (Georgopoulos et al., 1984; Kalaska et al., 1989) or joint angles (Evarts, 1968, 1969). These, and many similar, findings gave rise to what is now known as the representational view (Yuste, 2015). Under the representational view, the variability in the activity of single neurons was sought to be explained as encodings, or representations of external variables.

Donald Hebb was among the first to suggest that brain function arises not from single neurons, but from "cell assemblies", groups of neurons for which functional connections to each other would be stronger than to other neurons (Hebb, 1949; Buzsáki, 2010). The activity of a neuronal assembly would form the functional unit of cognition, and neurons would congregate into assemblies and disband again, as the functions of the brain demanded. The existence of discrete assemblies as Hebb described them was difficult to prove experimentally (Gerstein et al., 1989; Singer, 1999) but the idea that the function of neurons should be studied on the level of groups, or populations, of neurons persisted.

The theory of single-neuronal task parameter representation was soon extended to population level representations. While macaques perform manual tasks, movement parameters like movement direction or velocity were encoded within the response of the neural population of hand-related areas of their brain. Such task parameters could be decoded more successfully from the simultaneous activity of the entire neural population than from single neurons (Georgopoulos et al., 1982; Schwartz et al., 1988). Neurons can contribute to multiple representations simultaneously. These findings suggested that the representation is present on the population level, rather than on the level of single neurons. Other more recent publications have found that task parameter representations were distributed randomly across the sampled neural populations (Machens et al., 2010; Raposo

et al., 2014), implying that at least a very large part of the neurons in the population contributes to the representation.

However, the representational perspective has its own limitations. When examining the neural population response of macaques to tasks, like the before mentioned center-out reach task, both single neurons and the population as a whole were found to display rich temporal dynamics that could not be interpreted as representations of a limited number of task parameters (Fetz, 1994; Churchland and Shenoy, 2007; Shenoy et al., 2013). Finally, the representational perspective does not try to explain behavior is generated. A new perspective on neural population activity was necessary, one that could explain both the complex dynamics of the neural population response and the strategy by which these dynamics drive behavior. This approach came to be known as the dynamical systems perspective (Shenoy et al., 2013; Yuste, 2015).

1.3   The dynamical systems perspective

The dynamical systems perspective is an interpretational approach to neural population activity that is based on the mathematical theory of the same name (Shenoy et al., 2013). The focus of the perspective lies on understanding how output is generated by the neural population, rather than discovery of representations of external variables (Shenoy et al., 2013; Yuste, 2015). Under the dynamical systems perspective, the evolution of the neural population activity can be understood as a function of the current activity, the inputs to the system and noise, in accordance to a set of dynamical rules. The dynamical rules also describe how population activity is transformed into the output of the dynamical system (Fig. *2*).
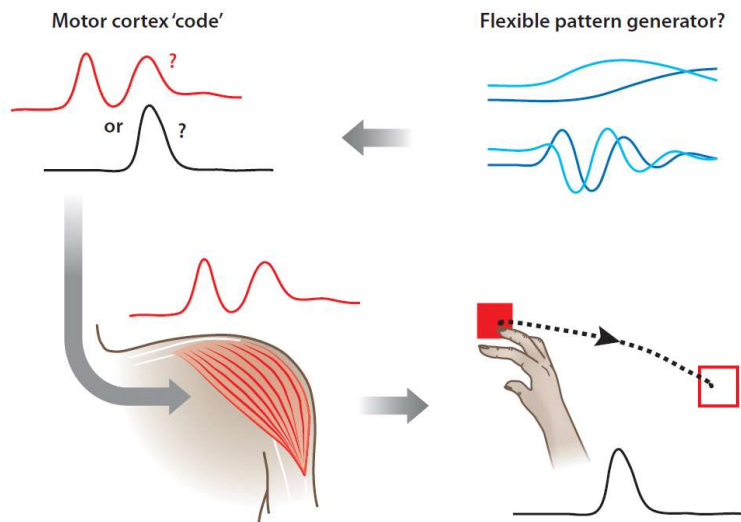
Figure 2: Schematic illustrating the relationship between the dynamical perspective and the representational view. The representational view focuses on discovering representations encoded in the population, for example muscle activity (red trajectory) or hand velocity (black trajectory) during a reaching task in the motor cortex (top left). The dynamical systems perspective is instead focused on how the final output is generated from the dynamics of the cortex (blue trajectories). Adapted from (Shenoy et Al. 2013).

It is useful to adopt a few terms and concepts from dynamical systems theory. Each neuron corresponds to a single dimension in the high-dimensional state space of population activity. The momentaneous population activity corresponds to a point, or state, in space of all possible states, and the set of all states over a given timespan traces a continuous trajectory through the same space. For a given set of inputs and initial states, the dynamical rules uniquely describe the shape of the resulting trajectories.

The dynamical systems perspective has developed simultaneously with the representational view and the two are not mutually exclusive. Representations of task variables can be interpreted as low-dimensional samplings of more complex, high-dimensional dynamics. However, as described before, the dynamical systems perspective can offer approaches to neural properties that can't be interpreted under the representational view. One of the earliest phenomena that could not be described under the representational perspective were central pattern generators (CPGs), described by Graham Brown in 1914. Brown observed rhythmic activity patterns in the spinal cord that drove locomotive behavior (Brown, 1914, 1915) in the absence of external inputs and proposed CPGs as the mechanism. Later studies examined the role of CPGs in body processes like digestion and respiration (Selverston, 1999) and sought to uncover the mechanics behind CPGs (Kopell and Ermentrout, 2002; Grillner, 2006). Spontaneous, rhythmical activity like CGPs does not represent any external variables and can be best interpreted as dynamical oscillators. In addition to rhythmic patterns, many studies have pointed out that the neuronal response is frequently more complex than would be necessary for representations. Response properties

of neurons in the motor cortex of macaques to movement parameters were not consistent across different movement conditions (Fetz and Cheney, 1980; Fetz, 1994; Churchland et al., 2006; Churchland and Shenoy, 2007; Shenoy et al., 2013) or even different epochs of the same task (Elsayed et al., 2016), implying that representations themselves change in response to external variables. Since the discovery of these phenomena, the mechanisms of their generation and the role they play in the functions of the brain have been the subject of many studies progressing the dynamical systems perspective (Fetz, 1994; Scott, 2004; Yuste et al., 2005; Cisek, 2006; Grillner, 2006; Churchland et al., 2012; Kaufman et al., 2014a).

The dynamical system perspective offers a framework to explain and conceptualize the complexity and heterogeneity of neural population activity. However, the neural spiking process is inherently noisy and the state space of population activity is extremely high-dimensional, and current recording methods capture only a fraction of this space. Therefore, the observation of the dynamics of the activity and the extraction the of dynamic rules that govern their evolution is complicated. Fortunately, there are reasons to believe that the dynamics of neural population activity do not occupy the entire state space, as will be shown in the next chapter.

1.4   The low-dimensional manifold of population activity

The true underlying dimensionality of the neural population response is the topic of many publications and is still controversially debated (Ganguli and Sompolinsky, 2012; Gallego et al., 2017; Stringer et al., 2018, 2019). It is assumed that the underlying connectivity of the neural network and the dynamics of population activity restrict the space of observed activity to a fraction of the theoretically possible states (Tsodyks, 1999; Sadtler et al., 2014; Okun et al., 2015). Results from many recent studies suggest that the dynamics of the neural population response occupy only a low-dimensional slice of the full state space, (Churchland and Shenoy, 2007; Churchland et al., 2010b; Mante et al., 2013; Kaufman et al., 2014a; Sadtler et al., 2014). Other authors counter-propose that the population response might be high-dimensional and population activity follows a power-law distribution across dimensions (Stringer et al., 2018, 2019). But, in any case, not every possible state of neural population activity is observed (Yuste, 2015). It is therefore helpful to define the 'manifold' of the state space of neural population activity as the volume that encompasses all states

that are hypothetically observable (Yuste, 2015; Gallego et al., 2017). To borrow terminology from mathematics, the manifold is said to be embedded, or realized, in the state space of the population response. Of course, we can record only a comparatively low-dimensional sampling of the entire high-dimensional population response. Even the before mentioned simultaneous recordings of the activity of over 10.000 neurons encompass only a tiny fraction of the total number of neurons of any given brain area. Furthermore, most recordings capture the response of the neuronal population to a trained, highly stereotyped task that does not necessarily engage the entire manifold (Gao et al., 2017).
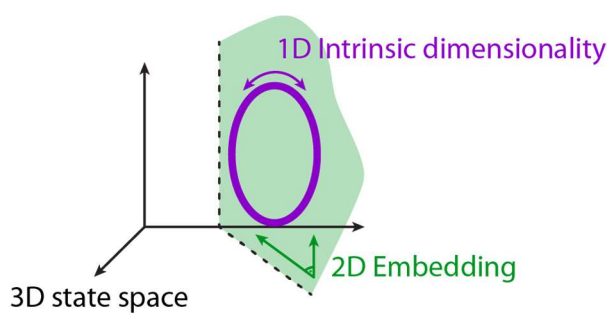
Figure 3: Illustration of the embedding space and the extrinsic and intrinsic dimensionalities on the example of a ring-shaped manifold embedded in 3d space. The violet ring is a closed curve with an intrinsic dimensionality of 1; its extrinsic dimensionality is 2, because it occupies a two-dimensional plane (green). The manifold is embedded in 3D space (black).

However, there are reasons to believe that this sampling is sufficient to draw conclusions about the dynamics of neural population activity as a whole. The dynamics of a neural population must be robust against disturbances and variability at the level of single neurons (Shenoy et al., 2013; Montijn et al., 2016). This necessitates a high degree of redundancy in the population, which in turn suggests that it is not necessary to capture the activity of every single neuron to obtain a complete portrait of the dynamics of the entire neural population. Additionally, as mentioned before, during sensory discrimination tasks, representations of task variables in the neural populations of macaques and rats were distributed randomly across individual neuron (Machens et al., 2010; Raposo et al., 2014). This can be, conversely, interpreted as evidence that single neurons are random samplings of the true structure of the state space of neural population activity. Theoretical studies of the geometry and dynamics of low-dimensional manifolds embedded in high-dimensional spaces show, that the shape and dynamics of such manifolds can be recovered from a small number of noisy, random samplings of the full space (Dasgupta and Gupta, 2003; Ganguli and Sompolinsky, 2012).

In general, the manifold is assumed to not occupy the entire dimensionality of the neural population space. The space occupied by the manifold is called the embedding space, and

its dimensionality is the extrinsic dimensionality of the manifold. Put more simply, the extrinsic dimensionality is the number of linear, orthogonal dimensions needed to capture the manifold. In contrast, the intrinsic dimensionality describes the dimensionality of the manifold itself, or the number of dimensions the manifold has within it, independent of how it is embedded (Fig. 3). It is the number of independent variables, or coordinates, needed to define a basis on the manifold[1]. In the simple case of a linear manifold, that is, a manifold that has no 'bends' anywhere, the intrinsic and extrinsic dimensionalities are equal.

Closely related to the dimensionality of the manifold is the concept of latent variables (Cunningham and Yu, 2014), also called latent factors (Chethan Pandarinath, K Cora Ames, Abigail A Russo, Ali Farshchian, Lee E Miller, Eva L Dyer, 2018). A manifold of lower dimension than the embedding space suggests that the individual observed variables, in our case single neuron activities, are not independent. Sets of correlated observed variables can be modeled as a function of a smaller set of independent, hidden latent variables and noise. In such a model, the true dimensionality of population activity is equal to the dimensionality of the set of latent variables and the latent variables form a basis for the manifold.

Per definition, latent variables are required to be independent of each other, but not necessarily to be linear combinations of observed variables. There are in general two different forms of latent variables (Fig. *4*): In the case of nonlinear latent variables, the set of latent variables is the smallest set necessary to uniquely define every point on the manifold. Since this is also the definition of the intrinsic coordinates of the manifold, the set of latent variables forms a basis on the manifold and their number is equal to the intrinsic dimensionality of the manifold (Gallego et al., 2017). In the case of linear latent variables, the set of latent variables form a basis for the embedding of the manifold and the number of latent variables is equal to the extrinsic dimensionality of the manifold. In both cases, the set of latent variables is not unique; Any set of orthogonal variables that spans the same space is in itself again a set of latent variables[2]. This has an important implication: the

---

[1] You can imagine the nonlinear manifold as a part of a linear space that has been bent and deformed. The nonlinear latent variables form a basis on the original, non-deformed manifold.

[2] For non-linear latent variables, orthogonality is defined locally; any set of non-linear variables that spans the same non-linear manifold and is approximately orthogonal in the area around each point on the manifold is another set of latent variables on the manifold.

manifold itself has no canonical or ground-truth set of latent variables. If we want to distinguish specific dimensions, we need additional information about the dataset.
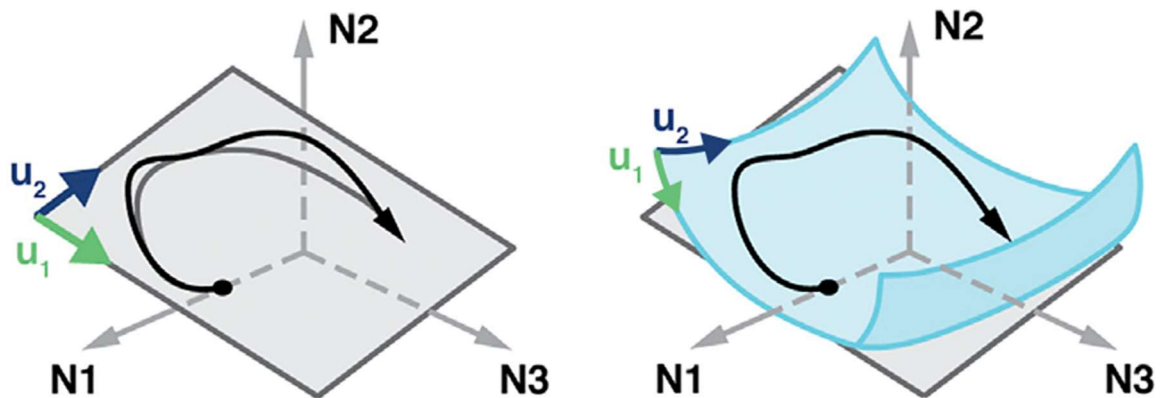
Figure 4: Examples of a linear and a nonlinear manifold latent variables. The black curve is an example trajectory in the state space of three neurons N1, N2 and N3. Left: the trajectory is mostly confined to the two-dimensional manifold (grey plane) spanned by the linear latent variables u1 and u2. Right: The trajectory fits the non-linear manifold (light blue) better, but the latent variables form no linear subspace anywhere. Adapted from (Gallego et Al. 2017).

A useful concept for distinguishing individual dimensions of the manifold are neural subspaces, also known as coding dimensions. Neural subspaces are dimensions, or groups of dimensions, of the neural manifold that can be associated with distinctly encoded aspects or features in the manifold, like cognitive and behavioral factors. For example, across several different wrist and grasping tasks, the population response of multiple macaques occupied two distinct subspaces, one capturing most of the activity related to temporal features of the task and the other containing a mapping of neural activity onto muscle activity (Gallego et al., 2018b). In another case, the neural population response of two macaques to a delayed center-out reaching task occupied two orthogonal subspaces, one containing activity during movement preparation and one during movement execution (Elsayed et al., 2016). The subspaces were linear, orthogonal and temporally separated, but with movement goal positions arranged in space in the same way to each other, suggesting that the subspaces were linked. Several studies of macaque reaching and grasping movements have shown the existence of a "movement trigger" dimension, that was reliably predictive of movement onset (Kaufman et al., 2014b; Sussillo et al., 2015; Michaels et al., 2018; Ames and Churchland, 2019).

Another concept that is important for the description of neural population activity under the dynamical systems perspective is the nullspace (Kraskov et al., 2009; Kaufman et al., 2014b); In the context of the output of a system, the nullspace describes activity that is reflected in the activity of most neurons, but cancels itself out at the level of the output. In terms of subspaces, the nullspace can be interpreted as a subspace of the manifold that is orthogonal to and independent of the output-related subspace. For example, activity in the motor cortex of a macaque during a reaching task was decomposed into two subspaces: one correlated to muscle activity during reaching and one independent. Muscle activity is considered the output of the motor cortex and therefore the independent activity that occupies the null space of muscle control (Kaufman et al., 2014b). The independent activity was assumed to be involved in the preparation and planning of movement execution.

Under the assumption that the manifold is indeed low-dimensional, the obvious question is: How can we extract the slice of the manifold that is contained in a recorded sample of neural population activity? How can we separate the manifold into meaningful subspaces? The search for the answer to this question has led scientists to try out a wide range of dimensionality reduction methods.

1.5   Dimensionality reduction methods

Dimensionality reduction is a basic problem of high-dimensional data analysis. The goal of dimensionality reduction methods is to reduce a high-dimensional dataset to a lower-dimensional representation that preserves some feature of interest, while discarding the remaining dimensions (Fig. *5*). In the context of neural population activity, the feature that we ultimately seek to preserve is the manifold that was introduced in the previous chapter. To this end, most (but not all) of the approaches presented in this chapter aim to reduce the dimensionality of the recorded neural population response to the extrinsic dimensions of the manifold discarding the remaining dimensions.

1.5.1   Singular-Value decomposition based methods

Most approaches focus on a particular feature of the manifold and seek to extract dimensions that exhibit that feature. The simplest feature of the manifold is the variance along single dimensions in the space of neural population activity. The assumption made here is that the dimensions occupied by the manifold capture the greatest variance in the

data. Therefore, reducing the dataset to the dimensions of greatest variance will separate the manifold from the remainder.

The most often-used approach for dimensionality reduction is principal component analysis (PCA) (Jolliffe, 2002; Cunningham and Yu, 2014). Because PCA can be calculated by applying singular value decomposition to the data matrix, PCA, and related methods, are often grouped under the term of 'singular value decomposition-based methods'. PCA is applied to high-dimensional datasets that can be expressed in the form of a $m \times d$ matrix where m is the number of variables and $d > m$ is the number of simultaneous, independent observations of all m variables. In the context of recordings of neural population response, variables are usually the recorded activities of single neurons or multi-units and observations are single time points. Note that although in the context of neural population recordings observations are usually sequential time points, and PCA can be used to analyze time series, this is not a necessary condition for PCA.

Principal component analysis centers and rotates the dataset onto an ordered set of orthogonal basis vectors, called principal components. Each principal component is a linear combination of the initial variables and all principal components are pairwise orthogonal and each successive principal component captures the maximal variance among the remaining dimensions. Because PCA consists only of a centering and a rotation of the data, the covariance structure of the data is preserved. Principal component analysis is used as a dimensionality reduction method by retaining a number of the first principal components and discarding the rest (Jolliffe, 2002; Cunningham and Yu, 2014). A significant number of studies uses PCA either as the dimensionality reduction method of choice, or as a first step, preparing the data for further, more specialized analyses (Gao et al., 2017; Ni et al., 2018; Stringer et al., 2018, 2019; Ebitz and Hayden, 2021).

SVD-based dimensionality reduction methods have a number of known issues that are relevant for their application to recordings of neural population activity. These issues have been first described in the context of PCA; Wood et Al. (Wood and McCarthy, 1984) and Beauducel et Al. (Beauducel and Debener, 2003) have applied PCA to models of event-related potentials (ERP) and found that individual principal components resembled individual components of the modeled ERP signal, but also contained 'misallocated' variance

from other components. In studies where PCA was used to analyze time series (Möcks and Joachim, 1986; Lakshmanan et al., 2015), it was sensitive to latency in the analyzed dataset; In the presence of stable delays between individual variables, PCA captured additional, spurious principal components. The activity profiles of the spurious components resemble derivatives of the activity profiles of the 'main' components. Additionally, the low-dimensional space identified by PCA captures both function-related neural variability and the random spiking and firing rate variability of single neurons (Cunningham and Yu, 2014). Finally, SVD-based methods do not offer a natural method for the estimation the extrinsic dimensionality of the manifold; The most common method of dimensionality estimation is to choose an arbitrary desired percentage of total variance that is to be retained and discard principal components beyond that percentage.
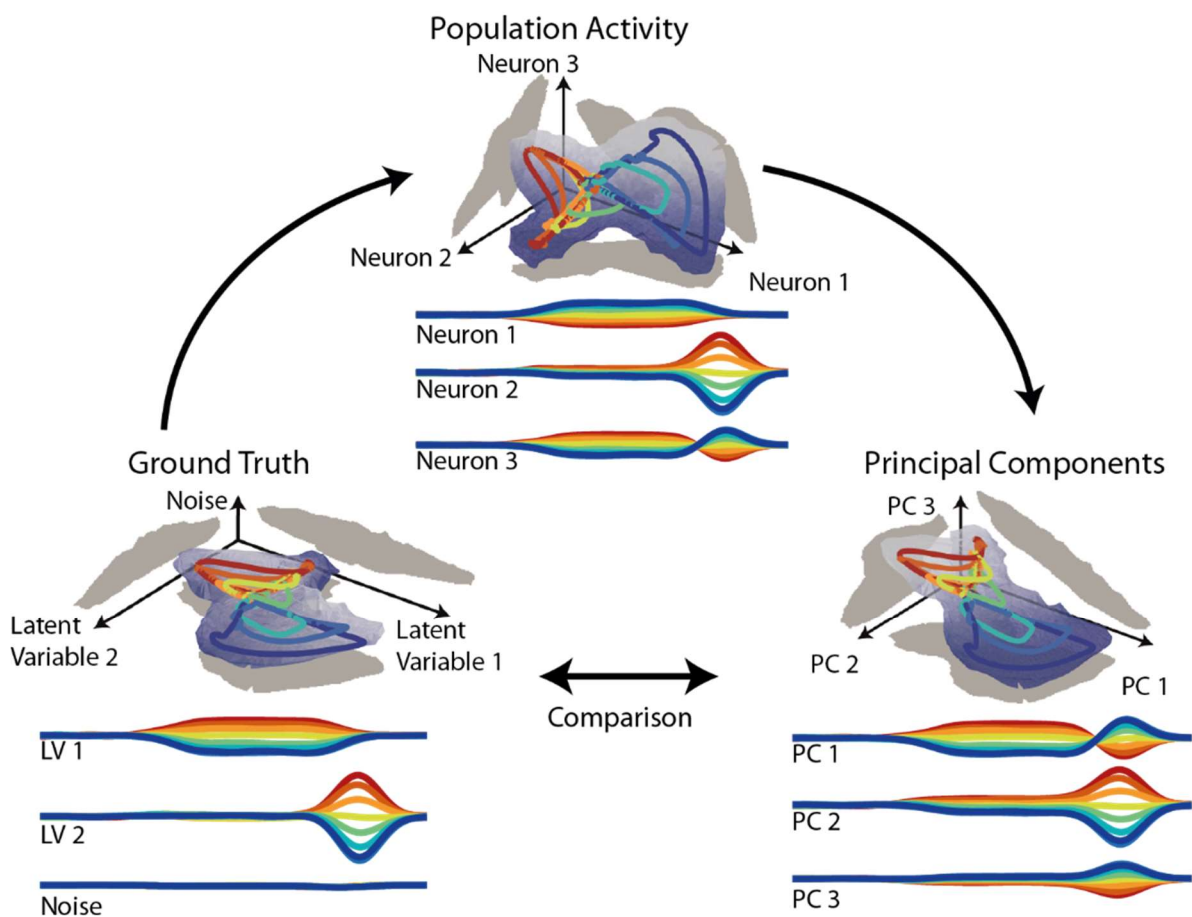
Figure 5: Illustration of PCA applied to a two-dimensional linear manifold embedded in population response space. Bottom left: the population response is structured by the trajectories of two latent variables and added noise. The variables are orthogonal in state space and jointly assume one of eight trajectories (color-coded). Top: the manifold is rotated randomly in the space of neurons and each neuron is a linear combination of latent variables and noise. Bottom right: PCA rotates the manifold onto components of maximal variance, but those do not necessarily coincide with the original latent variables.

Because of the mathematical simplicity of PCA, the method can easily be adapted to different requirements by imposing different constraints on principal components than to maximize the captured variance. One such adaptation is jPCA (Churchland et al., 2012; Shenoy et al., 2013). Individual components recovered using jPCA maximize rotational tendency rather than variance captured, and the low-dimensional representation recovered using jPCA highlights rotational dynamics of the dataset. Another method closely related to PCA is factor analysis (FA) (Churchland et al., 2010a; Cunningham and Yu, 2014). Factor analysis can be considered as PCA with an explicit model of the independent variance of each individual neuron. Discarding this individual variance allows FA to obtain a low-dimensional representation of the data that contains only variance that is shared across neurons. Because of this discarding step, and unlike the PCA variants presented so far, the underlying structure of the data is not preserved under FA. The method Gaussian process factor analysis (GPFA) (Yu et al., 2009) combines factor analysis with the additional constraint that the evolution of latent variables over time is to be a Gaussian process, smooth and continuous. This approach is based on the idea that processes in neural networks should be continuous, and latent variables are expected to change slowly and smoothly.

1.5.2 Model-based dimensionality reduction methods

A different approach to dimensionality reduction is chosen for model-based dimensionality reduction methods. In these methods, the activity of latent variables is explicitly modeled and a static function maps the activity of latent variables onto observed variables. A separate function models the independent noise of observed variables. The underlying model of latent variable activity, the mapping function and the noise function are estimated from observed variable activity. Unlike PCA, model-based methods are time-series based; The observations need to be sequential and continuous. Also, the number of latent variables for the model needs to be specified; Often, multiple models are constructed for different numbers of variables and the best-fitting model is selected.

The latent variables are not explicitly assumed to be embedded in the space of observed variables. However, in the case of a linear relationship between latent and observed variables, the linear weights between observed and latent variables can be interpreted as

projection vectors in the space of observed variables. These projection vectors span a subspace that is analogous to the low-dimensional representation that is the result of SVD-based methods. Thus, linear modelling of latent variables can be used as a dimensionality reduction method in the same manner as SVD-based methods.

One of the simplest models to be used in model-based dimensionality reduction methods is the hidden Markov model (HMM) (Seidemann et al., 1996; Jones et al., 2007; Kemere et al., 2008; Bollimunta et al., 2012; Ponce-Alvarez et al., 2012). In this model, latent variables form a vector that can assume discrete values, or states, and the current state depends only on the previous state and a transition probability. HMMs are comparatively easy to calculate, but the time-dependent trajectories of latent variables are not continuous, and thus HMMs can't model smooth dynamics of observed variables well.

A more complex family of models are the latent dynamical system models. In so-called linear latent dynamical system models (LDS) (Smith and Brown, 2003; Kulkarni and Paninski, 2007; Paninski et al., 2010; Buesing et al., 2012; Pfau et al., 2013), the latent variables are modeled as states of a dynamical system that can take continuous values. The dynamical rules that govern the evolution of the latent variables are linear transformations that can be represented by a static transformation matrix. The non-linear latent dynamical system models (NLDS) (Byron et al., 2006; Petreska et al., 2011) allow for non-linear latent variable state transition functions. Such non-linear functions can fit complex data better at the cost of increased computational complexity and the risk of overfitting.

1.5.3   Supervised dimensionality reduction methods

Both the SVD-based and the model-based methods presented so far extract the low-dimensional representation in an unsupervised fashion, utilizing only the matrix of observed values. The latent variables are estimated or discovered as part of the method. In contrast, supervised methods leverage external information to specify latent variables. In the context of neural population response to stereotypical tasks, external variables like task conditions, sensory cues or behavior are expected to have representations within the population response. These representations are used as latent variables (Churchland and Cunningham, 2014). Known non-representational features of population activity such as the condition independent temporal structure of the response can be also used as latent variables.

The simplest supervised methods of dimensionality reduction are linear discriminant analysis (LDA) and support vector machines (SVM). If the observations can be assigned to discrete classes, for example the neural population response to reaches towards the left vs. reaches towards the right, LDA and SVM find dimensions that maximize the separation between the classes, with LDA targeting dimensions which minimize the variance within each class while simultaneously maximizing the distance between classes and SVM selecting dimensions along which the margin of separation between the classes is maximal. For both methods, the number of dimensions discovered is at most equal to the number of classes minus one.

In cases where the external variables take continuous, rather than discrete, values, generalized linear models (GLM) (Mante et al., 2013; Cunningham and Yu, 2014; Raposo et al., 2014) are applicable instead. GLM are a form of multivariate linear regression, which estimates linear fits of all observed variables to each latent variable. Just as for the model-based dimensionality reduction methods introduced earlier, the weigths of the linear fits can be expressed as vectors that each span a dimension in state space. Along these dimensions, the observed variables and the latent variables are maximally correlated, and together the dimensions span the low-dimensional representation that we are interested in.

The three supervised dimensionality reduction methods presented so far share a weakness that, incidentally, mirrors the main weakness of the representational perspective: Not all neural population activity reflects some external variables (Shenoy et al., 2013; Kaufman et al., 2016). In fact, if we wish to understand the dynamics of the neural population response, we may be equally as interested in dimensions that contribute only to the dynamics without encoding any external variables. Demixed dimensionality methods, like the aptly-named demixed principal component analysis (dPCA) (Brendel et al., 2011; Kobak et al., 2016) combine supervised and unsupervised dimensionality reduction methods. dPCA first extracts representations of individual task parameters and interactions between task parameters from the population response in a supervised manner, then uses unsupervised SVD-based dimensionality reduction to discover dimensions of maximal variability for individual representations. However, dPCA fails to identify orthogonal subspaces that are occupied at different timespans over the course of the task.

Introduction

Another method that does leverage temporal separation was presented by Elsayed et al. In their 2016 paper, the authors used the temporal structure of the neural population response to a center-out reach task to identify two orthogonal subspaces for reach preparation and execution (Elsayed et al., 2016). Elsayed et al. exploited the fact that the population response occupied the two subspaces during two successive epochs of the task, by constraining PCA to maximize variance for each epoch independently, but simultaneously. However, this approach requires prior knowledge about the structure of the neural population response and is less applicable in the case of overlapping subspaces.

1.5.4 RNNs and non-linear dimensionality reduction methods

As a contrast to linear dimensionality reduction methods, we should also briefly touch upon nonlinear methods. Nonlinear dimensionality reduction methods typically transform the nonlinear embedding of a manifold into a linear one, onto which linear dimensionality reduction methods can be applied. The two most common methods are isometric mappings (ISOMAP) (Tenenbaum et al., 2000)and locally linear embeddings (Roweis and Saul, 2000; Stopfer et al., 2003; Brown et al., 2005; Carrillo-Reid et al., 2008). Both methods analyze small neighborhoods of each observed state to estimate the intrinsic dimensionality of the nonlinear manifold, then take the manifold apart and reassemble the neighborhoods into a linear manifold, while preserving nearest-neighbor relations. However, such methods are susceptible to the presence of noise (Boots and Gordon, 2012) and require a dense sampling of the manifold (Cunningham and Yu, 2014) in order to preserve the structure of the manifold.

In recent years, an increasing number of studies have also utilized artificial neural networks to investigate neural population response. Artificial neural networks like recurrent neural networks (RNNs) are inherently nonlinear and can be flexibly trained to perform a vast number of computations. Therefore, they are well-suited for the purpose of nonlinear dimensionality reduction. A recent dimensionality reduction method based on RNNs is latent factor analysis via dynamical systems (LFADS) (Sussillo et al., 2016; Chethan Pandarinath, K Cora Ames, Abigail A Russo, Ali Farshchian, Lee E Miller, Eva L Dyer, 2018; Pandarinath et al., 2018). LFADS is a so-called sequential auto-encoder: Two RNNs in sequence are simultaneously trained to first extract a low-dimensional latent factor

representation from a high-dimensional dataset, then reconstruct the original input from the latent factors as precisely as possible. The simultaneous training ensures that the low-dimensional representation captures as much information about the structure of the high-dimensional dataset as possible. LFADS can be applied to single-trial population responses, and because noise is not encoded in the low-dimensional representation, LFADS can be used to de-noise single-trial firing rates (Pandarinath et al., 2018). Note however, that unlike the representations obtained from the other methods presented so far, the low-dimensional representation generated by LFADS is not embedded in the state space of neural population response. Instead, if the set of latent factors discovered by LFADS is constrained to be as small as possible, the latent factors are analogous to the nonlinear latent variables that span the intrinsic dimensions of the manifold; both sets of variables form the minimum set of coordinates necessary to uniquely identify a point on the manifold. LFADS allows us to discover information about the dynamics of neural population activity that could not be obtained from linear methods.

Another frequently used application of RNNs is to construct a model of a recorded neural population and train the model to reproduce the recorded population response from recorded stimuli. The noise-free and fully sampled simulated population response of the RNNs can then be analyzed in place of the noisy and strongly sub-sampled recorded response. RNNs were used to model decision-making in a visual discrimination task (Mante et al., 2013), muscle activity in a sequential reaching task (Sussillo et al., 2015) and muscle kinematics in a delayed grasping task (Michaels et al., 2020). In each case, the RNNs could reproduce key dynamics of the recorded population response. Even though such RNNs are significantly smaller and less complex than biological neural networks, they still showcase the potential of models of neural population response to provide new insights into the dynamics of neural population activity.

Taken together, there are a variety of methods for dimensionality reduction, and each method has advantages and disadvantages. However, the performance of dimensionality reduction methods on neural data has not been systematically investigated so far. To evaluate the performance and success of any dimensionality reduction methods when applied to recordings of neural population activity, we would need to compare the low-dimensional representation obtained from the results of the method to the ground-truth

manifold of the recorded population. Unfortunately, we can't know the ground truth of the manifold; The presence of temporal delays between individual neurons, the inherent noisiness of the point process of spiking activity and the low signal-to-noise ratios of individual neural signals make it difficult to extract the dynamics of neural population response. However, we can use models of neural population response to draw conclusions about the modelled population, as long as the models are sufficiently accurate to reproduce key features of population activity.

1.6 The fronto-parietal grasping network

In the context of this thesis, we construct a model based on real recordings of the neural population response of two macaques to a visually guided reach-to-grasp task. The large-scale neural circuit in the macaque brain that the recordings were captured from is known as the fronto-parietal grasping network. This network, comprises the areas of F5 (the hand area of the ventral premotor cortex) and the area AIP (the anterior intraparietal area in the parietal lobe), shown in Fig 6. These areas are known to be strongly reciprocally connected (Luppino et al., 1999).
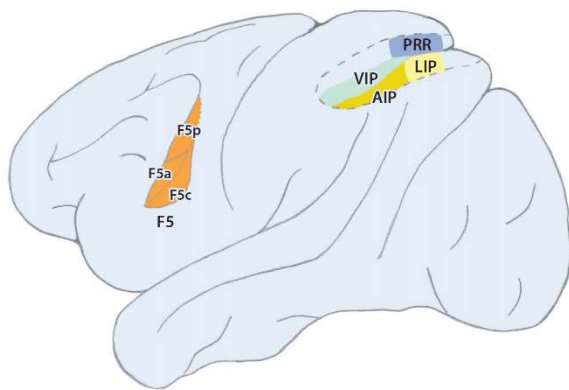
Figure 6: The locations of brain areas central to the fronto-parietal grasping network in the macaque brain. Left: Hand area of the ventral premotor cortex F5. Right: Anterior intraparietal area (AIP), together with the neighboring areas of Lateral and Ventral Intraparietal (LIP, VIP) and the Parietal Reach Region (PRR).

The involvement of areas of the fronto-parietal grasping network into grasping execution was shown in lesion studies. Temporary inactivation of the area of either F5 or AIP caused functional deficits in the pre-shaping of the hand during grasping movements (Gallese et al., 1994; Fogassi et al., 2001), showing that both areas are causally involved in the execution of grasping movements (Fluet et al., 2010).

Areas of the fronto-parietal network are also involved in other functions. The parietal reach region (PRR, shown in Fig. 6) and the dorsal premotor area (PMd) were shown to be involved in reaching motions (Gail and Andersen, 2006; Churchland et al., 2010b, 2012). Neurons in the areas of F5 and AIP were modulated in response to visual object fixation and

discrimination (Murata et al., 1997, 2000; Janssen and Scherberger, 2015). The neighboring lateral intraparietal area (LIP, see Fig. 6) and frontal eyefield (FEF) are similarly involved in saccadic eye movements (Freedman and Assad, 2006; Siegel et al., 2015). Finally, in the areas of AIP and F5 the neural population activity persisted through delays between object presentation and movement execution even if the object was not visible during this delay. This could be interpreted as a memory-related or movement preparation-related activity (Murata et al., 1996; Baumann et al., 2009; Fluet et al., 2010).

The presence of visual, preparatory and movement-related activity in the fronto-parietal network suggests an involvement in functional visuo-motor transformations, transformations of visual input into movement control signals (Janssen and Scherberger, 2015). This theory is supported by findings that neural selectivity of neurons in AIP for visual information was stronger than in F5. Conversely, selectivity for motor signals was weaker in AIP and stronger in F5, though not as strong as in areas of the motor cortex (Schaffelhofer et al., 2015; Schaffelhofer and Scherberger, 2016). This implies that the transformation of visual data into movement control takes place down the stream from AIP towards F5 and later the motor cortex.

The fact that the fronto-parietal grasping network is directly involved in reaching and grasping movements and that the neurons within the network are encoding visual information, movement preparation and movement execution in parallel makes the network an excellent candidate for the recording of the neural population response to visually cued, delayed reach-to-grasp tasks. We can reasonably expect the three different encodings to be independent of each other, and therefore to occupy three different, orthogonal subspaces of the full space of neural population activity (Kaufman et al., 2014b; Elsayed et al., 2016; Dann, 2017).

1.7   Simulation-based evaluation of dimensionality reduction methods

In this thesis, we present a simulation of neural population activity as a tool for the evaluation of dimensionality reduction methods. We design the simulation as a flexible toolbox, capable of generating single-trial recordings of spiking neural population activity and reproducing key properties of recorded neural population responses. The simulation can be used to evaluate a wide range of dimensionality reduction methods. Most

dimensionality reduction methods extract a linear representation, therefore our simulated population response is given structure by a low-dimensional, linear manifold. We do not simulate individual interneuron connections, as RNNs do, instead, correlations between neurons are explicitly and solely driven by the activity of latent variables.

Our simulation is based on a previously published recording of neural population response (Michaels et al., 2015; Dann et al., 2016) of two macaques to a visually cued, delayed reach-and-grasp task. The neural population response to this task was recorded from the fronto-parietal grasping network, including part of the ventral premotor (F5) and anterior intraparietal (AIP) areas. A supervised dimensionality reduction method based on SVM was used to analyze the recording and identify three subspaces related to visual cue information, movement preparation and movement initiation execution. The identified subspace decomposition is in accordance with previous studies that analyzed the fronto-parietal grasping network (Murata et al., 1996; Baumann et al., 2009; Fluet et al., 2010). This subspace structure is used as the ground truth of the simulation. We also ensured that the simulation accurately reproduces key population features of the recorded neural population, as explained in detail in the next chapter.

Of all dimensionality reduction methods, PCA is both the simplest and the one that is found most frequently in literature. As the PCA is the basis for several other dimensionality reduction methods, we can expect many findings obtained from evaluations of PCA to be applicable to these methods. We use the simulation to evaluate the performance of PCA as a dimensionality reduction method when applied to recordings of neural population activity. We focus on three central topics:

The first topic is the estimation of the true dimensionalities, both the intrinsic and extrinsic, of the manifold from the results of dimensionality reduction methods. SVD-based dimensionality reduction methods presented in literature are often accompanied by methods to estimate the dimensionality of the manifold (Jolliffe, 2002; Cunningham and Yu, 2014). Very often the method entails selecting an arbitrary amount of variance to be retained in the low-dimensional representation and rejecting dimensions that exceed that number. It is difficult to evaluate whether the dimensionality was estimated correctly, because the true dimensionality of the manifold is unknown.

The second topic is the capture of the true manifold by the low-dimensional representation. If we can accurately estimate, or maybe even generously overestimate the dimensionality of the manifold, will all subspaces and dimensions, coding and non-coding, of the manifold be captured by the representation? If not, what information is lost?

The final topic is the interpretation of the components discovered by SVD-based dimensionality reduction methods. As mentioned before, in the case of unsupervised dimensionality reduction methods the choice of latent variables is to a degree arbitrary. However, SVD-based methods return unique sets of components, like principal components in PCA or factor loadings in FA and these components are known to contain mixtures of activity from different subspaces (Möcks and Joachim, 1986; Lakshmanan et al., 2015). The extent of this mixing, and whether it can be reduced, has not been investigated yet.

2 Methods

2.1 Experiment

The task described here was performed by Benjamin Dann and published before (Michaels et al., 2015; Dann et al., 2016; Dann, 2017). Two monkeys, S and Z, were trained to perform a delayed reach-and grasp task; The monkey was cued to grasp a handle with either a pinching precision grip or a full-hand power grip. The task was structured into four epochs: In the fixation epoch the monkey held its hand still on a hand rest, fixated a dot on a monitor for a variable period of time (800-1000ms). In the cue epoch (300ms) one of two disks were displayed on a monitor left or right of the fixation disk and instructed the monkey to plan and to perform one of two associated grip types. Next, in the planning epoch the cue was turned off and the monkey had to memorize and prepare the corresponding grip type for a variable amount of time (1300-1500ms). Finally, in the movement phase, the fixation dot vanished, cuing the monkey to execute the planned movement within a limited amount of time (max 750ms). Note that the task also included two other contexts. In the free-choice context, both cues were displayed and the monkey was free to choose between the two grip types. In the delayed instructed context, the monkey received a second cue during the memory period following a free-choice cue (in 50% of cases), instructing the monkey to perform one of the two grip types. The free-choice and delayed instructed contexts were not used for all following analyses and are therefore not described in further detail here. For the analyses of this thesis, six recordings of monkey S and three recordings from monkey Z were used. The monkeys were successful in 95% and 96% of trials on average, with 730±106 and 722±167 trials per recording day for S and Z, respectively. Only successful trials of the instructed context were used for all further analyses.

2.1.1 Data recording and preprocessing

All methods and procedures in this section were performed by Benjamin Dann (Dann, 2017). The neural population response to the task was recorded from the cortical areas F5 and AIP as part of the fronto-parietal grasping network. In each area two floating microelectrode arrays with 32 electrodes per array were implanted. For spike detection, the signal was first high-pass filtered (median filter, cutoff frequency: 333Hz) and then low-pass filtered (non-causal Butterworth filter, 4th order, cutoff frequency: 5000Hz). Next, the common noise

sources were eliminated using principal component artifact cancellation (Musial et al., 2002; Dann et al., 2016). Offline spike sorting was performed using a modified version of the offline spike sorter Wave_clus (Quiroga et al., 2004; Kraskov et al., 2009; Dann et al., 2016). Only well-isolated neurons that had an average firing rate above 1Hz across all contexts (75±7% of all single units per recording) were used for further analysis. This corresponded to 55±14 neurons per recording session (63±5 for monkey S, 37±4 for monkey z). Spike times were binned into time bins of 1ms width.

The recorded single-neuron, single trial activity was smoothed by convolution with a Gaussian kernel, with a σ of 60 ms. Next, a square root filter (Yu et al., 2009; Afshar et al., 2011; Michaels et al., 2015) was applied to normalize the distributions of firing rates under two aspects: First, to reduce the impact of neurons with high overall firing rates when compared to slow-firing neurons. Second, to account for non-linear changes in the firing rate of individual neurons (Yu et al., 2009; Buzsáki and Mizuseki, 2014). Neural activity was aligned to multiple behavioral events because of the variable length of the fixation and memory epoch and the variable reaction time of movement initialization. Neural activity was aligned to the cue onsets and movement onset (cue onset: -800 to 1500 ms; and movement onset: -300 to 500 ms) (Dann et al., 2016; Dann, 2017). Alignments were then recombined to produce continuous, fixed-length trials.

2.1.2 Manifold identification

The following dimensionality reduction method was developed and executed by Benjamin Dann (Dann, 2017). The method uses support vector machines to identify orthogonal subspaces in the population response to trials (Raposo et al., 2014; Kaufman et al., 2015). Single-trial data of the instructed context was subsampled by selecting time bins spaced in 10 ms intervals. Due to the smoothing with a Gaussian kernel described before, each bin effectively corresponds to a Gaussian window with a σ of 60 ms. For every time bin, two support vector machines (SVMs) were trained independently; A condition-dependent SVM was trained to optimally separate the single-trial activity within the time bin that corresponded to different grasp conditions. Additionally, a condition-independent SVM was trained to separate the single-trial activity in each time bin from the baseline activity during pre-trial fixation.

Methods

In the next step, the SVMs were used to extract a low-dimensional representation from the high-dimensional population response in the following. A maximal random set of close to orthogonal SV-projections was selected and the low-dimensional representation was constructed as the space occupied by these SV-projections. As mentioned in the introduction chapter, low-dimensional representations are optimized to preserve a certain feature of the full space while reducing dimensionality. For continuous activity, such as trial-averaged activity, variance is a useful feature to preserve. Not so for single trials spiking activity, which can be considered as a point process. For this reason, the independent variance of individual neurons is high, and would cause the low-dimensional representation to rotate onto neurons with high firing rates. Therefore, the correlation between the single-trial activity captured by all SV-projections and the selected set of orthogonal SV-projections was optimized instead. First, the selection of close to orthogonal SV-projections was optimized up to a correlation of $R^2 = 0.95$ with the full space of SV-projections. Second, if the correlation threshold was not reached, an additional SV-projection was added and the first step repeated. Note that only one additional SV-projection had to be added in all cases. The selected set of SV-projections was QR-orthogonalized to obtain a true basis for the low-dimensional representation (Dann, 2017). The SV-projections could be grouped into three subspaces for visual, movement preparatory and movement execution related processing, shown in Fig. 8b and c. (visual subspace: one condition dependent and one condition independent projection; preparatory subspace: one condition dependent and one condition independent projection; movement execution subspace: one condition dependent and two condition independent projection) For the instructed context, occupancy of the three subspaces was separated over the time course of the task with brief transition periods. Individual subspaces are referred to as $U_i, i = 1, 2, 3$. The subspaces are spanned by the normal vectors of the hyperplanes of maximum separation $\vec{\beta}_{SVM,j}$ (Dann, 2017).

2.2   Simulation of a spiking neural population: Methods

2.2.1   Low dimensional structure of population firing rates

We want to simulate spiking neural population activity that is constrained to a low-dimensional manifold in the high-dimensional space of all neurons. This simulation will later be used to evaluate the performance of PCA when applied to neural data. We simulate the simultaneous activity of a population of I spiking neurons in response to the behavioral task

described in the last section. Each neuron is defined by its time-dependent firing rate $n_i(t)$. We use the firing rate to generate M individual trials m in T time steps t of 1 ms each of spiking activity $\hat{n}_{i,m}(t)$. We do not simulate individual inter-neural connections like an artificial neural network would; rather, the activity of the high-dimensional set of observed neurons reflects the trajectories of a smaller set of orthogonal, linear latent variables (LVs) $l_j(c,t)$, where c is the task condition.

The latent variables completely describe the dependence of the neural population response on time t and task conditions c and are therefore the only variable that explicitly depends on the two variables. We model latent variables as products of a time-dependent profile $l_j(t)$ and a condition-dependent scale factor $\gamma_j(c)$ as follows:

$$l_j(c,t) = \gamma_j(c) \cdot l_j(t) \tag{1}$$

In short, we use the temporal profiles of the projections of full activity onto individual subspaces, as described in the last section, and scale them differently for different conditions. E.g, for the recorded task $\gamma_j("power")$ was equal to 1 and $\gamma_j("precision")$ was equal to -1. The profiles $l_j(t)$ used in this thesis are shown in fig. 10d in the results section 3.1.3. The condition scaling factors $\gamma_j(c)$ are shown in fig. 17 in section 3.1. To improve readability, we leave out the condition c from equations that do not directly concern themselves with different conditions.

Each neuron has a fixed baseline firing rate $n_{i,0}$ which reflects the probability of the neuron to spike at any given moment. The baseline firing rate $n_{i,0}$ of each neuron is multiplicatively modulated by a linear sum of the weighted activity of each LV l_j at time t to obtain the time-dependent firing rate $n_i(t)$. The modulation strength is given by the constant linear coefficients $\beta_{i,j}$. Each neuron $n_i(t)$ has an individual temporal delay $\Delta t_{i,j}$ to each latent variable $l_j(t)$;

$$n_i(t) = n_{i,0} \cdot \left(1 + \sum_j \beta_{i,j} \cdot l_j\left(t + \Delta t_{i,j}\right) \right) \tag{2}$$

Methods

Out of the firing rates, we construct the population firing rate vector $\vec{N}(t)$, which has the individual $n_i(t)$ as its entry in position i. This is represented by the union operator: $\vec{N}(t) = \bigcup_i n_i(t)$. Because the $\Delta t_{i,j}$ are individual to each neuron and LV, the one-dimension latent variable trajectory $l_j(c,t)$ also needs to be expressed as a vector:

$$\vec{l}_j(c,t) = \bigcup_i l_j(t + \Delta t_{i,j}) \tag{3}$$

The entries of $\vec{l}_j(t)$ are time-shifted copies of $l_j(t)$. $\odot$ represents component-wise multiplication, also known as the Hadamard product.

$$\bigcup n_i(t) = \vec{N}(t) = \vec{N}_0 \odot \left(1 + \sum_j \vec{\beta}_j \odot \vec{l}_j(t)\right) \tag{4}$$

This is the same as equation 2, only expressed in terms of the population activity $\vec{N}(t)$.

2.2.2 Simulation of spiking neural activity

We transform the firing rate $n_i(t)$ of each individual neuron into M single trials m of spiking activity, using a weighted inhomogeneous Poisson process. As spiking activity can't be negative, only the non-negative firing rate $n^+_i(t)$ is used:

$$n^+_i(t) = \begin{cases} n_i(t), n_i(t) \geq 0 \\ 0, n_i(t) < 0 \end{cases} \tag{5}$$

For each neuron i, each trial m is divided into T time bins t_n of 1 ms length and the spiking activity is denoted $\hat{n}_{i,m}(t_n)$. Each bin n of trial m can contain either 0 or 1 spike, denoted as $\hat{n}_{i,m}(t_n) = 0$ and $\hat{n}_{i,m}(t_n) = 1$ respectively. The probability of the n-th time bin t_n to contain a spike is given by the squared average firing rate during that bin.

$$P\{\hat{n}_{i,m}(t_n) = 1\} = \left(\overline{n^+_i\left([t_n, t_{n+1})\right)}\right)^2 \tag{6}$$

The time-dependent firing rate is squared here to better approximate the true log-normal distribution of firing rates, as explained in Methods 2.1. To simplify computations, we simulate the smooth firing rate $n_i(t)$ with the same resolution of 1 ms as the bins of spiking activity. This way, each bin corresponds to exactly one firing rate value: $n^+_i\left([t_n, t_{n+1})\right) =$

$n^+_i(t_n)$ In the limit of infinite trials $m \to \infty$, the averaged spiking activity is equal to the nonnegative firing rate.

$$\langle \hat{n}_{i,m} \rangle \big|_{M \to \infty} = n^+_i \tag{7}$$

If we want to simulate spiking activity without the nonlinear effect of the cut-off of firing rates at zero, we can simulate negative spikes by modifying eq. 5:

$$n^-_i(t) = \begin{cases} 0, n_i(t) \geq 0 \\ -n_i(t), n_i(t) < 0 \end{cases} \tag{8}$$

$$P\{\hat{n}^-_{i,m}(t_n) = 1\} = \left(\overline{n^-_i\left([t_n, t_{n+1}]\right)} \right)^2 \tag{9}$$

2.2.3   Reconstruction of simulated PSTHs

The spiking activity $\hat{n}_{i,m}(t_n)$ is averaged across trials and smoothed with a Gaussian kernel K (Eq. 10) with a σ of 60 ms, as was done for the recorded data (Chapter 2.1: Data preprocessing). The square root transform is applied to the smoothed, trial-averaged spiking activity to obtain the smoothed single-neuron spiking activity $\tilde{n}_i(t)$, which, in turn, is combined to form the smoothed spiking population response $\tilde{N}(t)$.

$$K(t) = \left(\frac{1}{\sigma_{filter}\sqrt{2\pi}} \cdot e^{\frac{t^2}{2 \cdot \sigma_{filter}^2}} \right); \;\; \sigma_{filter} = 60 \tag{10}$$

$$\hat{n}_{smoot\,,i,m}(t) = \hat{n}_{i,m}(t) * K(t) \tag{11}$$

$$\tilde{N}(t) = \bigcup \tilde{n}_i(t); \; \tilde{n}_i(t) = \sqrt{\langle n_{smooth,m}(t) \rangle \big|_m} \tag{12}$$

Here, $*$ denotes convolution and $\langle N_{smoot\,,m}(t) \rangle \big|_m = \overline{N_{smooth}(t)}$ denotes averaging over trials. The square root transform reverses the earlier squaring of firing rates.
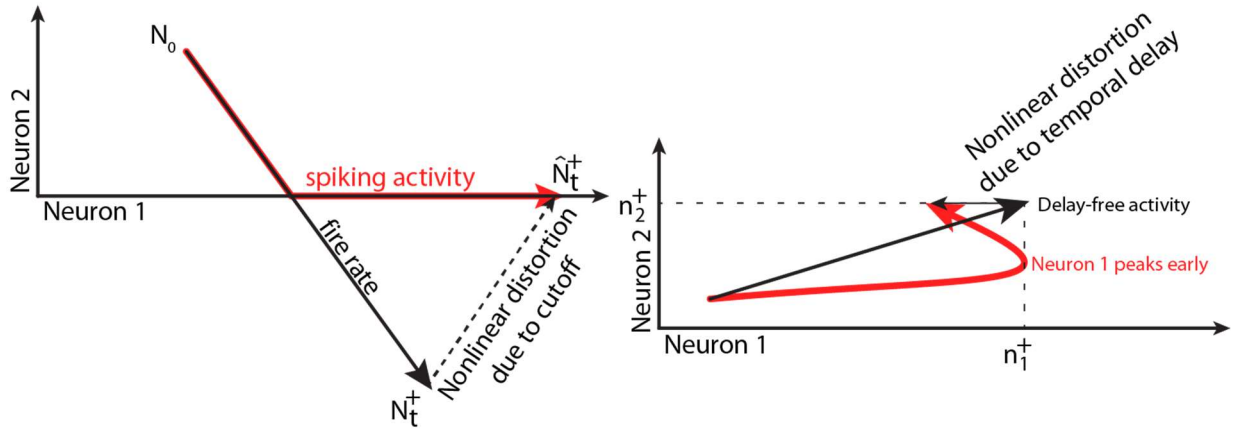
Methods

Figure 7. Left: Illustration of nonlinear distortion of spiking activity due to cutoff of firing rates of neuron 2 at 0. Right: Nonlinear distortion due to temporal delay. Neuron 1 peaks ahead of neuron 2, resulting in a distortion of the trajectory.

The variables have been chosen so that, similar to Eq. 7 and in the limit of large trial numbers $m \to \infty$, the trial-averaged PSTHs $\tilde{n}_i(t)$ are equal to the non-negative firing rates $n_i^+(t)$ smoothed by the same Gaussian smoothing kernel;

$$\lim_{M \to \infty} \tilde{n}_i(t) = n_i^+(t) * K(t) \tag{13}$$

As a corollary of this equation, we can estimate the activity attributable to the Poisson noise of the spiking process $\tilde{N}_{noise}(t) = \bigcup_i \tilde{n}_{i,noise}$ by subtracting the trial-averaged PSTHs from the firing rates:

$$n_i^+(t) * K(t) - \tilde{n}_i(t) = \tilde{n}_{i,noise} \tag{14}$$

2.2.4   Projection of activity onto subspaces of latent variables

For every latent variable $l_j(t)$, the weight vector $\vec{\beta}_j$ spans a dimension in the space of neural population activity that contains most of the activity related to that latent variable. We define the projection of the full smoothed and trial-averaged population space activity onto $\vec{\beta}_j$ as the projected latent variable trajectory $\tilde{l}_j(t)$:

$$\tilde{l}_j(t) = \tilde{N}(t)\,\vec{\beta}_j \tag{15}$$

We can successively expand eq. 15, using eqs. 2 and 5, to obtain the relationship between $\tilde{l}_j(t)$ and $l_j(t)$:

$$\tilde{l}_j(t) \approx K(t) * \left[\vec{N_0} \odot \left(1 + \sum_k \vec{\beta_k} \odot \vec{l}_k(t) \right) \right]_{N(t) \geq 0} \cdot \vec{\beta_j} \tag{16}$$

Eq. 16 contains three nonlinearities; the first is the cutoff of firing rates at 0, represented by $[...]_{N \geq 0}$, the second is the vector of latent variables with temporal delays $\vec{l}_j(t)$ and the third is the Poisson noise of spiking activity. If we assume the absence of cut-off and no temporal delay, we can cancel a few terms using the orthogonality between $\vec{\beta}$: $\vec{\beta}'_i \vec{\beta}_j = \delta_{ij}$.

$$N(t) \geq 0 \ \forall \ N(t), \Delta t_{i,j} = 0 \ \forall \ i,j \implies \tilde{l}_j(t) = \vec{\beta}'_j \cdot K(t) * N_0 \odot \left(1 + \sum_k \vec{\beta}_k l_k(t) \right) = N_0 + K(t) * l_j(t) \tag{17}$$

It is easy to see that in the absence of nonlinear effects the latent variable profiles can be recovered from the projection of population activity onto the weight vector:

$$\widetilde{N}_i(t)\,\vec{\beta}_j = \tilde{l}_j(t) \approx K(t) * l_j(t) + N_0 \tag{18}$$

The nonlinearities are illustrated in fig. 7; the presence of temporal delays between observed variables shifts variability out of the mutually orthogonal subspaces (Wood and McCarthy, 1984; Beauducel and Debener, 2003; Lakshmanan et al., 2015) (Fig. 7, right). Additionally, the cut-off of spiking activity at zero introduces nonlinear distortions to neural population trajectories (Fig. 7, left). Poisson noise is independent of all other effects and variables; it is represented in the inequality sign in eq. 21. If we assume that the other two effects are small, we can represent them as linear vector terms:

$$N(t)|_{N \geq 0} = N(t) + \vec{\varepsilon}_{cutoff}(t) \tag{19}$$

$$\vec{l}_j(t) = I * l_j(t) + \vec{\varepsilon}_{j,delay}(t) \tag{20}$$

Where I is the identity matrix. We can now expand Equation 16 into four terms:

$$\tilde{l}_j(t) \approx \vec{\beta}'_j \cdot K(t) * \left(\vec{N_0} \odot \left(1 + \sum_k \vec{\beta}_k l_k(t) + \sum_k \vec{\beta}_k \odot \vec{\varepsilon}_{k,delay}(t) \right) + \vec{\varepsilon}_{cutoff}(t) \right) \tag{21}$$

$$... = \vec{\beta}'_j \vec{N_0} + K(t) * l_j(t) + \vec{\beta}'_j K(t) * \sum_k \vec{N_0} \odot \vec{\beta}_k \odot \vec{\varepsilon}_{j,delay}(t) + \vec{\beta}'_j \cdot K(t) * \vec{\varepsilon}_{cutoff}(t) \tag{22}$$

Methods

Of eq. 22, the first term is constant, the second is the original latent variable smoothed with a Gaussian kernel and the final two are nonlinearity-related terms. The Poisson noise is reflected in the approximal equality sign. We can assume that in the case where nonlinear effects are small, $\tilde{l}_j(t)$ is a smoothed approximation of the ground truth latent variable $l_j(t)$.

2.3   Matching of modeled with recorded biological parameters

2.3.1   Matching of single-neuron distributions

First, we matched a large variety of population parameters of recorded data as accurately as possible in our model to maximize relevance to recorded data. The first parameters matched is the distribution of average firing rates of recorded neurons from both areas during the task. Target firing rates $\vec{N}_{target}$ were randomly sampled from the recorded distribution. We constrain the weighted Poisson process to draw exactly the number of spikes necessary to fulfill the target firing rates $\vec{N}_{target}$ over the course of a trial of length t_{trial}. Additionally, we need Eq. 7 to be valid. We need to adjust firing rates so that the integral over the firing rate for all task conditions c is equal to the target firing rate. The condition to be fulfilled is:

$$n_{i,target}\, c \cdot t_{trial} = \sum_c \int_0^{t_{trial}} \left[n_{i,0}\left(1 + \sum_k \beta_{i,k} l_{i,k}(t)\right)\right]_{N \geq 0} dt \qquad (23)$$

In other words, if we concatenate the firing rates of a single neuron over all conditions, the area between the curve the zero line needs to be equal to the target firing rate multiplied by the length of the trial times the number of conditions. We can leave out the smoothing kernel $K(t)$; The integral function is invariant under convolutions if the convolution kernel has an area of 1. Due to the nonlinear firing rate cutoff at 0, we can't simply scale the right side of the equation with a fixed number. We will instead iteratively adjust the scale together with the amplitudes of latent variables, as explained in detail in chapter 1.1.1.

The distribution of firing rate variances as a function of firing rates does not need further matching; The multiplicative modulation of firing rates by latent variables is sufficient to ensure that the distributions match, as is shown in the results section, fig. 10b.

We will also match the distributions of the LV contributions of individual neurons. The interaction of each neuron i to each LV j is given by the weight $\beta_{i,j}$, which corresponds to the contribution of each neuron to the activity of each LV, or in other words, the modulation of each neuron by each LV. The weight vectors $\vec{\beta}_j$ for each LV are normal, and $\beta_{i,j}$ can be positive or negative, corresponding to excitatory and inhibitory modulation. We thus consider the absolute values of $\beta_{i,j}$, divided by the square root of the number of neurons $\sqrt{I}$ as the contribution strength of neuron i to latent variable j:

$$contribution\ strength: neuron\ i \rightarrow LV\ j = \frac{|\beta_{i,j}|}{\sqrt{I}} \tag{24}$$

Dividing by the square root of the number of neurons normalizes contribution strengths across recordings with different numbers of neurons. We also use the contribution strength of neuron i to all LVs, given by the average of all individual contribution strengths:

$$contribution\ strength: neuron\ i \rightarrow all\ LVs = \frac{\sum_j |\beta_{i,j}|}{J \cdot \sqrt{I}} \tag{25}$$

Here, J is the number of latent variables j.

2.3.2 Estimation of latent variable profiles and amplitudes

Additionally to the before mentioned parameters, we want to match the temporal profiles $l_j(t)$ of LV activity in the recording. However, for the recorded data, we can only estimate $\tilde{l}_{recorded,j}(t)$. The latent variables $l_j(t)$ also depend on the task condition c, as mentioned before and given by eq. 1; $l_j(c,t) = \gamma_j(c) \cdot l_j(t)$. By extension, the projected latent $\tilde{l}_j(t)$ also depends on c. We can expand eq. 15 to:

$$\gamma_j(c) \cdot \tilde{l}_j(t) \approx \tilde{N}(c,t)\,\vec{\beta}_j \tag{26}$$

Eq. 26 is valid for every condition c separately. We can divide by $\gamma_j(c)$ and average across all C' conditions c', for which $\gamma_j(c)$ is not too small. We have chosen $|\gamma_j(c')| > 0.3$ as the arbitrary cutoff point.

$$\tilde{l}_j(t) = \frac{1}{C'}\sum_{c'} \frac{\tilde{N}(c',t)\,\vec{\beta}_j}{\gamma_j(c')}, C' \ni c' \rightarrow |\gamma_j(c')| > 0.3 \tag{27}$$

We only use conditions c where $\left| \gamma_j(c) \right|$ is large, because dividing by small values will amplify the noise in the term $\widetilde{N}(c',t)\, \vec{\beta}_j$. For the condition-dependent variables in the recorded data, $\gamma_j(c)$ assumes the values [-1, 1] and C is 2. Intuitively, we approximate $\tilde{l}_j(t)$ by flipping the trajectory for one of the conditions and averaging across both.

Now, we want to use the LV profiles of recorded data $\tilde{l}_{recorded,j}(t)$ as the ground truth profiles $l_j(t)$ in the simulation. However, the projections of population activity onto the individual subspace dimensions is based on data that has been filtered by two effects: For one, the activity of individual neurons has been smoothed by convolution with the Gaussian kernel K(t). Second, the presence of temporal delays between individual LVs and neurons results in additional filtering that can be approximated by convolution with a second Gaussian kernel, with a σ equal to the σ of the distribution of temporal delays. Therefore, we apply Richardson-Lucy deconvolution to the recorded LV profiles in order to obtain LV profiles for our simulation. Figure 9 shows, that the resulting projections of simulated LV activity closely resemble the recorded profiles.

2.3.3   PAIRS test

An important feature of neural population is the distribution of neural contributions to individual LVs across individual neurons. A random distribution is evidence against the existence of neural categories, and our model should reflect the distribution of recorded neural contributions. To test whether the contributions of individual simulated and recorded neurons to each LV are independently distributed, we used the PAIRS test (Raposo et al., 2014), using freely available code at: http://repository.cshl.edu/30912/. The PAIRS analysis tests for pairs of neurons in the neuronal population that would exhibit more similar response weights to their k nearest neighbors than expected by chance. For this purpose, PAIRS uses the per-neuron weights $\vec{\beta}_i$, which are composed of the contribution weights of a single neuron to all latent variables. First, we construct a matrix where the columns are the latent variable weight vectors $\vec{\beta}_j$, the per neuron weight vectors $\vec{\beta}_i$ are the rows of that matrix. For each neuron weight vector $\vec{\beta}_i$, PAIRS computes the average angle to the n nearest neighbors. Second, the distribution of all average nearest neighbor angles is compared against the same distributions of a large number of simulated surrogate datasets (10000 iterations) with the same number of neurons and dimensions. For the surrogate

datasets, the values of $\vec{\beta}_i$ drawn randomly from a Gaussian distribution. The PAIRS statistic is calculated as a function of the median nearest-neighbor angles $\tilde{\theta}$ as (Raposo et al., 2014):

$$PAIRS = \frac{\tilde{\theta}_{surrogate} - \tilde{\theta}_{data}}{\tilde{\theta}_{surrogate}} \qquad (28)$$

PAIRS is equal to one if the median nearest-neighbor angle of the recorded data is equal to zero, equivalent to each weight vector having at least n collinear neighbors, and zero if the weight vectors are distributed randomly. PAIRS can be negative, if the vectors in the dataset are distributed more randomly than expected by chance. A two-sided p-value for the PAIRS statistic was obtained by calculating the proportion of median average nearest neighbor angles of the 10000 surrogate datasets that was larger and smaller than the corresponding median value of the corresponding recorded or modeled distribution.

2.3.4 Procrustes Analysis-based Matching of Temporal Delays

An important factor to be matched between the recorded and simulated neural population activities are the temporal delays between individual neurons and latent variables, expressed as $\Delta t_{i,j}$ in eq. 3. The distribution of temporal delays cannot be inferred directly from the recorded data because of overlapping activity of different subspaces or latent factors, both in the population and for individual neurons. Instead, we assume a normal distribution with mean zero and distribution parameter σ, then find a value for σ to maximize the similarity between simulated and recorded population activity. To measure the similarity between simulated and recorded data, we used Procrustes analysis.

Procrustes analysis is a similarity assessment method that shift and rotates one dataset onto another while minimizing the difference between pairs of data points. In our case, let our simulated population activity be expressed as a matrix $X \in \Re^{CT,N}$, where C are the task conditions, T are the time points of the task and N is the number of simulated neurons. X is the matrix of single neuron PSTHs concatenated across conditions. Similarly, let the recorded activity be $Y \in \Re^{CT,M}$, where M is the number of recorded neurons. Observations that correspond to the same value of CT are assumed to be "similar", yet expressed in a different basis across the two matrices. Standard Procrustes analysis seeks a transformation $TR(X) = X \cdot R + b = Y$ that optimally aligns X and Y. R does not need to be a square

Methods

matrix; Procrustes analysis can compare data of different dimensionality, as long as the number of observations in each dataset is equal. We go a step further, centering both matrices and normalizing them by the Frobenius norm, yielding the new matrices $\hat{X}$ and $\hat{Y}$. The normalization by the Frobenius norm allows to compare matrices of different scales and sizes. The transformation TR reduces to an orthogonal projection R and we quantify the alignment between $\hat{X}$ and $\widehat{YR}$ as:

$$r_{similarity} = 1 - \left\| \hat{X} - \hat{Y} \cdot R \right\| \tag{29}$$

Here, $\|X\|$ denotes the Frobenius norm. $r_{similarity}$ is equal to 1 if the two matrices are aligned perfectly. Note that while this measure is conceptually related to the usually used mean squared error between Procrustes-matched datapoints, it is not mathematically identical.

2.4 Methods for the evaluation of PCA

In this chapter, we describe the methods used to analyze the performance of PCA when applied to simulated spiking activity of a neural populations in response to two behavioral tasks. We investigate, how well the true underlying dimensionality of the dataset can be estimated from the results of PCA, how well PCA captures individual subspaces of task-related activity and how individual principal components capture variance related to different individual subspaces.

2.4.1 PCA in general

Before we analyze the performance of PCA in application to neural data in detail, we need to first define some basic concepts of PCA (Jolliffe, 2002). PCA is a dimensionality reduction method that is applied to a $n \times i$ data matrix, which is in our case the matrix of smoothed spiking population activity $\widetilde{N}$. The rows n correspond to observations (time points) and columns i to individual observed variables (neurons). The covariance matrix of $\widetilde{N}$ is a matrix with real eigenvalues λ_i:

$$Eig\left(\widetilde{N}^T \cdot \widetilde{N}\right) = \lambda_1, \lambda_2, \ldots, \lambda_i \ \ \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_i \tag{30}$$

$\widetilde{N}$ must be column-mean subtracted; The mean of every column needs to be zero. PCA then finds a $i \times i$ rotation matrix W that rotates $\widetilde{N}$ onto a new basis:

$$T = \tilde{N} \cdot W;\ W = [\vec{w}_1 \quad \vec{w}_2 \quad \ldots \quad \vec{w}_i], T = [\vec{t}_1 \quad \vec{t}_2 \quad \ldots \quad \vec{t}_i] \tag{31}$$

W is called the coefficient matrix, and the vectors $\vec{w}_i$ are the coefficient vectors, called principal components. T is called the score matrix, the columns of T are the principal component scores. The vectors $\vec{w}_i$ each span a dimension in the space of observed variables, and the principal component scores can be interpreted as projections of the full dataset onto the principal component.

W and T fulfill the following conditions: All $\vec{w}_i$ are mutually orthogonal and unit length. All principal component scores are linear combinations of the original observations. The scores are furthermore mutually uncorrelated, the covariance matrix of the principal component score is a diagonal matrix whose entries are the eigenvalues of the original covariance matrix λ_i.

$$T' \cdot T = \begin{bmatrix} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_i \end{bmatrix} \tag{32}$$

The eigenvalues λ_i correspond to the variance of the corresponding score component:

$$var(\vec{t}_p) = var(\tilde{N} \cdot \vec{w}_i) = \lambda_i \tag{33}$$

Ordered by descending magnitude, the eigenvalues form the characteristic spectrum of variance per principal component.

2.4.2   Spectrum of variance captured per component

In figures 18 and 24a in the results section we show decompositions of the spectra of variance of the recorded and simulated datasets. The formula for variances per principal component is derived from Eq. 33: we divide the variance by the square root of the number of neurons in order to obtain the variance per neuron; this allows us to compare variances between populations of different sizes.

$$\lambda_{i,PCA} = \frac{var(\tilde{N} \cdot \vec{w}_i)}{\sqrt{I}} \tag{34}$$

Here, I is the number of neurons.

Methods

For the variance captured by individual subspaces, we adapt this formula by using the base vectors of the LV-related subspaces $\vec{\beta}_j$ instead of principal components:

$$\lambda_{j,subspace} = \frac{var(\tilde{N} \cdot \vec{\beta}_j)}{\sqrt{I}} \tag{35}$$

Unlike principal components, projections of activity onto individual subspaces are not necessarily uncorrelated. To calculate the spectrum of variance of the part of the population activity that lies outside the LV-related subspaces, we construct a base of random vectors $\vec{\xi}_k$ that are orthogonal and exclude the subspaces of LV-related activity and replace $\vec{\beta}_j$ in eq. 35:

$$\vec{\xi}_k \cdot \vec{\xi}_l = 0 \; \forall \, k \neq l \; \; \vec{\xi}_k \cdot \vec{\beta}_j = 0 \; \forall \, k, j \tag{36}$$

$$\lambda_{\xi,noise} = \frac{var(\tilde{N} \cdot \vec{\xi}_k)}{\sqrt{I}} \tag{37}$$

2.4.3 Decomposition of LV-related activity

To quantify the mixing of latent variables and noise in individual principal components, we would like to perform a linear decomposition of the activity captured by each PC into contributions attributable to individual latent variables j (and noise). The quantity that is usually measured in the context of PCs is the variance of the population activity that is captured by the n-th principal component; $var(\tilde{N}(t)\vec{w}_n)$. However, there is no natural linear decomposition of variance, as it is a non-linear measure; Instead, we measure individual amplitude contributions $\hat{\lambda}_{j,n}$ of each LV and of noise to each PC and normalize them by the total variance of each PC. We impose the following conditions on the decomposition;

1. Individual contributions to a PC should sum up to the variance along the PC:
   $var(\tilde{N}(t)\vec{w}_n) = \sum_i \hat{\lambda}_{j,i}$

2. The proportion of each contribution $\hat{\lambda}_{j,i}$ to the total variance $\hat{\lambda}_i$ should reflect the proportion of the activity of the single latent variable LV_j to all latent variables: $\frac{\hat{\lambda}_{j,i}}{\hat{\lambda}_i} :=$
   $\frac{LV_j \vec{w}_i}{\sum_j LV_j \vec{w}_i}$.

We aim to decompose the variance of the projection of the simulated neural population activity $\widetilde{N}(t)$. Eqs. [13] and [14] gives us a form for $\widetilde{N}(t)$: that depends only on known fire rates $N^+(t)$ and measurable Poisson noise of the spiking process $\widetilde{N}_{noise}(t)$, both given here for the whole population:

$$\widetilde{N}(t) = N^+(t) * K(t) + \widetilde{N}_{noise}(t), \qquad N^+(t) = \bigcup_i n_i^+(t), \qquad \widetilde{N}_{noise}(t) = \bigcup_i n_{i,noise}^+(t) \quad (38)$$

(Convolution with $K(t)$ is the smoothing step, see Eq. 10 ff). The population fire rates $N^+(t)$ are equal to the positive part of the sum of individual LV modulations (Eqs. 4 and 5). Unfortunately, the cut-off of fire rates at 0 (Eq. 5) is a nonlinearity that prevents us from constructing $N^+(t)$ as a linear sum of the pre-cutoff fire rates $N_i(t)$; Just as in the first paragraph of this section, we construct a representation $N^+(t) = \sum_i N_i^+(t)$ where each $N_i^+(t)$ depends only on the individual activity of a single LV and the amplitude of $N_i^+(t)$ in proportion of all $N_i^+(t)$ is proportional to the amplitude of the corresponding $N_i(t)$: $\frac{N_i^+(t)}{\sum_j N_j^+(t)} := \frac{N_i(t)}{\sum_j N_i(t)}$. The following form fulfills both conditions:

$$N_i^+(t) = \frac{N_i(t)}{\sum_j N_i(t)} \sum_j N_j^+(t) = N_i(t) \frac{N^+(t)}{\sum_j N_i(t)} \qquad (39)$$

Now, we can express the population PSTH $\widetilde{N}(t)$ as a linear sum (Recall that convolution is a linear operation):

$$\widetilde{N}(t) = \sum_i N_i^+(t) * K(t) + \widetilde{N}_{noise}(t) = \sum_i N_i(t) \frac{N^+(t)}{\sum_j N_i(t)} * K(t) + \widetilde{N}_{noise}(t) \qquad (40)$$

We use the formula for the variance of a sum of variables $var\left(\sum_j X_i\right) = \sum_j \sum_k cov(X_j, X_k)$ to calculate the variance of the projection of the population PSTH $\widetilde{N}(t)$ onto each individual PC $\vec{w}_n$;

Methods

$$var\big(\widetilde{N}(t)\vec{w}_n\big) = var\left(\sum_i N_i^+(t)\vec{w}_n * K(t) + \widetilde{N}_{noise}(t)\vec{w}_n\right) = \cdots$$

$$\sum_i var\big(N_i^+(t)\vec{w}_n * K(t)\big) + var\big(\widetilde{N}_{noise}(t)\vec{w}_n\big) + \cdots$$

$$\sum_{j \neq k} cov\left(N_j^+(t)\vec{w}_n * K(t), N_k^+(t)\vec{w}_n * K(t)\right) + \cdots$$

$$\sum_i cov(N_i^+(t)\vec{w}_n, \widetilde{N}_{noise}(t)\vec{w}_n) \tag{41}$$

It is reasonable to treat the terms related to noise in this equation in the same way as the terms related solely to LV-related activity because they are of the same form. We shall therefore append the noise term to the series of LV-related terms N_i^+; $N_{I+1}^+(t) :=$ $\widetilde{N}_{noise}(t)$. We separate each of the nonlinear covariance term in the last two lines into two summands $C_{j|k}$; $cov\big(N(t)_j^+ * K, N(t)_k^+ * K\big) := C_{j|k} + C_{k|j}$. As in the previous decompositions, each summand $C_{j|k}$ is to be proportional to the contributions of the two individual projections $N(t)_i^+\vec{w}_n$ to the covariance term. We calculate proportions using absolute values of $N(t)_j^+\vec{w}_i$:

$$\frac{\big|C_{j|k}\big|}{\big|C_{k|j}\big|} := \frac{\sum_t \big|N(t)_j^+\vec{w}_n * K - \overline{N(t)_j^+\vec{w}_n * K}\big|}{\sum_t \big|N(t)_k^+\vec{w}_n * K - \overline{N(t)_k^+\vec{w}_n * K}\big|} \tag{42}$$

For sake of readability, we shall substitute $K * N(t)_j^+\vec{w}_n = p_j(t)$. A form for $C_{j|k}$ that satisfies eqs. and is given by:

$$C_{j|k} = \sum_t \left(\frac{\big|p_j(t)\big|}{\big|p_j(t)\big| + \big|p_k(t)\big|} \cdot \frac{\big|p_j(t)p_k(t)\big|}{\sum_t \big|p_j(t)p_k(t)\big|}\right) \cdot cov\big(p_j(t), p_k(t)\big) \tag{43}$$

$C_{j|k}$ represents the portion of the covariance of latent variables j and k that we attribute to latent variable k. We now can sum all $C_{j|k}$ to obtain the portion of total variance that is attributable to latent variable j. However, because the covariance between two variables can be negative, $C_{j|k}$ can be negative as well. In order to properly compare positive and negative contributions, we once again use absolute values:

$$\hat{\lambda}_{j,i} = \sum_k |C_{j|k}| \cdot \frac{\left| \sum_k cov\left(p_j(t), p_k(t)\right) \right|}{\sum_k \left| cov\left(p_j(t), p_k(t)\right) \right|} \tag{44}$$

If j <=J, where J is the number of LVs, $\hat{\lambda}_{j,i}$ is the portion of total variance captured by PC i that can be attributed to LV j after considering the mutual amplification and cancellation of individual LVs and Poisson noise and the cut-off of individual neural firing rates at 0. Similarly, for j = J+1, $\hat{\lambda}_{J+1,i}$ is the portion of variance attributable to the Poisson noise of the simulated spiking process.

3 Results

3.1 Simulation of a Spiking Neural Population: Results

3.1.1 Electrophysiological Basis of Simulated Population Response

To systematically evaluate the accuracy of PCA in capturing the underlying population response structure, we need to know the ground truth of the structure of neural population response. Because we can't know the ground truth of a real recording, we evaluated the accuracy of PCA on a simulation. We performed the following steps: First, we designed a spiking model of the fronto-parietal population activity in response to two visually guided delayed movement tasks. The population response was structured as a set of multiple orthogonal subspaces; this structure has been found in many studies before (Elsayed et al., 2016; Dann, 2017; Michaels et al., 2018) Second, in order to increase the relevance of our model for recorded data, we matched a large number of neural population properties of the model to recorded population activity. Third, we generated a large variety of simulated neural population responses to cover a wide range of realistic recording conditions. For this purpose, we systematically varied the number of trials, neurons, and conditions. Finally, we compared the simulated population response structure with the corresponding response structure captured by PCA. Thereby we examined three aspects: (1) how well the dimensionality of the response structure can be estimated from the results of PCA, (2) how much of the activity of each subspace is lost when PCA is used to reduce the dimensionality of the dataset (see fig. *5*) and (3) to what degree individual PCs are aligned with the subspaces of the simulated response structure.

3.1.2 Model of spiking activity

We modeled the neural population response to two commonly used visually guided movement tasks; The delayed reach-to-grasp task (Townsend et al., 2011; Michaels et al., 2015; Dann et al., 2016) (Fig. 8a) and the center-out reach task (Georgopoulos et al., 1986; Shenoy et al., 2013; Sadtler et al., 2014; Gallego et al., 2018a)(Fig. 8d). The response to these tasks was structured into three orthogonal subspaces (Elsayed et al., 2016; Murray et al., 2017; Lara et al., 2018; Stringer et al., 2018), comprised of activity related to, respectively, visual cue information, movement preparation and movement execution. The
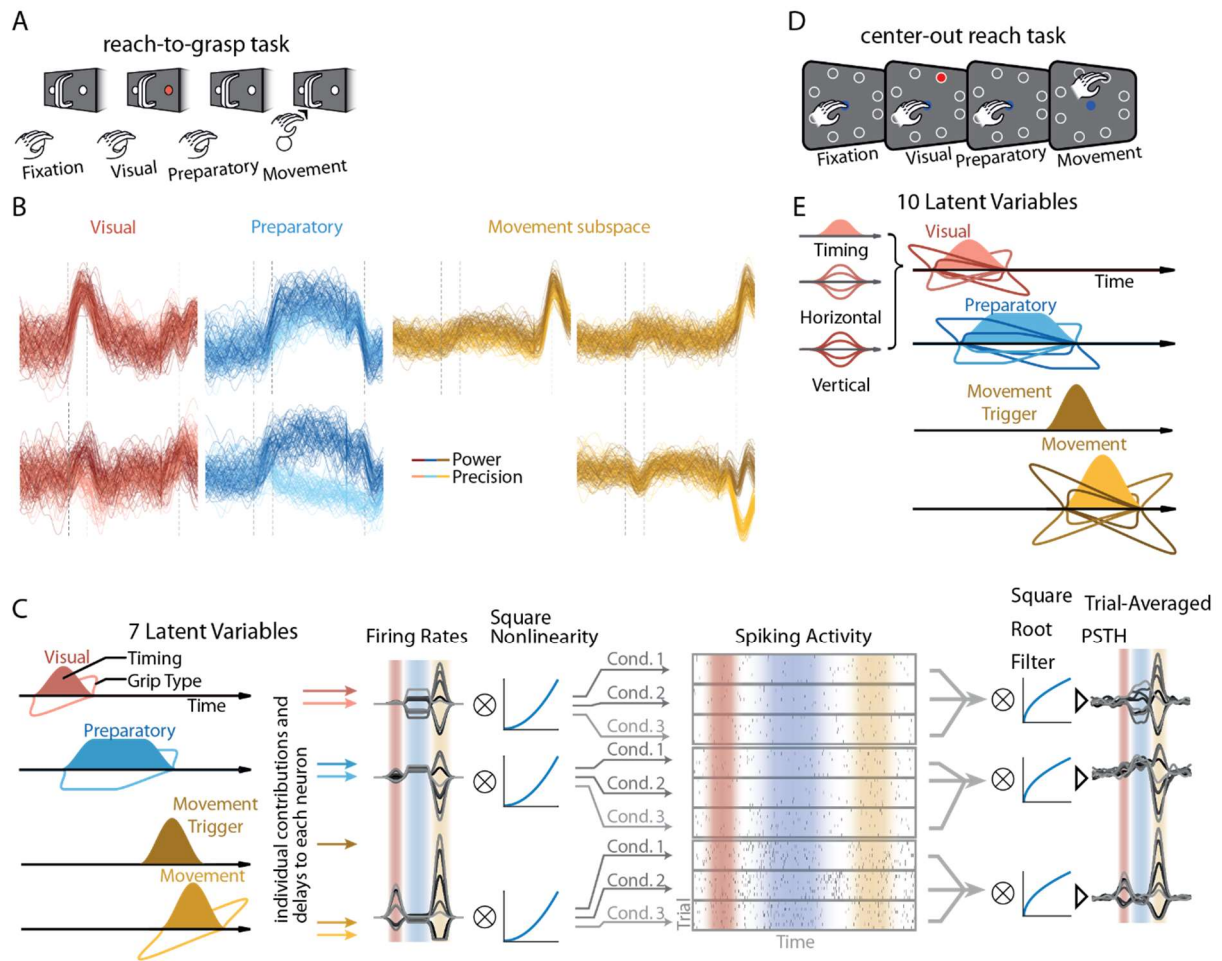
*Figure 8: Schematic illustration of the spiking model of neural population activity. **A**: The delayed two-condition reach-and-grasp task. After an initial fixation period, the monkey was visually cued to reach out and grip a handle with one of two grip types. Next, the cue vanished, and the monkey prepared its motion. Finally, the monkey was cued to execute movement. **B**: Recorded single-trial population response to the task in panel A projected onto seven latent variables (LVs) that can be grouped into three subspaces The subspaces were chosen to maximize the correlation between the activity profiles of the activity captured by them and the full activity, with an R^2 = 0.95. **C**: The recorded variables from B are used as the ground truth of the simulation The activity of each LV is distributed across neurons with random weights and temporal delays and preserving orthogonality between LVs, in the form of multiplicative firing rate modulations. The neuronal firing rates are fitted with a square non-linearity and used as weights in a weighted Poisson process to generate spiking activity. The spiking activity is then smoothed with a Gaussian kernel filter, the square nonlinearity is removed and the activity is averaged across trials to create Peri-Stimulus Time Histograms. **D**: The two-dimensional center-out reach task. The monkey is cued to reach towards one of eight goals. This task shares the temporal structure of the reach-and-grasp task. **E**: The LV structure of the response to the task in D. Each condition-dependent LV in the model of the previous task is replaced with two LVs that encode the horizontal and vertical position of the reach target.*

simulated population response structure is based on the observed population response structure of a dataset recorded in our lab. The dataset comprised multiple recordings of the neural response from the fronto-parietal network of two monkeys that were trained to perform three variants of a delayed grasping task. The monkey had to grasp a handle with

either a power and precision grip (Fig. 8a). The task was performed in one of three contexts: In the instructed context monkeys were visually cued to perform the associated grip type after a delay (Fig. 8a, CF methods). The full behavioral task also comprised two other contexts that were not used for this study. We recorded from the fronto-parietal grasping network, including part of the ventral premotor (F5) and anterior intraparietal (AIP). In each area, recordings were obtained from two floating microelectrode arrays (FMAs), for a total of 64 channels (32 per microarray) per area. 55±14 well isolated neurons with average firing rates above 1Hz (see Methods) were captured per session. The dataset comprised 9 recording sessions (6 and 3 sessions from monkey S and Z, respectively).

The underlying population response structure was extracted using a dimensionality reduction method based on support vector machines (SVM) that leverages orthogonal parts of the population response over time (see Methods 2.1.2). The extracted structure was consistent with the structure described in the literature (Elsayed et al., 2016; Murray et al., 2017; Lara et al., 2018; Stringer et al., 2018); We identified three subspaces for visual, preparatory, and movement information, each consisting of one condition-dependent and one condition-independent dimension (Fig. 8b). The subspace for movement execution comprised an additional condition-independent component preceding both other dimensions. In accordance with previous literature, this dimension was predictive of movement onset and is therefore also called "movement trigger" dimension (Kaufman et al., 2014b; Sussillo et al., 2015; Michaels et al., 2018; Ames et al., 2019). The seven identified subspace dimensions captured 79±1% of all task dependent linear separable variance.

We used the identified subspace structure as a basis for our simulation (Fig. 8c, left and 8e). The contributions of modeled subspace dimensions to all neurons were chosen randomly in accordance with empirical findings (Machens et al., 2010; Raposo et al., 2014)(Fig. 10e); the time-dependent, noise-free firing rate of each individual neuron is a linear combination of the activity of all LVs (Fig. 8c mid left). Neural firing rates were modeled using a square nonlinearity (see Methods 2.2.2) that reflects the nonlinear firing dynamics of real neurons (Yu et al., 2009; Afshar et al., 2011; Buzsáki and Mizuseki, 2014; Michaels et al., 2015). Next, the neural firing rates were used as weights in a weighted Poisson process to obtain spiking activity. To simulate multiple repetitions of a trial, we repeated the Poisson process multiple times for the same condition. The spiking activity was smoothed using a Gaussian kernel

with a σ of 60 ms and then a square root filter was applied to correct for the square nonlinearity (see. Methods 2.2.2). The activity was then averaged across multiple trials to obtain peri-stimulus time histograms (PSTHs). We modeled the delayed center-out reaching task (Fig 8d) as an extension of the reach-and-grasp task. The simulated neural population response to the center-out reaching task was based on the temporal dynamics and subspace structure of the first task, with an additional condition-dependent LV in each subspace (Fig. 8d). The circular arrangement of reaching targets in the center-out-reach task was modeled by a circular dependence of the two condition-dependent LVs in each subspace. This is a conservative approximation of the true structure of the neural population response (Georgopoulos et al., 1982).

3.1.3   Comparison of Biological Parameters between Recording and Simulation

To increase the relevance of our model for recorded data, we match a large number of biological parameters between the simulation and the recorded population response: the distributions and variations of neural firing rates (Buzsáki and Mizuseki, 2014; Dann et al., 2016), subspace contributions per neuron and across the population, the temporal profiles of subspace activity (Dann, 2017; Michaels et al., 2018), and the distributions of temporal delays between LV subspaces and individual neurons.
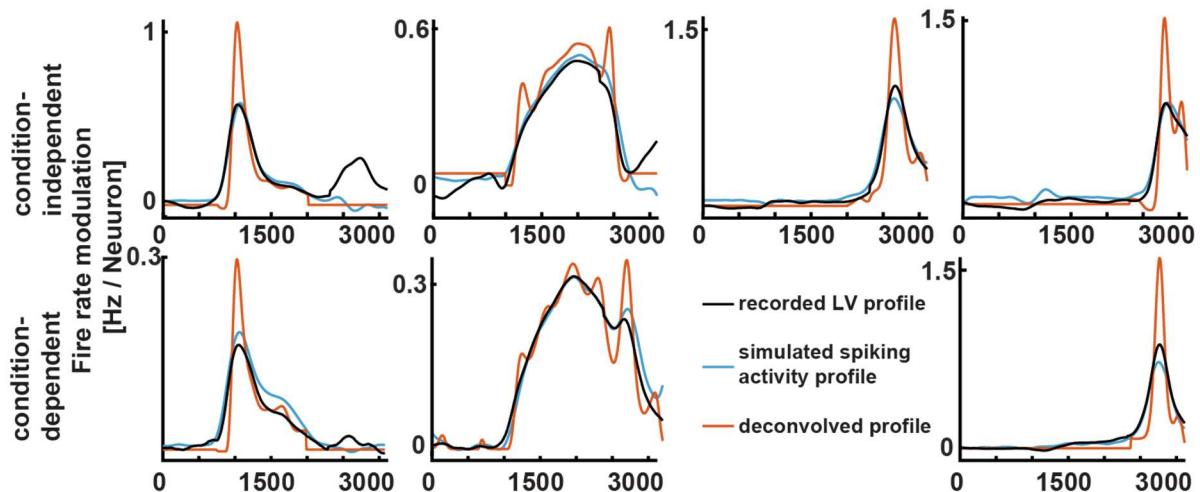

Figure 9: Recorded, de-convolved and simulated profiles of LV activity. Shown in black are the seven projections of population activity onto the LV subspaces. These projections are based on data that was, on the level of single neurons, filtered by convolution with a Gaussian Kernel and by the presence of temporal delays between neurons and LVs. The red profiles were obtained by applying Richardson-Lucy deconvolution to the black profiles to balance out the two aforementioned effects and setting secondary peaks in the profiles to zero. Using the secondary profiles as the ground truth LVs in an example simulation, we obtained the blue profiles as projections of simulated population activity onto the subspaces of LV-related activity.
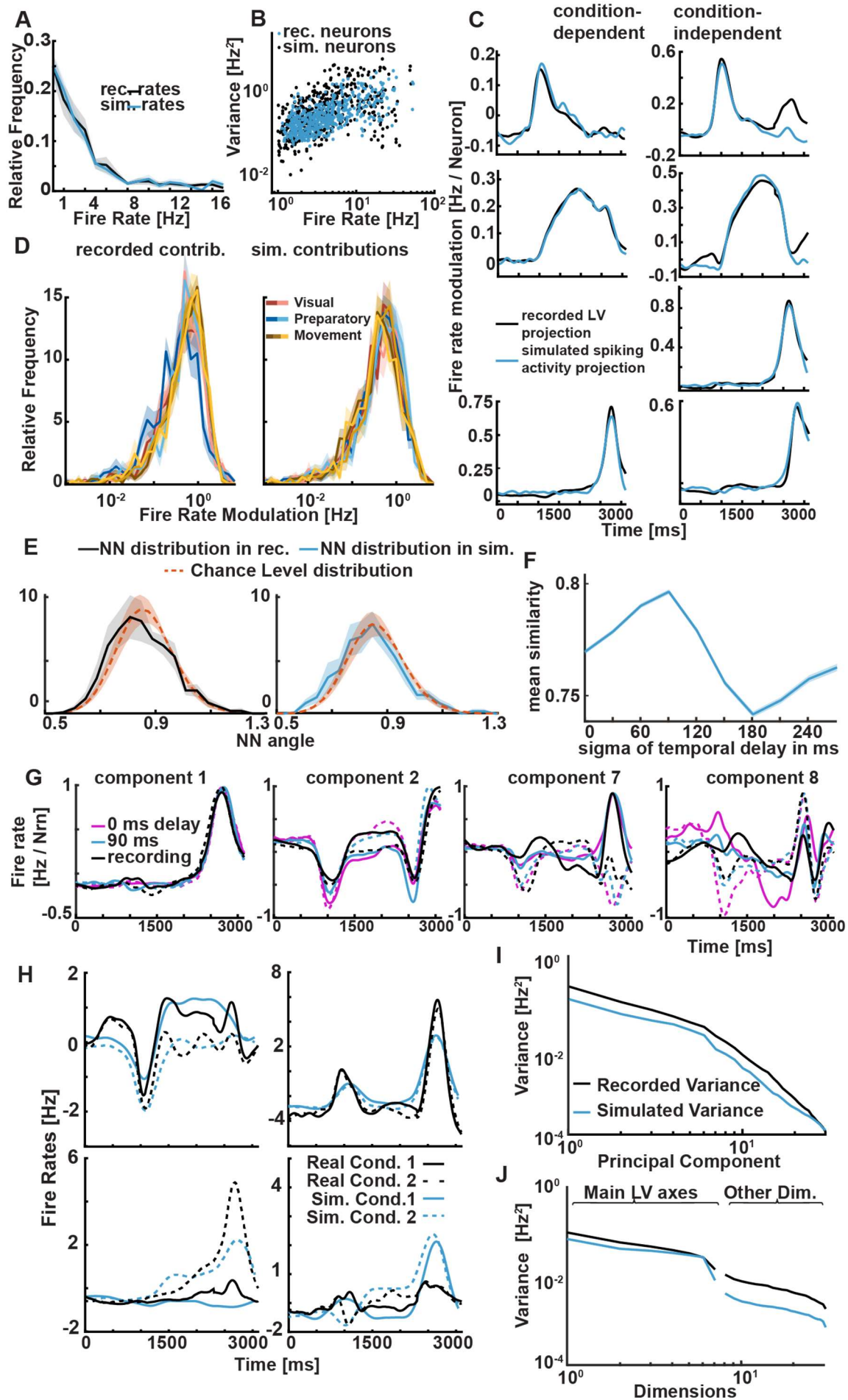
*Figure 10: Comparison of key parameters of the recording and simulation of the reach-to-grasp task. Unless stated otherwise, for each of the nine recordings a simulation with the same number of neurons was built. Each metric shown was calculated individually for each pair of recording and simulation, then averaged across the nine pairs. Data are shown with standard error. **A**: Histogram of firing rates of single neurons in simulation and recording. **B**: Scatter plot of variances of square-root transformed single-neuron firing rates plotted against the firing rates. **C**: Projections of the population response onto individual LVs. The recorded, smoothed spiking population activity (black) was projected onto dimensions discovered using SVM-based dimensionality reduction (see Methods 2.1.2). Simulated firing rate activity (red) and simulated smoothed spiking activity (green) are projected onto LV weight vectors $\vec{\beta}_j$ (see Methods 2.2.1). For condition-independent LVs (top row), activity was averaged across all trials. For condition-dependent LVs, activity was averaged across trials for both conditions, then the conditions were subtracted to obtain the difference between conditions (see Methods 2.2.1, eq. 1). The projected activity was normalized by the square root of the number of neurons. **D**: Histogram of multiplicative firing rate modulation strengths in response to the activity of each LV for each individual neuron. The modulation strengths were normalized by the square root of the number of neurons for each recording and simulation. **E**: Distributions of nearest neighbor angles for recording and simulation, as well as chance level distributions, as calculated for the PAIRS test (See Methods 2.3.4). The distributions of the median nearest neighbor angles of the recorded (black) and simulated (green) neurons are plotted alongside the surrogate distribution (red). **F**: Mean Procrustes similarity of simulated to recorded neural population responses. The Procrustes similarity is plotted against different values of the temporal delay distribution parameter σ (see Methods 2.3.2). **G**: Four principal components of a single recording (black) plotted together with the Procrustes-matched components of two simulations with different σ. Both simulations were individually rotated onto the same recording, then all three were rotated onto the principal components of the recording. **H**: PSTHs of four individual recorded neurons (black) and four simulated neurons that were matched to the firing rates, lv contribution coefficients and LV amplitudes of the recorded neurons (green). **I**: Spectrum of variance captured per principal component of recording (green) and simulation (black). This spectrum is equivalent to the spectrum of eigenvalues of the covariance matrix. **J**: Decomposed spectrum of variance. Solid lines: Variance captured by the subspaces in recording (black) and simulation (green), sorted by magnitude. Dashed lines: Variance of the remainder space (see Methods 2.4.2).*

We need to ensure that biological parameters are matched consistently across a wide range of simulated experimental parameters. We generate four sets of simulations with different simulation parameters and compare simulated and recorded parameters for each set. For the first set, we generate one simulation of the 7 LV reach-to-grasp task with the matching number of neurons and amplitudes of LV activity for each single recording (Fig. 10). This is the 7-LV single-recording matched set of simulations. For the second set, we average the amplitudes of LV activity across the nine recordings and generate 100 simulations of 100 neurons that match these averaged LV amplitudes (Fig. 11). This is the 7-LV, 100-Neuron set. Finally, we generate both the single-recording matched set and the 100-neuron set for the 10 LV center-out reach task (Figs. 12 and 13).

3.1.4 Single-neuron statistics

To match the distribution of individual firing rates, we sampled individual firing rates for every simulated neuron from recorded firing rates. This allowed us to accurately reproduce the firing rate distribution of the recording in all four sets of simulations (Figs. 10a, 11a, 12a and 13a). The distribution of firing rates against firing rate variances is the same in the recorded data and in all sets of simulations (Figs. 11b, 12b and 13b).
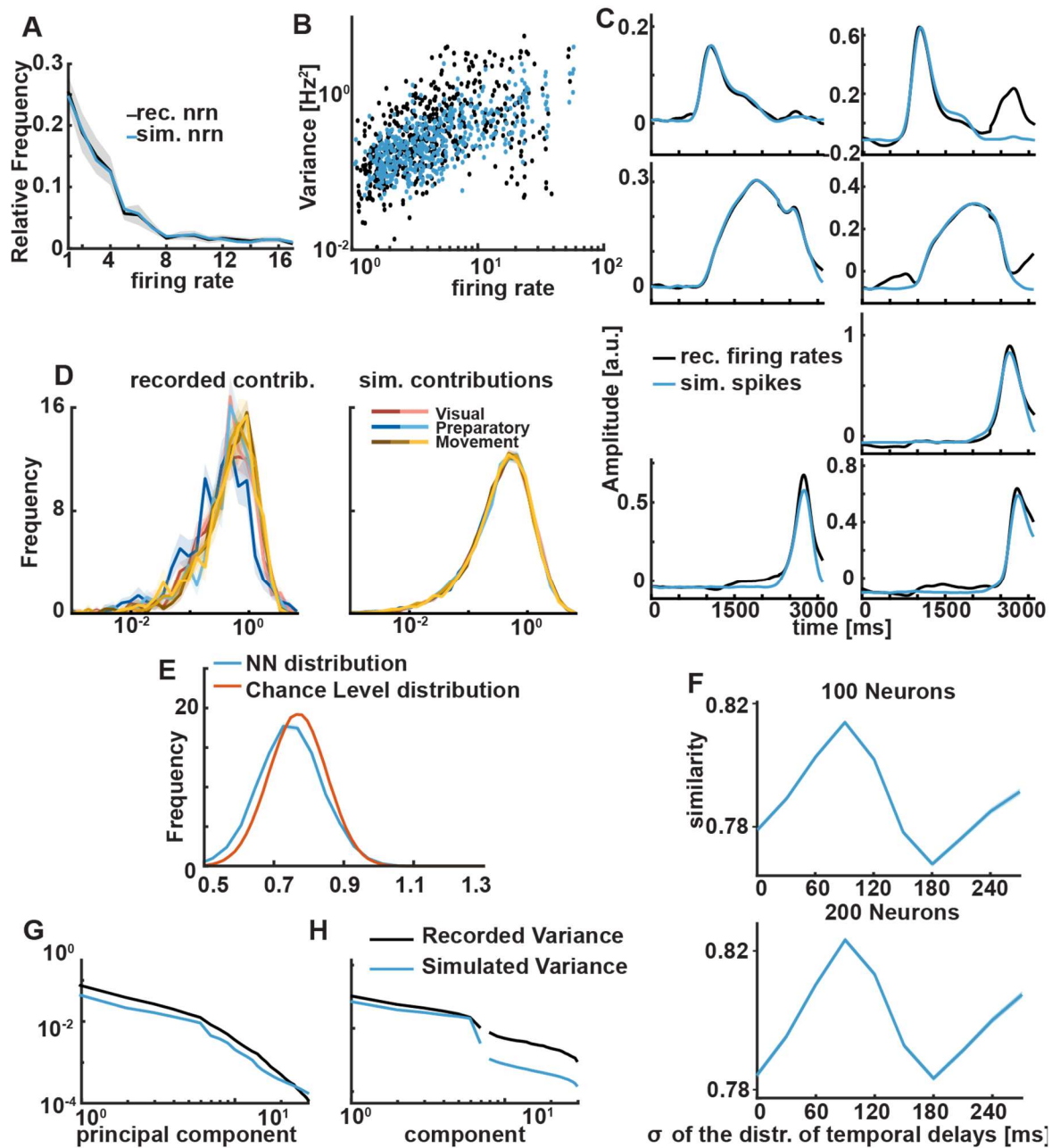
*Figure 11: Same as fig. 10, but the recorded parameters are averaged across all nine recordings. For all panels but F, we generate 100 simulated populations of 100 neurons each and average metrics across populations. All plots with standard error. **A, B, C, D:** as fig. 10 A,B,C,D. **E:** As fig. 10 E right half. **F:** As fig 10 F, except: For each σ, the activity of 20 populations of 100 neurons (top) and 200 Neurons (bottom) each was Procrustes-matched to each of the nine recordings and averaged across 180 pairings. **G,H:** As As fig 10 I,J.*

The contributions of individual neurons to each subspace reflect to what degree each individual neuron is modulated by the activity of each LV, and were also sampled from a recorded distribution. To ensure that dimensions capture independent variance, we orthogonalized them using Gram-Schmidt orthogonalization. Note that the orthogonalization procedure had little to no effect on the individual and average LV contribution distribution because contributions were already close to orthogonal.
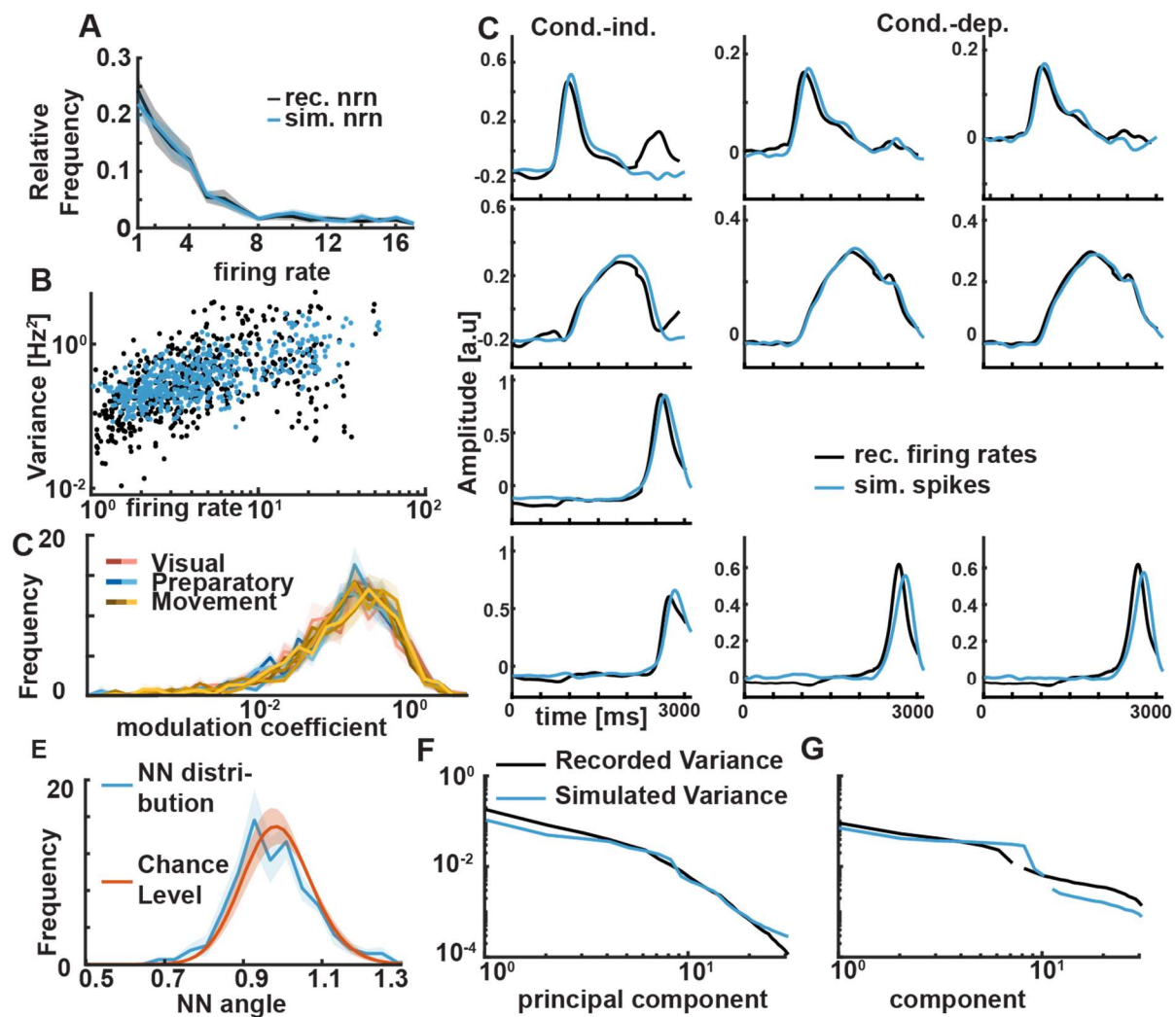
*Figure 12: Same as fig. 10, except we compare parameters of the simulated 10-LV center-out reach task (Fig. 8 D) to the recorded parameters. **A, B, C** as fig. 10 A,B,C. **D,** as fig. 10 D; The amplitudes of condition-dependent components were calculated using eq. X in the methods section X. For each subspace, the two simulated condition-dependent components are plotted against the same recorded component. **E:** As fig. 10 E right half. **F,G:** As fig 10 I,J.*

The task-dependent modulations of individual neuron firing rates correspond to the contributions of that neuron to the subspace activity. To ensure that the dynamics of neural firing rates of the simulation resemble the recording on the population level, it is important to match the distributions of recorded and simulated contributions. We sampled the simulated contributions from a distribution of all recorded contributions to each individual subspace (cf. Eq.1, Methods X.x). The distributions of simulated contributions to each individual subspace resembles the recorded distributions for both tasks.(Figs. 10d, 11d, 12d and 13d).

Our model accurately resembles the statistics of single-neuron contributions to individual latent variables. This does not yet tell whether the contributions are distributed randomly across the neural population, or whether categories of neurons with similar contributions

exist; the randomness on population level will be tested using the PAIRS test in chapter

3.1.6. This does however mean that the magnitudes of LV contributions are distributed
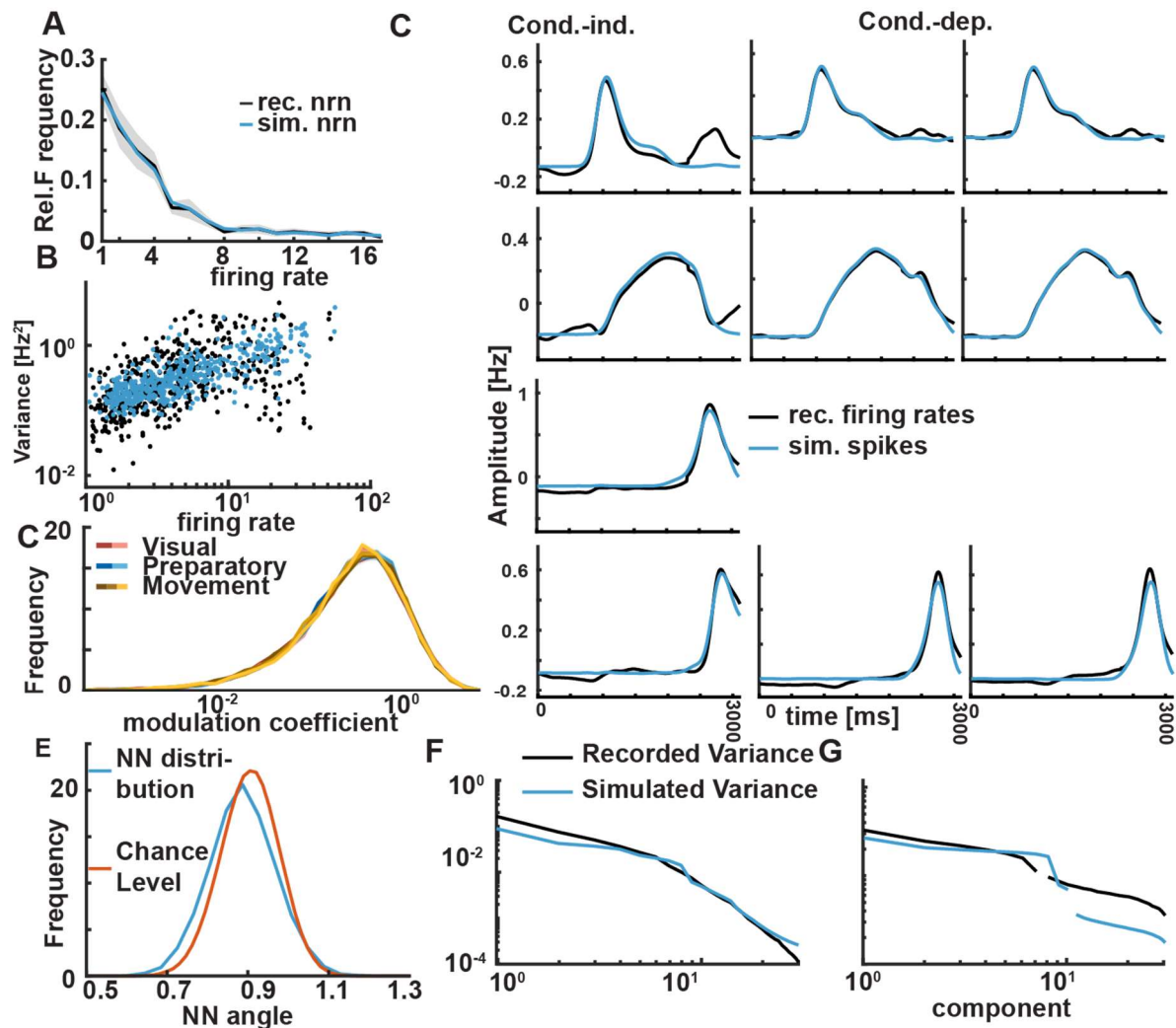
Figure 13: As fig.11, but for the simulated 10-LV center-out reach task. The recorded parameters are averaged across all nine recordings. We generate the response of 100 simulated populations of 100 neurons each to the center-out reach task. Metrics are averaged across 100 simulations. All panels otherwise as fig. 12.

equally in the recorded and simulated neural populations.

We tested whether the individual firing rate dynamics of simulated neurons resembled

recorded neurons; We matched both firing rates and LV contributions of several simulated

neuron to several individual recorded neuron (Fig. 10h); Matching these two parameters

and the latent variable profiles presented in the next section was sufficient to reproduce

recorded single-neuron PSTHs.

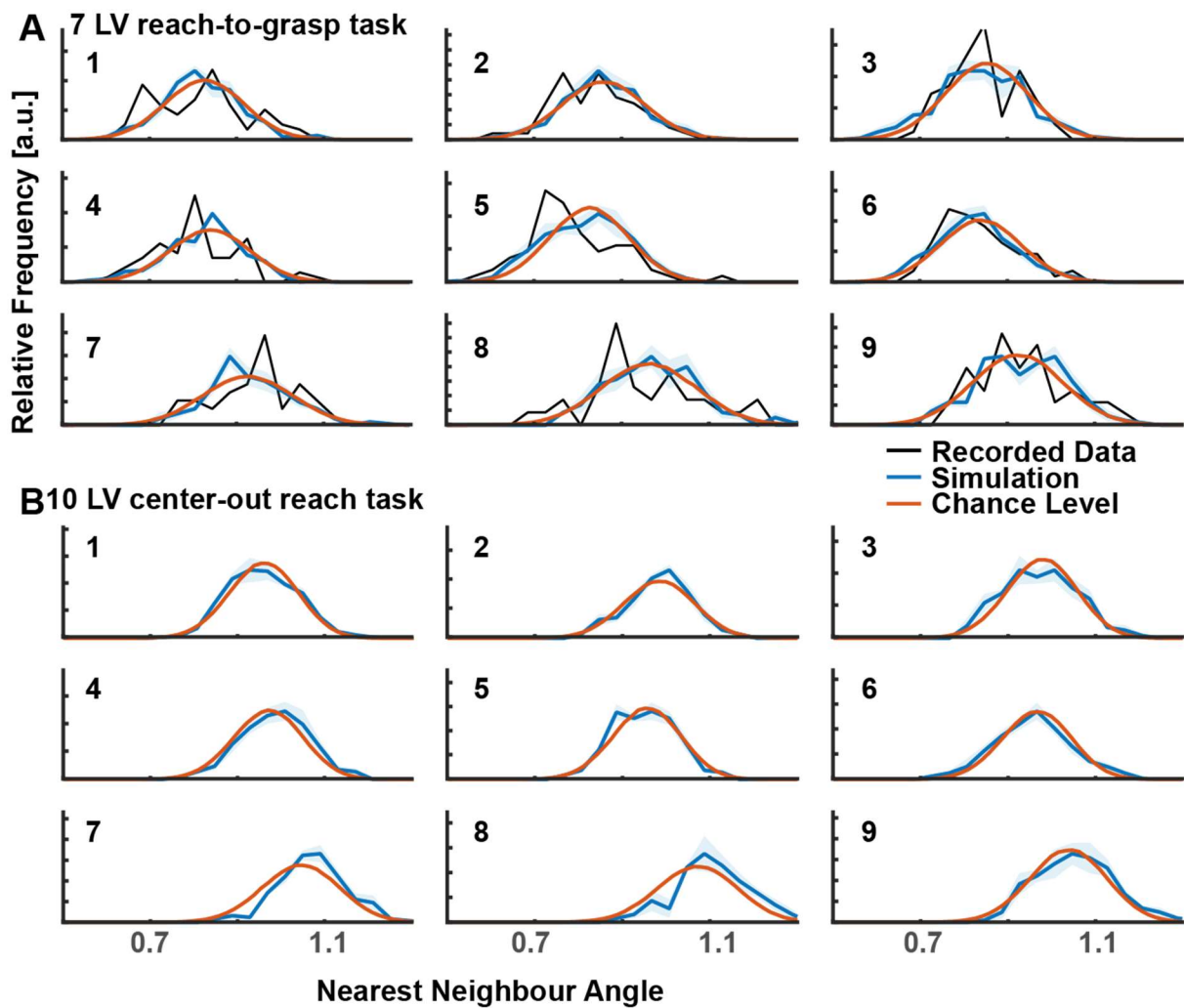*Figure 14: Results of the PAIRS test: Histograms of the median nearest neighbor angle for individual recordings (black) and simulations with the same number of neurons (green) plotted against 10 000 surrogate Histograms (red). Top half: the 7 LV, single-recording matched set of simulations. Bottom: 10 LV, single-recording matched set. Recordings are numbered 1-9.*

3.1.5   Profiles of latent variable activity

The recorded profiles of latent variable activity are projections of the smoothed, trial-averaged spiking activity onto the main axes of the corresponding latent variables (Methods 2.2.4, eq.15). The recorded profiles were deconvolved using Richardson-Lucy deconvolution to balance out the smoothing of single-neuron firing rates and the effect of temporal delay on the profiles of LV activity (Cf. Figure 9, Methods 2.3.2) and used as the LV profiles in the simulation. This was done simultaneously with the selection of the optimal temporal delays, as explained in detail in the paragraph after next. In all recording sets and for each LV, the amplitudes of the profiles of the simulated LVs either match or slightly exceed the profiles of the recording (Figs. 10c, 11c, 12c and 13c). This was done to ensure that the signal-to-noise ratio between variance explained by LVs and by Poisson noise is greater or equal in the simulation, and the conclusions we draw from analyses of the simulation are conservative.

Results

3.1.6 Randomness of latent variable contributions on population level

Latent variable contributions of individual neurons to individual LVs were distributed randomly across the recorded neural population. This was controlled using the PAIRS test (Raposo et al., 2014) (see Methods 2.3.3) We tested whether the same is true in the simulated population. The PAIRS test examines, whether neurons in a population have similar contribution coefficients more often than expected by chance. For each neuron, the contribution coefficients form a vector with a dimensionality equal to the number of LVs and similarity is defined by the angle between a vector and its k nearest neighbors. Displayed in Fig. 10e left in black is the distributions of nearest neighbor angles in the recording, averaged across recordings. The distribution of a set of the same number of random, uniformly distributed vectors, averaged across 10000 repetitions, serves as the chance level distribution (Fig. 10e left, red). The chance level is calculated on randomly distributed vectors; If the distribution of nearest neighbors of the simulated or recorded data is not significantly different from the chance level distribution, then it is also considered randomly distributed.

We apply the PAIRS to each set of simulations (Fig. 10e right, Figs 11e, 12e and 13e). In all four simulated sets and in the recorded data, distributions of nearest neighbor angles showed no significant difference to chance level; the recordings showed a PAIRS index of 0.01 ± 0.03 and a p-value of 0.45 ± 0.32, and the simulation a PAIRS index of -0.02 ± 0.03 with a p-value of 0.27 ± 0.29.

The distributions of nearest neighbor angles for each individual recording, and for each individual simulation from the 7-LV and 10-LV single-recording matched sets, are also not significantly different from chance distributions (Fig. 7). These results suggest that both recorded and simulated neural populations do not contain neurons that are categorical with respect to LV contributions, or alternatively, that LV contributions are distributed randomly across the neural population. This finding is consistent with results reported in the literature (Machens et al., 2010; Raposo et al., 2014).

3.1.7 Influence of temporal delays on population response structure

Changes in the activity of latent variables are not necessarily simultaneously reflected in the individual neurons. In neural populations, the activity of individual neurons can peak earlier

or later than the population average (Westendorff et al., 2010; Michaels et al., 2016). So far, several studies have investigated the effect of such delays on the results of dimensionality reduction methods (Möcks and Joachim, 1986; Yu et al., 2009; Lakshmanan et al., 2015; Kobak et al., 2016); The addition of temporal delays between observed and latent variables caused dimensionality reduction methods to capture additional components that resembled temporal derivatives of the latent variable profiles.

However, it is difficult to estimate the distribution of temporal delays between neurons and LVs in the recorded data. Instead, we estimated which temporal delay results in the highest similarity between simulated and recorded neural population responses. We assumed normally distributed temporal delays between individual neurons and individual LVs with mean of zero and scale parameter σ. σ was systematically variated between 0 and 270 ms for the simulated delayed reach-to-grasp task. To ensure that σ is estimated independently of population size, we estimated the optimal value for σ using several simulation datasets: the 7-LV, single-recording matched set resembling the same numbers of neurons as each individual recording session, the 7-LV, 100-neuron set, and an additional set with 200 neurons. Note that for each value of σ, we re-aligned the LV profiles of the simulation (Fig. 10d, 11d) with the recorded profiles, as described earlier. Next, we used Procrustes analysis to optimally align the simulated and recorded neural populations. Finally, we computed the similarity between pairs of Procrustes-matched dimensions to estimate the value of σ that results in the maximal similarity between simulated and recorded population responses (see Methods 2.3.4). The correlation was highest for a σ of 90 ms for all three simulated datasets: 7-LV, single-recording matched simulations (Fig. 10e) and for 7-LV simulations with a neural population size of 100 and 200 neurons (Fig. 11f). Therefore, we use σ = 90 ms throughout this paper, unless stated otherwise.

To illustrate the effects that temporal delay has on the Procrustes match of simulated and recorded components, we displayed selected Procrustes-matched dimensions of a single recording and the corresponding simulation from the set of 7-LV single-recording matched simulations (Fig. 10f). The simulation was repeated with the same LV contribution coefficients and neural firing rates, for σ of 0 and 90 ms. The profiles of the Procrustes-matched dimensions of the 90 ms simulation resemble the recorded components more than that of the 0 ms case (Fig. 10g). This was particularly pronounced in later components (Fig.

10g left). The effect of temporal delay on the latent variable structure of the simulated population activity is investigated in more detail in the next chapter and in Fig. 16.

3.1.8 Spectra of variance

Many studies have focused on the spectrum of variance of neural population responses such as variances captured per principal components in PCA (Jolliffe, 2002; Stringer et al., 2018, 2019), which corresponds to the eigenvalues of the covariance matrix of the neural population response. We computed the spectra of eigenvalues of the covariance matrix for recorded and simulated neural populations. The shapes of the spectra are very similar for the single-recording matched and the recording-averaged simulated 7-LV datasets (Figs. 10i, 11g). The spectra of the two corresponding simulated 10-LV datasets are still of similar shape (Figs. 12f and 13f). However, as expected, the total variance of the 10-LV datasets is higher.

In order to further compare the simulated with the recorded population activity, we separated the activity projected into the main dimensions of the 7 LVs from the activity outside of these 7 dimensions and studied the individual variance. Because we do not know the ground truth structure of the recorded population response, we can't separate variance within individual dimensions into variance attributable to single LVs and noise. However, we know the main dimension of each individual LV; it is the dimension discovered by applying SVMs-based dimensionality reduction to the recorded dataset (Methods 2.1.2) and the dimension spanned by the LV weight vectors $\vec{\beta}_j$ in the simulation. We project the full, simulated and recorded, population space activity onto the main dimensions of individual LVs (Note, that PCA is not applied in this step, we examine the unmixed projections onto individual LV dimensions). The resulting variance spectra, ordered by magnitude, were highly similar for the two simulated 7-LV and recorded datasets (Figs. 10j, 11h); The spectrum of variance of the main LV dimensions of the simulation resembled the recorded LV dimensions, regardless of the size of the simulated population. The distribution of the
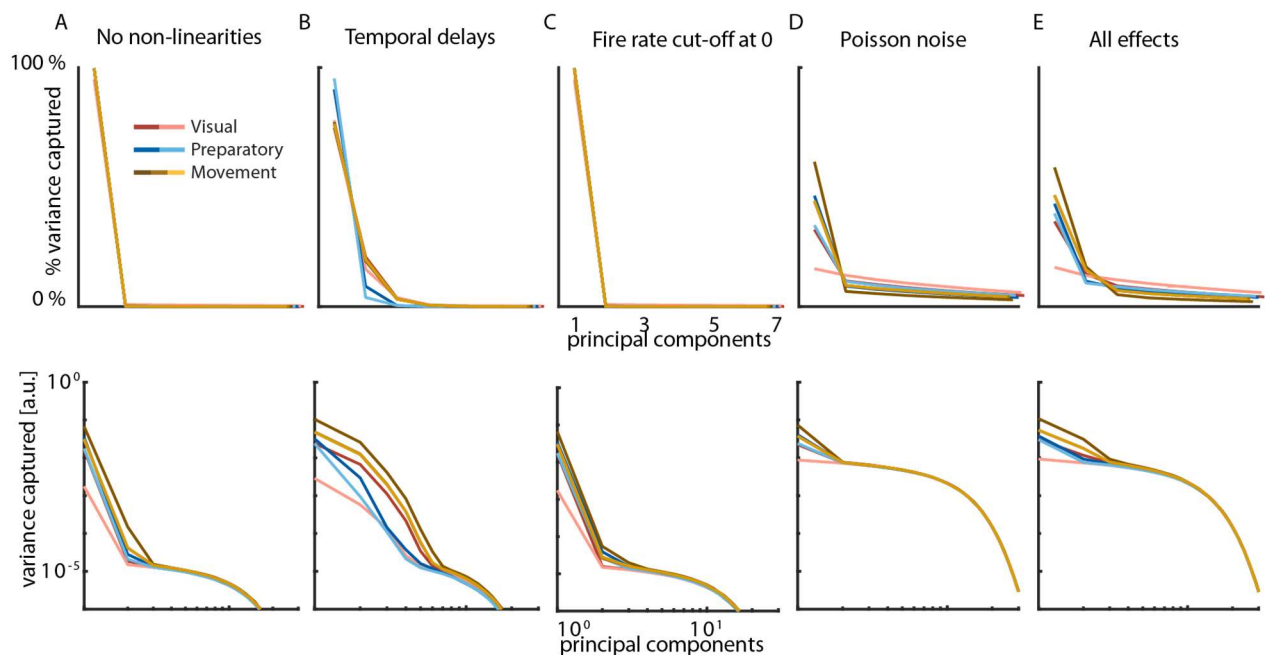
Results

*Figure 16: The impact of individual non-linear effects on the spectra of variance of individual, isolated LVs. For each column, we simulated the activity in response to the 7-LV center-out reach task for 1000 populations of 100 neurons. The activity of individual LVs was isolated for each populations as per eq. X, methods section X. PCA was applied to each population activity and the spectra of variance per component were averaged across populations. Note, that the recovered principal component axes were different for each LV and simulated population; variance per component was averaged across components with the same ordinal number. Top: Linear plot of variance percentages per principal component with standard error. Bottom: Log-log plot of the absolute values of the same variances. **A**: No temporal delays, no cut-off of firing rates at zero and Poisson noise was suppressed by averaging over 10 000 trials: For each LV, variance was almost completely captured by a single component. **B**: The addition of temporal delays with σ = 90ms shifts 5% to 20% of variance out of the first component into the second and later components. **C**: introducing only the cut-off of firing rates at 0 changes the spectra only minimally. **D**: Increasing only Poisson noise by averaging over 20 trials instead of 10 000 results in variance being shifted out of the first PC and distributed across all PCs. **E**: All three nonlinear effects present, the plots are very similar to the plots in D, but with more variance in the second and third component.*

variance along the main LV dimensions of the simulated 10-LV datasets was qualitatively similar (Figs. 12g and 13g).

As a complement to the variance of the main LV dimension described in the last paragraph, we next examine variance outside of the main dimensions. Again, we do not apply PCA, we calculate the variance of random dimensions that are orthogonal to the main dimensions and to each other. The variances of these random dimensions are of the same shape for the recording and for all four simulated datasets (Figs. 10j, 11h, 12g and 13g), but the spectra of the simulations exhibit less variance in total. However, low frequent variations of overall firing rates due to global changes in attention, satiation or other factors were not considered in our model and could explain the differences in variance between simulations and recording. Importantly, the goal of PCA as a dimensionality reduction method is to separate the main LV dimensions from the other dimensions; variance in the main dimensions can be considered the signal, and the variance outside of these dimensions the

noise. The ratio of signal to noise is greater in our simulation than in the recorded data, thus our simulation is conservative in that regard.

We have shown in the last chapter (Results 3.1.7) that the inclusion of temporal delays between latent variables and neurons results in a greater similarity between recorded and simulated population responses to the 7-LV task. Now, we will examine the influence of temporal delays, as well as two other nonlinear effects on the spectra of variance of individual LVs. The set of latent variables is low-dimensional: each LV is a single dimension. When we embed LV activity in the spiking activity of a population of neurons, three non-linear effects shift variance out of the LV subspaces: First, temporal delays between neurons and LVs, second, the cut-off of firing rates at zero and third, the Poisson noise of the spiking activity of individual neurons.

We built five additional sets of simulated responses of a population of 100 neurons to the 7-LV task; One without nonlinearities, one each with only a single nonlinearity and one with all three. Temporal delay was trivial to remove by setting σ = 0, Firing rate cutoff at zero was removed by simulating 'negative' spikes (see methods 2.2.2, eq. 9) and Poisson noise was reduced by simulating 10 000 trials and averaging. We decomposed the spiking population activity into latent variables (Methods 2.4.3, eq.), then applied PCA to obtain variance spectra that correspond to the activity of individual LVs. The spectra are shown in fig. 16.

In absence of nonlinear effects, the variance of each subspace is mostly contained in the first component. (Fig. 16a). Temporal delays reduces the variance of the first principal component, increasing the variance of the second and third components (Fig. 16b), whereas increasing Poisson noise increases variance across all PCs, reducing the relative amplitude of LV-related variance (Fig. 16d). The cutoff of firing rates at 0 results in very little changes to the spectra in comparison to the spectrum without nonlinearities (Fig. 16c). The increase of variance due to Poisson noise across all components can mask the additional components caused by temporal delays; components that are distinguishable from noise in the Poisson-noise free spectrum (Third and later component in Fig. 16b) are not distinguishable anymore in the presence of noise (Fig. 16e). The presence of temporal delays between individual latent variables and neurons results in latent variable activity that occupies additional
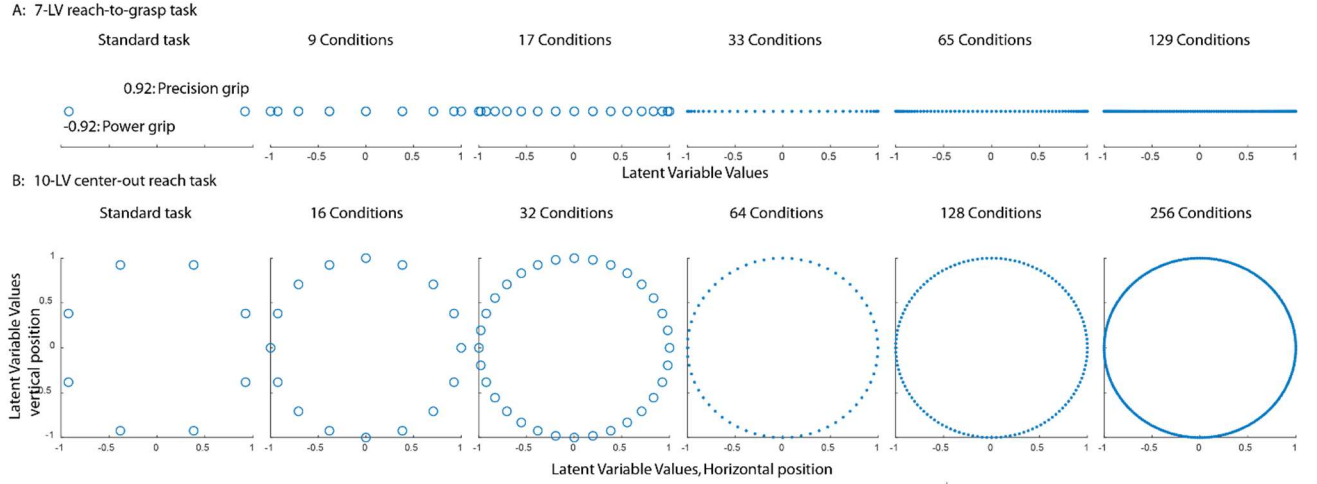
Results

Figure 17: Arrangement of latent variable scaling factors for different conditions in the two tasks. Each circle or dot corresponds to one condition, and the values shown are multiplied onto the profiles shown in fig. 10D to scale the LV profiles for that condition. **A**: The 7-LV reach-to-grasp task. For each condition, corresponding to a single circle or dot, each of the three condition-dependent LV profiles is multiplied with the value on the x-axis. Leftmost: The two values for power and precision grip are -0.92 and 0.92. This values were chosen so that the average of the absolute value of all conditions is 0.92, no matter how many conditions are simulated. Other panels: additional conditions are distributed sinusoidally between and around the initial two. **B**: For the center-out reach task, condition scaling factors are distributed circularly, with one condition-dependent latent variable per epoch being scaled with the X-value and the other with the Y-value. See also methods 2.2.1.

dimensions of the space of population activity; this effect is more visible if the noise of the spiking Poisson process is reduced.

In summary, despite not simulating individual connections between neurons, our simulation can accurately recreate important population features of a real recording. For the evaluation of the applicability of dimensionality reduction methods to recordings of neural population activity, the high similarity between the simulated and recorded population responses suggests that the results obtained by applying dimensionality reduction to the simulated data also relevant for the recorded data.

3.2 Application of PCA to Simulated population response

In this section, we use the simulation of neural population response to evaluate the performance of PCA as a dimensionality reduction method when applied to recordings of neural population activity. The simulation methods are described in chapter 2.2.

As our basic set of simulation parameters, we analyze a simulated population of 100 neurons that responds to the 7-LV reach-to-grasp task, with the response averaged over 20 trials for each condition. The smoothed, trial-averaged population response is represented

as a matrix $N \in \mathbb{R}^{ct \times n}$, where c are conditions, n are neurons and t are timepoints; Different conditions are concatenated behind each other. We apply PCA to N. We repeat this process 100 times and calculate the mean and standard error of each metric presented in this chapter.

We systematically varied individual parameters of this baseline simulation to show the effect of the parameters on the result of PCA: The first systematic variation is the simulation of the 10-LV center-out reach task. Second, for either task, the number of trials per condition is varied from 1 to 10 000. Next, for either task and 20 trials per condition, we vary the number of neurons in the population from 100 to 1000. Finally, we also model additional conditions; For the 2-condition reach-to-grasp task, the condition-dependent LVs assume sinusoidally distributed values between, and slightly exceeding, the two values of the two-condition case. For the center-out reach task, the addition conditions are arranged between the existing conditions (see Methods 2.2.1 eq. 1). See fig. 17 for an illustration.

Results

3.2.1   Estimation of Dimensionality from the Spectrum of Variance

Various methods have been developed to estimate the underlying dimensionality of neural

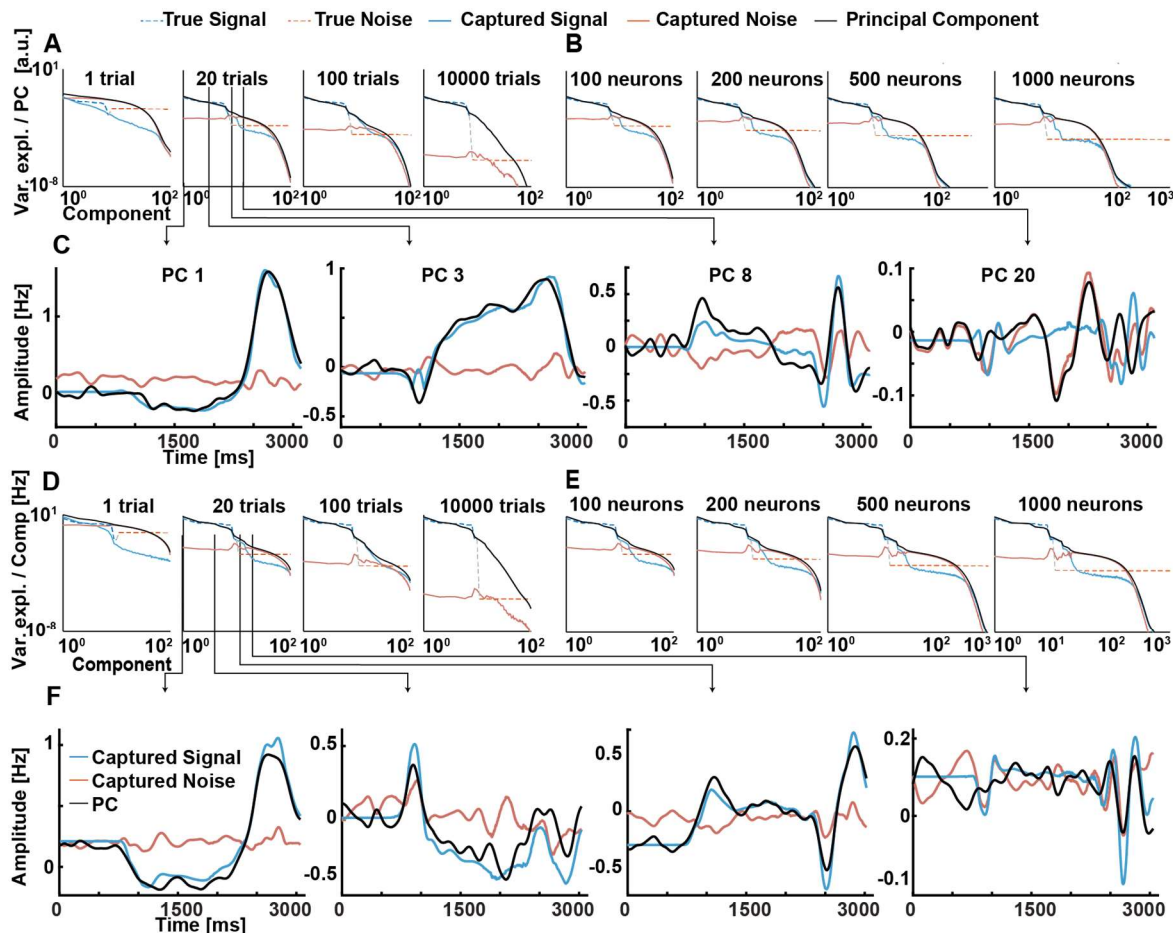population activity during performed tasks. For PCA, methods such as scree plot (Cattell,

*Figure 18: **A**: Spectra of variance of the response of a simulated population of 100 neurons to the 7-LV reach-to-grasp task. Gray: PCA spectra. Blue and orange: Decomposition of individual principal components into LV-related and noise-related variance (methods 2.4.3, eq. 63). Dotted lines: True LV and noise variances (eqs. 60 and 53). **B**: Spectrum of variance for varying population sizes, 7-LV task with 20 trials per condition. **C**: Time profile of selected principal components. Leftmost: While the profile of this PC is similar to the true profile of the trigger LV, it is of much higher amplitude. Fig. 20 shows that this PC captures a mixture of the trigger and the condition-independent movement component. Second from left: This PC captures a derivative-like component at 1500 ms. Other plots: further principal components increasingly capture noise, but still contain LV-related activity. **D,E,F**: same as A,B,C for the 10-LV center-out reach task. All plots averaged across 100 repetitions, shown with standard error.*

1966) and parallel analysis (Horn, 1965) have been developed, that use the spectrum of

variance to estimate dimensionality. This spectrum is the spectrum of variance captured by

individual principal components, sorted in descending order (Jolliffe, 2002).

We plot the spectrum of variance of the simulated delayed reach-to-grasp and center-out

reach tasks while varying the number of trials per condition and the number of simulated

neurons. Displayed in Fig 18, in grey is the variance of each PC against the ordinal number of

the PC (Methods 2.4.2, eq. 37). To study, whether LV-related variance and noise are mixed or separated in the spectrum of variance, we separate the variance into two portions: the variance attributable to the activity of all LVs (blue, see methods 2.4.3 eq. 44) and the Poisson noise (purple). We see for all sets of parameters a smooth transition from early components that are dominated by LV-activity over components that contain mixtures of LV-related variance and noise, to noise-dominated components.

As a comparison to the variance of the projections of LV-related activity onto individual PCs, we also plot the true variances of individual latent variables (Fig. 18a, b, d and e, blue dashed line, see methods section 0, eq.). Unlike the projections onto individual PCs described in the last paragraph, or the projections of population activity onto the main dimensions of LV-related activity described in section 3.1.5, these are not variances that are calculated along a single dimension. Rather, these variances are distributed across the entire space of neural population activity. The LVs are ordered by magnitude. Similarly, we plot the spectrum of the Poisson noise, obtained by simulating no LV activity and projecting the simulated neural activity onto random orthogonal axes of the population space. The projection of the noise onto random axes prevents random noise correlations from accumulating in the early components of the spectrum, as it does with PCA, The variance captured by these axes was, again, ordered by magnitude (Fig. 18a, b, d and e, red dashed line). The plots show, that for 20 trials per condition and above, the first seven PCs resemble the true spectrum of LV variances. After 7 PCs, the spectrum obtained from the results of PCA exhibits a smooth transition into a bump in the spectrum of noise-related variance. This bump is not present in the true spectrum of noise-related variance and can be attributed to random noise correlations.

The number of trials recordable during experiments is limited, resulting in different levels of Poisson noise for different numbers of trials per condition. Additionally, we would like to see if increasing trials per condition to unrealistically high numbers can improve the estimation of dimensionality. Therefore, we vary the number of trials per condition between 1 and 10 000 (Fig. 18a and d). Additionally, we test if varying the number of neurons in the simulated population (Fig. 18b and e) or task conditions (Fig 24a) has an effect on the spectrum of variance. As expected, the portion of variance attributable to Poisson noise is reduced with rising numbers of trial per condition (Fig. 18a and d). For reasonable numbers

of trials per condition (around 20), the first components of both the spectrum of total variance and the part attributable to LV activity closely reflect the true spectrum of variance per LV. However, there is no clear drop-off or "knee" between components that capture predominantly LV-related variance and components dominated by noise, as is described in the literature (Cattell, 1966). Instead, we see a smooth transition from LV-related variance to noise that does not depend on the number of trials per condition (Fig. 18a and d), the number of simulated neurons (Fig. 18b and e) or the number of task conditions (Fig. 24a). Thus, it is impossible to estimate the dimensionality of the underlying neural population activity from a drop-off in the spectrum of covariance eigenvalues.

For both tasks, the simulation of larger numbers of neurons has no effect on the transition between LV-related variance and noise (Fig. 18b and e). With increasing numbers of simulated neurons, a bump in the spectrum of noise-related variance becomes more pronounced. One possible explanation for this bump are random noise correlations between neurons, which increase with increasing numbers of neurons; Because such correlations increase the variance of the components that contain them, PCA aggregates the variance of such random correlations at the beginning of the spectrum, after LV-related variance.

To illustrate the transition from LV-related variance to noise, we examine a few individual example PCs (Fig. 18c and f). The profiles of individual PCs transition from what appear to be an isolated LV in the first PC, to a mixture of different LVs, to mixtures of LV and noise to, finally, completely noise-dominated components. While the first PC in each example appears to be a single LV from the movement epoch (Fig. 18c and f, leftmost panels), the amplitude of the component exceeds the highest amplitude among individual LVs, plotted in dashed blue in the same panel. A more detailed examination (Fig. 20c) reveals that this PC contains mixtures of the trigger and condition-independent movement components.
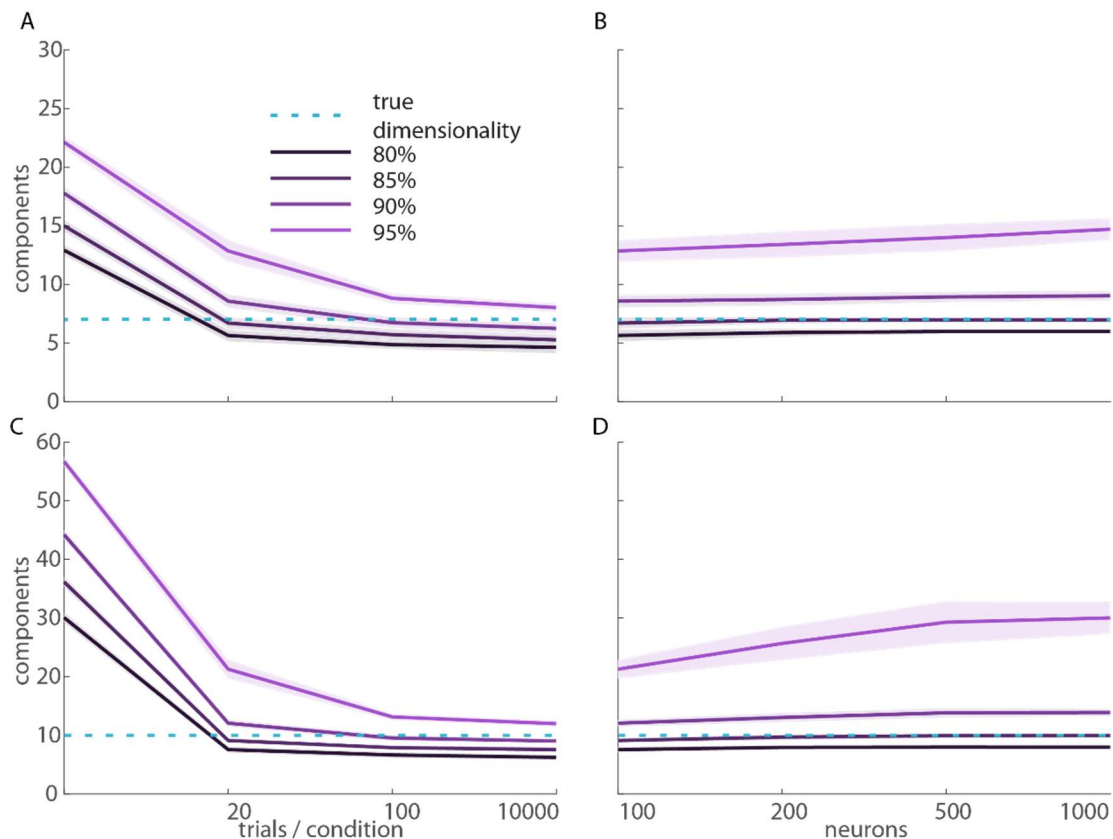
*Figure 19: Plots of the number of principal components necessary to capture a percentage of total variance. **A**: Number of Principal components necessary to capture 80% 85% 90% 95% of the total variance of the simulated neural population response to the 7-LV reach-to-grasp task, plotted against the number of trials per condition. Population size was 100 neurons. **B**: Same as A, plotted against varying numbers of neurons. 20 trials per condition were simulated. **C, D**: Same as A and B for the 10-LV center-out reach task. All plots averaged across 100 repetitions, shown with standard error.*

A frequently used criterion for the estimation of the dimensionality is the number of PCs needed to capture a certain percentage of variance. (Jolliffe, 2002; Stringer et al., 2019) In fig. 19, we plot the number of PCs needed to capture different percentages of total variance against varying numbers of trials per condition and simulated neurons. For all percentages and for both tasks, the number of PCs decreases with increasing numbers of trials per condition and decreases slightly less with increasing numbers of neurons. For large numbers of trials per condition at 100 neurons, the 90% criterion is the closest to the true number of latent variables, while for > 200 neurons at 20 trials per condition the 85% criterion is the closest. None of the criteria is reliable across the entire range of parameters and the true dimensionality of the set of latent variables can't be reliably estimated from these criteria.

Taken together, it is difficult to estimate the dimensionality of the set of latent variables from the spectrum of variance of a spiking, neural recording. The presence of temporal delays between latent variables and individual neurons and the Poisson noise of the spiking

process cause a smooth transition between components that capture predominantly LV-related variance and components that capture noise, and this smooth transition is present regardless of the number of trials per condition, number of neurons in the simulated population or number of task conditions.

3.2.2 Extraction of Dimensions of Task-Related Activity

The goal of PCA as a dimensionality reduction method is to separate the manifold spanned by the main LV dimensions, which contain the most LV-related variance (Fig. 10d, 16), from
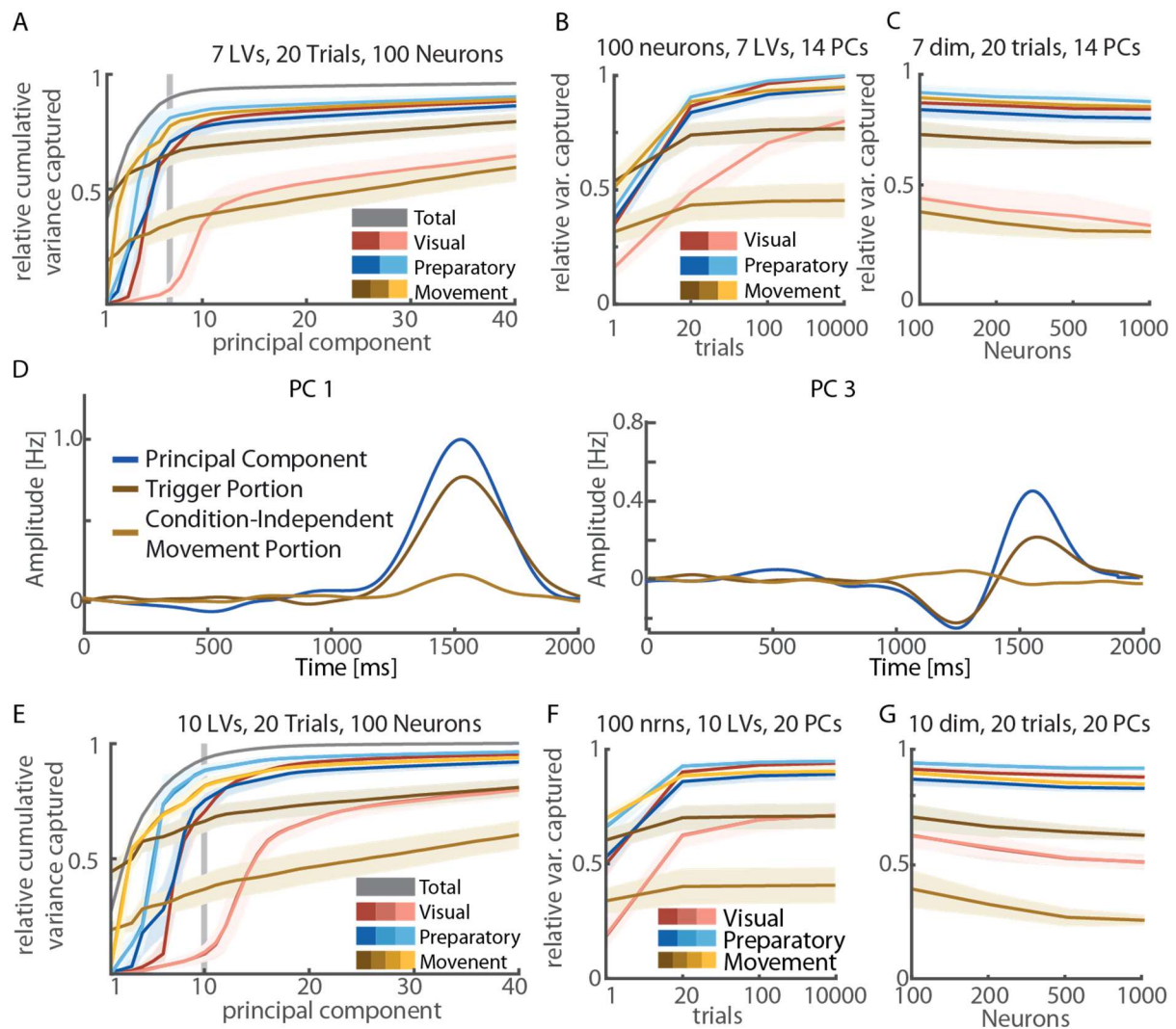

*Figure 20: **A**: The relative proportion of variance of the simulated population response of 100 neurons to the 7-LV reach-to-grasp task captured by the first N principal components, plotted against N. In grey, the total variance, in color the variance of individual LVs. The vertical grey line is the true number of LVs. **B**: The proportion of variance captured by the first 14 PCs (Twice the number of LVs) for varying numbers of trials per condition. **C**: As B, for varying numbers of neurons. **D**: Decomposition of an example PC from a simulation of the 7-LV task, 20 trials per condition, 100 neurons, for a single condition. Blue: Projection of the full activity onto a single PC. Dark brown and brown: Projections of the portion of neural activity attributable to the movement trigger and the condition-independent movement components. These two components contributed the most variance to PC 1 and 3. **E, F, G**: As A, B, C, for the 10 LV center-out reach task. F and G display the proportion of variance captured by the first 20 PCs. All plots averaged across 100 repetitions, shown with standard error.*

the other dimensions. Even if the first few PCs can capture a large percentage of the variance of the manifold, they do not necessarily capture all subspaces equally. To investigate this possibility, we decomposed the neural population activity into individual LV activities (methods section 2.4.3, eq.). Next, we projected the individual LV activities onto individual principal components and calculated the variance captured. We systematically varied the number of considered principal components, trials per condition, conditions and simulated neurons to test, how the percentage of cumulative variance captured by the first condition changes.

For every individual LV, as well as for the overall LV-related variance, we estimated the proportion of variance captured by the first N PCs against N (Fig. 20). We first explored the effect of considering different numbers of principal components based on our default simulation parameters (100 neurons and averaged each condition across 20 trials). For the simulated reach-to-grasp task, the population response was structured by seven LVs and the equal number of PCs captured more than 90% of the total LV-related variance (Fig. 20a). Furthermore, for all but one of the LVs, the percentage of variance captured was saturated and did increase only marginally with additional PCs. However, variance captured by the visual condition-dependent LV (bright red) started to rise only at 10 PCs and variance captured by both the visual condition-dependent LV and the condition-independent movement component (light brown) increased only slowly up to only slightly above 50% even when considering 40 PCs. For the simulated center-out reach task, the results are qualitatively similar (Fig. 20e). Ten PCs captured >90% total variance, but variance captured by both condition-dependent visual LVs started to rise only at around 15 PCs. In summary, to capture a sufficient amount of variance from all LVs, we needed between 1.5 and 2 principal components for each LV in these tasks.

To explore the effect of number of trials on the cumulative variance captured per LV, we held the number of considered PC constant at twice the number of LVs (14 for the reach-to-grasp and 20 for the center-out reach task). We decided to use the conservative number of twice the number of LVs because at that number a substantial amount of variance is captured from all LVs default simulation parameters and only little additional variance is captured when considering additional components. For most LVs, variance captured was low for only 1 trial, but increased sharply to over 80% at 20 trials, then increased only slowly

until up to the unrealistic number of 10000 trials, for both tasks. The only exceptions were the visual condition-dependent LV for the reach-to-grasp task, for which more variance was captured with increasing number of trails, and the condition-independent movement LV; No more than 50% of variance was captured at any number of trails for both tasks (Fig. 20b and f) for the latter. To test the effect of the number of neurons on the amount of variance captured per LV, we additionally fixed the number of trials per condition at 20, and systematically varied the number of simulated neurons. Surprisingly, the number of simulated neurons had no observable effect on the amount of variance captured for any LV (Fig. 20c and g). These results suggest that even when recording large numbers of trials and neurons, a significant amount of variance of task-related LVs is lost during dimensionality reduction with PCA.

From Fig. 20 it is apparent that the visual condition-dependent and movement condition-independent LVs behave differently from the other components and from each other. The behavior of the visual condition-dependent LVs can be explained by the fact that this LV is of smaller amplitude and thus explains less variance than the other LVs (see Fig. 10 d and j). Therefore, PCA could not separate the activity of these LVs from noise. The movement condition-independent LV on the other hand is one of the largest LVs. A possible cause for the poor capture of this LV is, that the temporal activity profile of the movement condition-independent LV is overlapping with the profile of the trigger LV, which is of even greater amplitude (Fig. 20c). PCA possibly rotates preferentially on the larger component and fails to capture the components smaller.

Finally, we held the number of neurons and principal components fixed and varied the number of conditions for both tasks between 4 and 256 (Fig. 24b). The percentage of variance captured by the first 14 resp. 20 PCs changed very little with increasing numbers of conditions.

Taken together, these results suggest that even when considering many PCs, or large amounts of variance captured, significant amounts of the variance of individual LVs can be lost. Thus, the activity occupied by the reduced space spanned by PCs may be a biased or incomplete reflection of the task-related population response.

3.2.3 Recovery of unmixed LV components

Finally, we examined whether and to what degree individual PCs correspond to the ground truth LVs and subspaces. If PCs are aligned with the ground truth LVs or subspaces, PCA would be a useful method to investigate the population response structure of a neural population. In contrast, if PCs are rather random mixtures of the ground truth LVs or
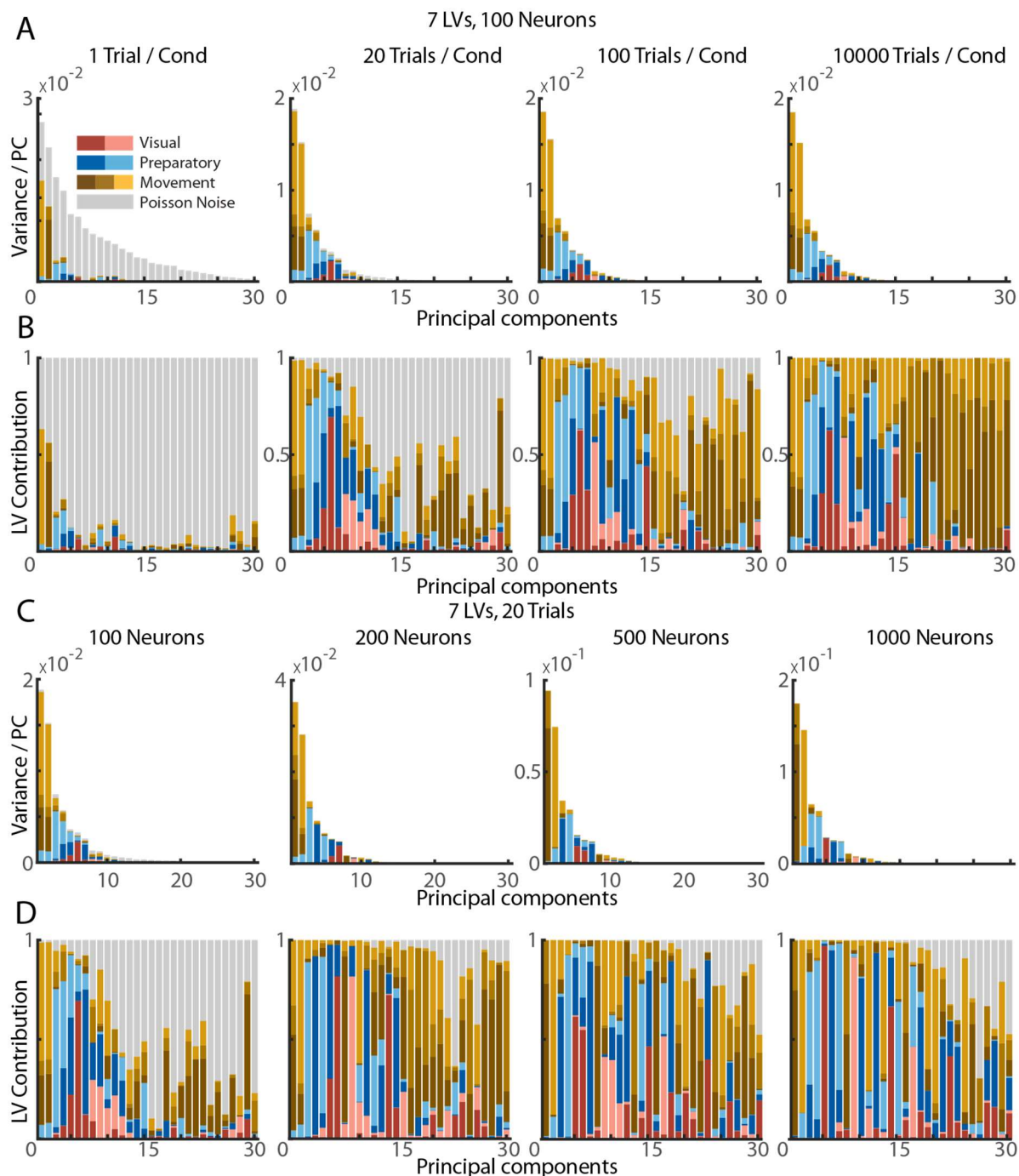
*Figure 21: Decomposition of the full variance captured by each PC into portions attributable to each individual LV. Each plot shows the decomposition of the variance of a single simulated population response. Decomposition was performed according to eq. 70 in the methods section 2.2.4. **A**: Variance decomposition of the response of a 100 neuron-population to the 7 LV center-out reach task for each PC plotted against PC number. **B**:as a, but each column normalized to height 1. **C, D**: as A, B, but for varying numbers of neurons*

subspaces, no direct conclusions should be drawn from individual or combinations of PCs and only the combined space of many PCs should be considered with the aforementioned restrictions.

Figure 22: as fig. 21, for the 10-LV center-out reach task.

Previous work has shown that the projection of simulated event related potentials onto single PCs contained mixtures of the activity of single LVs, suggesting PCA as a non-ideal method for LV reconstruction (Wood and McCarthy, 1984; Beauducel and Debener, 2003).
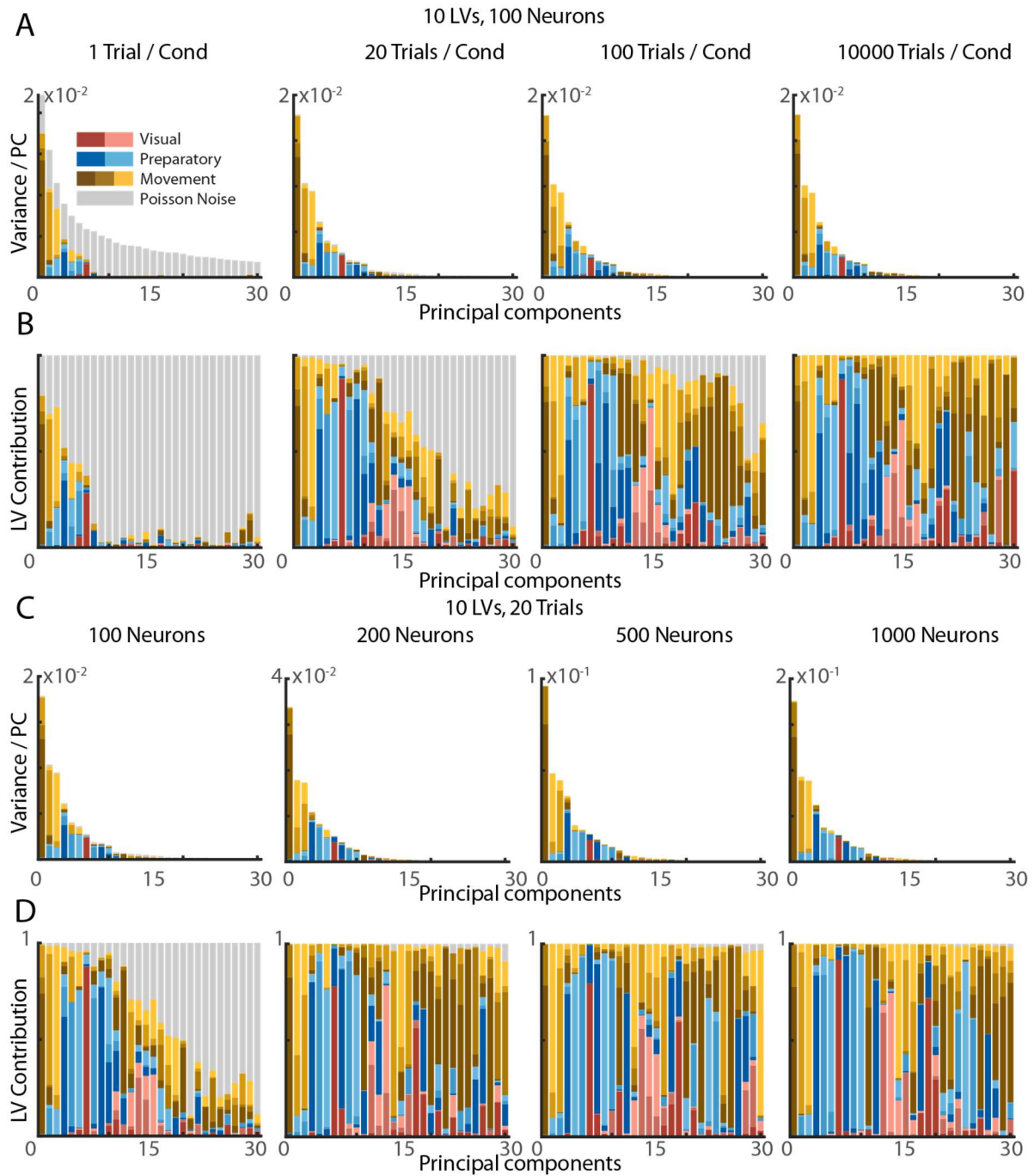
However, to which degree PCA mixes underlying LVs and subspaces of neural population activity has not been studied so far. Especially, the effect of Poisson noise on LV reconstruction remains elusive. We investigate to what degree individual PCs exhibit mixing and how mixing changes if we vary the number of recorded trials, neurons and task conditions.

For all following analyses we used the same simulated datasets as in the previous section (Figs 18, 20 and 24a and b). First, to explore the effect of different numbers of trials on LV mixing, we used the simulated datasets of a population of 100 neurons for both tasks and applied PCA to the response. We decomposed the captured variance per PCs into the individual contributions of each individual LV and Poisson noise (See Methods 0). Displayed in Fig. 21a and b is the decomposition of the first 30 PCs of an example population for varying numbers of trials per condition and for both tasks. For visibility, we displayed both the absolute values of variance per LV and PC, as well as the relative variance normalized by the total variance per PC. For both tasks and all numbers or trials all PCs captured mixtures of the ground truth LVs as well as subspaces. With increasing number of trials per condition, the variance attributable to noise was reduced, but with no effect on the level of LV and subspace mixing.

To quantify these observations, we first calculated the signal-to-noise ratio (nSNR; defined as the ratio of variance attributable to noise to the total variance of the PC) across all 100 simulations of the datasets. As expected, the nSNR rose with greater numbers of trials per condition, but it rose faster in earlier PCs (Fig. 23a). Next, to test whether individual LVs are predominately captured by individual PCs, we computed the ratio of the largest LV contribution relative to all LV contributions across all 100 simulated populations (Fig. 21b and 22b). Apart from a slight predominance of individual LVs for the early PCs, the LV-mixing per PC was at chance level. The chance level of mixing is calculated by drawing random vectors in the space of neural population activity and calculating the ratio of the largest component to the sum of all components. Different numbers of trials had only a small effect on LV-mixing and LV-mixing changes very little for more than 20 trials per condition.

Next, we examined the effect of different numbers of neurons on LV mixing. For this purpose, we varied the population size between 100 and 1000 neurons. As expected, the

absolute value of the variance captured by each PC increased approximately linearly with the number of neurons (Figs. 21c and 22c). Yet, The nSNR does not increase with larger number of neurons (Fig. 23c). For some PCs, the nSNR even decreased with larger population sizes (PC Nr.15 and greater in the 7LV case and PC Nr.25 and greater in the 10-LV case).
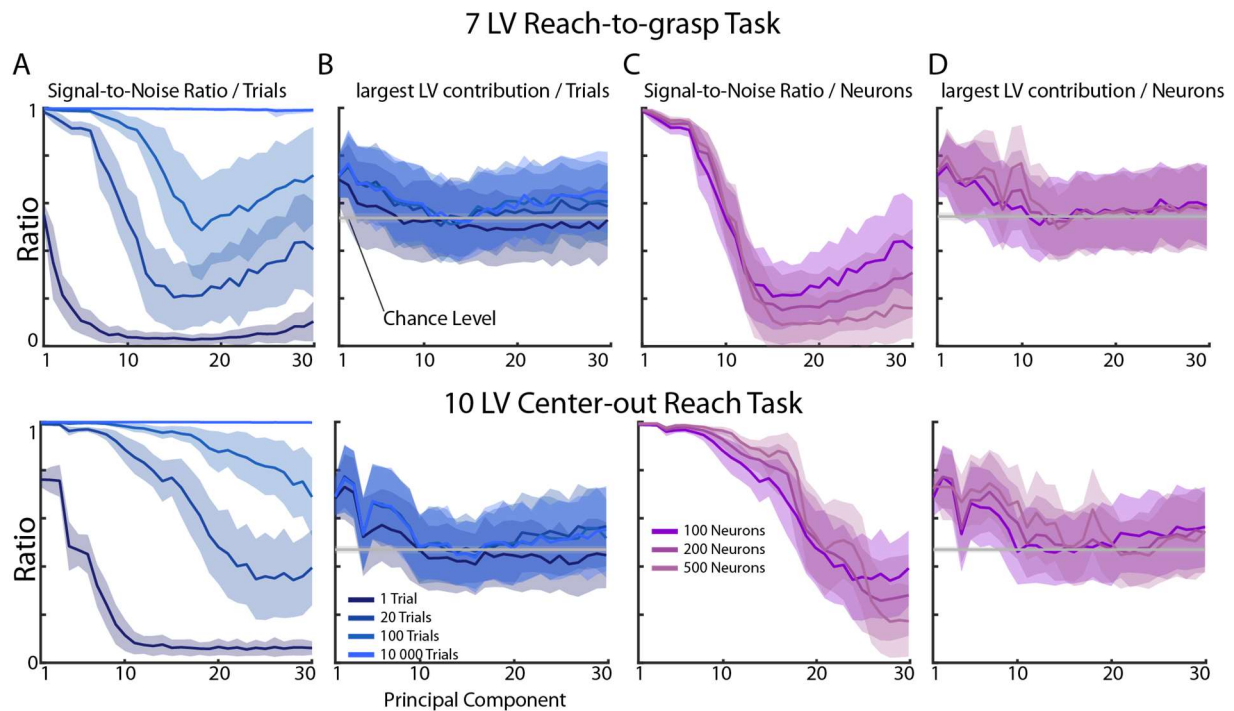

*Figure 23: Key values of fig. 21, averaged across 100 repetitions with standard error. **A**: The signal-to-noise ratio, here defined as the ratio of LV-related variance to total variance, plotted against PCs for different numbers of trials per condition. Top: for the 7-LV task, bottom for the 10-LV task. **B**: the ratio of the largest LV contribution, equivalent to the largest colored bar in the plots in fig. 21, to the other LV contributions. **C, D**: Same as A, B, but plotted for different numbers of neurons in the simulated population.*

The relative value of variance captured by individual PCs (Figs. 21d and 22d) show, that mixing between LVs does not qualitatively change for different population sizes and the proportion of noise contributions changes only marginally. To quantify this observation, we again calculated the ratio of the largest LV contribution relative to all LV contributions across all 100 simulated populations (Fig. 23d). The ratio was independent of the number of neurons.

Finally, we examined the effect of different number of conditions for both tasks (Fig. 24c and d), for a population size of 100 and 20 trials per each individual condition fixed. When examining the response of a single simulated neural population to either task, the mixing between individual LVs and the mixing between LV activity and noise does not qualitatively

change with varying number of trials per condition. The average SNR (Fig. 24c left) changes only marginally with the number of conditions for either task and the largest LV contribution ratio (Fig. 24c right) does not change at all.

In conclusion, individual PCs are essentially random mixtures of latent variables and noise. While the contribution of noise to individual PCs could be reduced by increasing the number of trials per condition, PCs contained mixtures of individual latent variables across all numbers of trials per condition, neurons and conditions. Therefore, individual PCs, or even combinations of PCs should not be interpreted as representations of latent variables. Given the results of previous subchapters, only the complete, combined space of a large number of principal components should ever be used to draw conclusions about the structure of the neural population response.
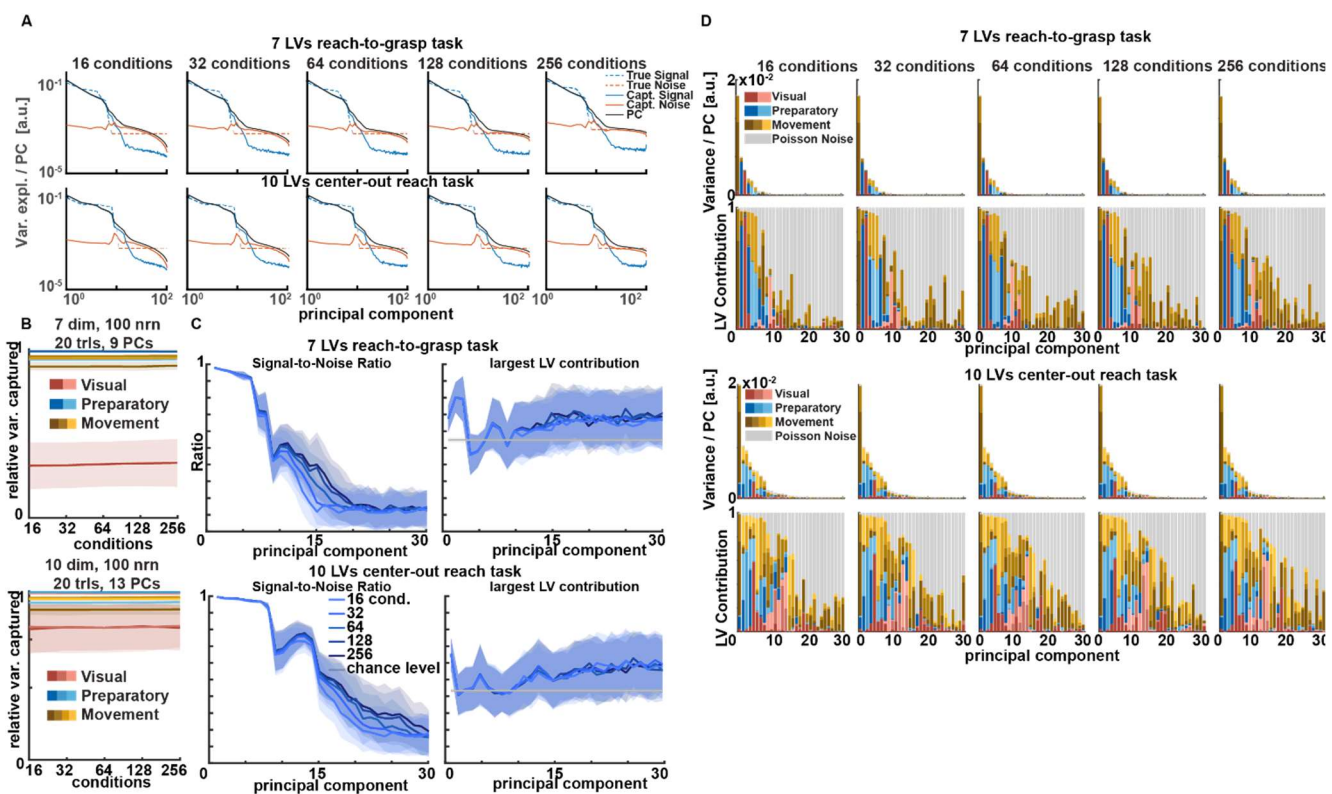
Fig. 24: **A**: *as fig. 18, for varying number of conditions.* **B**: *Same, for fig. 20.* **C**: *Same, for fig. 23.* **D**: *Same, for figs. 21 and 22.*

3.2.4 Theory of principal component mixing

In the previous section, we showed that Principal components contain mixtures of latent variables and noise. We can describe mixing mathematically, using the definition of principal components as eigenvectors of the covariance matrix of the dataset. We assume a two-

dimensional dataset with two orthogonal, uncorrelated latent variables, with a 2x2 diagonal covariance matrix V and variances λ_1 and λ_2: $\lambda_1 > \lambda_2$.

$$V = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \qquad (X)$$

The eigenvectors of V, and thus principal components of the dataset are $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$; the latent variables are unmixed. Now, we introduce a covariance term to the matrix:

$$V' = \begin{bmatrix} \lambda_1 & \varepsilon_{12} \\ \varepsilon_{12} & \lambda_2 \end{bmatrix} \qquad (X)$$

ε_{12} is the covariance between the temporal profiles of the two latent variables. In our model of neural population activity, ε_{12} encompasses both the covariance between the temporal profiles of individual LVs and the random correlations due to noise. We can calculate the new eigenvalues λ'_1 and λ'_2 using the general formula for eigenvalues of a 2x2 matrix:

$$\lambda'_{1,2} = \frac{\lambda_1 + \lambda_2}{2} \pm \sqrt{\frac{(\lambda_1 - \lambda_2)^2}{4} + \varepsilon_{12}{}^2} \qquad (X)$$

And the corresponding eigenvectors $\vec{v}_{1,2}$:

$$\vec{v}_1 = \begin{bmatrix} \lambda_2 - \lambda'_1 \\ -\varepsilon_{12} \end{bmatrix} = \begin{bmatrix} \dfrac{\lambda_2 - \lambda_1}{2} - \sqrt{\dfrac{(\lambda_1 - \lambda_2)^2}{4} + \varepsilon_{12}{}^2} \\ -\varepsilon_{12} \end{bmatrix} \qquad (X)$$

$$\vec{v}_2 = \begin{bmatrix} -\varepsilon_{12} \\ \lambda_1 - \lambda'_2 \end{bmatrix} = \begin{bmatrix} -\varepsilon_{12} \\ \dfrac{\lambda_1 - \lambda_2}{2} + \sqrt{\dfrac{(\lambda_1 - \lambda_2)^2}{4} + \varepsilon_{12}{}^2} \end{bmatrix}$$

Note, that $\vec{v}_{1,2}$ are not unit vectors. It is easy to control that for $\varepsilon_{12} \to 0$, $\frac{\vec{v}_{1,2}}{|\vec{v}_{1,2}|} \to \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\lambda'_{1,2} = \lambda_{1,2}$; in absence of correlations between LVs, the eigenvectors and eigenvalues are unchanged. If, on the other hand, $\varepsilon_{12} \gg \lambda_1 - \lambda_2$, $\vec{v}_{1,2} \approx -\varepsilon_{12} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \varepsilon_{12} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. The covariance between the latent variables drives mixing.

An important observation is that the new eigenvectors $\vec{v}_{1,2}$ depend only on the difference of original eigenvalues $\lambda_1 - \lambda_2$, not on the amplitudes of the eigenvalues. Therefore, latent variables that explain similar variances mix more readily.

In conclusion, mixing between latent variables is driven by the covariance between the temporal profiles of the LVs, and is independent of the actual variances captured by individual LVs; Only the difference between variances influences mixing, in that latent variables with similar variances mix more easily.

4 Discussion

4.1 Summary

We have presented a highly accurate model of the spiking population responses during two standard behavioral tasks in monkey motor neurophysiology. Based on this model, we evaluated the performance and accuracy of the most commonly used dimensionality reduction method PCA (Jolliffe, 2002; Cisek, 2006; Cunningham and Yu, 2014; Stringer et al., 2018, 2019), which was missing so far. The simulated population activity occupied a low-dimensional manifold that consisted of orthogonal latent variables. The LVs could be grouped into behavior related subspaces for visual, movement preparatory, and movement execution related processing, in agreement with previous results on recorded neural populations (Elsayed et al., 2016; Murray et al., 2017; Lara et al., 2018; Stringer et al., 2018). To increase the relevance of our model for recorded data, we matched a large number of biological parameters of our simulation to a recording of the neural population response of two macaques to a visually cued delayed grasping task. Despite not simulating individual connections between neurons, our simulation could accurately resemble key properties of recorded neural population responses (see Results 3.1).

We used the model to evaluate the performance of principal component analysis for a very wide range of realistic task and recording conditions. Thereby, we focused on three important aspects: The first aspect is the estimation of the intrinsic and extrinsic dimensionalities of the manifold of population activity from the results of PCA. For this purpose, we varied the numbers of trials per condition, numbers of conditions and numbers of neurons for two widely used tasks in the community: the reach-to-grasp task and the center-out reach task.

We first examined, whether the spectrum of variance captured per principal component could be used to estimate the dimensionality of the set of latent variables. However, for all sets of parameters, the spectrum of variance transitioned smoothly from early components dominated by LV-related variance to later components dominated by noise; This transition made the estimation of the dimensionality of the set of latent variables difficult.

The second aspect was the extraction of the manifold by the low-dimensional representation obtained from the results of PCA. We examined, how well principal components could capture individual subspaces under different task and recording conditions. We calculated the variance captured by the first N PCs for varying numbers of PCs N, trials per condition, numbers of conditions and numbers of neurons for both tasks. Surprisingly, only the number of trials per condition affected the amount of variance captured of latent variables. In contrast, the numbers of neurons and conditions had little effect on the amount of variance captured per LV. We varied the number of principal components N that had to be considered until further PCs did not capture any significant task-related activity. By this criterion, the number of PCs needed was approximately 1.5 to 2 times as high as the number of latent variables, and even then only small amounts of variance were captured of some LVs. Very high numbers of PCs were required to capture a significant amount of the variance of every single LV. Additionally, LVs that had overlapping temporal profiles, which is expected for many task designs (Dann, 2017; Murray et al., 2017), were only partially captured by even large numbers of PCs. Furthermore, the typically used criterion for the selection of the number of PCs to be retained is the percentage of cumulative total variance captured by the first PCs. This criterion proved unreliable in our simulation; For the typically used thresholds of variance captured, very little variance was captured from individual LV or even entire subspaces.

Finally, we examined whether and to what degree individual PCs were mixtures of individual LVs and the Poisson noise of the spiking process. If PCs are random mixtures of the ground truth LVs or subspaces, no direct conclusions can be drawn from individual PCs, or combinations of PCs, and only the combined space of many PCs should be considered. We found that individual PCs were indeed random mixtures of latent variables and noise. We could reduce the contribution of noise to individual PCs by increasing the number of trials per condition, but not by increasing the number of conditions or neurons. More importantly, and regardless of increasing numbers of trials per conditions, numbers of neurons or task conditions, individual PCs captured mixtures of variance from different subspaces. Therefore, individual PCs, or combinations of PCs should not be interpreted as representations of latent variables. Therefore, these results suggest that if PCA is used as a dimensionality reduction method, only the combined space of all retained principal

components should be used to draw conclusions about the structure of neural population response.

4.2 Simulation of fronto-parietal activity in response to a behavioral task

We systematically evaluated the accuracy and performance of PCA in capturing the underlying response structure of simulated spiking neural population activity. To ensure that conclusions drawn from the evaluation of PCA on a simulation are relevant for recorded population activity, we matched a large number of biological parameters of the simulation to recorded population activity. In this section, we will compare our model to other relevant models of neural population activity that were used in literature.

4.2.1 Overview of models of neural activity used in literature

Recurrent neural networks (RNN) are one of the most frequently used models of neural population activity (Mante et al., 2013; Sussillo, 2014; Sussillo et al., 2015; Nayebi et al., 2018; Michaels et al., 2020). Therefore, we should compare RNNs to our model. RNNs can reproduce some dynamics of neural population response, but they do not reproduce important population features (Bengio et al., 1994; Sussillo, 2014): RNN are most often non-spiking firing rate models, the firing rates range between -1 and 1 (or 0 and 1) and the distributions of firing rates, firing rate variances and LV contributions are different from those observed in recordings. Our model accurately resembles features of recorded neural population data. Thus, our model is more suitable than RNN-based models for the goal of this thesis.

Other studies have examined the performance of PCA or related methods on simulations before; we compare the models used in these studies to our model. Wood and McCarthy (Wood and McCarthy, 1984) and Beauducel and Debener (Beauducel and Debener, 2003) studied the misallocation of variance between principal components on a model of event-related potentials. Möcks (Möcks and Joachim, 1986) studies the effects of temporal delay on the results of PCA on similar models. The models of event-related potentials are very simple and have little relevance to real dynamics of spiking population activity; Lakshmanan et al. (Lakshmanan et al., 2015) examines two dimensionality reduction method related to PCA, called GPFA (Yu et al., 2009) and TD-GPFA, and applies them to a model of population activity to evaluate the performance of the two dimensionality methods in the presence of

temporal delays. The model used by Lakshmanan et al. is a non-spiking model of population activity that does not recreate important features of real recordings of population activity.

All the models presented so far are too simple to be relevant to spiking neural population activity, and should therefore not be used to draw detailed conclusions about the latter. Our model on the other hand accurately reproduces multiple features of spiking population activity and the results are applicable to real recordings. This thesis is therefore, to my knowledge, the first study to use a model of spiking population activity with high relevance to real recordings. We examine the performance of PCA when applied to recordings of spiking neural population activity, in detail and under realistic conditions.

4.2.2   The ground truth structure of simulated population activity

The ground truth structure of simulated population activity in our simulation is a low-dimensional set of orthogonal latent variables. A low-dimensional population structure is in accordance with the result of many recent studies that argue for a low dimensionality of the underlying structure of neural population response (Churchland and Shenoy, 2007; Churchland et al., 2010b; Mante et al., 2013; Kaufman et al., 2014b; Sadtler et al., 2014; Gallego et al., 2017). Other publications suggest a population structure based on individually-tuned neurons, in accordance with the representational view (Georgopoulos et al., 1984; Kalaska et al., 1989; Yuste, 2015) or a high-dimensional manifold with a power-law distributed variance spectrum (Stringer et al., 2018, 2019). However, PCA is used as a dimensionality reduction method to extract a linear, low-dimensional representation from the neural population. Therefore, a low-dimensional set of LVs is the best-suited ground truth for the simulated population response.

Another important feature of our simulation are the temporal delays between latent variables and neurons. The presence of temporal delays has been described in literature before (Baumann et al., 2009; Fluet et al., 2010; Westendorff et al., 2010). In the presence of temporal delays, individual LVs in our simulation occupy more than one dimension in the space of neural population activity (fig. 16). Because of the increased dimensionality, the manifold of neural population activity occupies dimensions outside of the subspaces of LV-related activity. The distribution of temporal delays was matched to the recording using Procrustes analysis; the distribution of temporal delays estimated that way results in the

best fit between recorded and simulated datasets for different simulated population sizes, and the spectra of variance between simulation and recording are similar. The conclusion to be drawn here is, that the manifold of population activity also occupies further dimensions outside of the subspaces of LV-related activity. However, an increased dimensionality due to the presence of temporal delays is not the only possible explanation for the dimensionality of the recorded population response; The variance could also be caused by latent variables not considered in our model, for example low frequent variations of overall firing rates due to global changes in attention, satiation or other factors. Nonetheless, the inclusion of temporal delays in the simulation results in a better fit between simulated and recorded population activity.

In summary, we designed our model and matched a large number of population features with the specific goal of the evaluation of linear dimensionality reduction methods. The high relevance of our model to real recordings of population activity allowed us to evaluate the performance of PCA under realistic conditions and draw conclusions about the applicability of PCA to neuronal recordings.

4.3 Dimensionality of population response

The true dimensionality of recordings of neural population response, specifically whether the underlying manifold of latent variable activity is low-dimensional or not, is still controversially debated (Ganguli and Sompolinsky, 2012; Gallego et al., 2017; Stringer et al., 2018, 2019). There are different definitions of the dimensionality of the manifold; the intrinsic dimensionality is given by the smallest number of latent variables needed to uniquely define a point or a trajectory on the manifold. It can be argued that the 7-LV reach-to-grasp task has an intrinsic dimensionality of one: there are two distinct, stereotypic trajectories that correspond to the two grip types, and distinguishing between the two grip types is sufficient to define the trajectory during a trial. However, we can discriminate additional dimensions from external information; The recorded response to the 7-LV task occupies three different subspaces over the course of a trial, which coincided with three different task epochs. Each task epoch is related to a function of the fronto-parietal network: visual processing (Murata et al., 1997, 2000; Janssen and Scherberger, 2015), movement preparation (Murata et al., 1996; Baumann et al., 2009; Fluet et al., 2010) and

movement execution (Gallese et al., 1994; Fogassi et al., 2001; Fluet et al., 2010). Each subspace is spanned by one linear condition-dependent and one condition-independent latent variable, plus one extra latent variable during movement execution that is related to movement onset (Kaufman et al., 2014b; Sussillo et al., 2015; Michaels et al., 2018; Ames and Churchland, 2019).

The extrinsic dimensionality on the other hand is defined as the number of dimensions in the space of population activity occupied by the manifold. The extrinsic dimensionality is often estimated from the spectrum of variance of the dataset (Jolliffe, 2002; Cisek, 2006; Cunningham and Yu, 2014; Stringer et al., 2018, 2019). According to recent studies (Stringer et al., 2018, 2019), a high number of dimensions is needed to capture task-related variance; In the context of the recordings examined in these studies, the manifold of neural population activity appears not to be contained in the first few components of the spectrum, but is rather high-dimensional.

Our model of spiking population activity reveals two effects that confounded the estimation of dimensionality from the spectrum of variance. The first effect is the presence of temporal delays between individual neurons and LVs: as mentioned in the previous section, the manifold of population activity also occupies further dimensions outside of the subspaces of LV-related activity in the presence of temporal delays. The second effect is the Poisson noise of the spiking process: If we increase Poisson noise in the simulated response, the variance increases across all dimensions (results, Fig. 16d). Random noise correlations are sorted towards the front of the spectrum where they form a bump in the spectrum of noise-related variance (Result figs. 18, 24a, purple) at the transition from LV-related variance-dominated to noise-dominated components. The overall increase in variance can mask the spurious components caused by temporal delays, as can be seen in fig. 8e. These two effects result in smooth transitions in the spectrum of variance, from LV-dominated components, to mixtures of LV-related variance and noise, to noise-dominated components (Figs 18, 24a). The spectra of variance are smooth for all numbers of trials per condition, neurons in the simulated population and conditions, and there is no fall-off point, or "knee" in the spectrum that would reveal the true dimensionality of the set of latent variables. There is also no saturation point, after which additional dimensions would not capture additional variance; In literature, such spectra that fulfill these conditions are said to correspond to a

high-dimensional manifold (Cattell, 1966; Jolliffe, 2002; Stringer et al., 2019). It is not possible to estimate the dimensionality of the set of latent variables from the shape of the spectrum of variance.

Another commonly used way of estimating the dimensionality of a dataset from the spectrum of variance uses the percentage of variance captured by the first N principal components as a criterion (Jolliffe, 2002; Cunningham and Yu, 2014; Stringer et al., 2018, 2019). This is one of the easiest methods of estimating the dimensionality of the dataset from the spectrum of variance, other methods like scree plot (Cattell, 1966) or parallel analysis (Horn, 1965) entail additional calculations. For different commonly-used variance percentages, the number of principal components varied strongly with the number of trials per condition (Results, fig. 20). This criterion, too, could not reliably estimate the dimensionality of the set of latent variables.

In summary, the combined effects of temporal delays and Poisson noise makes the spectrum of variance unsuitable for the estimation of the dimensionality of the set of latent variables. The spectrum of variance is not a measure that is specific to PCA. It is a part of all singular value decomposition-based methods and the results that concern the spectrum of variance are likely to be relevant for all SVD-based dimensionality reduction methods.

4.4   Extraction of dimensions of task-related activity

We also have investigated how many principal components are necessary to extract the dimensions of the manifold. A sufficient number of principal components is reached, when further components do not capture significant LV-related variance (Stringer et al., 2019). For a realistic number of 20 trials per condition, we need approximately 1.5 to 2 times as many principal components as there are latent variables (~14 for the 7-LV task. ~20 for the 10-LV task) to fulfill this criterion (Results figs. 20a and e). In consequence, underestimating the number of PCs will cause very little variance to be captured from individual LVs or even entire subspaces.

Additionally, the variance of the movement condition-independent LV was not captured well by a small number of PCs; the activity of that LV was rather distributed across all principal components. This does not mean that the activity of that LV is truly distributed

across all dimensions of the space of neural population activity in a non-linear manner; PCA simply failed to extract the dimension that contained most of the activity related to that LV.

For 20 or more trials per condition, neither the number of task conditions nor the number of neurons had any significant effect on the capture of variance by the first N principal components. Increasing the number of trials from 20 to 100 or even a very generous 10 000 trials per condition improved the capture of LVs that explained little variance, but did not improve the capture of other LVs. In conclusion, a large number of principal components is needed to extract a sufficient portion of task-related variance from the high-dimensional space of neural population activity, and even then, PCA is not guaranteed to extract all subspaces equally well. The activity that occupies the low-dimensional representation spanned by PCs should be considered a biased or incomplete representation of the task-related population response.

4.5   Mixing of latent variable activity in principal components

Previous studies of PCA have described the misallocation of variance between principal components in the context of event-related potentials (Wood and McCarthy, 1984; Beauducel and Debener, 2003). This thesis is the first study to examine misallocation, or mixing, between principal components in detail in the context of spiking neural population activity. We found, that the activity captured by individual principal components consist of mixtures of activity attributable to individual LVs and Poisson noise (Results, figs 21, 22, 23 and 24c and d). While the portion of variance attributable to Poisson noise could be reduced by increasing the number of trials per condition, mixing between latent variables did not change for any set of parameters.

Mixing between latent variables is not completely random; A theoretical investigation of mixing (Results 3.2.4) shows that mixing is driven by the covariance, or correlation, between temporal profiles of LV activity. Principal components are sorted by the amount of variance they explain: Early PCs that contain mixtures of LV activity likely contain LV activities that are positively correlated, increasing the total variance of mixed LV activities. This is illustrated in Fig. 20d: The principal component is of higher amplitude than the constituent LVs, and the LVs are positively correlated. By extension, later PCs contain negatively correlated mixtures of LVs. This has an important consequence for the extraction of dimensions of task-related

activity, as described in the previous section: If the number of PCs is underestimated, PCA will retain components that contain correlated LV activity. Thus, the activity of latent variables in the low-dimensional representation obtained from PCA is biased to be more strongly correlated than the true latent variables.

In conclusion, individual PCs, or even combinations of PCs should not be interpreted as representations of latent variables. The low-dimensional representation obtained from PCA should be regarded as a biased representation of the task-related population response, that does not accurately reflect the true variances or the correlations between temporal profiles of the set of latent variables.

4.6 Outlook

We have presented a highly accurate model of the spiking population activity and used it to evaluated the performance and accuracy of PCA as a dimensionality reduction method for recordings of neural population activity. This thesis has highlighted several effects that make the interpretation of the results of the application of PCA to spiking recordings of neural population activity difficult.

PCA is the standard singular value decomposition-based dimensionality reduction method, and many results of this paper are likely relevant for other, SVD-based dimensionality reduction methods, such as dPCA (Brendel et al., 2011; Kobak et al., 2016), factor analysis (Churchland et al., 2010a; Cunningham and Yu, 2014), and maybe even to dimensionality reduction methods in general. Next, we should use the model of spiking population activity presented in this thesis to evaluate further dimensionality reduction methods. We should examine, whether other methods share the issues of PCA. We expect in particular the estimation of dimensionality to be difficult regardless of the method used. The high dimensionality of the manifold was a feature of the structure of neural population response, rather than a result of PCA. On the other hand, the dimensionality reduction method used by Dann (Dann, 2017) that was described in the methods part of this thesis (Methods 2.1.2) proves, that supervised dimensionality reduction methods can perform the capture of individual, unmixed LV components more successfully than PCA did. Our model can be used to evaluate these, and other, dimensionality reduction methods.

An important effect examined in this thesis was the presence of temporal delays between latent variables and individual neurons. Temporal delays are known to exist in real recordings of neural population activity (Westendorff et al., 2010; Michaels et al., 2016). In our simulation, temporal delays were shown to be the main cause of the increased dimensionality of the manifold of neural population response and complicated the estimations of dimensionality from the spectrum of variance. We can assume that temporal delays play the same role in recorded data, but it would be premature to assume that it is the main complicating effect in real recordings of neural population activity. Therefore, real recordings need to be studied to determine, whether temporal delays are the main factor that causes the perceived high dimensionality of the manifold of real recordings, or whether high dimensionality can be explained by other factors that were not part of our model.

It could be possible to reduce or remove these effects from recordings of neural populations. The most impactful effect is the presence of temporal delay between LVs and neurons. A method for removal or mitigation of temporal delays could reduce the complexity of a recorded neural population activity. Lakshmanan et al presented time-delay gaussian-process factor analysis (TD-GPFA) (Lakshmanan et al., 2015). TD-GPFA is a modification of GPFA (Yu et al., 2009) that explicitly models temporal delays. Another important feature of spiking neural population activity is the Poisson noise of the spiking process. Poisson noise can be reliably reduced by increasing the number of trials per condition, which is an increased demand on the experimental procedure. The final feature is the correlation between the temporal profiles of individual LVs. Correlation between LV profiles causes PCs to contain mixtures of activities of different LVs. It can be avoided by de-coupling individual subspaces and identifying the structure of single subspaces in isolation. This would require tasks designed specifically to engage individual subspaces.

We have presented a highly accurate model of spiking neural population activity. Based on this model, we evaluated the performance and accuracy of PCA when applied to neural population activity. As the next step, we will make the simulation publicly available, in the form of a flexible toolbox for the evaluation of dimensionality reduction methods.

Discussion

5 References

Afshar A, Santhanam G, Byron MY, Ryu SI, Sahani M, Shenoy K V (2011) Single-trial neural correlates of arm movement preparation. Neuron 71:555–564.

Ames KC, Churchland MM (2019) Motor cortex signals corresponding to the two arms are shared across hemispheres, mixed among neurons, yet partitioned within the population response. bioRxiv:1–45.

Ames KC, Ryu SI, Shenoy K V. (2019) Simultaneous motor preparation and execution in a last-moment reach correction task. Nat Commun 10:1–13 Available at: http://dx.doi.org/10.1038/s41467-019-10772-2.

Baumann MA, Fluet M-C, Scherberger H (2009) Context-specific grasp movement representation in the macaque anterior intraparietal area. J Neurosci 29:6436–6448.

Beauducel A, Debener S (2003) Misallocation of variance in event-related potentials: Simulation studies on the effects of test power, topography, and baseline-to-peak versus principal component quantifications. J Neurosci Methods 124:103–112.

Bengio Y, Simard P, Frasconi P (1994) Learning long-term dependencies with gradient descent is difficult. IEEE Trans Neural Networks 5:157–166.

Bollimunta A, Totten D, Ditterich J (2012) Neural Dynamics of Choice: Single-Trial Analysis of Decision-Related Activity in Parietal Cortex. J Neurosci 32:12684–12701 Available at: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.5752-11.2012.

Boots B, Gordon G (2012) Two-manifold problems with applications to nonlinear system identification. arXiv Prepr arXiv12064648.

Brendel W, Romo R, Machens CK (2011) Demixed principal component analysis. Adv Neural Inf Process Syst 24:2654–2662.

Brown SL, Joseph J, Stopfer M (2005) Encoding a temporally structured stimulus with a temporally structured neural representation. Nat Neurosci 8:1568–1576.

Brown TG (1914) On the nature of the fundamental activity of the nervous centres; together

References

with an analysis of the conditioning of rhythmic activity in progression, and a theory of the evolution of function in the nervous system. J Physiol 48:18–46.

Brown TG (1915) On the activities of the central nervous system of the un-born fœtus of the cat; with a discussion of the question whether progression (walking, etc.) is a "learnt" complex. J Physiol 49:208–215 Available at: https://onlinelibrary.wiley.com/doi/10.1113/jphysiol.1915.sp001704.

Buesing L, Macke JH, Sahani M (2012) Learning stable, regularised latent models of neural population dynamics. Netw Comput Neural Syst 23:24–47 Available at: https://www.tandfonline.com/doi/full/10.3109/0954898X.2012.677095.

Buzsáki G (2010) Neural Syntax: Cell Assemblies, Synapsembles, and Readers. Neuron 68:362–385.

Buzsáki G, Mizuseki K (2014) The log-dynamic brain: How skewed distributions affect network operations. Nat Rev Neurosci 15:264–278 Available at: http://dx.doi.org/10.1038/nrn3687.

Byron MY, Afshar A, Santhanam G, Ryu SI, Shenoy K V, Sahani M (2006) Extracting dynamical structure embedded in neural activity. In: Advances in neural information processing systems, pp 1545–1552.

Carrillo-Reid L, Tecuapetla F, Tapia D, Hernández-Cruz A, Galarraga E, Drucker-Colin R, Bargas J (2008) Encoding network states by striatal cell assemblies. J Neurophysiol 99:1435–1450.

Cattell RB (1966) The Scree Test For The Number Of Factors. Multivariate Behav Res 1:245–276 Available at: https://doi.org/10.1207/s15327906mbr0102_10.

Chemla S, Chavane F (2012) Voltage-sensitive dye imaging. In: Handbook of Neural Activity Measurement (Brette R, Destexhe A, eds), pp 327–361. Cambridge: Cambridge University Press. Available at: https://www.cambridge.org/core/product/identifier/CBO9780511979958A067/type/book_part.

Chethan Pandarinath, K Cora Ames, Abigail A Russo, Ali Farshchian, Lee E Miller, Eva L Dyer JCK (2018) Latent Factors and Dynamics in Motor Cortex and Their Application to Brain–Machine Interfaces. :36–50.

Churchland MM et al. (2010a) Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nat Neurosci 13:369–378 Available at: http://www.nature.com/articles/nn.2501.

Churchland MM, Afshar A, Shenoy K V. (2006) A Central Source of Movement Variability. Neuron 52:1085–1096 Available at: https://linkinghub.elsevier.com/retrieve/pii/S0896627306008713.

Churchland MM, Cunningham JP (2014) A dynamical basis set for generating reaches. In: Cold Spring Harbor Symposia on Quantitative Biology, pp 67–80.

Churchland MM, Cunningham JP, Kaufman MT, Foster JD, Nuyujukian P, Ryu SI, Shenoy K V. (2012) Neural population dynamics during reaching. Nature 487:51–56 Available at: http://www.nature.com/articles/nature11129.

Churchland MM, Cunningham JP, Kaufman MT, Ryu SI, Shenoy K V. (2010b) Cortical Preparatory Activity: Representation of Movement or First Cog in a Dynamical Machine? Neuron 68:387–400 Available at: http://dx.doi.org/10.1016/j.neuron.2010.09.015.

Churchland MM, Shenoy K V. (2007) Temporal complexity and heterogeneity of single-neuron activity in premotor and motor cortex. J Neurophysiol 97:4235–4257.

Cisek P (2006) Preparing for Speed . Focus on "Preparatory Activity in Premotor and Motor Cortex Reflects the Speed of the Upcoming Reach." J Neurophysiol 96:2842–2843 Available at: https://www.physiology.org/doi/10.1152/jn.00857.2006.

Cunningham JP, Yu BM (2014) Dimensionality reduction for large-scale neural recordings. Nat Neurosci 17:1500–1509 Available at: http://www.nature.com/articles/nn.3776.

Dann B (2017) Encoding, coordination, and decision making in the primate fronto-parietal grasping network.

References

Dann B, Michaels JA, Schaffelhofer S, Scherberger H (2016) Uniting functional network topology and oscillations in the fronto-parietal single unit network of behaving primates. Elife 5:1–27.

Dasgupta S, Gupta A (2003) An elementary proof of a theorem of Johnson and Lindenstrauss. Random Struct Algorithms 22:60–65 Available at: https://onlinelibrary.wiley.com/doi/10.1002/rsa.10073.

Desimone R, Albright T, Gross C, Bruce C (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. J Neurosci 4:2051–2062 Available at: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.04-08-02051.1984.

Ebitz RB, Hayden BY (2021) The population doctrine revolution in cognitive neurophysiology. Available at: http://arxiv.org/abs/2104.00145.

Einevoll GT, Linden H, Tetzlaff T, Leski S, Pettersen KH (2013) - Local Field Potentials: Biophysical Origin and Analysis. In: Principles of Neural Coding, pp 54–77. CRC Press. Available at: https://www.taylorfrancis.com/books/9781439853313/chapters/10.1201/b14756-7.

Elsayed GF, Lara AH, Kaufman MT, Churchland MM, Cunningham JP (2016) Reorganization between preparatory and movement population responses in motor cortex. Nat Commun:13239 Available at: http://www.nature.com/doifinder/10.1038/ncomms13239.

Evarts E V (1968) Relation of pyramidal tract activity to force exerted during voluntary movement. J Neurophysiol 31:14–27 Available at: https://www.physiology.org/doi/10.1152/jn.1968.31.1.14.

Evarts E V (1969) Activity of pyramidal tract neurons during postural fixation. J Neurophysiol 32:375–385 Available at: https://www.physiology.org/doi/10.1152/jn.1969.32.3.375.

Fetz EE (1994) Are movement parameters recognizably coded in the activity of single neurons? In: Movement Control, pp 77–88. Cambridge University Press. Available at: https://www.cambridge.org/core/product/identifier/CBO9780511529788A013/type/b

ook_part.

Fetz EE, Cheney PD (1980) Postspike facilitation of forelimb muscle activity by primate corticomotoneuronal cells. J Neurophysiol 44:751–772 Available at: https://www.physiology.org/doi/10.1152/jn.1980.44.4.751.

Fluet M-C, Baumann MA, Scherberger H (2010) Context-specific grasp movement representation in macaque ventral premotor cortex. J Neurosci 30:15175–15184.

Fogassi L, Gallese V, Buccino G, Craighero L, Fadiga L, Rizzolatti G (2001) Cortical mechanism for the visual guidance of hand grasping movements in the monkey: A reversible inactivation study. Brain 124:571–586.

Freedman DJ, Assad JA (2006) Experience-dependent representation of visual categories in parietal cortex. Nature 443:85–88.

Gail A, Andersen RA (2006) Neural dynamics in monkey parietal reach region reflect context-specific sensorimotor transformations. J Neurosci 26:9376–9384.

Gallego JA, Perich MG, Chowdhury RH, Solla SA, Miller LE (2018a) A stable , long-term cortical signature underlying consistent behavior. bioRxiv:1–33.

Gallego JA, Perich MG, Miller LE, Solla SA (2017) Neural Manifolds for the Control of Movement. Neuron 94:978–984 Available at: http://dx.doi.org/10.1016/j.neuron.2017.05.025.

Gallego JA, Perich MG, Naufel SN, Ethier C, Solla SA, Miller LE (2018b) Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. Nat Commun 9:1–13 Available at: http://dx.doi.org/10.1038/s41467-018-06560-z.

Gallese V, Murata A, Kaseda M, Niki N, Sakata H (1994) Deficit of hand preshaping after muscimol injection in monkey parietal cortex. Neuroreport An Int J Rapid Commun Res Neurosci.

Ganguli S, Sompolinsky H (2012) Compressed Sensing, Sparsity, and Dimensionality in Neuronal Information Processing and Data Analysis. Annu Rev Neurosci 35:485–508

References

Available at: http://www.annualreviews.org/doi/10.1146/annurev-neuro-062111-150410.

Gao P, Trautmann E, Yu BM, Santhanam G, Ryu SI, Shenoy K V., Ganguli S (2017) A theory of multineuronal dimensionality, dynamics and measurement. bioRxiv:214262 Available at: https://www.biorxiv.org/content/early/2017/11/05/214262.

Georgopoulos AP, Caminiti R, Kalaska JF (1984) Static spatial effects in motor cortex and area 5: Quantitative relations in a two-dimensional space. Exp Brain Res 54 Available at: http://link.springer.com/10.1007/BF00235470.

Georgopoulos AP, Kalaska JF, Caminiti R, Massey JT (1982) On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. J Neurosci 2:1527–1537.

Georgopoulos AP, Schwartz AB, Kettner RE (1986) Neuronal population coding of movement direction. Science (80-) 233:1416–1419.

Gerstein GL, Bedenbaugh P, Aertsen AMHJ (1989) Neuronal assemblies. IEEE Trans Biomed Eng 36:4–14 Available at: http://ieeexplore.ieee.org/document/16444/.

Grillner S (2006) Biological Pattern Generation: The Cellular and Computational Logic of Networks in Motion. Neuron 52:751–766 Available at: https://linkinghub.elsevier.com/retrieve/pii/S0896627306009020.

Hebb DO (1949) The Organization of Behavior. Wiley: New York.

Hodgkin AL, Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. J Physiol 117:500–544 Available at: https://onlinelibrary.wiley.com/doi/10.1113/jphysiol.1952.sp004764.

Horn JL (1965) A rationale and test for the number of factors in factor analysis. Psychometrika 30:179–185 Available at: https://doi.org/10.1007/BF02289447.

Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J Physiol 160:106–154 Available at:

https://onlinelibrary.wiley.com/doi/10.1113/jphysiol.1962.sp006837.

Janssen P, Scherberger H (2015) Visual Guidance in Control of Grasping. Annu Rev Neurosci 38:69–86 Available at: http://www.annualreviews.org/doi/10.1146/annurev-neuro-071714-034028.

Jolliffe IT (2002) Principal Component Analysis. Second Edition.

Jones LM, Fontanini A, Sadacca BF, Miller P, Katz DB (2007) Natural stimuli evoke dynamic sequences of states in sensory cortical ensembles. Proc Natl Acad Sci 104:18772–18777 Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.0705546104.

Jun JJ et al. (2017) Fully integrated silicon probes for high-density recording of neural activity. Nature 551:232–236 Available at: http://www.nature.com/articles/nature24636.

Kalaska J, Cohen D, Hyde M, Prud'homme M (1989) A comparison of movement direction-related versus load direction- related activity in primate motor cortex, using a two-dimensional reaching task. J Neurosci 9:2080–2102 Available at: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.09-06-02080.1989.

Katzner S, Nauhaus I, Benucci A, Bonin V, Ringach DL, Carandini M (2009) Local Origin of Field Potentials in Visual Cortex. Neuron 61:35–41 Available at: https://linkinghub.elsevier.com/retrieve/pii/S0896627308010064.

Kaufman MT, Churchland MM, Ryu SI, Shenoy K V. (2014a) Cortical activity in the null space: Permitting preparation without movement. Nat Neurosci 17:440–448 Available at: http://dx.doi.org/10.1038/nn.3643.

Kaufman MT, Churchland MM, Ryu SI, Shenoy K V. (2014b) Cortical activity in the null space: Permitting preparation without movement. Nat Neurosci.

Kaufman MT, Churchland MM, Ryu SI, Shenoy K V. (2015) Vacillation, indecision and hesitation in moment-by-moment decoding of monkey motor cortex. Elife 4:1–21.

Kaufman MT, Seely JS, Sussillo D, Ryu SI, Shenoy K V., Churchland MM (2016) The Largest

References

Response Component in the Motor Cortex Reflects Movement Timing but Not Movement Type. eNeuro.

Kemere C, Santhanam G, Yu BM, Afshar A, Ryu SI, Meng TH, Shenoy K V. (2008) Detecting Neural-State Transitions Using Hidden Markov Models for Motor Cortical Prostheses. J Neurophysiol 100:2441–2452 Available at: https://www.physiology.org/doi/10.1152/jn.00924.2007.

Kobak D, Brendel W, Constantinidis C, Feierstein CE, Kepecs A, Mainen ZF, Qi XL, Romo R, Uchida N, Machens CK (2016) Demixed principal component analysis of neural population data. Elife 5:1–36.

Kopell N, Ermentrout GB (2002) Mechanisms of Phase-Locking and Frequency Control in Pairs of Coupled Neural Oscillators* *Work partially supported by NIH grant RO1-MH47150 to NK and GBE, NSF grant 9706694 to NK and a NSF grant to GBE. We thank J. Ritt and S. Epstein for careful reading. In, pp 3–54 Available at: https://linkinghub.elsevier.com/retrieve/pii/S1874575X02800224.

Kraskov A, Dancause N, Quallo MM, Shepherd S, Lemon RN (2009) Corticospinal Neurons in Macaque Ventral Premotor Cortex with Mirror Properties: A Potential Mechanism for Action Suppression? Neuron 64:922–930 Available at: http://dx.doi.org/10.1016/j.neuron.2009.12.010.

Kulkarni JE, Paninski L (2007) Common-input models for multiple neural spike-train data. Netw Comput Neural Syst 18:375–407 Available at: https://www.tandfonline.com/doi/full/10.1080/09548980701625173.

Lakshmanan KC, Sadtler PT, Tyler-Kabara EC, Batista AP, Yu BM (2015) Extracting Low-Dimensional Latent Structure from Time Series in the Presence of Delays. Neural Comput 27:1825–1856 Available at: http://www.mitpressjournals.org/doi/10.1162/NECO_a_00759.

Lara AH, Elsayed GF, Zimnik AJ, Cunningham JP, Churchland MM (2018) Conservation of preparatory neural events in monkey motor cortex regardless of how movement is initiated. Elife 7:1–34.

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A (2001) Neurophysiological investigation of the basis of the fMRI signal. Nature 412:150–157 Available at: http://www.nature.com/articles/35084005.

Luppino G, Murata A, Govoni P, Matelli M (1999) Largely segregated parietofrontal connections linking rostral intraparietal cortex (areas AIP and VIP) and the ventral premotor cortex (areas F5 and F4). Exp Brain Res 128:181–187.

Machens CK, Romo R, Brody CD (2010) Functional, but not anatomical, separation of "what" and "when" in prefrontal cortex. J Neurosci 30:350–360.

Mante V, Sussillo D, Shenoy K V., Newsome WT (2013) Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature 503:78–84 Available at: http://dx.doi.org/10.1038/nature12742.

Michaels JA, Dann B, Intveld RW, Scherberger H (2015) Predicting Reaction Time from the Neural State Space of the Premotor and Parietal Grasping Network. J Neurosci.

Michaels JA, Dann B, Intveld RW, Scherberger H (2018) Neural Dynamics of Variable Grasp-Movement Preparation in the Macaque Frontoparietal Network. J Neurosci 38:5759–5773 Available at: http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.2557-17.2018.

Michaels JA, Dann B, Scherberger H (2016) Neural population dynamics during reaching are better explained by a dynamical system than representational tuning. PLoS Comput Biol 12:e1005175.

Michaels JA, Schaffelhofer S, Agudelo-Toro A, Scherberger H (2020) A goal-driven modular neural network predicts parietofrontal neural dynamics during grasping. Proc Natl Acad Sci U S A 117:32124–32135.

Möcks, Joachim (1986) The Influence of Latency Jitter in Principal Component Analysis of Event-Related Potentials. Psychophysiology 23:480–484 Available at: http://doi.wiley.com/10.1111/j.1469-8986.1986.tb00659.x.

Montijn JS, Meijer GT, Lansink CS, Pennartz CMA (2016) Population-Level Neural Codes Are

References

Robust to Single-Neuron Variability from a Multidimensional Coding Perspective. Cell Rep 16:2486–2498 Available at: http://dx.doi.org/10.1016/j.celrep.2016.07.065.

Murata A, Fadiga L, Fogassi L, Gallese V, Raos V, Rizzolatti G (1997) Object representation in the ventral premotor cortex (area F5) of the monkey. J Neurophysiol 78:2226–2230.

Murata A, Gallese V, Kaseda M, Sakata H (1996) Parietal neurons related to memory-guided hand manipulation. J Neurophysiol 75:2180–2186.

Murata A, Gallese V, Luppino G, Kaseda M, Sakata H (2000) Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. J Neurophysiol 83:2580–2601.

Murray JD, Bernacchia A, Roy NA, Constantinidis C, Romo R, Wang X-J (2017) Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. Proc Natl Acad Sci 114:394–399 Available at: http://www.pnas.org/lookup/doi/10.1073/pnas.1619449114.

Musallam S, Bak MJ, Troyk PR, Andersen RA (2007) A floating metal microelectrode array for chronic implantation. J Neurosci Methods 160:122–127 Available at: https://linkinghub.elsevier.com/retrieve/pii/S016502700600450X.

Musial PG, Baker SN, Gerstein GL, King EA, Keating JG (2002) Signal-to-noise ratio improvement in multiple electrode recording. J Neurosci Methods 115:29–43.

Nayebi A, Bear D, Kubilius J, Kar K, Ganguli S, Sussillo D, DiCarlo JJ, Yamins DLK (2018) Task-Driven Convolutional Recurrent Models of the Visual System. :1–14 Available at: http://arxiv.org/abs/1807.00053.

Ni AM, Ruff DA, Alberts JJ, Symmonds J, Cohen MR (2018) Learning and attention reveal a general relationship between population activity and behavior. Science (80-) 359:463–465 Available at: https://www.sciencemag.org/lookup/doi/10.1126/science.aao0284.

Nicolelis MA., Ghazanfar AA, Faggin BM, Votaw S, Oliveira LM. (1997) Reconstructing the Engram: Simultaneous, Multisite, Many Single Neuron Recordings. Neuron 18:529–537 Available at: https://linkinghub.elsevier.com/retrieve/pii/S0896627300802950.

95

Nicolelis MAL, Dimitrov D, Carmena JM, Crist R, Lehew G, Kralik JD, Wise SP (2003) Chronic, multisite, multielectrode recordings in macaque monkeys. Proc Natl Acad Sci 100:11041–11046 Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1934665100.

Okun M, Steinmetz NA, Cossell L, Iacaruso MF, Ko H, Barthó P, Moore T, Hofer SB, Mrsic-Flogel TD, Carandini M, Harris KD (2015) Diverse coupling of neurons to populations in sensory cortex. Nature 521:511–515.

Pachitariu M, Steinmetz N, Kadir S, Carandini M, Harris K (2016) Fast and accurate spike sorting of high-channel count probes with KiloSort. Adv Neural Inf Process Syst:4455–4463.

Pandarinath C, O'Shea DJ, Collins J, Jozefowicz R, Stavisky SD, Kao JC, Trautmann EM, Kaufman MT, Ryu SI, Hochberg LR (2018) Inferring single-trial neural population dynamics using sequential auto-encoders. Nat Methods 15:805–815.

Paninski L, Ahmadian Y, Ferreira DG, Koyama S, Rahnama Rad K, Vidne M, Vogelstein J, Wu W (2010) A new look at state-space models for neural data. J Comput Neurosci 29:107–126 Available at: http://link.springer.com/10.1007/s10827-009-0179-x.

Perin R, Berger TK, Markram H (2011) A synaptic organizing principle for cortical neuronal groups. Proc Natl Acad Sci 108:5419–5424 Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1016051108.

Petreska B, Yu BM, Cunningham JP, Santhanam G, Ryu S, Shenoy K V, Sahani M (2011) Dynamical segmentation of single trials from population neural data. Adv Neural Inf Process Syst 24:756–764.

Pfau D, Pnevmatikakis EA, Paninski L (2013) Robust learning of low-dimensional dynamics from large neural ensembles. In: Advances in Neural Information Processing Systems (Burges CJC, Bottou L, Welling M, Ghahramani Z, Weinberger KQ, eds). Curran Associates, Inc. Available at: https://proceedings.neurips.cc/paper/2013/file/47a658229eb2368a99f1d032c8848542-Paper.pdf.

References

Ponce-Alvarez A, Nacher V, Luna R, Riehle A, Romo R (2012) Dynamics of Cortical Neuronal
Ensembles Transit from Decision Making to Storage for Later Report. J Neurosci
32:11956–11969 Available at:
https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.6176-11.2012.

Quiroga RQ, Nadasdy Z, Ben-Shaul Y (2004) Unsupervised spike detection and sorting with
wavelets and superparamagnetic clustering. Neural Comput 16:1661–1687.

Raducanu BC, Yazicioglu RF, Lopez CM, Ballini M, Putzeys J, Wang S, Andrei A, Rochus V,
Welkenhuysen M, Helleputte N van, Musa S, Puers R, Kloosterman F, Hoof C van, Fiáth
R, Ulbert I, Mitra S (2017) Time Multiplexed Active Neural Probe with 1356 Parallel
Recording Sites. Sensors 17:2388 Available at: http://www.mdpi.com/1424-
8220/17/10/2388.

Raposo D, Kaufman MT, Churchland AK (2014) A category-free neural population supports
evolving demands during decision-making. Nat Neurosci 17:1784–1792 Available at:
http://www.nature.com/articles/nn.3865.

Rousche PJ, Normann RA (1998) Chronic recording capability of the Utah Intracortical
Electrode Array in cat sensory cortex. J Neurosci Methods 82:1–15 Available at:
https://linkinghub.elsevier.com/retrieve/pii/S0165027098000314.

Roweis ST, Saul LK (2000) Nonlinear dimensionality reduction by locally linear embedding.
Science (80-) 290:2323–2326.

Sadtler PT, Quick KM, Golub MD, Chase SM, Ryu SI, Tyler-Kabara EC, Yu BM, Batista AP
(2014) Neural constraints on learning. Nature 512:423–426 Available at:
http://dx.doi.org/10.1038/nature13665.

Schaffelhofer S, Sartori M, Scherberger H, Farina D (2015) Musculoskeletal Representation
of a Large Repertoire of Hand Grasping Actions in Primates. IEEE Trans Neural Syst
Rehabil Eng 23:210–220.

Schaffelhofer S, Scherberger H (2016) Object vision to hand action in macaque parietal,
premotor, and motor cortices. Elife 5:1–24.

Schwartz AB, Kettner RE, Georgopoulos AP (1988) Primate motor cortex and free arm movements to visual targets in three-dimensional space. I. Relations between single cell discharge and direction of movement. J Neurosci 8:2913–2927.

Schwartz JH, Kandel ER, Jessell TM, Siegelbaum SA, Hudspeth A (1991) Principles of neural science.

Scott SH (2004) Optimal feedback control and the neural basis of volitional motor control. Nat Rev Neurosci 5:532–545 Available at: http://www.nature.com/articles/nrn1427.

Seidemann E, Meilijson I, Abeles M, Bergman H, Vaadia E (1996) Simultaneously recorded single units in the frontal cortex go through sequences of discrete and stable states in monkeys performing a delayed localization task. J Neurosci 16:752–768 Available at: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.16-02-00752.1996.

Selverston A (1999) Chapter 22 General Principles of Rhythmic Motor Pattern Generation Derived from Invertebrate CPGs. In, pp 247–257 Available at: https://linkinghub.elsevier.com/retrieve/pii/S0079612308628617.

Shenoy K V., Sahani M, Churchland MM (2013) Cortical control of arm movements: A dynamical systems perspective. Annu Rev Neurosci 36:337–359.

Sherrington CS (1906) Observations on the scratch-reflex in the spinal dog. J Physiol 34:1–50 Available at: https://onlinelibrary.wiley.com/doi/10.1113/jphysiol.1906.sp001139.

Siegel M, Buschman TJ, Miller EK (2015) Cortical information flow during flexible sensorimotor decisions. Science (80-) 348:1352–1355.

Singer W (1999) Neuronal Synchrony: A Versatile Code for the Definition of Relations? Neuron 24:49–65 Available at: https://linkinghub.elsevier.com/retrieve/pii/S0896627300808211.

Smith AC, Brown EN (2003) Estimating a State-Space Model from Point Process Observations. Neural Comput 15:965–991 Available at: https://direct.mit.edu/neco/article/15/5/965-991/6734.

References

Stopfer M, Jayaraman V, Laurent G (2003) Intensity versus identity coding in an olfactory system. Neuron 39:991–1004.

Stosiek C, Garaschuk O, Holthoff K, Konnerth A (2003) In vivo two-photon calcium imaging of neuronal networks. Proc Natl Acad Sci 100:7319–7324 Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1232232100.

Stringer C, Pachitariu M, Steinmetz N, Carandini M, Harris KD (2019) High-dimensional geometry of population responses in visual cortex. Nature Available at: http://www.nature.com/articles/s41586-019-1346-5.

Stringer C, Pachitariu M, Steinmetz N, Reddy C, Carandini M, Harris KD (2018) Spontaneous behaviors drive multidimensional , brain-wide neural activity. bioRxiv:1–26.

Sussillo D (2014) Neural circuits as computational dynamical systems. Curr Opin Neurobiol 25:156–163 Available at: http://dx.doi.org/10.1016/j.conb.2014.01.008.

Sussillo D, Churchland MM, Kaufman MT, Shenoy K V. (2015) A neural network that finds a naturalistic solution for the production of muscle activity. Nat Neurosci 18:1025–1033 Available at: http://dx.doi.org/10.1038/nn.4042.

Sussillo D, Jozefowicz R, Abbott LF, Pandarinath C (2016) Lfads-latent factor analysis via dynamical systems. arXiv Prepr arXiv160806315.

Tanaka K (1996) Inferotemporal cortex and object vision. Annu Rev Neurosci 19:109–139.

Tenenbaum JB, De Silva V, Langford JC (2000) A global geometric framework for nonlinear dimensionality reduction. Science (80-) 290:2319–2323.

Townsend BR, Subasi E, Scherberger H (2011) Grasp Movement Decoding from Premotor and Parietal Cortex. J Neurosci 31:14386–14398 Available at: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.2451-11.2011.

Tsodyks M (1999) Linking Spontaneous Activity of Single Cortical Neurons and the Underlying Functional Architecture. Science (80-) 286:1943–1946 Available at: https://www.sciencemag.org/lookup/doi/10.1126/science.286.5446.1943.

Westendorff S, Klaes C, Gail A (2010) The cortical timeline for deciding on reach motor goals. J Neurosci 30:5426–5436.

Whittingstall K, Logothetis NK (2013) Physiological foundations of neural signals. In: Principles of Neural Coding.

Wood CC, McCarthy G (1984) Principal component analysis of event-related potentials: Simulation studies demonstrate misallocation of variance across components. Electroencephalogr Clin Neurophysiol Evoked Potentials 59:249–260 Available at: https://linkinghub.elsevier.com/retrieve/pii/0168559784900649.

Wood F, Black MJ, Vargas-Irwin C, Fellows M, Donoghue JP (2004) On the Variability of Manual Spike Sorting. IEEE Trans Biomed Eng 51:912–918 Available at: http://ieeexplore.ieee.org/document/1300782/.

y Cajal SR (1888) Estructura de los centros nerviosos de las aves.

Yu BM, Cunningham JP, Santhanam G, Ryu SI, Shenoy K V., Sahani M (2009) Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity (Journal of Neurophysiology (2009) 102, (614-635) DOI: 10.1152/jn.90941.2008). J Neurophysiol 102:2008.

Yuste R (2015) From the neuron doctrine to neural networks. Nat Rev Neurosci 16:487–497 Available at: http://dx.doi.org/10.1038/nrn3962.

Yuste R, MacLean JN, Smith J, Lansner A (2005) The cortex as a central pattern generator. Nat Rev Neurosci 6:477–483 Available at: http://www.nature.com/articles/nrn1686.

References

6 Curriculum Vitae

Andrej Filippow

Deutsches Primatenzentrum GmbH

Kellnerweg 4

37077 Goettingen, Germany

+49 1577 159 37 88

afilippow@dpz.eu

Born: 17.08.1988 in Jekaterinenburg, Russia

Citizenship: German

Education:

PhD: 2017 Present

Neurobiology Lab, German Primate Center, Goettingen

Master: 2013 –2015

Master of Physics, Specialisation Biophysics and Complex Systems at

Georg- August-University Goettingen

Bachelor: 2009 –2013

Bachelor of Physics, Georg-August-University Goettingen

2008 Abitur, Käthe-Kollwitz-Gymnasium Hannover.

Employment: 2016-2017

Research Assistant, Neurobiology Lab, German Primate Center

2017 Present

PhD: Neurobiology Lab, German Primate Center, Goettingen

Curriculum Vitae

Oral Presentations:	A. Filippow[1,2], B. Dann[1],H. Scherberger[1,2]

Neural population response dimensions identified with PCA contain mixed content and are only meaningful in case of high signal-to-noise ratio

Annual Meeting of the Society for Neuroscience (2019) Chicago

A. Filippow

"Adaptation in Neural Population dynamics during a Neural Interface Experiment"

Sensory and Motor Neuroscience Retreat (2018) Göttingen

Poster Presentations	A. Filippow, B. Dann,H. Scherberger

"A systematic evaluation of the usability of PCA on neuronal activity."

Primate Neurobiology Meeting (2019) Göttingen

A. Filippow, B. Dann,H. Scherberger

Simulation-based evaluation of PCA application to neuronal population activity

Annual Meeting of the Society for Neuroscience (2018) San Diego

In-progress publications:	A. Filippow, H. Scherberger, B. Dann

Simulation-based evaluation of the applicability of PCA to recordings of neural population

Languages:	Russian native speaker

German native speaker

English high level

French intermediate

Latin basic